

# BIG DATA VALUE cPPP

ANNEX 2 to the Monitoring Report 2019-20

Detailed explanations of KPIs

Version 1.0. September 2021

## ABOUT THIS DOCUMENT

This document is part of the Big Data Value contractual Public-Private Partnership (BDV cPPP) Monitoring report 2019-20 and complements the information provided in the main report and Annex 1 to the main report.

Annex 2 Includes detailed information per KPI: methodology, response rate, detailed input and assessment. It also provides information about the specific questionnaires used to gather information and the list of stakeholders who provided input.

### **Version control:**

**Version 1.0** (September 2021) Annex 2 to the first official version of the Big Data Value cPPP Annual Monitoring Report 2019-20.

**Disclaimer:** *This document has been prepared by the Big Data Value Association (BDVA) and it reflects the views only of its authors.*

## Table of Contents

About this DOCUMENT.....	2
<b>1 INTRODUCTION.....</b>	<b>5</b>
<b>2 Mobilised Private Investments (I.1) .....</b>	<b>6</b>
2.1 Methodology .....	7
2.2 Responses rates .....	10
2.3 Overall results.....	11
<b>3 New skills and/or job profiles (I.2, II.8) .....</b>	<b>14</b>
3.1 KPI I.2 (Number of jobs created).....	14
3.1.1 Job creation .....	14
3.1.2 Job Profiles established / Skills created. ....	17
3.2 KPI II.8 Higher establishment availability of big data value creation skills development.....	26
3.2.1 Number of training programs established arising from cPPP .....	26
3.2.2 Number of European training programs involving 3 different disciplines arising from cPPP .....	28
3.2.3 Number of Master and Phd students involved in PPP projects .....	28
3.2.4 Number of dissemination events, seminars, conferences organised in cPPP projects (including number of participants) .....	29
<b>4 Impact of the BDV cPPP on SMEs (I.3, II.18) .....</b>	<b>29</b>
4.1 KPI I.3: Impact of a cPPP on SMEs (in Euros/Qualitative Analysis).....	29
4.1.1 Methodology .....	29
4.1.2 Map of responses and responses rates.....	30
4.1.3 Overall results: .....	31
4.2 KPI II.18: Participation and benefits for SMEs .....	33
<b>5 Innovations emerging from projects (I.4, II.7, II.17) .....</b>	<b>33</b>
5.1 KPI I.4 Number of significant innovations to market.....	33
5.2 New economic viable services of high societal value (KPI II.7) .....	172
5.3 New Systems and Technologies (KPI II.17).....	176
<b>6 Support major sectors and major domains by Big Data technologies and applications (II.13) .....</b>	<b>178</b>
<b>7 Experimentation (II.11, II.12, II.14).....</b>	<b>180</b>
7.1 II.11: Large Scale experiments conducted in cPPP projects and i-Spaces involving closed data.....	180
7.2 II.12: Uptake of BDV use cases and experiments.....	182
7.3 II.14: Amount of data that has been made available for experimentation (cPPP projects and i-Spaces)	189

<b>8</b>	<b>SRIA update and implementation (II.9, II.10)</b> .....	<b>193</b>
8.1	<i>KPI II.9: Ensure efficiency, transparency and openness of the cPPP's consultation process</i> .....	193
8.2	<i>KPI II.10: Ensure that technology progress is in line with multi-annual roadmap of SRIA</i> .....	193
<b>9</b>	<b>Technical Projects (II.4, II.15, II.16)</b> .....	<b>197</b>
9.1	<i>KPI II.4: Privacy and Security KPI</i> .....	197
9.3	<i>KPI II.15: Availability of metrics for measuring the quality, diversity and value of data assets</i> .....	199
9.4	<i>KPI II.16: Increase the speed of data throughput compared to 2014</i> .....	199
<b>10</b>	<b>Macro-Economic KPIS (KPI II.1, II.2, II.3, II.5)</b> .....	<b>201</b>
10.1	<i>KPI II.1: Market share of the European Union in the global Big Data Market</i> .....	202
	<b><i>KPI II.2: Growth of the number of Data Companies in the EU</i></b> .....	<b>207</b>
	<b><i>KPI II.3: Growth of Data Companies revenues in the EU</i></b> .....	<b>212</b>
	<b><i>KPI II.5: Increase of Data Professionals in the European Union</i></b> .....	<b>215</b>
<b>11</b>	<b>Contributions to environmental challenges (II.6, II.19, II.20, II.21)</b> .....	<b>220</b>
11.1	<i>Contribution to the reduction of energy use (KPI II.6)</i> .....	220
11.2	<i>Contribution to Energy use reduction (KPI II.19)</i> .....	221
11.3	<i>Contribution to the reduction of CO<sub>2</sub> (KPI II.19)</i> .....	222
11.4	<i>Contribution to the reduction of waste (KPI II.20)</i> .....	224
11.5	<i>Contribution to the reduction in the use of material resources (KPI II.21)</i> .....	224
<b>12</b>	<b>Standardisation activities (III.2)</b> .....	<b>225</b>
<b>13</b>	<b>QUESTIONNAIRES TO COLLECT INFORMATION</b> .....	<b>228</b>
13.1	<i>Questionnaire for PPP projects running in 2018</i> .....	228
13.2	<i>Technical priorities table for projects - Questionnaire</i> .....	235
13.3	<i>Significant Innovations to Market – Questionnaire</i> .....	236
13.4	<i>Investments questionnaire</i> .....	237
13.5	<i>Questionnaire for BDVA members</i> .....	239

# 1 INTRODUCTION

When referring to key performance indicators (KPIs) we use the following scheme: A Roman number refers to the annex in which the KPI is listed, i.e. I for common, II for specific KPIs and III for contribution to programme level KPIs. This is followed by an Arabic number indicating the number of the indicator in the respective table, e.g. KPI II.1 refers to the first KPI specified in the list of specific KPIs.

The list of KPIs for this cPPP, description, and target values are defined by the following documents:

1. KPI\_note\_for\_annual\_reporting-cPPPs.pdf released by DG RTD in February 2018, defining, describing and providing a methodology for the 4 common KPIs to all the cPPPs.
2. BDV cPPP contractual agreement<sup>1</sup>.
3. BDV cPPP SRIA<sup>2</sup>.

BDVA leads the production of the Monitoring Report 2019-20 of the Big Data Value cPPP as part of its contractual obligations in the cPPP. The work is developed by the BDVA Office and the BDVA TF2 (impact). To produce this report the Association has targetted input from:

1. **All the running and selected Big Data cPPP projects** during years 2019 and 2020 (**15 projects** from the Call 1, 2016, **17 projects** from Call 2, 2017, **14 projects** from the calls 2018, including the calls ICT-11-a (combined HPC-Data projects), **15 projects** from the calls 2019 and **10 projects** from the calls 2020). **In total 71 projects** have been addressed for the Monitoring Report 2019-20. They all were active during the reporting period except 3 projects from the calls ICT-51 that started in 2021.
2. **All for-profit project partners** (including those ones receiving cascade funding) from the projects active during 2019-20: **519 for-profit companies** from which **340 are SMEs (205 through cascade funding or In-kind support** including the **29 start-ups/SMEs** funded by the DataPitch project, **76 start-ups** funded by EDI during 2018, using the modality of cascade funding and **100 start-ups/SMEs** receiving in-kind services by Data Market Services).
3. **Members of BDVA: 235 organisations**<sup>3</sup> from which **89 are private for-profit** companies.
4. BDVA Task Forces and BDVA Office.
5. EC DG CNECT G1.
6. The European Data Market Monitoring Tool<sup>4</sup> from which this report has gathered information for the Macro-Economic KPIs.

The information was collected in the following way:

For projects active in 2019-20:

- Calls 2016 (15 projects): AEGIS BDVe BigDataOcean Data Pitch DataBio e-Sides euBusinessGraph EW-Shopp FashionBrain MH-MD QROWD SLIPO SODA SPECIAL TT

---

1 [http://www.bdva.eu/sites/default/files/BDVPPP\\_Contractual\\_Arrangement\\_.pdf](http://www.bdva.eu/sites/default/files/BDVPPP_Contractual_Arrangement_.pdf)

2 BDV cPPP SRIA was updated to version v4 in November 2017 and therefore it does not reflect the new re-arrangement of KPIs defined by DG RTD. The KPI catalog has been updated and it will be added to the new version of the SRIA planned for Sep-Oct 2018.

3 BDVA members at the time of submitting the report

4 SMART 2016/0063 – Study “Update of the European data marketMonitoring Tool”, IDC and Lisbon Council

- Calls 2017 (17 projects): BigDataGrapes BigDataStack BigMedilytics BODYPASS BOOST 4.0 CLASS Cross-CPP DataBench E2Data EDI FANDANGO I-BiDaaS ICARUS Lynx TheyBuyForYou Track and Know TYPHON
- Calls 2018 (14 projects): DeepHealth, Lexis, Cybele, Evolve, DataMarketServices, CloudButton, Elastic, Examode, ExtremeEarth, INFORE, SmartDataLake, Safe-Deed, Musketeer, MOSAlcROWN
- Calls 2019 (15 projects): IoTwins, Infinitech, Platoon, the BD4OPEM, Synergy, BD4NRG, SmasHIT, KRAKEN, PIMCITY, TRUSTS, i3-Market, DataVaults, the FSM, OpertusMundi, DataPorts
- Calls 2020 and active in 2020 (even only for a few months) (7 projects): EuHubs4Data, Reach, MediaFutures, I4-Trust, Daphne, Everest, More

The following tools were used to collect information:

- **Online questionnaire** with 80 questions to gather information for each of the KPIs and specific sections of the document,
- **Excel questionnaire** to report on delivered, ongoing and planned **contributions to the different technical priorities** to assess coverage of the SRIA implementation. (Technical priorities table for projects - Questionnaire)
- **Excel questionnaire** to collect information in relation to **innovations to market** (*Significant Innovations to Market – Questionnaire*)

Projects selected and funded as result of the H2020 calls 2019 (November) or 2020 (January), which are starting in early 2021: This includes: VesselAI, SELMA, Marvel (3 projects)

- Information was collected through an **Excel questionnaire** to report on planned **contributions to the different technical priorities** to assess coverage of the SRIA implementation. (Technical priorities table for projects - Questionnaire).

**For all Industrial partners** (for-profit organisations) from projects active during the reporting period (including cascade funding) and BDVA members, information was collected through an **online questionnaire** to gather information for the private investments and SME mapping. (*Add*)

**For BDVA members (all members):** Information was gathered through an **Online questionnaire** with 40 questions mainly covering aspects of job creation, job profiles and skills development, Innovations and technical results, contribution to Standards, contribution to Big Data Ecosystem, and success stories. (*Questionnaire for BDVA members*).

Task Force and Subgroup leads from BDVA, BDVA\_office and the European Commission (DG CNECT G1) provided direct input to specific parts of the report and the annexes.

All questionnaires are available in section 13 of this annex.

## 2 MOBILISED PRIVATE INVESTMENTS (I.1)

Through this KPI we attempt to understand and capture/show the level of industrial engagement within the BDV cPPP.

All **for-profit organisations** participating in **cPPP projects** active during 2019-20 and all for-profit organisations **members of BDVA** were called to provide input to this KPI through an investments survey or questionnaire.

In total **554 organisations** were requested to provide input:

- **519 for-profit companies** from the projects running in 2019-20 from which **340 are SMEs (205 through cascade funding or In-kind support** including the **29** start-ups/SMEs funded by the DataPitch project, **76** start-ups funded by EDI during 2018, using the modality of cascade funding and **100** start-ups/SMEs receiving in-kind services by Data Market Services).
- **89 for-profit** companies from **members of BDVA** (from which 58 are SMEs, 65% of the total). 54 of these 89 organisations participated in projects active in 2019-20 (therefore counted under the 519 companies represented in projects).

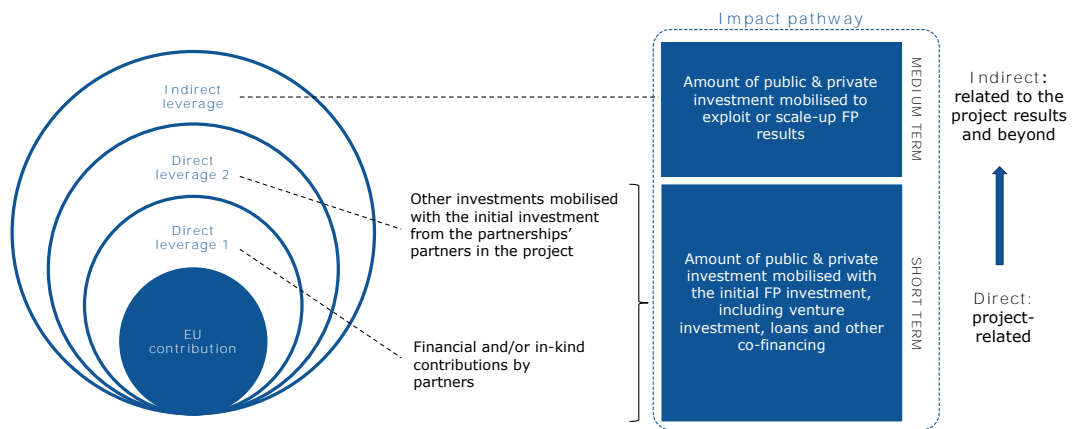
## 2.1 Methodology

For the Annual Monitoring Report 2018 the European Commission suggested a common structure and methodology for all the contractual PPPs. That methodology was used for MR2018 and It has been used for the MR2019-20.

This methodology identifies different categories of investments as follows:

- **EC:** EU contribution received for projects under the cPPP
- **Direct leverage** ( $A = A1 + A2$ ):
  - **A1:** Direct contributions from the industry to the cPPP projects, as predetermined in the projects (co-funding)
  - **A2:** Additional investment or resources spent by your company for the execution of cPPP projects, beyond the direct contributions under A1 (these are typically due to shortfalls in fixed overheads for a project, infrastructure costs not covered by direct funding, consumables and maintenance costs, equipment costs, additional personnel costs, other costs).
- **Indirect leverage** ( $B = B1 + B2$ )
  - **B1:** Additional investment or resources spent as a follow-up to the activities and results related to cPPP projects (e.g. to exploit or scale-up results) and that are not supporting directly the execution of these projects. This can include additional funding mobilised thanks to the cPPP projects such as venture capital, other public funding, etc.
  - **B2:** Additional investments or resources spent that support the objectives of the cPPP and that have been triggered/encouraged by the existence of the cPPP, but that are not following up activities or results of the cPPP projects as such (hence, not counted in A1, A2 and B1).

Figure 1 visually shows the structure of this KPI as proposed by the EC.



**Figure 1:** Methodology and KPI structure proposed by EC for MR2018 (cPPPs) and used in MR2019-220

This structure to present leverage investments is in line with the structure and methodology used by BDVA in previous annual Monitoring reports (before 2018), allowing consistency in presenting the evolution of numbers over the years. The mapping in between the sub-KPIs (A1, A2, B1 and B2) from this year report and previous reports can be seen in Table 1.

New Leverage investments Categories (MR 2018)	Leverage investment categories used in MR2017 (and before)
<p>B2: Additional investments or resources spent that support the objectives of the cPPP and that have been triggered/encouraged by the existence of the cPPP, but that are not following up activities or results of the cPPP projects as such (hence, not counted in A1, A2 and B1).</p>	<p>Amount of R&amp;D expenses that are related to the Big Data PPP but are not related to EC funded projects. (This <b>excludes</b> any expenses that are funded by the EC by definition)</p> <hr/> <p>Amount of R&amp;D expenses resulting from follow-up investments of projects funded by the EC that are topic-wise related to the Big Data PPP however initiated outside the Big Data PPP (in FP7 or in H2020). (This <b>excludes any expenses that are funded by the EC by definition</b>)</p>

<p>B1: Additional investment or resources spent as a follow-up to the activities and results related to cPPP projects (e.g. to exploit or scale-up results) and that are not supporting directly the execution of these projects. This can include additional funding mobilised thanks to the cPPP projects such as venture capital, other public funding, etc.</p>	<p>Amount of R&amp;D expenses resulting from follow-up investments of Big Data PPP projects (This excludes any expenses that are funded by the EC by definition)</p>
<p>A2: Additional investment or resources spent by your company for the execution of cPPP projects, beyond the direct contributions under A1 (these are typically due to shortfalls in fixed overheads for a project, infrastructure costs not covered by direct funding, consumables and maintenance costs, equipment costs, additional personnel costs, other costs)</p>	<p>Additional investments in execution of PPP projects (in BDV CPPP projects). (This excludes any expenses that are funded by the EC by definition)</p>

Table 1: Mapping KPI structure proposed by EC for MR2018, MR2019-20 and the one used in previous years

This mapping allows to present evolution of figures over the years and also to calculate cumulative figures and overall leverage and leverage factor.

An online questionnaire was sent out to the 554 companies<sup>5</sup>. The questionnaire requested the following relevant information<sup>6</sup>:

- Name of the company and cPPP projects where the company is involved (if any)
- Turnover, employment and overall R&D Expenses (regional coverage Europe)
- Estimation of the total amount of R&D expenses that are related to Big Data and Data-drive AI (BDV SRIA and roadmap).
- Estimation of the amount of R&D expenses that are related to the Big Data and Data-drive AI (BDV SRIA and roadmap) but are not related to EU-funded projects. **Answer from companies to this question is aggregated to calculate B2.**
- Estimation of the amount of R&D expenses resulting from follow-up investments of projects funded by the EC that are topic-wise related to the Big Data PPP however initiated outside the Big Data PPP (in FP7 or in H2020) (excluding any expenses that have been funded by the EC). **Answer from companies to this question is aggregated to calculate B2.**
- Estimation of the amount of R&D expenses resulting from follow-up investments of Big Data PPP projects (Excluding expenses that are funded by the EC) (only for participants in cPPP projects). **Answer from companies to this question is aggregated to calculate B1.**
- Calculation of additional investment or resources spent by the company for the execution of cPPP projects, beyond direct contributions or co-funded reported to EC (in BDV CPPP projects). (only for BDV CPPP project partners): **Answer from companies to this question is aggregated to calculate A2.**

<sup>5</sup> Questionnaire can be found in section 12 or this annex 2.

<sup>6</sup> This same questionnaire was used to gather input about SME evolution.

The data was mainly requested for year 2019-20 and companies were also asked to estimate values 2021-2024 when possible (mainly for the follow up investments in cPPP projects).

The companies were further asked to provide the following information for their figures:

- the regional coverage of the figures;
- whether or not their organisation is an SME according to the Horizon 2020 participation rules.

SMEs were asked to provide turnover and employment data from 2014, and age of the company, but this historical data was used only to input I.3 (impact in SMEs).

The questionnaire sent to companies can be found in section 12 of this annex 2.

No extrapolations are done, nor external sources used (e.g. investments in the sector). Investments in Big Data Value does not relate only to the companies developing technologies, but those ones integrating them and bringing them to the different verticals, industrial user communities adopting these new technologies and data providers (industrial users or public sector). The BDV cPPP incorporates all this kind of companies throughout the different projects and BDVA members, however it is difficult to find external sources that combine investments of all involved stakeholders in many different sectors. In future reports, and once the number of companies increases and so do the maturity of the projects it might be relevant to assess investment figures per sector or per role in the Big Data Value chain.

The figures for EU funding and A1 values (in relation to the specific public investments and private co-funding for the year 2018-20) is provided by the EC<sup>7</sup>.

## 2.2 Responses rates

In total **86 for-profit organisations provided input** to this KPIs. 97,7% of these organisations are partners or are linked to projects running in 2019-20. None of the startups/SMEs linked to the Incubators DataPitch, EDI or DMS replied to this questionnaire. This applies to 205 organisations. The contribution to the Investments part is normally very low therefore the purpose of this exercise they are not counted. We consider the total target number for this KPI of **349 for-profit organisations**.

The **response rate is 24,6%** (much lower than in 2018 with a response rate of 40,9%).

Not all companies provided input to all the questions needed to calculate aggregated figures for the different categories and therefore the responses rates for each of the sub-KPIs is different.

Leverage investments Categories (MR 2019-20)	Response rate
<b>B2:</b> Additional investments or resources spent that support the objectives of the cPPP and that have been triggered/encouraged by the existence of the cPPP, but that are not following up activities or results of the cPPP projects as such (hence, not counted in A1, A2 and B1).	12%
<b>B1:</b> Additional investment or resources spent as a follow-up to the activities and results related to cPPP projects (e.g. to exploit or scale-up results) and that are not supporting directly the execution of these	6%

<sup>7</sup> Current version of the document works with estimates as those numbers have not been provided to the Association by the EC.

projects. This can include additional funding mobilised thanks to the cPPP projects such as venture capital, other public funding, etc.	
<b>A2:</b> Additional investment or resources spent by your company for the execution of cPPP projects, beyond the direct contributions under A1 (these are typically due to shortfalls in fixed overheads for a project, infrastructure costs not covered by direct funding, consumables and maintenance costs, equipment costs, additional personnel costs, other costs)	6%

**Table 2:** Response rate per investment sub-KPI

**The response rate is very low.** Although 25% of the targetted companies responded to the questionnaire only 12% or 6% (depending on the sub-KPI) provided Input to the questions related to the specific sub-KPIs. This report only uses only provided data and does not extrapolate totals. Therefore, It Is Important to highlight that the real values of this KPI are most likely much bigger than the ones provided by this report.

## 2.3 Overall results

Table 3 shows the EU contribution granted by the EC in calls for all the years since the launch of the cPPP and the committed private contribution as per Grant agreements. The projects active in 2019-20 are result of the calls 2016-2020.

In total the **EC has allocated 380,8 M€ in projects of the Big Data Value cPPP.**

KPI	Description	Unit	2015	2016	2017	2018	2019-2020	Total
EC (committed)	Total EU contribution for projects granted during reporting period (starting the year after)	MEUR	N/A	69,90	88,90	42,50	179,50	380,80
A1 (committed)	Total Private contribution for projects granted during reporting period (starting the year after)	MEUR	N/A	13,70	10,40	0,00	29,70	53,80

**Table 3:** Total EU and private contribution for projects granted since the launch of the PPP

To calculate the amount corresponding to the execution of projects during the year 2019 and 2020, we estimate a 2,5-year project duration average, and a flat distribution of funding along that time. Considering the contribution granted in 2016 to 2020 we estimate the EU contribution received for projects of the cPPP during 2019-20 is **144,88 MEuros.**

Table 4 shows the overall results of private investments for 2019-20

Description	2019-20 (Meuros)	Response rate (%)
EC: EU contribution received for projects under the cPPP	144,88	Calculated over 100%. Estimation based on running projects 2019-20

<b>Direct leverage (A = A1 + A2)</b>	<b>25,07</b>	
A1: Direct contributions from the industry to the cPPP projects, as predetermined in the projects (co-funding)	19,50	Calculated over 100%. Estimation based on running projects 2019-20
A2: Additional investment or resources spent by your company for the execution of cPPP projects, beyond the direct contributions under A1 (these are typically due to shortfalls in fixed overheads for a project, infrastructure costs not covered by direct funding, consumables and maintenance costs, equipment costs, additional personnel costs, other costs)	5,57	6%
<b>Indirect leverage (B = B1 + B2)</b>	<b>601,94</b>	
B1: Additional investment or resources spent as a follow-up to the activities and results related to cPPP projects (e.g. to exploit or scale-up results) and that are not supporting directly the execution of these projects. This can include additional funding mobilised thanks to the cPPP projects such as venture capital, other public funding, etc.	21,77	6%
B2: Additional investments or resources spent that support the objectives of the cPPP and that have been triggered/encouraged by the existence of the cPPP, but that are not following up activities or results of the cPPP projects as such (hence, not counted in A1, A2 and B1).	580,17	12%
<b>Total leverage (A+B)</b>	<b>627,01</b>	

Table 4: Investment results 2019-20

### **Direct leverage (A = A1 + A2)**

Direct leverage includes the direct contributions from the industry to the cPPP as predetermined by the projects (co-funding) and the additional investments or resources spent by companies for the execution of the projects beyond the known direct contribution.

**A1 - Direct contributions from the industry to the cPPP projects, as predetermined in the projects (co-funding).** This number is known by the EC, but not to the Association that has calculated a number based on the following assumptions.

Total amount of private funded committed for projects active in 2019-20 is 53,8 MEuros (13,7M€ in projects granted 2016 and 10,40 M€ in projects granted in 2017, 29,70 M€ in projects granted in 2019-20). Considering an average duration of 2,5 year per project, the amount of private contribution during reported period 2019-10 is **19,5 MEuros**.

**A2 - Additional investment or resources spent by companies for the execution of cPPP projects,** beyond the direct contributions under A1 (these are typically due to shortfalls in fixed overheads for a project, infrastructure costs not covered by direct funding, consumables and maintenance costs, equipment costs, additional personnel costs, other costs).

This input is provided through the questionnaires. Companies have reported **5,57 M€** of additional investment in project execution (with a response rate of 6%). Extrapolation to the 100% of the companies would take us to an additional private investment of **92,8 M€**. However, extrapolations are not considered in this exercise but only aggregated numbers.

### Indirect leverage (B = B1 + B2)

**B1 - Additional investment or resources spent as a follow-up to the activities and results related to cPPP projects** (e.g. to exploit or scale-up results) and that are not supporting directly the execution of these projects. This can include additional funding mobilised thanks to the cPPP projects such as venture capital, other public funding, etc.

**21,77 M€** have been reported for this KPI, with a 6% of response rate. Even considering the low response rate this number is much higher than the one reported in 2018. This is logical as the programme is more advanced, many projects have reached the end, and therefore follow-up investments are on their way. Follow-up investments reported for 2021+ by these group of companies is **42,93 M€**. Considering some input from previous years a 6% of the addressed companies have reported total of **73,6 M€ in follow-up investments to the cPPP projects**. If this value is extrapolated to the 100% of the companies, the total value of follow-up investments would be very similar in value to the EC contribution to the cPPP projects overall (370 M€).

**B2 - Additional investments or resources spent that support the objectives of the cPPP and that have been triggered/encouraged by the existence of the cPPP**, but that are not following up activities or results of the cPPP projects as such (hence, not counted in A1, A2 and B1).

The total based on aggregated numbers is **580,17 M€**:

- **38,77 M€** relate to as follow-up investments for EC-funded projects topic-wise related to Big Data but not part of the cPPP
- **541,41 M€** as investments related to the Big Data PPP, but not related to EC-funded projects.

Considering the response rate, we can assume this number is much higher, but as in previous years we only report on aggregated numbers based on real individual inputs with no extrapolation. As the methodology used this year to calculate this KPIs is the same that the one used in previous Monitoring Reports exercises we can show the evolution.

Table 5 shows the evolution of reported numbers in private investments from 2015 to 2020.

KPI	Description	Unit	2015	2016	2017	2018	2019-2020	2021+	Total
B2	Estimate the amount of R&D expenses that are related to the Big Data PPP but are not related to EC funded projects. (This excludes any expenses that are funded by the EC by definition!)	MEUR	245,80	289,10	388,67	370,33	541,41	-	1.835,31
B2	Estimate the amount of R&D expenses resulting from follow-up investments of projects funded by the EC that are topic-wise related to the Big Data PPP however initiated outside the Big Data PPP (in FP7 or in H2020). (This excludes any expenses that are funded by the EC by definition!)	MEUR	35,10	49,40	83,55	70,40	38,77	-	277,22
B1	Estimate the amount of R&D expenses resulting from follow-up investments of Big Data PPP projects (This excludes any expenses that are funded by the EC by definition!)	MEUR	N/A	N/A	1,09	7,57	21,77	42,93	73,36
A2	Additional investments in execution of PPP projects (in BDV PPP projects). (This excludes any expenses that are funded by the EC by definition!)	MEUR	N/A	N/A	3,46	9,52	5,57	-	18,55
A1 (estimated)	Annual private contribution (estimated for reporting period 2019-20)	MEUR			5,48	9,64	19,50	19,18	53,80

**Table 5:** Evolution private investments on BDV cPPP over time

Aggregated to the numbers reported in 2015 (280,9 M€), 2016 (338,5 M€), 2017 (482,25 M€), 2018 (467,47M€) and the remaining co-funded committed in cPPP projects, and some estimations on follow-up investments provided the amount of mobilised private investments since the launch of the cPPP is **2.258,2 M€ (2,26 B€)**. Considering the amount of EU funded allocated to the cPPP overall (380,8 M€) the BDV cPPP ends with a **leverage factor of 5,93** higher than the leverage factor of 4 committed contractually.

## 3 NEW SKILLS AND/OR JOB PROFILES (I.2, II.8)

### 3.1 KPI I.2 (Number of jobs created)

This KPI considers new skills and job profiles created or forecasted within a reasonable timeframe in the future. This KPI is split in 2 different sub-KPIs:

- Job creation (quantitative value and qualitative description of causality)
- Job profiles established and skills created by 2024 (quantitative value and qualitative description).

BDV cPPP has 3 additional specific KPIs contributing to the job creation, profiles and skills:

- II.2 (Number of Data Companies) (Macro-economic)
- II.5 (Increased number of Data workers) (Macro-economic)
- II.8 (High establishment availability of big data value creation skills development)

A combination of the input provided in the 4 KPIs (I.2, II.2, II.5 and II.8) is needed to have a full overview about how job profiles and skills are being created within the activities and sectors covered by the BDV cPPP, the number of workforce affected in the European Union and forecasts of its evolution on the employment markets<sup>8</sup>.

#### Target and response rates

All Big Data cPPP projects active in 2019-2020 (67 projects) and all the BDVA members were asked to contribute with input to these KPIs. Responses rate:

- 40 over 67 projects provided input to these KPIs (**59,7%**)
- 22 over 2229 BDVA members provided input to these KPIs (**9,9%**)

#### 3.1.1 Job creation

**70% of the BDV cPPP projects that replied to the survey indicated their project will contribute to job creation by the end of 2023**, with an estimation in accumulated numbers of thousands.

Projects indicated in a qualitative way the causality in between the project activities and the job creation, and some provided an estimation in numbers (although several indicated that a serious estimation is difficult

---

<sup>8</sup> KPI\_note\_for\_annual\_reporting-cPPPs.pdf

<sup>9</sup> Total number of members accounted in May 2021

to do at this stage). Job creation in discrete numbers is not possible to calculate precisely at this stage but **estimated numbers go over 8.158 new jobs created until 2024 linked to project activities** and many more considering indirect effects.

BDV cPPP projects contribute to job creation in Europe by 1) **Increase market share and benefit big data economy actors**; 2) developing **new job profiles** that generate new jobs; 3) **increasing the actual number of IT jobs in Europe**, in fields such as data science, software engineering and machine learning and, to a less extent, non-IT jobs (financial, communication, insurance, marketing and business) 4) **developing new opportunities for entrepreneurs and startups** in the new Data Economy; 5) generating job opportunities by **increasing data sharing**; 5) creating new **jobs already during the life time of the project**; 6) forecasting new jobs created as a **follow up of project results**; and 7) **creating jobs in sectors such as: IT, energy, healthcare, journalism, finance, insurance, biomedicine** etc.

**1) Increase market share and benefit big data economy actors**, as for example:

- **SmashHit:** With the growing ability of CPP (e.g. cars) to generate, gather and share data with third parties among different data-sharing platforms, there will be a general need for flexible and easily manageable procedures to handle data owner consent and legal rules, to achieve effective and traceable contracting. The challenges and complexities of the General Data Protection Regulation (GDPR) directives make cumbersome mechanisms necessary to gain, record and manage consent. Also, consumers are wary about improper use of their data. The combination of understanding and relating to the value proposition, consumer trust, and cumbersome consent processes results in a low opt-in rate for connected CPP data exchange and prevent the creation of innovative services (e.g. connected insurance programs, or smart city solutions).

Since smashHit is developing solutions for above mentioned problems, it contributes to the rapid growth of the European data economy. Such an innovative inter-data-platform framework solution will have several positive effects on the European Data Economy, as e.g.

- To overcome existing barriers, as e.g. the complicated and time-consuming consent/contract processes
  - Increase trust for all stakeholders in the data market, due to strong data traceability mechanism
  - Increase willingness of Data Providers (OEM) to provide CPP data for 3rd party use
  - Open new opportunities for value-creation by creation of innovative services using data from multiple platforms
- **Synergy:** The big data-driven AI-enabled innovative solutions introduced in SYNERGY are expected to bring significant benefits not only for the electricity data value chain stakeholders (focal target groups of the project), but also for businesses working on the Big Data and AI domain, thus, on the one hand enabling resilience and improving green, sustainable growth, while at the same time, disrupting existing business models and introducing new business roles (around big data and AI) across the electricity data value chain

**2) Increase because of the new job profiles developed**, as for example:

- **AEGIS:** reduces the barrier for organisations to start using big data analytics but requires from them to have data analysts – the target users of the AEGIS platform. Therefore, jobs of two categories were generated: 1) Data scientists designing analytic processes using AEGIS 2) Marketing / Business personnel promoting added value services built on top of AEGIS.
- **DeepHealth:** Besides, the enhanced platforms and new biomedical applications generated in the project to address specific health pathologies and purposes could derive in new business lines for industrial partners that will require new personnel to be deployed commercially. Finally, deploying and updating other HPC related new developments, tested within the project, will require expert personnel to use it and evolve it.

- 3) **Increase the actual number of IT jobs in Europe**, in fields such as data science, software engineering and machine learning and, to a less extent, non-IT jobs (financial, communication, insurance, marketing and business)
- **Boost 4.0:** The partners of the project have estimated the direct impact of the project in job creation in around 600 direct new jobs created until 2024. Most of these jobs will be directly related with IT (programmers, communication experts, app developers, etc.), and many of them specifically to data management, analytics, visualization, etc
  - **DataPorts:** The expected technological advancements (covering solutions related to IoT, AI, etc.) that can be directly implemented to port related solutions such as supply chain integration, transportation, tracking, storage and more, would create more job positions. Big Data technology (analytics, data management, processing architectures) is expected to be in place, so more IT staff will be required, such as developers, data analysts and software engineers, to support it.
  - **CloudButton:** Considering the huge potential of Python community and the inherent complexity of transition to cloud technologies, we think that the results of the CloudButton project, that greatly simplify the transition to cloud for Python data scientists, will impact in considerable job creation in the near future.
  - **TT:** can forecast the creation of jobs by 2023 in 200 technical positions mainly focused on Big Data Developers (related to technologies such as Hadoop, elastic search, and Apache among others...), IT Architectures, Data Scientist, Product Managers, etc
- 4) **Creation, support and new opportunities for entrepreneurs and startups** in the new Data Economy. For example:
- **BigMedylitics:** part of the project's results are currently aiding the start-up of a venture within the company. Due to the data-driven proposition of the start-up, the venture will employ more people with a data science background. There will also be other data science workers employed outside the venture to support the activities of the venture
  - **EuHubs4Data:** The involvement of SMEs / start-ups / web entrepreneurs in the different activities of the project (training activities, data driven innovation programme, cross-border experiments, events) will help them to identify needs around data driven innovation that could be addressed with new jobs. Specifically, training activities will directly tackle this aspect, and AI / Big Data maturity assessment services from the catalogue will support them on the needs identification and most appropriate type of profile.
- 5) **Job opportunities generated by increasing data sharing**, as for example:
- **Cross-CPP:** The most important project outcome should be the third party services based on CPP data. These services can be created in cross domains with different objectives in mind, such as the CPP data can be used for weather forecast and at the same time can be used for energy management in buildings etc. Thus, new service providers can take advantage of vast amount of data gathered from cross domain CPPs and can enhance their services.
- 6) **New jobs created already during the life-time of the project**, as for example:
- **DataMarketServices:** Considering 25 startups x3 people x 3 editions= 225 jobs (estimate of jobs created in the cohorts of DMS). Considering the same steady growth rate in 50% of the DMS companies, in the coming years after DMS 2021, we could estimate that DMS will lead to the creation of +900 jobs. However, we expect that some of the companies will growth exponentially, or get acquired by bigger companies, increasing our estimation over 1500 jobs.
- 7) **New jobs created as a follow up of project results**, as for example:

- **BigMedylitics:** Very likely as a result of the project main outcomes, it is expected to create more positions for either PhD students and/or FTE to expand further the developments of the project in the area of healthcare by leveraging the innovations developed that might be applied to other scenarios.
  - **DataPitch:** Employment generated by Data Pitch startups is estimated to rise to 717 (an increase of 407) by 2022. This figure is based on impact projections of an independent impact assessment of Data Pitch. London Economics, who conducted the impact assessment, used a simulation-based approach with random model inputs. The model defined calculation of employment growth based on a set of input assumptions, drawing inputs at random from a pre-defined distribution. This exercise was repeated in total 10,000 times, drawing different random inputs at every repetition, and the average over these 10,000 simulations provides the final estimate for employment growth. The model combined data on job growth from both Data Pitch cohorts.
- 8) Creating jobs in sectors such as: IT, energy, healthcare, journalism, finance, insurance, biomedical etc**
- **BigMedylitics:** The project will open the door to new services and projects in the context of Health Sector using Big Data and Artificial Intelligence solutions. Currently, there is not a relevant number of job offers for data scientists in this sector, as it is complex to access and process data. Bigmedylitics addressed specifically the latter challenge
  - **MORE:** The goal of MORE is to provide the data management and AI tools to unlock the potential of data from the Renewable Energy Sector in Europe. This potential would lead to increased efficiency of energy production, reduced maintenance costs, better forecasting and increased competitiveness. We expect that a successful completion of MORE and the adoption of its results from the industry would lead to at least 1% of additional jobs in the industry due to increased competitiveness and an equal number of jobs create inside and outside the industry in the data management field to use and expand the results of MORE. Given that the wind industry employs around 110k persons in the EU and the solar industry employes another 133K, we expect that we project can increase employment by 2400 new jobs in these industries and another 2400 in the data management industry.
  - **Infinitech:** This project will have impact in all skill jobs from IT to Data Scientists to Financial / Insurance personnel
  - **Fandango:** the project explores the possibility to use Big Data and Artificial Intelligence in the disinformation fight, this expresses a large possibility in defining a new professional profile between technologies, journalism and fact checking

Moreover, **80%** of the PPP projects that have replied to the questionnaire has stated that their project contributes to **increase the number of data professionals in Europe**, either directly by the project (big data architects, data models designers, developers, computer scientists, researchers) or by enabling new service providers and business opportunities (medical experts with big data skills, data scientists, data-driven entrepreneurs, Big Data administrators, Big Data application developers).

Almost **41%** of the members stated that **their participation in the BDVA/BDV CPPP has already contributed directly or indirectly to job creation**, with a total number of **3255 new jobs created** in the years 2019-2020 (2150 for 2019 and 1105 for 2020). The types of jobs created listed referred mainly to the hiring of new experts to develop H2020 projects (Data architects, Data scientists, specialists of data visualisation and machine learning specialists, big data analysts and programmers), new PHDs and postDoc researchers.

### 3.1.2 Job Profiles established / Skills created.

**35% of the projects which answered the questionnaire and that were running in 2019-2020, reported contribution to the creation of new job profiles.** Projects reported **114 job profiles** were established in 2019,

and **74 new job profiles** have reportedly been created in 2020. Moreover, almost 1000 new job profiles are estimated to be created by the end of the projects<sup>10</sup> and linked to the project activities.

Projects directly contributing to new job profiles are BigMedilytics, MORE, DataMarketServices, SYNERGY, Boost 4.0, DataPorts, EUHubs4Data, IoTwins, EDI, TT, Lynx, QROWD.

Specific Job profiles identified/created (by 2020):

In the Healthcare domain (BigMedilytics project):

- Data curation, anonymization and treatment
- To manage the cluster installed as well as the secure channel to access the cluster
- Data scientist profiles
- Big data analytics, cloud computing and distributed real-time streaming systems, etc
- Project management, ethics and regulatory submissions, patient contact, consenting patients, patient follow up etc.
- Knowledge Engineers: experts in knowledge representation, ontologies, controlled vocabularies in the biomedical domain, rule mapping languages
- Knowledge Extraction and Mining: experts in techniques for extracting and annotating knowledge from textual and non-textual material
- Knowledge Management and Representation: experts in techniques for integrating knowledge extracted from heterogeneous data sources, and representing and exploring all these knowledge in knowledge graphs
- Data scientists with special expertise in health sector. Skills in data analytics, big data, data processing, databases, programming, administration of databases, business, economic aspects of healthcare and healthcare processes in a hospital.
- Machine Learning: experts in machine learning, natural language processing, statistical analysis
- Data scientist, big data engineer, and data ethics specialist
- Researcher track and trace, track and trace analyses

In the manufacturing domain (BOOST4.0, IoTwins and MORE projects):

- Mobility Expert: knowledge in autonomous driving, traffic data analysis
- NLP-Expert: knowledge in NLP specific methods, text processing and analysis.
- Data engineer / integrator that is able to understand and translate machine communication protocols, both old and new, and make the data usable for other implacations. Typical skills include undstanding of PLC programming, PLC logic, connection interface, industrial protocols (CodeSys, EtherCAT, Modbud, Serial, etc.) and translation to modern data structures (OPC-UA, Euromap, JSON etc).
- Front-end designer, capable of making quick and easy visuals for many different types of users. Skills including good understanding of design rules, programming for web-interfaces, knowledge of web technologies and design and understanding of WEB-API's.
- Technical Expert & Evangelist - Know, apply and communicate broad technical skills and expertise
- Machine Learning Engineer & Evangelist - Deeper Knowledge of ML Algorithms, systems and processes
- Data Modeling Expert & Technical Evangelist - Leading the Smart Data Models activities
- Data Scientist (Skills in data anlytics like procedures and advanced mathematic skills or programming e.g.in Python)

---

<sup>10</sup> The end of each project may different according to the starting date

- Manager Industry 4.0 (Coordination of data analytics projects with a high level of know how in effort & benefit of data analytic projects)
- Researcher in IoT (domain knowledge, TSN, edge computing, containerization, real-time, robotics)
- IoT developer, web technologies languages and programming, agile development
- Data analysts, developers, system administrators, machine learning practitioners, experts in the artificial intelligence area
- "Vertical integerators" (as per definition of the IoTwins project: "we identified that the skills required to build the entire framework from the IoT level to the cloud (passing through the edge) are typically distributed among different people, with diverse competences (e.g., attaching physical sensors to the IoT components, gathering data on the edge from different data sources and execute lightweight computing tasks, analyse large amount of data using cloud resources, etc.). Having a single person capable of handling all these aspects would greatly benefit SMEs and data centres, as they could perform the integration among the three layers in a more holistic manne")
- Data engineers for compressed streams: data management expert, stream management, compression and modeling techniques for streams
- Data engineers for pattern extraction in time series: data management expert, time series management, pattern extraction from time series

EDI and DataPitch incubators have identified the following new job profiles:

- Several startups have reported to have increased their teams during the programme. Nonetheless, the job profiles are mainly data workers, and sales
- Technical people, who have knowledge about data-driven solutions
- Product managers who oversee the development, roadmap and strategy for the development of a product
- Marketing and sales staff to allow the newly developed products to reach the market

On a similar note, the EUHubs4Data project will create new jobs profiles with the involvement of SMEs / start-ups / web entrepreneurs in the different activities of the project (training activities, data driven innovation programme, cross-border experiments, events) and will help them to identify needs around data driven innovation that could be addressed with the creation of new job profiles. Specifically, training activities will directly tackle this aspect, and AI / Big Data maturity assessment services from the catalogue will support them on the needs identification and most appropriate type of profile.

In the energy sector (Synergy):

- The energy domain is currently adapting to the digital transformation era and a data-driven mentality focusing on extracting value from the data. To this end, new job profiles that are expected to be created include data scientists and data engineers that will be able to effectively use the SYNERGY Big Data Platform & AI Marketplace offerings to acquire missing data from external stakeholders and gain insights into their organisation's and external data.

In the legal sector (Lynx project)

- There is a new job profile created and defined as Legal AI Scientist: a person who is a data scientist combined with knowledge in the legal domain.

In the logistics domain (TT and DataPorts projects) will contribute with:

- Data analysts for each transport domain need very specific skills and knowledge constituting different profiles:
  - Big Data analyst specialised in the airports domain
  - Big Data analyst specialised in the highways/urban mobility domain
  - Big Data analyst specialised in the logistics domain
  - Big Data analyst specialised in the ports domain
  - Big Data analyst specialised in the automotive/connected car/autonomous vehicle domain
  - Big Data analyst specialised in the railway domain

- Big Data Infrastructure Managers for the following transport domains (each category with very specific requirements):
  - Infrastructure Managers specialised in the airports domain
  - Infrastructure Managers specialised in the railway domain
  - Big Data analyst specialised in the ports domain
  - Infrastructure Managers specialised in the highway/urban mobility domain
- Senior IoT analyst: Skills in IoT Platforms, Semantics, IT.
- Senior IoT developer: Skills in Data Access Agents, Interoperability Broker, IoT Platforms, Semantics.
- Semantic Engineer: Skills in ontologies, data models, semantic tools and frameworks.
- Data-driven Seller: A job title that could a member of IT or ICT Sales Units. This profile will require both selling and technical skills.
- Cognitive Services Developer: IDS (International Data Spaces) data consumer.
- IDS engineer: IDS Data Owner, IDS data consumer, and IDS software architect.

Additionally, the project QROWD contributes with the following new job profile:

- Crowdsourcing Engineer: A professional that is skilled in the design, implementation and deployment of Crowdsourcing pipelines, from interaction with crowdworkers, to the analysis of the obtained data.
- A Crowdsourcing engineer should have the following skills:
  - Web development
  - Data analytics
  - UI design

**67,5%** of the responding projects to the questionnaire reported **contribution to the generation of new skills** by end of the project.

In addition to the skills linked to the new job profiles reported above new skills are expected to be developed in cross sectorial domains (e.g new skills in machine learning, data analysis, data management e.g. Kraken, DataBench and Synergy projects) and in specific sectors (e.g in the health sector the project BigMedilitycs is fostering skills in machine learning techniques for processing Electronic Health Records or Examode project is enhancing the digital skills of the histopathology healthcare workers).

Regarding the BDV cPPP incubators and accelerators, DataPitch provided mentorship and training in areas including data science and skills; open data; skills for digital businesses, including marketing and cyber-security; business innovation, and the legal aspects of sharing data. EDI has focussed on Data Managers, since it has reportedly stated that data corporates are increasable realizing about the importance of incorporating professionals with skills that can turn Industrial Data from a resource used internally, into a marketable asset from which additional revenues that can be obtained from companies. DataMarketServices focuses on offering free support services to data-centric startups. Thanks to DMS, the companies can learn about fundraising strategies, acceleration, data skills, IPR, GDPR and standardization, and marketing.

Details of activities developed by the projects, and specific skill sets develops show the **richness of the programme in skills development**. In particular:

- The project **BigMedilitycs** enhanced skills on machine learning/deep learning, Hadoop, Spark, application of machine learning techniques for processing Electronic Health Records and Infrastructure deployment for trustful EHR analysis and data sharing
- **Daphne** is helping new system infrastructure (usage, development, maintenance) and integrated data analysis pipelines
- **Everest**, to better exploit the EVEREST programming environment, will empower designers with more hardware/software skills for application development (with domain-specific languages), application descriptions (with orchestration frameworks), and system integration (with minimal knowledge of the hardware infrastructure)
- In the **EW-Shop** skills are developed to use wheather and event data to improve predictive models in retail, e-commerce, digital advertising
- **INFORE** contributes to deep and multi-source analysis of stock- and commodity markets

- **MosaiCrown** worked on novel techniques and solutions for enabling selective and sanitized data disclosure, providing for efficient and scalable privacy-aware collaborative computations. Such techniques and tools clearly play a major role in the development of competences and skills related to data confidentiality and control over data
- **Boost 4.0** by boosting the value of data in the industry, is contributing in a very significant way to the development of new skills, mostly related with data analysis and data treatment. Some examples our partners have identified are: Skills in semantics and graph data, ETL processes dedicated for time-series data, Machine learning for root-cause analysis, Interactive dashboard design, IoT technologies, Engineering management, Big data analysis skills, Knowledge in sensor technology.
- **DataPorts** is focussing on Kubernetes expert, Semantics technologies, Big Data Analytics, Data Access Agents, and blockchain technologies
- On a different note, **DeepHealth** will be organizing a Summer School in 2022, where it is planned to provide specific training in the following areas: Deep Learning, Computer Vision, HPC and Cloud environments, GPUs & CUDA, FPGAs & HLS, Usage of the DeepHealth toolkit and libraries, Usage of GPUs, Usage of FPGAs, Intro to COMPSs + examples on HPC and Cloud, Intro to StreamFlow. Also, partners have involved junior researchers in their teams and they are been developing new skills in these areas learning by doing in the project.
- **Safe-Deed**, due to the focus of its project is working on enhanceing skills in the field of privacy-preserving data analytics.
- The **IoTwin**s project contributed to the creation and development of new skills via the exploration of the requirements of the test-beds and the development and deployment of the platform. The firstly created new skill consists in the capability of integrating the multiple layers of the platform (from IoT to cloud). This macro-skill is broken down into the following sub-categories: Data gathering on the edge with low impact on the monitored systems (low-overhead data collection), Data processing on the edge (merging of heterogeneous data flows, visualization, preliminary data analysis), Data transfer from edge to cloud and back, Service transfer from edge to cloud and back, Big data analysis tailored for SMEs (simulation using high-performance hardware resources), Development and deployment of Machine Learning models for predictive maintenance (ML models on data streams, training using cloud resources, inference performed on the edge), Support for the SMEs which may lack knowledge to understand the usage of the platform (training programs), Support for SMEs and data centres for the identification of useful tasks and in the data preparation (e.g., identification of labels to be used in Machine Learning models), Support to SMEs in identifying the right components to be adopted and the correct services to be employed, Tailoring generic services to the specific use-case, Creation of light-weight services to be used on the edge with low overhead
- CloudButton has allowed the researchers involved in the project to improve their scientific skills validating their research in industrial experimental settings grounded with real production environments and workloads. Also, the researchers have improved their communication skills presenting their results in international workshops and conferences. Furthermore, the developers involved in the project have developed new skills (open source contributions, cloud infrastructure).
- The project **PimCity** aims at creating new technical solutions that require specific skill, from privacy preserving analytics, to data provenance certifications, from distributed data stores to advanced analytics to process raw data
- **Fandango** project explores the possibility to use Big Data and Artificial Intelligence in the disinformation fight, this expresses a large possibility in defining a new professional profile between technologies, journalism and fact checking
- **SmartDataLake** focuses on deep and multi-source analysis of (agricultural) stock- and commodity markets.

Among the **BDVA members**, **50%** of organisations responding the questionnaire have reported contribution to the **creation of new Job Profiles** (with a total of 391 new job profiles) and **54% contribute to the creation of new skills linked to the Big Data Value**.

Some examples of job profiles and skills developed by BDVA members:

- **TNO**: AI researchers that have a background in topics like bias (detection, mitigation) and explainable AI (including transparency)
- **IT4Innovations, VSB – Technical University of Ostrava**: New profiles will support us with research and development activities in data analysis, distributed data management, and activities related to easy access to HPC resources with related management software solutions.
- **SatCen**: Different vacancies where published in 2019 for 'technology' profiles. Knowledge in Big data, AI and similar technologies where required skills
- **Know-Center GmbH**: software development team, Data Analysts, specialized Doc and Post-Doc positions working on different AI challenges.
- **Supsi**: PhD students in AI algorithms and applications
- **GMV**: Data Architect, Data Engineer, Data Scientist, specialist for data visualization and Machine learning applied to the processing of large data volumes from satellites.
- **Leonardo S.p.a.**: Engineers with expertise on AI, Big Data (referring to profiles in line with key theme for BDVA and BDV PPP)
- **ITI - INSTITUTO TECNOLÓGICO DE INFORMÁTICA**: In 2020, we have created a profile for Data Leader, expert in the creation of Data Spaces, very oriented to technical aspects and building blocks. Experience related to corporate Data Centres, tools for accelerating Data Analysis, Data Storage, Security. In addition, a new profile related to non technical aspects of Data Spaces, dealing with Data Governance, Data Sharing, Sovereignty. Also an specialization in monetising data has been added as a profile in the business development department.
- **IBM Research**: Mainly analysis of data in cloud, AI automation of cloud systems, NLP AI algorithms, AI computer vision
- **Universidad Politécnica de Madrid**: Data Scientist, Data Analyst, Data Engineer. All three profiles with specialization in different domains: health, finance, manufacturing.
- **Department of Economics and Management, University of Pavia**: 3 Post doc researchers: data scientists for fintech risk management. Required skills: statistics, computer programming, finance
- **LORTEK**: Data Scientist with Mathematics background, Strong knowledge on statistics and applied science. proficiency in python, knowledge of state-of-the-art tools for machine learning.

Some interesting initiatives by members of the BDVA include (non-exhaustive list):

- **IT4Innovations, VSB – Technical University of Ostrava**: Develops the skills in operation of Cloud software stack together with high-level support in the domain of convergence HPC/Cloud/BigData was made
- The **Know-Center GmbH** is upskilling scientists, and providing tools and initiatives to support people in their lifelong learning pathways
- **UCPH** educates a large number of students /Ph.D. students (+1,000) related to BDVA/DVV PPP activity areas, e.g. HPC, Actuarial Mathematics, Bioinformatics, Social Data Science, Computer Science, IT, and Cognition
- **GMV** collaborates with higher education organizations such as the following from Madrid, Sevilla, Valencia, Valladolid and Utrecht, respectively:
  - URJC's Data Science Master, where GMV offers internships in diverse application fields such as precision agriculture, prevention of banking fraud, cybersecurity, digital surveillance, detection of datacenter anomalies, monitoring and analysis of internet publicity campaigns, management of clinical and epidemiological data and evidence-based clinical rehabilitation.
  - Uni. Sevilla's Data Science and Big Data Master, where GMV offers applied Data Science training as well.
  - Uni. Valladolid's Master in Big Data Science, where GMV is involved in Fundamentals and Applications of Data Science and Big Data.
  - Utrecht Network summer school at University of Valencia, where GMV deals about machine learning's impact on society, and the growing concerns about its transparency and fairness.
- **Engineering** is organising internal training courses on skills related to the entire context of big data, from system level to data modelling and analytics.

- Following our mission of knowledge and technology transfer to the market, we have adapted the training courses in 2019 and 2020 to the needs of those markets.
- **ITI - INSTITUTO TECNOLÓGICO DE INFORMÁTICA** launched in 2020 a new course titled "Development of Secure Big Data Analytics solutions". It also launched others like:
  - NLP: From Zero to Hero: Natural Language Processing introduction, use cases, pre-trained models and NLP on cloud
  - Introduction to Anomaly Detection Methodologies with PYTHON: Anomaly detection problem, Classification Methodologies, Assembling models, Open source tools
  - Introduction to Machine Learning with Python: Introduction to Python, Introduction to Data Science, Exploratory nData Analytics, Data Processing, Introduction to Machine Learning, Machine Learning in Python
  - Big Data implementation strategy for Managers: Introduction to Big Data; Big Data Technologies; Case Studies; Big Data Strategy
- **Universidad Politécnica de Madrid** has contributed to the development of new skills by offering the following master programmes:
  - Master in Data Science
  - EIT Master in
    - Digital Innovation, specialising in Data Science / Itinerary Data Science
    - Digital Innovation, specialisation in Data Science / Health Itinerary
    - Digital Innovation, specialising in Design and Human-Computer Interaction
    - Artificial Intelligence
  - Degree in Data Science and Artificial Intelligence
  - Degree in Data Systems and Engineering
  - Degree in Computer Engineering
  - Management and organisation of information and communications systems dsic
  - Video game design and development
  - Big data projects
  - Postgraduate courses:
  - Big data applied to decision-making.
- **TU Wien:** Training of multiple doctoral students in big data areas and it is also running a very successful Data Science master's course with about 50 new students starting every semester
- **Paluno/UDE** strengthened the educational offerings of the university's applied computer science and business information systems programmes in the following directions:
  - Offering specialized Master courses and labs on the topic of data-driven AI. Examples in the reporting period include a master course on using online reinforcement learning for building highly adaptive software systems.
  - Augmenting and strengthening existing Master courses with relevant elements linked to big data and data-driven AI. Examples include the master course on "Software Quality Assurance", which offers dedicated units related to data quality, quality of prediction models and how to exploit explainable machine learning for facilitating the quality of the data-driven systems.

Finally, **45% of the projects in 2019 and 67,5% in 2020** as well as **63% of the BDVA members** who replied to the questionnaires have reported **contributions to the to the Skills Agenda for Europe**<sup>11</sup>. Some examples of specific contributions by projects:

- **BigMedilytics:** Emphasis on AI- based solutions Education of digital health experts Experts on legal aspects of AI

---

<sup>11</sup> <http://ec.europa.eu/social/main.jsp?catId=1223>

- **EDI:** Action 3 - EU support for strategic national upskilling action, Action 5: Rolling out the European Universities Initiative and upskilling scientists, Action 7: Increasing STEM graduates and fostering entrepreneurial and transversal skills
- **More:** The project made its beginning and employed the first IT specialists, developers, MSc Students and PhD candidates. All our new personnel are data management experts who will develop the following new skills: - Domain knowledge of the renewable energy industry, thus contributing to strengthening sustainable competitiveness contributing to Action 6: skills to support the twin transitions of the Skills Agenda - Increase their expertise in data analysis and AT, thus contribution to Action 2: Strengthening skills intelligence of the Skills Agenda
- **DataMarketServices:** Action 2: Strengthening skills intelligence, Action 3: EU support for strategic national upskilling action, Action 5: Rolling out the European Universities Initiative and upskilling scientists, Action 7: Increasing STEM graduates and fostering entrepreneurial and transversal skills
- **Class:** 5.1 Solving technical problems 5.2 Identifying needs and technological responses 5.4 Identifying digital competence gaps
- **INFORE:** Analysis, evaluation, impact assessment Skills and qualifications
- **KRAKEN** as a research and innovation project helps people, project consortium participants, to build / improve their skills making them better equipped for life and future work
- **MosaiCrown** thanks to the activities of the members of the consortium in sharing and disseminating the project results, and in giving lectures, seminars and keynotes at wide and diverse audiences, the project is contributing in promoting scientific and technical skills, for instance contributing to Actions 3 and 7 (STEM studies and dissemination of research results)
- **SmashHit** is establishing national digital skills coalitions connecting public authorities, business, education, training and labour market stakeholders and developing concrete measures to bring digital skills and competences to all levels of education and training, supporting teachers and educators and promoting active involvement of business and other organisations
- **IoTwins** has contributed to Action 3 (EU support for strategic national upskilling action) of the Skills Agenda for Europe, in that it has brought new digital skills into the manufacturing sector which will be useful to face the challenges posed by the Industry 4.0 transition

From the BDVA members some examples of specific contributions:

- **IT4I** supported as PRACE training centre the development of new and deepened skills intelligence, including at regional and sectoral levels focused on HPC/Cloud/BigData and AI (Action 2, Action 3) and through training activities provided as centre for digital innovations in HPC, high performance data analysis (HPDA), and artificial intelligence identified by EC. IT4Innovations together with the Moravian-Silesian Innovation Centre have merged their know-how and established the Digital Innovation Hub Ostrava (DIH Ostrava) to enable companies primarily from the Moravian-Silesian Region to assess as well as meet their needs in the field of digitalization (Action 6).
- **Know-Center GmbH** is upskilling scientists, and providing tools and initiatives to support people in their lifelong learning pathway
- The **University of Copenhagen** contributes to digital skills on a national level by educating candidates in a number of areas where there a great societal need for them.
- **GMV** is full partner in EMIMEO (Erasmus Mundus on Innovative Microwave Electronics and Optics) which is focused on Information Science and Engineering fields of study. It is aligned with Skills Agenda for Europe regarding actions to ensure that people have the right skills for jobs.
- **Engineering IT & Management School** is providing training courses on different ICT related topics, including Big Data and Artificial Intelligence. It also joined the SME digital skill programme.
- **University of Duisburg Essen (paluno)** is getting the BDV EducationHub (<http://bigdataprofessional.eu/>) recognized as a pledge of the European Digital Skills and Jobs Coalition

In relation to the forecasted contribution to the Skills Agenda for Europe, the projects have highlighted the following:

- **Daphne:** Action 1: A Pact for Skills - collaboration among research groups from DM, ML systems, HPC  
Action 2: Strengthening skills intelligence - new system infrastructure with new abstractions and techniques  
Action 3: EU support for strategic national upskilling action - ML systems at core of many national AI strategies  
Action 7: Increasing STEM graduates and fostering entrepreneurial and transversal skills - support for interdisciplinary research, and thus, potential for broader STEM interest  
Action 12: Improving the enabling framework to unlock Member States' and private investments in skills - system infrastructure as basis for potential startups and products in existing organizations
- **DeepHealth** will organize a summer school for training students and data scientists in general on AI, DL, BigData and HPC applied to the Health domain, and a hackaton challenge for health data scientists to train and validate anonymization techniques for health imaging data and data-lakes generation.
- **Everest:** The project is expected to contribute mostly to Action 7 (Increasing STEM graduates and fostering entrepreneurial and transversal skills), with new technologies that we will be disseminated through inter-disciplinary courses.
- **Infinitech:** Contribution to Reskilling and upskilling of Professionals in the Finance sector through the INFINITECH academy and Preparation of the Finance Sector for the Future of Work
- **KRAKEN** will contribute to the European Skills Agenda in the context of the European Digital Strategy. The European Skills Agenda is linked to the European Digital Strategy which encompasses among other pillars the European data strategy. KRAKEN shares a specific objective with this strategy: the creation of a single market for data and will contribute to create the expertise to support it
- **Synergy:** Strengthening sustainable competitiveness, as set out in the European Green Deal. Ensuring social fairness, putting into practice the first principle of the European Pillar of Social Rights. Building resilience to react to crises, based on the lessons learnt during the COVID-19 pandemic.
- **EUH4D** aims at upskilling human resources that are essential for widespread digitalisation and successful adoption of big and smart data technologies, especially among SMEs, in two dimensions: (i) expanding the supply of data literate workforce and data analytics practitioners by providing data analytics and training programmes that address necessities of targeted audiences, (ii) raising the competences of data scientists through a knowledge-sharing cross-border mobility programme. More specifically, the project will coordinate and improve awareness of Data Analytics training offerings across partners, leading to share and disseminate training best practices and organise transversal training and upskilling services for different stakeholders EUHubs4Data ecosystem. It will also organise transversal training and upskilling services for industrial workforce and develop knowledge sharing mechanisms within the federation by temporarily relocating DIH staff

The members, on the other hand, have stated the following:

- According to its strategic plan for the period 2021–2024, **IT4I** will continue in supporting the development of new and deepened skills intelligence, including at regional and sectoral levels focused on HPC/Cloud/BigData and AI (Action 2, Action 3) through training activities provided as centre for digital innovations in HPC, high performance data analysis (HPDA), and artificial intelligence identified by EC. IT4Innovations as Digital Innovation Hub Ostrava (DIH Ostrava) will continue in enabling companies primarily from the Moravian-Silesian Region to assess as well as meet their needs in the field of digitalization (Action 6).
- **Engineering IT & Management School** will continue the training program improving the offering on Big Data and Artificial Intelligence in relation to the improvement of research activities in that topic
- **Leonardo S.p.a.** will contribute to the European Agenda for Skills in line with the Group's strategy, and the future needs in terms of competences and skills, also to support the business' needs. Innovation, technology and sustainability are the factors underlying Leonardo's strategy, which are integrated with each other and on which its competitiveness and future growth are founded
- **ITI - INSTITUTO TECNOLÓGICO DE INFORMÁTICA** has created a demo site, called DATAROOM, where students (not only from companies but high school students) can come and test data technologies in action

- **University of Duisburg Essen (paluno)** is continuing the maintenance of the EduHub jointly with BDVA

### 3.2 KPI II.8 Higher establishment availability of big data value creation skills development

This KPI is linked and complements the common KPI I.2 related to the creation of new job profiles and skills.

This KPI is composed by different sub-KPIs:

- Number of training programs established arising from cPPP.
- Number of European training programs involving 3 different disciplines arising from cPPP.
- Number of Master and Phd. students involved in PPP projects.
- Number of dissemination events, seminars, conferences organised in cPPP projects (including number of participants).
- Number of other actions that contribute to this topic and not reported in as part of I.2 or II.8.

#### 3.2.1 Number of training programs established arising from cPPP

47,5% (2019) and 50% (2020) of the cPPP projects which replied to the questionnaires have developed training programs arising from the cPPP in 2019-2020 and have reported **375 training activities during 2019-2020** involving **6.421 participants**.

Some examples:

- **BigMedilytics:**
  - Intro to ML, Company internal
  - Seminars describing possibilities and training in RTLS
  - Seminars and training in data anonimization
  - Seminars and training of GDPR
  - Seminars and training in analytics
  - medication quality circle, attended by physicians
  - Individual training on software and background of involved radiologists
  - AI/ML intro for consortium members
  - See the agenda of the tutorial at <https://events.tib.eu/dils2018/programme/pre-conference/>
  - Workshop with partners and clinicians to define the clinical journey of the breast cancer patient
  - Speech on AI event, which participants were University students and Healthcare professionals. (Jyväskylä 16.1.)
  - Southampton university hospital
  - Rules and regulations for BIG data Technologies (Channel: workshops, Content: Rules and regulations for BIG data Technologies (ethics, legal, policy), Audience: participants of BIG medilytics)
  - Defining KPIs to measure productivity gains (Channel: workshops, Content: balance score card, KPI's, data management, Audience participants of BIG medilytics)
  - Business models for BIG data technologies (Channel: workshops, Content: Canvas, financial models, Audience participants of BIG medilytics)
- **AEGIS:** Organization of 1 hackathon, 3 demonstrator-specific workshops and 1 platform-specific workshop. Delivery of three leaflets, one banner, one video, four newsletters and twelve blog posts.
- From **DataMarketServices** some examples are:

- H2020 and FP9 Introduction to the programme in the EU and basic structure on how funding works.
- Alternative EU funding opportunities emerging from other programmes different than H2020 and FP9. Based on the leveraged knowledge from the H2020-AltFinator project.
- How to pitch in front of investors (events view). Recipe for a successful business, how to present your start-up when you onboard an investor.
- How to pitch in front of investors (MVP approach). Training on how to present MVP to investors.
- Finance skill to access VC or CVC (version 1). Train and support startups in developing financing skills to collect VC or corporate capital.
- On-line B2B sales package. Giving tools and skills to distribute startups services/products to larger B2B entities or customer will be provided.
- IP Business. Covered IP related contract agreements and transfer of assets
- Basic GDPR principles. privacy and data protection
- Data protection strategies.
- Basic data standardisation. SMEs learn how to get or complement the semantics needed for their data value chains. Performed by W3C
- MOOC GDPR. W3C
- Creativity & Online advertising.
- Effectiveness & Digital Design
- **EW-Shopp:**
  - Data Science Summer School, live, Lesson on Data Enrichment, students
  - ESWC 2019: Tutorials, live, tutorial on Semantic Data Enrichment, researchers
- **INFORE:**
  - BDVA Webinar: SIMPLIFY - Data Analytics and Machine Learning Made Simple. Audience: Data scientists, data analysts, researchers in Big Data and ML
- **DataBench:** organised 3 conference sessions where was described how to use DataBench resources in 2019 with essentially BDVA participants. Also, it organised 4 Webinars where we described how to use DataBench resources in 2019-2020 with essentially BDVA participants and projects in the Data area

**BDVA members** have reported additional **4000 training activities** with over **139.000 participants** involved during the years 2019 and 2020 (out of the context of the cPPP projects but linked to the BDV SRIA topics). Some examples include:

- **Software AG:** provides online as well as in-person trainings by internal and external trainers. We use up-to-date training formats including short video tutorials, MOOCs and TV formats that are recorded and broadcasted (live) by our own TV studio. The trainings can be booked by our employees at any time. We had 1800 (1700) different courses in 2019 (2020), most of them being e-learning with a total of 44838 (59085) course participations
- **TNO:** An estimate of 5 bigger events (>50 attendees) is given. Mostly these events are addressing (potential) partners from business and government. The content involves the latest achievements in Big Data, AI and related topics such as Secure Data Analytics. In the estimates given, only the ones that are given in the larger research programmes.
- **ITI - INSTITUTO TECNOLÓGICO DE INFORMÁTICA:** Training activities are covering technical aspects and also no technical aspects around the data cycle. They all are oriented to professionals from companies and, in exceptional cases, to individuals. Some of the covered areas are:
  - Computer Science Subject in PGPE Master of the UV
  - Workshop on 'Software tools for surveillance and intelligence'
  - Good practices in the use of Entity Framework to improve performance
  - Advanced Excel Seminar in the Official Master in Planning and Management of Business Processes (PGPE) of the University of Valencia
  - Introduction to Machine Learning with Python
  - Master in Big Data Analytics 2020
  - Design Thinking

- Programming of algorithms in Python in the Computer Science subject of the Master PGPE of the UV
- UX course student recruitment to improve your reports with QA

### 3.2.2 Number of European training programs involving 3 different disciplines arising from cPPP

These interdisciplinary programmes are meant to contribute to developing the knowledge and skills needed to deal with the complexity of Big Data. 7 projects which answered the questionnaire (17,5%) have contributed to this KPI and in total **62 developed interdisciplinary programs arising from the cPPP during 2019-2020** (27 for 2019 and 35 for 2020), with **around 1800 participants** engaged in such programs. In particular:

- **BigMedilytics:** Examples include Innovation in Healthcare, Data Science for Healthcare, Constitutional problems related to new technologies
- **Boost 4.0:**
  - Seminar session as part of university master's degree in Connected Industry in UC3M: Lectures in BOOST4.0 key concepts and big data application for industry.
  - RISC AI Academy – Workshop „Applying Data Analysis and Machine Learning in Industrial Environments“
  - IDSA Starter Kit development and workshops - digital, live - Technicians, Business, Certification
  - DIH2 - Welcome Camp - digital -Technicians, Business, Machinists, Data Scientists
  - FIWARE Industrial day - digital - Technicians, Business, Certification, Coder
  - FIWARE Technical Steering Committee - digital - Technicians, Business, Certification, Coder
  - Training events for companies and professionals. Physical presence during 2019, online during 2020: Experis, LIUC, Università Cattolica, Innovation Forum, Università di Pavia, Università di Torino, HBR, Università Bicocca, Tor Vergata. Audience was mainly university or master students, professionals and employees.
- **CloudButton:**
  - Distributed systems course at URV, audience: bachelor's degree students
  - Architecture of Distributed systems course, URV, audience: master students
  - Scalable systems and Data course, Imperial College London, audience: master students
  - CNCF Live Webinar: Toward Hybrid Cloud Serverless Transparency with Lithops Framework, audience: IT practitioners.
  - Tutorial at IBM T.J. Watson Research Center, audience: IBM researchers.
  - Tutorial at StrataData Conference NY, audience: data scientists
- **MUSKETEER** organised a remote hackathon - coding, federated ML and security.

**4 BDVA members** have reported 118 training programmes with a total participation of 685 attendees. Some examples that were provided are:

- **TNO:** The multidisciplinary for events organized by TNO is almost always high, since we usually address use cases that come from a specific domain, and then present our research findings that are almost always both technical and ELSE-oriented (specifically human machine interaction).
- **Leonardo** is member of the Italian National Competence Center within the EuroCC project. The Italian National Competence Center is aimed at supporting the activities on high performance computing (HPC), High Performance Data Analytics (HPDA) and Artificial Intelligence (AI), for the benefit of companies, with particular regard to Small and Medium Enterprises (SMEs).

### 3.2.3 Number of Master and Phd students involved in PPP projects

With this KPI we intend to measure the way the cPPP projects contribute to skills transfer between industry and academia by involving more Master and Phd students with industrial players in developing industry-

driven solutions and deploying experimentation testing scenarios. During 2019-2020 52,5% for 2019 and 67,5% for 2020 of the 40 **projects** that replied the questionnaire reported involvement of **Master students** and 55% for 2019 and 70% for 2020 the involvement of **PhD students** involved in project development. Overall, **786 equivalent FTEs** of the mentioned profiles (503 Master and 283 Phd).

### 3.2.4 Number of dissemination events, seminars, conferences organised in cPPP projects (including number of participants)

To mobilise stakeholders the projects have developed intensive communication and dissemination campaigns, with a result of **708 events in total organised by the responding projects during 2019 and 2020** (including own events, seminars, conferences, etc.) outreaching 114.112 participants with the objective of raising awareness about their different activities, to engage new stakeholders, and to keep expectations about future upcoming results.

In relation to the number of people outreached and engaged in dissemination activities developed by the projects during 2018, projects have reported **10,6 Million people** outreached (over 114.000 have been outreached through events, conferences, etc; over 40.000 have been engaged in online activities such as webinars, online training, etc; over 5,2 Million people have been outreached via mass media, online or physical format; and over 113.000 people have been outreached via other ways of communication and dissemination).

A full report with all the details is provided in Annex 4.

## 4 IMPACT OF THE BDV CPPP ON SMES (I.3, II.18)

### 4.1 KPI I.3: Impact of a cPPP on SMEs (in Euros/Qualitative Analysis)

Through this KPI we attempt to understand the impact of the cPPP activities on SMEs.

The European Commission offered the cPPPs the possibility to address this KPI in 2 different ways:

- Calculating the actual turnover of cPPP participants, and in particular the evolution on the development of the turnover of SMEs should span throughout Horizon 2020, with the year 2014 as the baseline (The link between the activities of a cPPP and this evolution should be explained, in particular if it differs from developments at a macroeconomic level).
- Alternative option of mapping the SME community represented within the cPPP including data on the evolution of the number of SMEs, size distribution, and company age.

The BDV cPPP has worked with both approaches, as the combination of both approaches can give a more accurate indication of the impact of the BDV cPPP on SMEs.

#### 4.1.1 Methodology

All SMEs participating in cPPP projects active during 2019-20<sup>12</sup> and all SMEs members of BDVA were called to fill a questionnaire including the following relevant information<sup>13</sup>:

- Name of the company
- Age of the company
- Country of headquarters

---

<sup>12</sup> Project partners, linked third party of cascade funding

<sup>13</sup> Same questionnaire was used to gather information related to private investments (KPI I.1) so those fields are explained under I.1

- Annual Turnover from 2014 to 2020 (in €)
- Number of employees (annual figure from 2014 until 2020)
- cPPP projects they were involved in (if any)

Based on the Input to those questionnaires all the KPIs have been calculated.

#### 4.1.2 Map of responses and responses rates

In total 43 SMEs provided input to the Monitoring report 2019-20 for this KPI (88% non BDVA members and 12% are BDVA members):

- 19% are medium-size companies (8),
- 47% are small companies (20) and,
- 35% are micro companies (15)<sup>14</sup>.

All SMEs replying to the questionnaire were either full project beneficiary or BDVA members (or both). None of the SMEs, Start-ups funded under FSTP budget (open calls) have replied to the questionnaire.

The response rate considering only full project beneficiaries and BDVA members is of 24,7%. If we would consider the overall number of SMEs and Start-ups engaged through FTSP budget during 2019-20 (205 companies) the response rate would be of 11%. However, as no SME/Start-up funded by the Incubators has replied to the Investments and SME questionnaire, we have decided to leave this group out of the statistics.

**Average age of the SMEs is 18,2 years**, and in particular:

- 19% of the SMEs (0 to 4 years old)
- 21% of the SMEs (5 to 10 years old)
- 60% of the SMEs (> 10 years old)

Geographical distribution of the SMEs **providing input to this KPI** is large covering the following European Countries: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Portugal, Romania, Slovenia, Spain, Switzerland and UK. The geographical distribution of SMEs Involved In the ecosystem (projects and BDVA members) is larger including Estonia, Latvia, Bulgaria, Czech Republic, Cyprus, Romania, The Netherlands, Poland, Norway, Luxembourg and Serbia.

42 out of the 43 respondents were SMEs linked to cPPP projects running in 2019-20<sup>15</sup>: None of the respondents was part of the Incubation programmes provided by DataPitch, EDI and DataMarket during 2019-20.

Summary of Responses rates:

- **Overall response rate** (respondents vs all SMEs in projects and Association): **24,7%**
- **Response rate by SMEs partners in projects: 31%**
- Response rate in DataPitch, EDI and DMS start-ups: **0%**
- Response rate in Association members: **9%** (33% If we consider SMEs that are members and have received funding in projects)

<sup>14</sup> Criteria for classification following EC rules: [http://ec.europa.eu/growth/smes/business-friendly-environment/sme-definition\\_en](http://ec.europa.eu/growth/smes/business-friendly-environment/sme-definition_en)

<sup>15</sup> Project partners, linked third party of cascade fundinhg

### 4.1.3 Overall results:

**Total turnover** reported by SMEs in 2019-20 is **268,1 M€ in 2019** and **259 M€ in 2020**<sup>16</sup>. Table 6 shows the evolution of turnover over the years, considering the SMEs that provided input for this report (2019-2020).

	2014	2015	2016	2017	2018	2019	2020
<b>Turnover (totals)</b>	84.560.333,46 €	94.515.012,82 €	111.352.801,40 €	184.595.753,61 €	216.924.485,39 €	268.154.888,20 €	258.998.590,88 €
<b>Turnover evolution (% baseline 2014)</b>	0	11,77%	31,68%	118,30%	156,53%	217,12%	206,29%
<b>Turnover evolution (% annual increase)</b>		11,77%	17,81%	65,78%	17,51%	23,62%	-3,41%

Table 6: SME Turnover evolution

In terms of **turnover evolution** there is an **increase in aggregated data of 217% in 2019 with respect to 2014** (beginning of H2020) and 206% increase in 2020 with respect to 2014. Looking at the annual increase, there was a 23,6% increase in 2019 with respect to 2018, and a 3,41% decrease in 2020 with respect to 2021.

Segmented results considering age of the SME, and size (Medium, small, micro) are reflected in Table 7

	2014	2015	2016	2017	2018	2019	2020
<b>Companies (&lt;5 years)</b>	- €	- €	- €	30.300,00 €	1.201.744,75 €	4.939.582,88 €	5.280.947,00 €
Turnover evolution (% baseline 2018)						311,03%	339,44%
Turnover evolution (% annual increase)						311,03%	6,91%
<b>Companies (5 to 10 years)</b>	529.621,36 €	1.318.287,28 €	3.167.744,84 €	4.449.304,05 €	7.565.861,11 €	9.389.023,88 €	13.066.830,52 €
Turnover evolution (% baseline 2014)		148,91%	498,12%	740,09%	1328,54%	1672,78%	2367,20%
Turnover evolution (% annual increase)		148,91%	140,29%	40,46%	70,05%	24,10%	39,17%
<b>Companies* (&gt;10 years)</b>	84.030.712,10 €	93.196.725,54 €	108.185.056,56 €	180.116.149,56 €	208.156.879,53 €	253.826.281,44 €	240.650.813,36 €
Turnover evolution* (% baseline 2014)		10,91%	28,74%	114,35%	147,72%	202,06%	186,38%
Turnover evolution* (% annual increase)		10,91%	16,08%	66,49%	15,57%	21,94%	-5,19%
<b>Medium Enterprises</b>	51.686.712,24 €	59.164.409,60 €	76.010.394,44 €	102.467.253,75 €	109.372.768,65 €	114.716.284,60 €	87.424.957,38 €
Turnover evolution (% baseline 2014)		14,47%	47,06%	98,25%	111,61%	121,95%	69,14%
Turnover evolution (% annual increase)		14,47%	28,47%	34,81%	6,74%	4,89%	-23,79%
<b>Small Enterprises</b>	30.187.836,22 €	32.867.013,22 €	31.962.378,63 €	78.656.513,28 €	103.736.592,05 €	148.325.821,53 €	166.939.331,25 €
Turnover evolution (% baseline 2014)		8,88%	5,88%	160,56%	243,64%	391,34%	453,00%
Turnover evolution (% annual increase)		8,88%	-2,75%	146,09%	31,89%	42,98%	12,55%

<sup>16</sup> Aggregated total of the companies

	2014	2015	2016	2017	2018	2019	2020
Micro Entreprises	2.685.785,00 €	2.483.590,00 €	3.380.028,33 €	3.471.986,58 €	3.815.124,69 €	5.112.782,07 €	4.634.302,25 €
Turnover evolution (% baseline 2014)		-7,53%	25,85%	29,27%	42,05%	90,36%	72,55%
Turnover evolution (% annual increase)		-7,53%	36,09%	2,72%	9,88%	34,01%	-9,36%

Table 7: Turnover evolution - segmented (per size and age)

Table 8 shows the evolution of FTEs in SMEs who provided input to the monitoring report.

	2014	2015	2016	2017	2018	2019	2020
Evolution in FTEs	383,5	454,8	793,6	1.309,9	1.565,7	1.705,4	1.530,9
FTEs evolution (% vs baseline 2014)		18,59%	106,94%	241,57%	308,28%	344,72%	299,22%
FTEs evolution (% annual increase)		18,59%	74,51%	65,05%	19,53%	8,92%	-10,23%

Table 8: Evolution of FTEs in SMEs

## Assessment

### Turnover:

We can see a general increase in turnover in the companies' part of the cPPP with reported numbers of up to 217% increase in turnover with respect to 2014 and 23,6% in 2019 with respect to 2014. In 2020 there was a decrease with respect to 2019 reflecting the effect of Covid-19 in companies (in particular medium-sized companies). Young SMEs (5 and 10-year-old) show in average the biggest growth in turnover in relation to 2014.

It is also important to mention that not all the SMEs involved in BDV cPPP projects are technology companies but Data users or providers, and the overall results and trend clearly indicate an **ongoing growth of turnover along the whole value chain**.

### Employment:

The trend is positive in all companies' part of the PPP with an average in Increase in employment for the SMEs part of the PPP is of 300% with respect to 2014. Year 2020 shows a decrease in employment with respect to 2019 (effect of Covid-19 ?). We can also observe that the major increases take place in 2017, year of the implementation of the PPP.

Results show the BDV cPPP is benefiting a broad range of SMEs considering geographical distribution, age and size of the company. There is also variety roles in the data value chain<sup>17</sup>. SMEs participating in cPPP projects clearly show a trend of increase in turnover and increase in number of employees.

<sup>17</sup> This information has not been specifically gathered through a questionnaire but it is known because of the role of the companies in the projects. In future exercise it could be useful to request this information also.

## 4.2 KPI II.18: Participation and benefits for SMEs

This KPI is included in the contractual agreement of this PPP. It is partially replaced by the new common KPI I.3 but not completely, so we consider this KPI complementary for that common KPI.

### Number of SMEs participating in cPPP projects and Share of participation of SMEs in cPPP projects

Annex 3 provides details of the results of the call for proposals from 2019-20 with a result of **88 SMEs** participating in projects selected for funding in 2019-20 **22,9% of the partners are SMEs** that are taking **18,95% of the total budget** of the call). These values are not considering the FTSP cascade funding.

# 5 INNOVATIONS EMERGING FROM PROJECTS (I.4, II.7, II.17)

## 5.1 KPI I.4 Number of significant innovations to market

This KPI concerns all developed items that have a marketable value, including products, processes, instruments, methods, technologies, datasets, etc. It should involve all items directly linked to the cPPP projects as developed foreground, as well as any items beyond the scope of the project that is linked to the foreground.

The BDV cPPP has 2 additional specific KPIs (part of article 7 of the contractual agreement and described in the SRIA) contributing to innovation coming from projects (more details for each KPI can be found below):

- II.7: New economic viable services of high societal value
- II.17: New Systems and technologies

A combination of the input provided in the 3 KPIs (I.4, II.7 and II.17) is needed to have a full overview about all innovation created in the BDV cPPP projects that has an exploitable and marketable value, is a new system/technology beyond state-of-the-art and, as some of them, address also a high societal value.

### Target and response rates:

All Big Data cPPP projects active in 2019 and 2020 (67 projects) were asked to contribute with input to these KPIs. **Response rate:** 38 over 67 projects replied (56,7%). This percentage is much lower than the one of MR2018 (close to 100%) because 30 of those projects (42%) had already ended at the time the data was gathered (beginning of 2021).

It must be pointed out that Innovation might not come from the CSA Projects like e-SIDES (involvement of stakeholders to increase citizen's confidence in Big Data Technologies), and BDVe (the Support Action for the whole BDV CPPP Ecosystem). This means that **65 projects** were the target group to create exploitable and marketable results. Therefore, the response rate is **58,4%**.

The low response rate affects the value of this KPI, as the data provided is based on actual information provided and never extrapolated to the total of the projects. For this reason, the values provided need to be considered as a minimum (in total values).

**Questionnaires** to collect information for this KPI can be found in Annex 2.

As In previous years the projects were asked to report not only on Innovations delivered during the reporting period (2019 and 2020) but also expected In 2021 and after, and they are also asked to report on Innovations delivered or reported earlier but further developd during 2019. The replies to this last point was Interpreted In different ways by the project and assessment case by case has been necessary. 8 over 38 projects that replied, listed a number of Innovations as reported in 2018 but further developd In 2019. BigMedilytics, BodyPass, Boost4.0 and CLASS reported Innovations were Indeed reported In 2018 but as to be delivered In 2019+. Therefore, these Innovations are quantified In totals as delivered In 2019. Reported Innovations In this category by DataPitch were not Included in the report 2018 and therefore added to 2019. All Innovations reported by QROWD and TT projects under this category were already counted as Innovations delivered in 2018 and although Improvements took place In 2019 the Innovations are not counted again In the totals. Same occurs to EW-Shopp, where 10 of the 12 Innovations listed as reported in 2018 but advanced in 2019 were actually counted as delivred In 2018 while 2 of them were not (and therefore counted as 2019).

**Results:**

In total, **328 Innovations of exploitable value** are reported as **delivered in 2019 and 2020** or delivered earlier but further developd during that period. The number of unique Innovations (not reported as delivered In 2018 or before) delivered during the reporting period Is **274** (81 during 2019 and 193 during 2020) and 54 were reported as delivered in 2018 and just Improved during the reporting period.

Considering the Innovations reported as delivered in 2017 and 2018 **the cPPP has delivered 406 Innovations of exploitable value by end of 2020**. Projects have reported **162 Innovation of exploitable value to be delivered during 2021 or after** (38 projects of the cPPP are still running during the period 2021-2024)

Based on the reported information we can say that the partnership will deliver over **560 Innovations of exploitable value**. Figure 2 offers a visualisation of the evolution.

The decrease In value In between the year 2018 and 2019 Is due to the % of the sample. While 100% of the projects Intended to report on this KPI provided Information for the MR2018, only 60% have done It for the period 2019-20. With a higher reponse rate that decrease would mostly likely not exist.

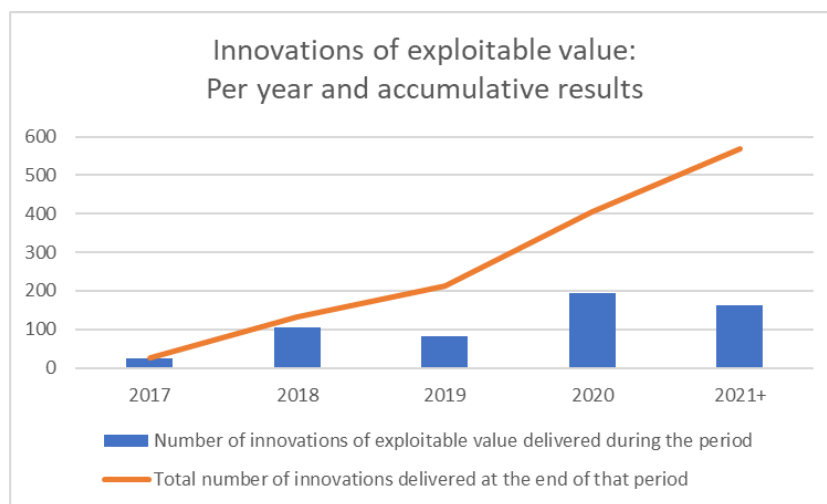


Figure 2: Innovations of exploitable value per year and accumulative results

Focused on the reporting period of 2019-2020, from the 328 innovations reported as delivered or enhanced during this period, **73%** have **medium impact** and **27%** are considered innovations of **significant impact**.

Considering 4 types of innovation (incremental, architectural, disruptive and Radical), **62,2%** of the innovations are **incremental innovations**, **12,5%** are **architectural**, **25%** are **disruptive innovations**, and **0,3%** is **radical innovation**). The unclassified Innovations have been considered as Incremental.

Innovations to Market	Total	Medium Impact	Significant impact	Incremental innovation	Architectural innovation	Disruptive Innovation	Radical Innovation
Reported/Delivered in 2018 but progress in 2019-2020	96	79	17	65	14	16	1
Delivered in 2019	39	23	16	19	2	18	0
Delivered in 2020	193	137	56	120	25	48	0
Expected in +2021	162	97	65	79	16	60	7
Total delivered during 2019-20 (reporting period)	328	239	89	204	41	82	1
Total expected 2021+	162	97	65	79	16	60	7

Table 9: Innovations for market 2019-20 classified by Type of innovation

**86% of the innovations** delivered or enhanced in 2019-20 have **economic impact** and **30% have societal impact**<sup>18</sup>. 11% of the innovations reported are not labelled under any of these categories.

Innovations to Market	Total	Economic impact	Societal impact	Other or not assessed
Reported/Delivered in 2018 but progress in 2019-2020	96	80	38	11
Delivered in 2019	39	37	18	2
Delivered in 2020	193	165	43	22
Expected in +2021	162	143	88	14
Total delivered during 2019-20 (reporting period)	328	282	99	35
Total expected 2021+	162	143	88	14

Table 10: Economic and Societal input of delivered innovations

55% are technologies (including platforms), 33% are systems, 15% are services, 14% products, 8% are methods, 8% are systems, 5% is software, 4% are components or/and modules, 2% are frameworks and architectures and 9% others, processes, tools and toolkits, spin-offs, datasets, ontologies, patents and knowledge.

Innovations to Market	Total	Framework / Architectures	Technology (including platforms)	Service	Product	method	system	Software	Components and modules	Others
Reported/Delivered in 2018 but progress in 2019-2020	96	3	14	25	13	10	5	9	8	18
Delivered in 2019	39	0	18	4	13	4	1	4	0	2
Delivered in 2020	193	3	149	19	20	13	102	4	4	11
Expected in +2021	162	0	54	82	39	15	30	1	0	16
Total delivered during 2019-20 (reporting period)	328	6	181	48	46	27	108	17	12	31
Total expected 2021+	162	0	54	82	39	15	30	1	0	16

Table 11: Innovations to market - Type of Innovation

<sup>18</sup> Note that many innovations are labelled both as having economic and societal impact.

### Sectors covered:

Innovations delivered by the partnership are providing solutions to many different sectors and application domains, demonstrating the cross-sectorial nature of the Big Data Value cPPP. Looking at the Innovations delivered or enhanced during 2019-2020:

- **27% to the Finance sector:** mainly Influenced by the Infinittech project that has reported quite a large number of exploitable assets during 2020, and start-ups funded by the EDI Incubator.
- **23% of the innovations are fully cross-sectorial**
- **12%** provide solutions to the **Transport, mobility and logistics sector** mainly delivered TT, DataPorts, QROWD and EDI projects
- **10%** relate to **health and healthcare** delivered by BigMedilytics, BodyPass, DeepHealth, Examode, and EDI but also but others such cloudbutton, DataPitch, DMS and Musketeer.
- **9%** to **industry and manufacturing** due to the lighthouse projects Boost4.0, IoTwins and some Innovations linked to the DataPitch Incubator
- **4%** of the innovations relate to **Public services and Smart Cities**
- **4%** to the **ICT market**
- **3%** to the **energy sector**
- **3%** to **commerce**
- **2%** are linked to **Marketing activities**
- **1%** to the **bio-economy**
- **5%** to **others** (including media, business services, and retail mainly)

Innovations to Market	Total	Cross-sectorial	Transport, mobility, logistics	Industry (manufacturing, production, automotive, ..)	Bio-economy	Health and healthcare	ICT	Marketing	geospatial	Public services and Smart Cities	Telecommunications	Commerce	Energy	Finance	Others
Reported/Delivered in 2018 but progress in 2019-2020	96	14	28	9	0	17	14	6	0	11	0	10	0	0	1
Delivered in 2019	39	17	1	11	1	1	0	0	0	1	0	0	3	1	3
Delivered in 2020	193	45	10	9	3	15	0	2	0	1	0	0	8	89	13
Expected in +2021	162	41	5	28	0	6	6	15	0	9	0	0	58	4	3
Total delivered during 2019-20 (reporting period)	328	76	39	29	4	33	14	8	0	13	0	10	11	90	17
Total expected 2021+	162	41	5	28	0	6	6	15	0	9	0	0	58	4	3

**Table 12:** Innovations to market – sectors

We can not draw any detailed conclusions on the percentages of coverage of the different sectors in terms of innovations of exploitable value as we are missing Information of over 40% of the projects however some general conclusions can be reached also looking at **Table 13** (so evolution over time from 2017):

- Cross-sectorial nature of around 15-25% of all Innovations of exploitable value delivered stimulating the use of Big Data technologies in many sectors of activity.
- Coverage of multiple sectors In the programme, not only with specific pilots addressing that vertical but delivering also sector specific assets
- The Impact of the lighthouse projects or large/close-to-market sectoral demonstrators over time, with solid coverage of those sectors.

Sectors (coverage of innovations per sector per reporting period)	2017	2018	2019-20	2021+
Cross-sectorial	15,4%	16,0%	23,2%	26,5%
Transport, mobility, logistics	84,6%	74,5%	11,9%	3,1%
Industry (manufacturing, production, automotive,...)		18,9%	8,8%	17,3%
Bio-economy		14,2%	1,2%	
Health and healthcare		7,5%	10,1%	3,7%
Finance			27,4%	2,5%
Energy			3,4%	35,8%
ICT		8,5%	4,3%	3,7%
Marketing		12,3%	2,4%	9,3%
geospatial		6,6%		
Public services and Smart Cities		19,8%	4,0%	5,6%
Telco		13,2%		
Commerce	23,1%	4,7%	3,0%	
Others		2,8%	5,2%	1,9%

**Table 13:** Innovations to market – sector: Percentage over the total per reporting period. Evolution.

In relation to the Maturity levels and TRLs, 2% of the innovations delivered are TRL 3 (experimental proof of concept), 10% are TRL 4 (technology validated in lab), 16% is TRL5 or TRL 6 (technology validated in relevant environment (industrially relevant environment in the case of key enabling technologies), or technology validated in relevant environment (industrially relevant environment in the case of key enabling technologies), 15% is TRL 7 (system prototype demonstration in operational environment), 4% is TRL 8 (system complete and qualified) and 1% TRL 9 (actual system proven in operational environment (competitive manufacturing in the case of key enabling technologies; or in space).

Innovations to Market	Total	TRL3	TRL4	TRL5/TRL6	TRL7	TRL8	TRL9
Reported/Delivered in 2018 but progress in 2019-2020	96	4	10	18	28	4	0
Delivered in 2019	39	0	7	7	0	6	2
Delivered in 2020	193	4	16	29	21	2	2
Expected in +2021	162	5	15	57	25	26	2
Total delivered during 2019-20 (reporting period)	328	8	33	54	49	12	4
Total expected 2021+	162	5	15	57	25	26	2

**Table 14:** Innovations to market – Maturity levels or TRLs

168 (over 50%) of the listed Innovations delivered or enhanced in 2019-20 did not Indicate any TRL level. Over 50% of the listed Innovations delivered or enhanced in 2019-20 did not Indicate any TRL level (not applicable as 45% of the innovations don't relate to know technologies) but indicated a maturity level of the innovation. Not reflected in the percentages above most of them were considered TRL 4-6 (so low-medium maturity level) and some (when indicated marketable ready) TRL 9% (high maturity model). This common view TRL + Maturity level Is Integrated in the overall picture of significant Innovations to market.

### Overall picture of the BDV cPPP Significant innovations to Market

Projects submitted their list of Innovation categorising them according to several criteria. Type of Innovation was one of them, split if 4 categories defined In the following way:

- Incremental Innovation: Improving existing technology to increase customer value (features, design changes, etc) within an existing market
- Disruptive Innovation: Applying a new technology or processes in an existing market
- Architectural Innovation: taking existing skills and technology and applying it in a different market
- Radical Innovation: New technology that creates new industries

Projects were also asked to self-assess If the Innovations delivered could have a significant Impact (regardless of the type of Innovation), based on the real Impact they would expect their Innovation to have. Projects were also asked to justify Innovations labelled as having a significant Impact.

Figure 3 provides the full overview of the innovations delivered by the cPPP during the period 2019-20 combining level of significance, type of innovation (incremental, disruptive, architectural and radical) and the TRLs. For the purpose of this exercise the Innovation to market not labelled with any TRL number are labelled as TRL 3-6 (low-medium) except on those cases where the project has specified in the time-to-market a marketable solution (and therefore a high TRL has been applied) or where TRL or time-to-market may be not relevant or clear (and in those cases the TRL value was left empty and not counted).

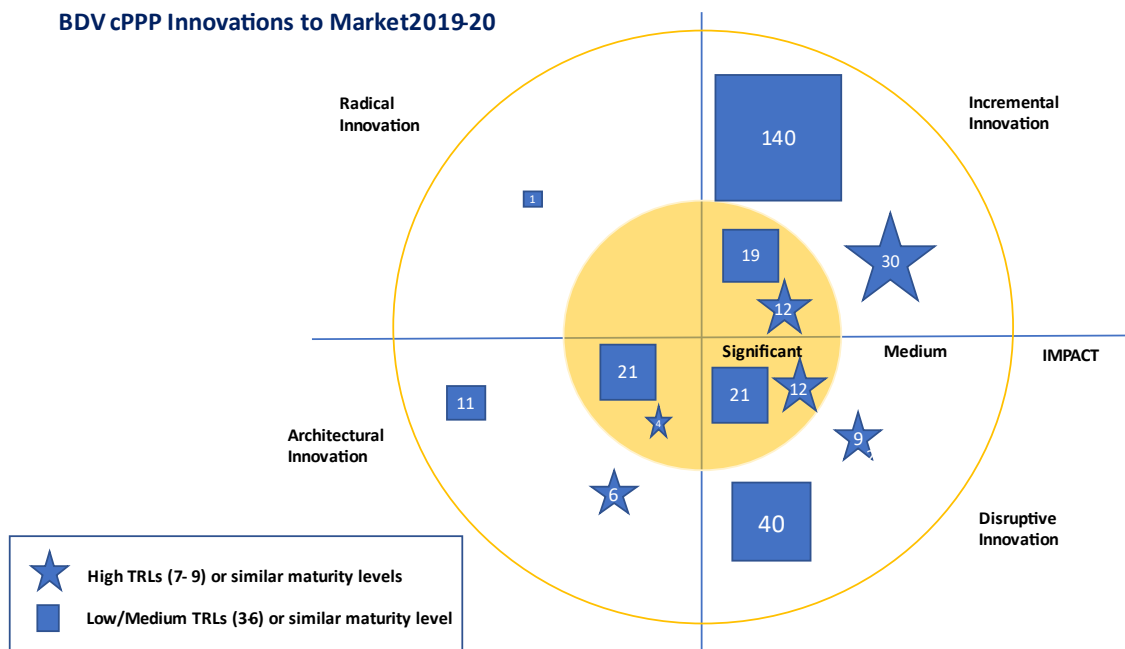


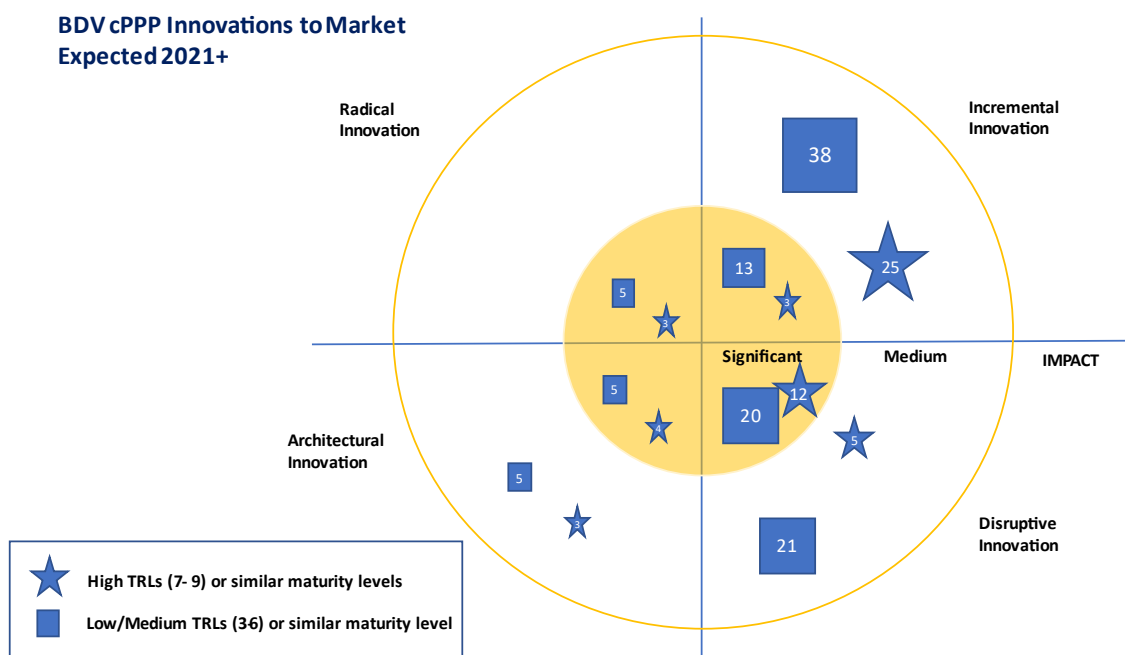
Figure 3: Innovations to market 2019-20

We can observe:

- Innovations of **significant impact** are equally distributed among Incremental, disruptive and architectural. and around a third of them have high TRLs or maturity levels: **9% of the Innovations delivered combine Impact and close-to-market characteristics at the same time.**
- The largest number of Innovations are classified as Incremental, with medium Impact and low-medium TRLs (43% of the total).

Figure 4 provides an overview of the expected evolution towards +2021. Following the general trend, it is important to highlight the expectations to deliver quite a large number of disruptive Innovations of significant Impact In the upcoming years.

For additional Information and details on the Innovations delivered or expected to be delivered by the projects the reader can consult the **Table 15** that provides all the raw materials as delivered by the projects.



**Figure 4:** BDV cPPP innovations expected to be delivered in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
<b>BD4OPEM</b>											
BD4OPEM	Impact study PV, EV & new loads	Service	Real time state estimation of one point of the low voltage network which is not monitored in real time	Significant	LV networks are highly polluted and subject to a large number of aggressions, new load and new small and large generation facilities. In this context specific simulation services for grid disturbances will be available across BD4OPEM as well as a service for the impact study of DER and loads in the LV networks	Economic and societal	5	Prototype or demonstrator	Energy	Disruptive	expected in +2021
BD4OPEM	Asset and investment planning	Service	Development of optimal investment strategies that contribute to the long-term planning using traditional assets combined along with flexible assets	Significant	Automated grid planning strategies to effectively support the decision making process in real life distribution systems planning. This service develops optimal investment strategies that contribute to long-term planning using traditional assets combined along with flexible assets to drive into an optimal decision-making, minimizing the capital expenditure and operational expenditure costs (CAPEX/OPEX).	Economic and societal	5	Prototype or demonstrator	Energy	Disruptive	expected in +2021
BD4OPEM	Asset estimation optimization for microgrids	Service	Optimal scheduling of a microgrid assets that determines the least-cost schedule of available resources while minimising the cost of distribution transformer loss of life	Medium		Economic and societal	5	Prototype or demonstrator	Energy	Incremental	expected in +2021
BD4OPEM	Topology	Service	Retrieval of the topology of the low voltage network and estimate grid parameters based on smart meter data	Significant	Enables to identify the LV grid topology from power and or voltage measurements from meters and with or without substation measurements. New functionalities referring to the distance between electrical components of the network will be evaluated for topology and potentially incorporated just in case it was feasible. These distances will allow to detect GIS inconsistencies	Economic and societal	6	Prototype or demonstrator	Energy	Disruptive	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
BD4OPEM	Observability	Service	Real time state estimation of one point of the low voltage network which is not monitored in real time	Significant	The estimation of the state of low voltage networks is of major relevance for distribution system operator. Historically, this was performed through load flow calculation. Thanks to the roll out of smart meters and the data they bring, new methods are used to estimate the state of the networks. New functionalities referring to the distance between electrical components of the network will be evaluated for observability and potentially incorporated just in case it was feasible. These distances will allow to detect GIS inconsistencies	Economic and societal	5	Prototype or demonstrator	Energy	Disruptive	expected in +2021
BD4OPEM	Measurement errors detection	Service	Identification, detection and eventually correction of anomalies, errors or missing values coming from data sources, in order to ensure data quality and usability	Medium		Economic and societal	5	Prototype or demonstrator	Energy	Incremental	expected in +2021
BD4OPEM	Inconsistencies in energy balance and power-voltage	Service	Location and quantification of the non-technical losses which occurs on a given low voltage network based on power-voltage inconsistencies	Medium		Economic and societal	5	Prototype or demonstrator	Energy	Incremental	expected in +2021
BD4OPEM	Fraud patterns detection	Service	Quantification of the non-technical losses which occurs on a given low voltage network based on energy balance inconsistencies and identification of the type of grid user who may be responsible of them in order to carry out a field inspection	Medium		Economic and societal	5	Prototype or demonstrator	Energy	Incremental	expected in +2021
BD4OPEM	Service KPIs	Service	Management and calculation of service KPIs to evaluate their performance	Medium		Economic and societal		Prototype or demonstrator	Energy	Disruptive	expected in +2021
BD4OPEM	Predictive maintenance in	Service	Prediction of the failure probability of different grid assets and their remaining useful life	Medium		Economic and societal	5	Prototype or demonstrator	Energy	Incremental	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
	electrical power system										
BD4OPEM	Grid disturbance simulations	Service	Identification of possible congestions or other PQ problems in a MV/LV distribution grid in long-term, day-ahead, or real-time scenarios	Medium		Economic and societal	5	Protoype or demonstrator	Energy	Disruptive	expected in +2021
BD4OPEM	Flexibility Forecast	Service	Aggregated flexibility prediction of consumers and DER resources flexibility in the system	Medium		Economic and societal	5	Protoype or demonstrator	Energy	Incremental	expected in +2021
BD4OPEM	Flexibility aggregated services for BRPs	Service	Flexibility request calculation to be used by the BRP for day-ahead and self-balancing optimization	Medium		Economic and societal	5	Protoype or demonstrator	Energy	Incremental	expected in +2021
BD4OPEM	Flexibility aggregated services for DSOs	Service	Flexibility request calculation to be used by the DSO for congestion management, grid capacity management and voltage control	Medium		Economic and societal	5	Protoype or demonstrator	Energy	Incremental	expected in +2021
BD4OPEM	EV to Grid	Service	Flexibility services provided to the DSO by an electric vehicle fleet with V2G capability in order to avoid grid congestions	Significant	The aim of this service perform flexibility services through bidirectional EV charging for the local DSO in order to avoid grid congestions. Flexibility from an EV fleet will be pooled in order to assist a balanced operation of the grid. As a result, this service can help to postpone and reduce the need of grid expansion investment.	Economic and societal	7	Protoype or demonstrator	Energy	Disruptive	expected in +2021
BD4OPEM	Energy management at household or at community level	Service	Energy management system to allow optimal scheduling of the electrical appliances at end-user premises to utilize their own flexibility behind-the-meter	Medium		Economic and societal	5	Protoype or demonstrator	Energy	Disruptive	expected in +2021
BD4OPEM	Energy forecasting	Service	Prediction of baseline demand, flexible assets consumption and generation curves at household or community level within a short-term horizon	Medium		Economic and societal	5	Protoype or demonstrator	Energy	Incremental	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
BD4OPEM	P2P trading	Service	Local trading of energy between end-users in a trusted and automated fashion by initiating respective smart contracts on the blockchain	Medium		Economic and societal	5	Prototype or demonstrator	Energy	Incremental	expected in +2021
BD4OPEM	Innovation platform	System	Enables large-scale multiparty data exchanges and real-time processing of energy-related data in smartdistribution grids. Will connect energystakeholders' needs directly with services to enhance energy asset management, promote consumer participation and create new datadriven business models and value chains for energy stakeholders.	Significant	The BD4OPEM platform will share data and provide Data Analytics services. It will represent a strategic Hub that collects all the information in a single toolbox to provide useful information and a list of interconnected services to key target groups. A big data architecture is designed to handle the ingestion, processing, and analysis of data that is too large or complex for traditional database systems. This will enable the design, development and deployment a marketplace for large-scale multi-party data exchanges, management, governance and real-time processing (including distributed/edge processing) of energy-related data in smart distribution grids.	Economic and Societal	3	Prototype or demonstrator	Energy	Architectural	expected in +2021
<b>BigMedilitycs</b>											
BigMedilitycs	Hospital workflow optimization	spin-off	Spin-off providing solution using real-time location information of tagged entities in a hospital (e.g. patients, assets, staff). Real-time location information is used to optimize workflows within a hospital	Significant	Can help improve the workflow efficiency of hospitals by up to 20%. That has a significant impact on operational costs, patient and staff satisfaction	Economic and Societal	8	demonstrator installed	Healthcare	Disruptive	Already reported in 2018 but progress in 2019, 2020
BigMedilitycs	Risk stratification for population health	methods	Methods to identify high risk patients in primary care	Medium	Will help reduce the no. of patients who end up unnecessarily in secondary care which is much more expensive	Economic and Societal	6	prototype	Healthcare	Incremental	Already reported in 2018 but progress in 2019, 2020
BigMedilitycs	Analytics for Prostate cancer	product	Solution to integrate diverse medical datasets for decision making for care of prostate cancer patients	Significant	Will help in better decision making when performing surgeries on prostate cancer patients. Can have an a significant impact on long term care of patients, e.g. avoid the need to spend	Economic and Societal	8	demonstrator installed	Healthcare	Disruptive	Already reported in 2018 but progress in 2019, 2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
					EUR50k on adult diapers over a 10-year period due to better surgery						
BigMedilitycs	Multiparty Computation (MPC)	technology	1) Multiparty Computation (MPC): a method/tool to analyse sensitive data from multiple parties (who do not wish to share data with each other) without revealing the data itself, but only revealing the outcome of the (data) analysis, i.e. the coefficients of the learned model.	Significant		Economic	6	prototype		Radical	Already reported in 2018 but progress in 2019, 2022
BigMedilitycs	Predictor for HF comorbidities	method	1) A LASSO based regression that allows for prediction of comorbidities of hospital admissions out of > 4.000 features in a Health Insurance database.	Significant		Economic	6	prototype		Architectural	Already reported in 2018 but progress in 2019, 2023
BigMedilitycs	PRO Heart Failure Revalidation	process	1) A smart watch based protocol to increase % of HF patients enrollment for Heart Failure Revalidation	Significant		Societal	6			Incremental	Already reported in 2018 but progress in 2019, 2024
BigMedilitycs	Adhere Heart Failure Revalidation	process	1) A smart watch based protocol to increase adherence of HF patients to increased exercise post Heart Failure Revalidation	Significant		Societal	6			Incremental	Already reported in 2018 but progress in 2019, 2025
BigMedilitycs	Predicting inpatient treatments of patient with known heart failure with health insurance data	Method	Using all the available data about for insured members with known heart failure with complex analytical procedures (machine learning) to predict the probability of inpatient treatment for each individual. When high risk patients can be identified, actions might be possible to prevent inpatient treatments	Medium		Economic and Societal	6	Medium	Healthcare	Incremental	Already reported in 2018 but progress in 2019, 2026
BigMedilitycs	Forta Score	Algorithm	The Forta score is based on an algorithm that can be computed based on routine clinical data and informs health care providers of	medium	Signals improvement actions to health care providers, requires clinical judgement to assess specific actions in individual patients	economic	7	Prototype in deployment	Health care	incremental	Already reported in 2018 but progress in 2019, 2027

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			potential over- and underuse of medications in persons aged 65 or older								
BigMedilitycs	SEARCH@MUW	service	on site deployed search system in a radiological routine environment. This enables the validation, testing of the search software as part of clinical routine	Medium	Initial deployment of the existing search engine, will be extended throughout project	Economic and Societal	8	demonstrator installed	healthcare	Architectural Innovation	Already reported in 2018 but progress in 2019, 2028
BigMedilitycs	GDM system (integrated solution)	System	Overall architecture design used for the project. It can be exported for any other medical solution, and to any other hospital	Significant	The hospital was previously using manual measurements. The time savings for the medical team, along with the reduce in human errors result in a significant impact	Economic and Societal	7	Prototype	Medicine	Disruptive	Already reported in 2018 but progress in 2019, 2029
BigMedilitycs	GDM mobile App	Product	Mobile App used by the patients to collect information. Can be exported to other hospitals, slightly modified for different use cases	Medium	Makes easier for the users to collect the data	Societal	7	Prototype	Medicine	Architectural	Already reported in 2018 but progress in 2019, 2030
BigMedilitycs	GDM medical portal	Product	Online dashboard where the medical team can monitor the patients. The design can be ported with minor modifications	Medium	Eases the medical team the monitoring of the patients	Economic	4	Prototype	Medicine	Architectural	Already reported in 2018 but progress in 2019, 2031
BigMedilitycs	GDM analytics	Service	Analytics for the process. It's a model based on real data that can be exported to other hospitals	Medium	Gives more personalized diagnosis, and helps medical team to adapt	Economic	4	Prototype	Medicine	Architectural	Already reported in 2018 but progress in 2019, 2032
BigMedilitycs	Biomedical Framework	technology	Framework for integrating and deploying medical imaging algorithms (specifically DL) into applications. The framework supports multi-modal analytics pipelines and exposes them as microservices in docker containers. The framework enables transparent run of polyglot pipelines on spark cluster by translating the descriptive pipeline flow to efficient spark application.	Medium		Economic		Supported by a prototype	Healthcare and Lifesciences	Incremental	Already reported in 2018 but progress in 2019, 2033
BODYPASS											

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
<b>BODYPASS</b>	Hyperledger 3D data exchange network	System	The Hyperledger 3D data exchange network is a blockchain-enabled framework to access to data from several providers from different sectors (tested in the health and consumer goods sectors). The framework allows to connect with scattered data sources such as databases (accessed using PRESTO) or APIs. Hyperledger allows to make off-chain calls to retrieve data objects, while managing the data access and auditing without storing any data from the providers.	Significant	This approach allows to break data silos among different organizations from the same or different sectors, allowing data exchange in a secure and traceable manner	Economic	7	Currently prototype	Health / Consumer goods	Incremental	Already reported in 2018 but progress in 2019
<b>BODYPASS</b>	Anonymization of face information in full body 3D scans	methods	This methods will enable to replace the actual face of a 3D full body scan by an anonymous one. Deep learning technology will be used to detect the face traits of the person in the raw 3D pointcloud or mesh. Different algorithms such as smoothing or data-driven alternatives will be used to anonymise or create the anonymous face content.	Medium	All the individual 3d data sets will follow this process to make possible the anonymous extraction of individual or aggregated data from BodyPass data pools.	Societal	7	prototype	Consumer goods /Health	Incremental	Already reported in 2018 but progress in 2019
<b>BOOST 4.0</b>											
<b>BOOST 4.0</b>	Digital Twin AGV	product	AGV virtualized simultaneously connected to the real AGV and to a simulation tool. Revenues generated by the sales of this new AGV type.	Significant	The connection with simulation tools and the use of Data Technologies will allow improvements of the models that are being used to calculate the number of AGVs for installation, as well as optimize the distance travelled by the AGVs, the energy, the material resources. The virtualization of the AGV will allow offer the AGVs as a service.	Economic	6	demonstrator	Industrial	Incremental	Already reported in 2018 but progress in 2019
<b>BOOST 4.0</b>	IDS connected AGV	product	AGVs connected to the IDS architecture. Revenues generated by data as a service pricing with data	Medium		Economic	5	demonstrator	Industrial	Incremental	Already reported in 2018 but progress in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
<b>BOOST 4.0</b>	RISC framework Data Analytics	technology, methods, process	Newly developed methods of big data processing as well as methods from the fields of data engineering and data visualization or special methods of machine learning are tested, extended and integrated.	Significant	The Boost 4.0 project enables RISC to deepen its knowledge of big data technologies and to further develop its existing expertise in data science and artificial intelligence. The business processes focus primarily on research and development so that the existing RISC framework Data Analytics can be further developed.	Economic		prototype	Industry	Architectural	Already reported in 2018 but progress in 2019
<b>BOOST 4.0</b>	IDS connector 4 Obelisk	technology	Hightly scalable and optimized data storage platform for distributed and heterogeneous sensor-based data sources	Medium			4	prototype, frequently used in research projects	smart city, Industry 4.0	architectural	Already reported in 2018 but progress in 2019
<b>BOOST 4.0</b>	UNINOVA Big Data Architecture	Architecture/Framework	The UNINOVA BDA enables fast and easy configuration and deployment of known open-source Big Data technologies via a containerized architecture in a multi-node environment	Medium		Economic	NA	Supported by a prototype/demo	ALL	Architectural	Already reported in 2018 but progress in 2019
<b>BOOST 4.0</b>	UNINOVA Data Harmonization / Transformation IDS Big Data App	service	IDS-based Big Data Application enabling Data Harmonization / Transformation using defined schemas	Medium		Economic	NA	Supported by a prototype/demo	ALL	Incremental	Already reported in 2018 but progress in 2019
<b>BOOST 4.0</b>	Hybrid Twin approach for casting process	methods		Medium			4			Disruptive	Already reported in 2018 but progress in 2019
<b>BOOST 4.0</b>	Interactive visualization of large series	technology	Ability to visualize, explore, manipulate large series of data	Medium			4		Any	Incremental	Already reported in 2018 but progress in 2019
<b>BOOST 4.0</b>	redBorder Active Industrial Network Monitoring Platform	product/service	Technical Data Analytics, and Cybersecurity Big Data solution that is scalable up to the needs of Industry 4.0 cyber physical systems and industrial networks. redborder is the only platform that blends data from	Medium	Deployment of protected communication chain from third party interference, where a series of "Industry 4.0 business rules" will be applied to trigger new events, alarms, or actions to prevent risks associated with the cooperation value	Economic	7	Supported by a prototype/demo	Manufacturing	incremental	Already reported in 2018 but progress in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			wired & wireless devices activity and can scale to millions of flows, logs, events, and users per second from thousands of devices without practical limits, from single-server to multi-site deployments. Crowd-sourced threat intelligence allows the platform to analyse, exploit, secure, and manage data from IIoT (Industrial Internet of Things) vendor protocols for making decisions in real time.		creation network established between companies related to a smart factory. Visualisation of real-time status of industry 4.0 assets and risk level based on big data stream monitoring						
<b>BOOST 4.0</b>	M3 Big Data 3D Point Cloud Analytic Algorithms	product	Multisensor massive measurement platforms and algorithms for control of extremely large data sets comprised of 3D point clouds beyond 10 million points. Capability for fast processing and geometrical feature extraction and worker friendly visual analytics of multiple regions of interest with heterogeneous point cloud density. Particularly suitable for large parts in chassis and body in white applications in automotive	Medium	Easily store and share your dimensional information with great efficiency. Capability to share multi million 3D point cloud information anywhere and anytime, allowing easy access from different departments. Optimized to gain space without losing information and smart management	Economic	6	first prototype	Manufacturing	Disruptive	Already reported in 2018 but progress in 2019
<b>BOOST 4.0</b>	IDSA TestLab	service	Facility and validation and verification framework for assessment and certification of IDSA component implementation	Medium	Ensure interoperability of machines and big data platforms used for implementation of AI algorithms based on industrial Data Space specifications	Economic	5	Developing	Manufacturing	incremental	Already reported in 2018 but progress in 2019
<b>BOOST 4.0</b>	IDSA connected CMM	product	Coordinate Measurement Machines connected to the IDSA architecture following QIF data model specifications. Revenues generated by data as a service pricing with data	Medium	Interoperability of machining and quality control equipment in the implementation of zero defect manufacturing workflows	Economic	4	demonstrator	Manufacturing	Incremental	Already reported in 2018 but progress in 2019
<b>BOOST 4.0</b>	AI Tube Platform	Product	Zero defect manufacturing IIoT and AI-based analytics and decision support system	Medium	Multi dashboard visualization system for decision support based on process-centric information	Economic	5	demonstrator	Manufacturing	Disruptive	Already reported in 2018 but progress in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
					analysis. Rule based and event management support						
<b>BOOST 4.0</b>	i2Track element	instrument/product	<p>i2Track is our way to demonstrate i2CAT's capabilities in positioning technologies. In this case, we showcase the application of the new DWM1001 UWB module developed by decaWave®. Like previous UWB implementations it can be used to provide high precision localization in indoor environments, but this time it goes one step further into accelerating the design, pilot and implementation phases, delivering a new approach to positioning projects. As size, weight and power consumption also shrink, the range of industrial applications for UWB positioning technologies is vastly widened. The design of the solution for this pilot and its use case, has been taking advantage of a development owned by i2CAT that is on the way to be valued as a tool or product that can evolve as a commercial solution. We call this product i2track and there is information published in brochure format and explanatory document.</p> <p>i2tracking UWB + RFID application is background of i2cat. The operational algorithm is fully developed by i2CAT</p>	Medium	<p>1.- 3D positioning: Unlike other indoor positioning methods such as RF based ones, i2Track provides the position in the X, Y and Z coordinates. This enables using this technology in height sensitive use cases.</p> <p>2.- Optimization of UWB anchor deployment: Leveraging the capabilities of the new chipset and i2CAT expertise in UWB location technologies, a Proof of Concept can be rapidly installed.</p> <p>3.- Low power: Combination of wake up radio via BLE and motion sensor. If the asset is not moving, the system does not positioning it. With the wake up BLE system, it can be activated from a distance for periodic checks.</p> <p>4.- Low cost: a basic Proof of Concept location system will have a total cost of no more than 200€</p> <p>The i2CAT Foundation continues to bet on the potential of UWB in combination with RFID identification tags to build precise and cost-effective location systems. With this integration, a simple passive RFID tag is enough to set the exact position of parts, assets and other materials where until now it was not costefficient to install a location system. We do it by installing a simple UWB location tag in the equipment, personnel or asset around the warehouse together with an</p>	Economic	6	The design of the solution for this pilot and its use case, has been taking advantage of a development owned by i2CAT that is on the way to be valued as a tool or product that can evolve as a commercial solution. We call this product i2track and there is information published in brochure format and explanatory document. i2tracking UWB + RFID application is background of i2cat. The operational	INDUSTRIAL	Incremental	Already reported in 2018 but progress in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
					active RFID antenna. The i2CAT algorithm determines whether the part with the RFID tag is being moved and the final resting location.			algorithm is fully developed by i2CAT			
<b>BOOST 4.0</b>	Process model simulation library	Product	Simulation library to enable fast concept to deployment configuration of manufacturing processes.	Significant	Supports to the user to configure quickly the production set up in the virtual environment avoiding errors and allowing easy reconfiguration	Economic	8	Demonstrator	ALL	incremental	expected in +2021
<b>BOOST 4.0</b>	AML support	Product	Support to handle big data sets between different engineering tools serialized in AML.	Medium	Seamlessly communication of big data sets between engineering tools removing errors and improving performance.	Economic	7	Demonstrator	ALL	incremental	expected in +2021
<b>BOOST 4.0</b>	Virtual Commissioning tools	Product	Adding support to the simulation environment for virtual commissioning.	Significant	Validation of the production configuration environment before deploying in the real production line. The validation allows to discover configuration errors from a very early stage, reducing errors to zero during runtime operations and accelerating ramp up production	Economic	8	Demonstrator	ALL	incremental	expected in +2021
<b>BOOST 4.0</b>	Communication interface	Product	Adding Fiware support to the simulation environment.	Medium	FIWARE emerges as a new communication middleware in the smart factory. FIWARE support in the digital twin targets the level 3, operation management and data visualization.	Economic	7	Demonstrator	ALL	incremental	expected in +2021
<b>BOOST 4.0</b>	Intralogistics library	Product	The simulation library provides intralogistics capabilities to easily configurate new logistics layouts in the virtual environment.	Significant	The library offers an easy to use solution with tools to model the intralogistics resources inside the factory.	Economic	7	Demonstrator	ALL	incremental	expected in +2021
<b>BOOST 4.0</b>	Autonomous AGV Interface	Product	Simulation interface which allows to develop and validate autonomous AGV systems in the Digital Twin.	Medium	Utilizing the digital twin for validation and deployment of AGVs in the factory accelerates its commissioning at the same time that the ramp up is more efficient removing operation errors.	Economic	6	Prototype	ALL	incremental	expected in +2021
<b>BOOST 4.0</b>	CERTH Cognitive Analytics Platform	system/service	A complete platform for data analysis and visualization. It provides various methods for supervised or unsupervised learning, data	Medium	High quality real time anomaly detection based on various ML and DL models with comparisons and metrics for the selection of best available	Economic	5	Supported by a prototype/demo	Industry	incremental	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			preprocessing and visualization. It is able to connect with real-time data by using IDS trusted connector. The platform trains various models based on data characteristics and selects the best one to run for anomaly detection based on some metrics and comparisons have been applied on trained models. Furthermore, retraining of models is another feature was adopted in order to keep the accuracy of predictions in high levels.		model. Boosting algorithms have been implemented in the top of learners for even better predictions. The cognitive functionalities with continuous retraining mechanisms triggered by significant changes in imported data or by decreasing predictions' performance, provide a high level of performance and accuracy of the predictions.						
<b>BOOST 4.0</b>	GF Predictive Maintenance Application	Product and Service	Predictive maintenance application based on specific ontology model and linked to production information system as well as field data, using machine learning algorithms and web technologies to deploy an application through the GF web portal for anticipating failures and services to GF customers	High	Predicting failures can have have significant cost savings as unprevented failures for critical components require interrupting production and replacing equipment which leads to cots of the order tof 100 KEUR	Economic and Societal	5	prototype	Manufacturing	Disruptive	expected in +2021
<b>BOOST 4.0</b>	GF Intelligent Spindle Production Dashboard	service	Information Dashboard providing production KPIs in real time and integrating machine learning for predicting quality issues whcih help to anticipate interruptions and iterations in the process	High	Iterations in the manufacturing of machines are due to failed quality controls at the end of the production process, the Dashboard provides transparency on this process as well as decision making tools for improving productivity and reduce costs	Economic and Societal	6	Advanced prototype	Manufacturing	Disruptive	expected in +2021
<b>BOOST 4.0</b>	FDT	service	Fault Detection Tool (FDT) is a combination of unsupervised fault detection approaches applied on streaming data.	Medium		Economic	7	supported by a prototype or demonstrator	Industry	Disruptive	expected in +2021
<b>BOOST 4.0</b>	FPT	service	Fault Prediction Tool (FPT) is a combination of unsupervised and supervised Machine Learning techniques for the prediction of upcoming/prominent failures.	Significant	FPT assists the adoption of the Predictive Maintenance to the Industry. The benefits are multiple: 1. Equipment downtime is decreased and the	Economic and Societal	6	supported by a prototype or demonstrator	Industry	Disruptive	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
					<ul style="list-style-type: none"> <li>number of major repairs are reduced</li> <li>2. Better conservation of assets and increased life expectancy of assets, thereby eliminating premature replacement of machinery and equipment</li> <li>3. Reduced overtime costs and more economical use of maintenance workers due to working on a scheduled basis instead of a crash basis to repair breakdowns</li> <li>4. Timely, routine repairs circumvent fewer large-scale repairs</li> <li>5. Improved safety and quality conditions for everyone</li> </ul>						
<b>BOOST 4.0</b>	DSS	product	Decision Support System (DSS) enhanced with asset health status tracking via expert users' defined rules, mobile notification system for recommendation of maintenance action, optimisation of the recommendations based on maintenance experts' feedback.	Medium		Economic	7	supported by a prototype or demonstrator	Industry	Incremental	expected in +2021
<b>BOOST 4.0</b>	Fusion	product	Fusion is responsible for the merge of the data analysis outputs from multiple prediction or detection tools, based on user provided strategies or machine learning algorithms.	Medium		Economic	6	supported by a prototype or demonstrator	Industry	Incremental	expected in +2021
<b>BOOST 4.0</b>	Causality Driven Modelling of Manufacturing Lines	technology	Probabilistic ML-based Simulation (real manufacturing data) that can help optimize production lines and uncover existings production problems thorough statistical data analysis	Medium	Many non Factory 4.0 production lines are performing close to optimal production. The developed ML based model helps answer 'what if' scenario in case of production reordering, production type change and supply/machine modifications applied to a line without need to perform desing of experiment and much cost.	Economic	4	demonstrator	Manufacturing	Disruptive	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
<b>BOOST 4.0</b>	Big data approach	Product	Big data approach in order to identify optimal process parameter which largely reduced leakage problems in the cast part.	Medium	Diminution of the porosity per twice	Economic	7	demonstrator	Manufacturing	Disruptive	expected in +2021
<b>BOOST 4.0</b>	Nemak Assistant Systems in Production	Technology	Use data analytics results for assistant systems in production to support operators in their daily work.	Significant	Data analytics is an evolving field in the industrial real world enabled through ongoing connectivity between systems. Production assistance systems exploiting the results of data analysis approaches will have a significant impact in existing manufacturing systems.	Economic	3	Concepts and Prototype	Production	Incremental	expected in +2021
<b>BOOST 4.0</b>	IDS Connector	Software	Adaptation of the Fraunhofer Trusted IDS Connector to an Industry 4.0 pilot site of Gestamp. The connector makes available data collected from plant floors by UWB and RDIF tags, with the objective of optimizing Gestamp's production and manufacturing processes. The connector was extended with the necessary technologies to meet the specific requirements of the use case.	Medium	The IDS connector is still being tested in PoCs and pilot sites across Europe. Our adaptation serves as further evidence that the connector can be used in Industry 4.0 environments.	Economic	4	Supported by a prototype	Industry 4.0, Data Spaces	Incremental	new 2019
<b>BOOST 4.0</b>	Blockchain applied in the supply chain of a manufacturing industry	Technology, methods, processes	Blockchain supports the exchange of information in a network of several participants with a high security. In this pilot, blockchain is applied in a network with several actors participating in a supply chain transporting goods for manufacturing. The	Medium	It is an application of already existing technology to a new area. The impact lies in the manufacturing sector to receive a more reliable governance of their supply chain of goods and the exchange of information.	Economic		Developing	Manufacturing industry, supply chains, logistics	Disruptive	new in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			technical implementation will enable a higher transparency of the exchange of goods and information between the involved partners and will ensure a real-time tracking in the supply chain. This will provide value in terms of reliability in receiving goods in time, more effective material handling to mention some.								
<b>BOOST 4.0</b>	Generic Data Model for Injection Molding Machines	technology	We devised a the machine-readable data model meant to capture environment and process variables of an injection molding machine. This model is built upon on top of the object model provided by the Open Platform Communications Unified Architecture (OPC-UA), along with two companion specifications: EUROMAP 83 (which encapsulates the information regarding production processes of plastic and rubber machines) and EUROMAP 77 (which refer specifically to injection moulding machines).	Medium		Economic	6	supported by a prototype or demonstrator	Manufacturing	Incremental	new in 2019
<b>BOOST 4.0</b>	BOOST4.0 Advanced Mapping Tool	technology	We developed a Web-based interface for mapping live streams of sensor readings coming from Injection molding machines, to entities and parameters defined in the Generic Data Model we designed. The interface also allows for dynamically filtering of the behaviour of the parameters, like monotonicity, incremental and/or between certain ranges. If all parameters are mapped to the data model, an output-file is	Significant	This mapping process typically involves 8+ hours of error-prone manual input, requiring both process engineers and IT experts to partake, now demands about one hour and can be conducted entirely by process engineers.	Economic	6	supported by a prototype or demonstrator	Manufacturing	Incremental	new in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			generated that can be included in the operator dashboard.								
<b>BOOST 4.0</b>	Quality Dashboard	Product	The Capvidia "Quality Dashboard" is web-based application for storage, monitoring, visualization and alerting system for quality control process. In case an inconsistent (out of tolerance) entity is detected in the quality control process, it will automatically trigger a message (alert) to a responsible stake holder using web devices such as notepad, smart phone, or laptop. The responsible person can "dig" into the data and investigate the reported problem in detail, using extensive graphical presentation projected on a 3D model. The Capvidia "Quality Dashboard" can be deployed on internal company server to ensure privacy and confidentiality of the data. The application is based on QIF ANSI/ISO standard, which makes it "open" and compatible with the MBD downstream and upstream data flows.	Significant	The implementation of the above described MBD based digital work flow resulted in diminishing the amount of documents used to support the process to only 1. In addition, the data, being fully digital, are now traceable and organised in a standard (QIF) data format, guarantying open access and transparency. They are also uniquely linked with the physical part, which significantly simplifies the repair and maintenance activities. Cost savings exceeding 60% are reported, minimization of errors and increase of product quality are other reported benefits.	Economic	8	demonstrator	Quality Control / Manufacturing	Disruptive	new in 2019
<b>BOOST 4.0</b>	Spare Parts Augmented Forecasting	product + method	The innovation is a system able to produce a very precise demand forecast for spare parts. Though the system is conceived around an appliances use cases (mainly washing machines), it can be applied to almost any industry (discrete manufacturing) either based or not on connected products.	Medium	Demand forecasting based on big data (several time series included)	Economic	8	demonstrator	Manufacturing	incremental	new in 2019
<b>BOOST 4.0</b>	From sheep to shop with a	Technology	Textile value chain is characterized by the production and treatment of unfinished goods in B2B relationships in a very	Medium-large	Ability to track and trace the textile supply chain and support the authentication of related documentation via a blockchain network that	Economic and societal	8	prototype supported by demonstrator	Industry and supply chain	Disruptive	new in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
	single source of truth		fragmented and distributed network. We implemented a blockchain based solution that records the transformation of raw materials into fabrics and enables verification of EU Preferential Certification of Origin (PCO).		assures transparency, provenance, and tamper proof of the data along with anon-repudiation process.						
<b>BOOST 4.0</b>	IDS Connector for Obelisk	technology	IDS allows for a formal definition of an exchange of data between parties. Obelisk is a multitenant time series data broker mainly used for IoT sensor values. With an IDS connector other parties that have an agreement for exchanging data can be given access to query the gathered data	Medium	The IDS interface allows for an easier exchange of data between parties	Economic		concept, the current IDS model is incompatible with the Obelisk API demands	All	incremental	new in 2020
<b>BOOST 4.0</b>	Cognitive AI & Robot control	Technology	Combination of symbolic + statistical approaches to information and control systems	Medium	Easy to use approach to semantic technologies for digital transformation inspired by work in the cognitive sciences	Economic	3	spec + demonstrator	Manufacturing	Disruptive	new in 2020
<b>BOOST 4.0</b>	FIWARE EIDS CIM REST Connector	product	Connect the FIWARE ecosystem of components to the EIDS/IDS	Significant	Each FIWARE member/user, SME, city, REST API service can easily bring it's data into EIDS data spaces. It is the first implementation of a HTTPS PubSub service. All existing and future architectures gravitating around the FIWARE Context Broker (also a CEF Building Block) can easily be integrated into data spaces, without substantial changes to the individual setup. With this set of features it will be easy and quick to setup distributed, federated and reactive networks of smart applications. The multi-datacenter, multi-cluster and multi-tenant	Economic	7	Supported by a prototype/demo	Any	Architectural	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
					approach enables companies and organizations of each size to participate with either own on premise installations or hiring into service provider models. This scalable solution offers quick start and long-running opportunities to help the EIDS/IDS gain traction in the market. Planned extensions into the direction of edge-integrations will enlarge this set of features.						
<b>BOOST 4.0</b>	TeraLab AI marketplace	Service	showcase and access to AI resources (data, processing, models,...)	Medium	Handle (add item, browse, download..) AI resources Features resources descriptors, rating, search, demo ... Coming: IDS connector		7	prototype	Any		new in 2020
<b>BOOST 4.0</b>	Casting Process Data Analysis Models	technology	Data Analysis models for casting processes support the identification of symptoms and causes of problems and provide hints for solutions through the identification of optimal process parameter. The models are exploitable in the industrial casting domain.	Medium	Increase of process transparency; identification of symptoms and causes for problems in casting processes; identification of optimal process parameters usable for casting process optimisations.	Economic	6	Prototype	Productions	incremental	new in 2020
<b>BOOST 4.0</b>	Edge Computing Platform NERVE	Product	NERVE edge computing platform is a solution for the fast transfer of machine data via edge devices to a cloud instance for later analytics on premise or on public cloud infrastructure. Nerve Blue offers hard real-time capability, container virtualization and the use of open technologies such as Linux, OPC-UA or MQTT.	Significant	The Nerve platform provides a software infrastructure for the plant floor and the cloud that enables users to access data, manage devices and deploy applications remotely. It provides features like: * Secure data connectivity on the machine, in the factory and to the cloud * Develop and deploy "smart services" * Easily install new applications and software update * Quickly and easily transfer real-time data from machines to their predictive maintenance	new 201+[[Impact type (Economic/Societal)]]9+B60:F65	7	Prototype, Demonstrator	Manufacturing	incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
					solution * Run analytics application at the edge of the network * Offer new IIoT services to customers for new and already installed machines						
<b>CLASS</b>											
CLASS	EXPRESS	SW component	EXPRESS is a novel framework for landing cluster applications onto cloud platforms. Used in CLASS for enhancing OpenWhisk[1] with real-time / low-latency capabilities.	Medium		none	TRL3	EXPRESS is considered as a technological novelty and is undergoing proper patenting by IBM	Edge & Cloud Computing	Disruptive	new in 2020
CLASS	Owperf	SW component	Owperf is a performance benchmark tool for OpenWhisk. It is a unique end-to-end deep profiling of OpenWhisk invocations and rules (triggers) resulting from cross-referencing observed timing (clock reading) with OpenWhisk internal timing data exposed through	Medium		None	TRL6	Supported by a prototype/demo nstrator	Cloud Computing	Disruptive	new in 2020
CLASS	Lithops	SW component	Lithops is a map/reduce engine that executes on top of serverless platforms. Both simple to use and platform independent (Python library[1]) and cost-effective (because of serverless) – only pays for actual computation cost, not cluster cost.	Medium		None	TRL7	Supported by a prototype/demo nstrator	Big Data/Data Processing	Incremental	new in 2020
CLASS	COMPs	SW component	COMPs is a framework to facilitate the development and execution of applications for distributed infrastructures. COMPs offers a simple task-based programming model based on sequential development in which	medium		None	TRL6	Supported by a prototype/demo nstrator	Distributed infrastructures (clusters, grids, edge/fog/cloud architecture)	Incremental	already reported in 2018 but progress in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			the user is mainly responsible for identifying t								
CLASS	dataClay	SW component	dataClay is a distributed storage framework that ensures user-defined data consistency, while enforcing the desired visibility scopes for each piece of data (at any granularity). In addition, dataClay enables the execution of code next to the data. By mov	medium		None	TRL4	Supported by a prototype/demo nstrator	Distributed infrastructures (clusters, grids, edge/fog/cloud architecture)	Incremental	already reported in 2018 but progress in 2019
CLASS	OTA	SW component	OTA is a module based on ImageAI[1] that provides an analysis of the trajectories and movements of objects detected by cameras to predict potential impacts.	medium		none	TRL7	Supported by a prototype/demo nstrator	Video data analysis, data visualization	Incremental	new in 2020
CLASS	SLALite	SW component	The SLALite is an automatic tool that manages Service Level Agreements (SLAs) in Edge environments. The tool creates the digital contracts for the micro services, capturing the availability metric through the Monitoring System, and assessing the contrac	medium		none	TRL5	Supported by a prototype/demo nstrator	Edge & Cloud Computing	Incremental	already reported in 2018 but progress in 2019
CLASS	TK Fusion	SW component	TK Fusion allows to detect, track and identify an object and notify it to the infrastructure owner based on Long Term Evolution (LTE) 4G network	Medium		none	TRL7	Supported by a prototype/demo nstrator	Advanced Driver-Assistance Systems (ADAS)	Incremental	already reported in 2018 but progress in 2019
CLASS	CLASS-Edge: Camera object detection and tracking	SW component	Camera object detection and tracking: Allow the detection and tracking from a single fix camera installed on top of public lighting. Allow the possibility to track the objects identified.	Medium		none	TRL6	Supported by a prototype/demo nstrator	Smart Cities	Incremental	already reported in 2018 but progress in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
CLASS	CLASS-aggregator: Control room data fusion and records	SW component	Control room data fusion and records (DKB): Merge the data from the various infrastructure cameras and vehicles notification and record them together.	Medium		none	TRL6	Supported by a prototype/demonstrator	Smart Cities	Incremental	already reported in 2018 but progress in 2019
CLASS	CLAS SW architecture ecosystem for distributing big-data workloads along the compute continuum while providing real-time guarantees	Technology	The development and execution of big-data analytics in a distributed computing environment composed of edge and cloud resources requires a redesign of big data system stacks, ranging from the programming model to the data-analytics runtimes. This innovati	significant	This innovation was selected by the European Commission's (EC) Innovation Radar among the most highly innovative EU projects. CLASS' innovation was assessed as having a "High" level of Market Creation Potential. Only innovations that are showing multiple	economic and societal		Supported by the demonstrator of CLASS projects results	All sectors with big-data analytics needs in real time scenarios	Disruptive	new in 2020
CLASS	Development of data analytics workloads combining data-in-motion and data-at-rest that benefit from distribution capabilities	process	This innovation leverages the distribution capabilities of the innovation #1 to develop novel big-data analytics techniques that combine low latency and high throughput requirements (as needed by reactive and latent analytics respectively) within a unified	Medium		none		Supported by the demonstrator of CLASS projects results	All sectors with big-data analytics needs in real time scenarios	Incremental	new in 2020
CLASS	A novel distributed sensing/computing infrastructure for the development of advanced urban	other	This innovation tackles the enhancement of a real urban laboratory of 1 Km2 from city of Modena (Italy) named Modena Automotive Smart Area (MASA). MASA incorporates an innovative sensing and computing network capable of collecting and processing large and	Medium		none		MASA ( Modena Automotive Smart Area)	smart cities	Incremental	already reported in 2018 but progress in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
	mobility applications with data analytics and real-time requirements										
<b>Cloudbutton</b>											
<b>Cloudbutton</b>	Faasm	technology	Faasm is a novel runtime that can execute application components that are expressed using the universal WebAssembly representation in a serverless cloud context. This runtime therefore can act as an execution mechanism for functions in a function-as-a-service (FaaS) serverless cloud environment.	Medium	Compared to existing execution engines, which are all container-based, Faasm enables higher efficiency by allowing a high density of functions to execute on the same hardware, while maintaining their isolation. In addition, Faasm avoids the cold-start problem, which plagues existing serverless platform, and it also allows functions to efficiently share state with other without serialization overheads.	Economic	4	Supported by a prototype	IT	Incremental	New in 2020
<b>Cloudbutton</b>	Lithops	product	Lithops (previously known as PyWren-IBM) is a framework to massively scale the execution of Python code and its dependencies on serverless computing platforms and monitor the results. Lithops delivers the user's code into the serverless platform without requiring knowledge of how functions are invoked and run.	Significant	Lithops provides great value for a variety of uses cases, like processing data in object storage, running embarrassingly parallel compute jobs (e.g. Monte-Carlo simulations), and enriching data with additional attributes. Lithops has been adopted in production by several IBM business units around the world. Foreseeable impact in the general Python and data analytics communities.	Economic and Societal	6	Prototype in real world environment	IT	Architectural	New in 2019
<b>Cloudbutton</b>	Serverless Metaspaces	system	Spatial metabolomics is a field of research and technology focused on small molecules, metabolites, drugs, and lipids. It is being applied in biology and medicine, drug discovery, and diagnostics. METASPACE is an open-source cloud platform for spatial metabolomics developed by EMBL. Together	Medium	Using serverless solutions based on Lithops framework and IBM Cloud Functions allowed EMBL to address key challenges in spatial metabolomics and annotate dark molecular matter that will have impact for our biological research and more broadly for various biomedical	Economic	7	Prototype in real world environment	Bioinformatics	Incremental	New in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			with IBM Research, EMBL has developed a serverless version of METASPACE that uses a novel type of cloud computing, namely "serverless computing" to execute METASPACE without installation, with minimal infrastructure maintenance effort, yet with almost unlimited scalability.		applications of hundreds of METASPACE users. Currently running in production on top of Lithops.						
<b>Cloudbutton</b>	Infinispan AnchoredKeys	technology	Infinispan is an in-memory data grid which can be scaled horizontally by provisioning additional storage nodes. For high-availability and balancing purposes, adding new nodes to a cluster has always required moving data around to ensure the required number of copies of each data item, thus temporarily decreasing throughput until the state transfer is complete. Anchored keys is an alternative implementation of the scaling algorithm which avoids costly state transfer when adding new nodes so that throughput is maintained. It does so by temporarily relaxing some of the rules concerning high-availability and balancing.	Medium	It allows horizontal scaling without state transfer for the Infinispan in-memory data grid.	Economic	7	Pilot, demonstration and testing activities	IT	Incremental	New in 2020
<b>Cross-CPP</b>											
<b>Cross-CPP</b>	Cross CPP Big Data Marketplace	System, Service	<ul style="list-style-type: none"> <li>Brand independent concept, open for integration of diverse CPP data providers coming from different industrial areas, also providing a standardized cross industrial CPP data model which needs to be flexible enough to incorporate data coming from various industrial sectors.</li> <li>CPP Big Data marketplace providing to</li> </ul>	Significant	Cross-CPP is a first attempt to provide a Marketplace with integrated data streams from products from various sectors; data analytics will be simplified by: single access point for service providers to integrated data sets; agreed data model where all data streams from various products will be transformed; toolbox with data analytics solutions suitable for cross sectorial	Economic		full prototype and demonstrator		Radical	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			<p>service providers a single CPP data access point with just one interface (one-stop-shop), as well as support functionalities for easy data min-ing/analytics. By these means, data customers (Service Providers) just need to set-up and main-tain one interface to gather diverse CPP data from different CPP providers.</p> <ul style="list-style-type: none"> <li>Controlled access to diverse CPP data streams and optimal management of data ownership and data rights, applicable to various cross CPP data streams.</li> </ul>		<p>services</p> <p>By CPP Data Marketplace with integrated data streams from various sectors will allow for more efficient building of more personalised services</p> <p>Cross-CPP will motivate CPP vendors to open many more data from their products to service providers</p>						
<b>Cross-CPP</b>	Cross CPP Big Data Analytics Toolbox	System, Service	<p>The Big data Analytics toolbox will provide easy to use big data analytic functionalities for Service Providers with low big data expertise and knowledge. It simplifies the extraction of the latter, by providing a set of libraries and modules designed to satisfy most data-related needs, and based on the most recent concepts and algorithms developed by the scientific community. It is buttressed by a modular structure, in which new analytics services can be added to fulfil new requirements; and in which multiple algorithms can be chained together, to give answer to even more complex questions.</p>	Significant	<p>provide easy to use big data analytic functionalities for Service Providers with low big data expertise and knowledge</p>	Economic		full prototype and demonstrator		Disruptive	expected in +2021
<b>Cross-CPP</b>	CIDM - Common Industrial Data Model	Data Model	<p>The Common Industrial Data Model (CIDM) is an advaced version of the CIVM Model developed in previous project AutoMat. This model deals with cross industrial data streams comming from variuos CPPs and</p>	Significant	<p>The CIDM model is uniform format under which data providers, i.e. CPP manufacturers shall provide CPP data. This would help to achieve uniformity over interfaces and implemented solutions with the raw data.</p>	Economic		developing standard, full prototype		Radical	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			offers uniform approach of representing them in common format. This model should be flexible enough to incorporate data from various industrial sectors.								
<b>Cross-CPP</b>	Cross Sectorial Services	Service	The Cross sectorial services are one of the main outcome of Cross CPP project. These services are using data from various CPPs and provide optimised solutions to their end users.	Significant	The Cross domain services will provide optimum solutions to end users with CPP data. The envisaged services are: different weather services including the development of a 100x100m city-scale weather model, Real time drivers warning and weather-based navigation, and e-mobility with charging and parking solution	Economic		full prototype and demonstrator		Radical	expected in +2021
<b>Data Market Services</b>											
<b>Data Market Services</b>	Mimirium Surveys	Product, SaaS	a SaaS platform for internal surveys for companies that need more engaging way for interaction combined with 100% anonymity of the votes and rewarding points system.	Medium	Development of a new product within the company which will increase Mimirium's revenue.	Economic	-	NA	Human Resources	Disruptive	new in 2019
<b>Data Market Services</b>	Geabit hardware product	Product	Product/device to continuously monitor and analyse plant health, development and stress.	Medium	Development of a new product within the company which will increase Geabit's revenue.	Economic	-	NA	Agriculture	Disruptive	new in 2019
<b>Data Market Services</b>	PactCare MVP	Product	Digital solutions	Medium	Redefinition of the MVP	Economic and Societal	-	Prototype	Healthcare	Disruptive	new in 2019
<b>Data Market Services</b>	Vesputi Analytics Software	Product. Analytics Software	Intelligent systems for networked mobility	Medium	Definition of a new service. Analytics software	Economic	-	NA	Mobility	Incremental	new in 2019
<b>Data Market Services</b>	HOPU Urban Design Tool	Product	Urban design tool based on data services	Medium	Definition of a new product to increase HOPU's revenue.	Economic	-	NA	Smart Cities	Disruptive	new in 2019
<b>Data Market Services</b>	VAY AG Solution	Product, software	Mobile and cloud solution	Medium	Definition of a new product to increase VAY AG's revenue.	Economic	-	NA	Fitness & Health	Disruptive	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
Data Market Services	Builtrix product	Product	New product to predict the consumption of energy for clients	Medium	Definition of a new product capable of measuring energy consumption to increase Infinite Foundry's revenue.	Economic and Societal	-	NA	Energy	Disruptive	new in 2020
Data Market Services	Infinite Foundry new product	Product	Application of technology to sports.	Medium	New product to increase Infinite Foundry's revenue	Economic	-	NA	Sports	Disruptive	new in 2020
Data Market Services	Ester digital solution	Product	New digital solution	Medium	New product to increase Infinite Ester's revenue	Economic	-	NA	Food	Disruptive	new in 2020
Data Market Services	Advanced Infrastructure analysis platform	Product	Data-as-a-service energy data analysis platform	Medium	New product to increase Infinite Advanced Infrastructure portfolio and revenue	Economic	-	NA	Energy	Disruptive	new in 2020
Data Market Services	RENVIS mobile app	Product	mobile app	Medium	New mobile app to increase companie's revenue	Economic	-	NA	Multi sector (data solutions)	Disruptive	new in 2020
Data Market Services	Dynargy cloud service	Service	Cloud service	Medium	New cloud service to increase the companies portfolio and revenue.	Economic	-	NA	Energy	Disruptive	new in 2020
Data Market Services	Cittamap service	Service	Market research, and customer data development	Medium	Customized service for market research and data analytics.	Economic	-	NA	Market analysis. Multi sector.	Disruptive	new in 2020
Data Market Services	Inventori Solutions LLC platform	Product	Not applied for IPR, web based platform	Medium	Web based platform	Economic	-	NA	Agriculture	Disruptive	new in 2020
Data Market Services	AIspotter Oy cloud service	Service	Cloud service, no IPR requests raised. Video AI analysis service	Medium	Cloud service	Economic	-	NA	Sports & Other sectors	Disruptive	new in 2020
<b>DataPitch</b>											
DataPitch	Virtualized data repository for fast analytics	Technology	This component creates a distributed tabular view of the involved data sources, specifically designed to enhance the performance for analytics operations. Such view of the data will be provided according	Significant	This software component can be used by Data Owners or Providers, to provide a unified view of the data available for analysis purposes.	Economic	5	Prototype	Technology	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			to the ontologies or semantic mechanisms defined in the context of the project.								
DataPitch	Stockwatch (OBUU)	Product	Stockwatch is a software solution developed by OBUU for the optimisation of spare part stocks for industrial machinery. During Data Pitch and their partnership with Greiner, OBUU developed the Stockwatch solution, and found that it could achieve a reduction of a fixed asset investment of at least a double-digit percentage for their client.	Medium		Economic		Marketable product, currently being licensed to clients	Manufacturing, aerospace, energy	Architectural	new in 2019
DataPitch	Bliq	Product	Bliq create live parking maps for developers in mobility. Blig provide these maps and APIs for developers to create products that improve mobility and reduce traffic congestion and air pollution. One current user base is for Bliq's solution to be built into car entertainment systems to show the driver the location of available parking spaces near them. Bliq are in multiple European countries, and looking to expand into the US and Japan.	Medium		Economic and Societal		Marketable product	Automotive	Incremental	new in 2019
DataPitch	Energeo	Product	Energeo has a SaaS platform that delivers sustainable intelligence to create a low carbon economy. The platform uses a variety of geospatial inputs and integrated proprietary data to identify potential energy solutions, for example helping a local authority decide which rooftops are ideal for solar panels to be most effective to reduce carbon emissions.	Significant	If scaled, this solution could have a significant impact to reduce carbon emissions.	Economic and Societal		Minimum Viable Product	Energy	Incremental	new in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
DataPitch	Guardian (Predina)	Product	Guardian is a software package and the final product made up of two solutions developed by the startup Predina as part of Data Pitch. The first of these solutions was an optimised algorithm which was used to predict the risk of road accidents. The second of these solutions was a data layer which would connect to a sat nav product; this product would then be able to alert drivers to the risk of road accidents. Both solutions came together in a web-based tool which enabled users to plan their routes, along with the expected safety of their route. Predina have contracts with construction companies in Europe, and are looking to expand to the US.	Significant	Six months after Data Pitch, Predina could demonstrate 25% in accidents as a result of the implementation of their solution.	Economic and Societal		Marketable product	Transport, manufacturing, automotive	Incremental	new in 2019
DataPitch	Radiobotics	Product	Radiobotics automate X-ray analysis for hospitals, specifically creating machine learning algorithms that could detect osteoarthritis on x-ray images of knees. Radiobotics aims to sell the solution to hospitals, with a current focus on European (Denmark, Sweden and UK specifically) and American clients.	Medium		Societal		Marketable product with CE-marking	Healthcare	Disruptive	already reported in 2018 but progress in 2019
DataPitch	Recognai	Product	Recogn.ai's solution uses deep learning to match records across different data sources and to create new classifications and categories. The solution can identify customers across transaction datasets and other related customer datasets. Their customers include businesses such as	Medium		Economic		Marketable product	Retail	Incremental	already reported in 2018 but progress in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			retailers, banks, financial institutions, insurers and utility providers.								
DataPitch	Heptasense	Product	Heptasense develops video analytics software to recognise risky behaviours, such as people carrying dangerous objects or to predict vehicle collisions. Heptasense has customers in 8 countries, and they include Vodafone, BMW and Mercedes-Benz.	Medium		Economic and Societal		Marketable product	Retail, automotive	Incremental	new in 2019
<b>DataPorts</b>											
DataPorts	Virtualized data repository for fast analytics	Technology	This component creates a distributed tabular view of the involved data sources, specifically designed to enhance the performance for analytics operations. Such view of the data will be provided according to the ontologies or semantic mechanisms defined in the context of the project.	Significant	This software component can be used by Data Owners or Providers, to provide a unified view of the data available for analysis purposes.	Economic	5	Prototype	Technology	Incremental	new in 2020
DataPorts	Data Access Component	Technology	Implementation of the Data Access component on the DataPorts Platform, including the framework and dockerization of the component.	Significant	This component can be used for easier access to the platform	None	5	Prototype	Technology	Incremental	new in 2020
DataPorts	Data Access Manager	Technology	Implementation of the Data Access Manager component on the DataPorts Platform.	Significant	This component can be used for easier access to the platform	None	5	Prototype	Technology	Incremental	
DataPorts	Blockchain Data Governance	Technology	One of DataPorts Platform building blocks is a Blockchain component for trusted sharing of information concerning ports operations. All Blockchain code and infrastructure.	Significant	This component will ensure trust in data sharing	None	4	Prototype	Technology	Incremental	new in 2020
DataPorts	Data Semantic Interoperability Component	Technology	Implementation of a framework for semantic interoperability of diverse data platforms, autonomous systems and applications. The framework will be based on the semantic	Medium		None	4	Prototype	Technology	Architectural	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			linking of data streams that will stem from different systems.								
<b>DataPorts</b>	Data Abstraction and Virtualization Component	Technology	Implementation of the Virtual Data Container concept to share Data as a Service.	Medium		None	5	Prototype	Technology	Architectural	new in 2020
<b>DataPorts</b>	Automatic Model Training Engine	Technology	This component creates a distributed tabular view of the involved data sources, specifically designed to enhance the performance for analytics operations. Such view of the data will be provided according to the ontologies or semantic mechanisms defined in the context of the project.	Significant	This software component can be used by Data Owners or Providers, to provide a unified view of the data available for analysis purposes.	Economic	6	Prototype	Technology	Architectural	new in 2020
<b>DataPorts</b>	Ports KPIs cognitive application	Technology	This cognitive application provides predictions mainly related to ports KPIs, such as a container ETA. The application uses a set of pre-trained analytical models, created using the DataPorts Prediction Engine, to generate predictions according to the end Users' preferences: target objective, data to consider, frequency to deliver predictions etc. Predictions are exposed via a subscribe API to be used by an external app such as a dashboard.	Significant	This component will provide predictions mainly related to ports KPIs	Economic and Societal	4	Prototype	Transport/Logistic	Architectural	new in 2020
<b>DataPorts</b>	Predictive services development based on the physical Smart container data	Technology	Development of value-added services including different alerts and predictive services such as ETA (Estimated Arrival Time).	Significant	This component will provide predictions mainly related to ports KPIs	Economic	5	Prototype	Transport/Logistic	Architectural	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
DataPorts	AI application for Thessaloniki UC	Technology	AI driven algorithms, libraries and methodologies. They address the data and visual analytics at Thessaloniki pilot, to provide with intelligent solutions for CERTH's service consumer.	Significant	This component will provide valuable analytics for the port	None	4	Prototype	Transport/Logistic	Architectural	new in 2020
DataPorts	Predictive Process Monitoring using Deep Learning Ensembles	Technology	This technique can be used to generate continuous predictions of ongoing business processes, together with reliability estimates for each prediction, thereby supporting the process operators (logistics manager, terminal operator etc.) in taking proactive decisions. Underlying technology is deep learning (recurrent neural networks).	Medium		Economic	4	Prototype	Technology	Architectural	new in 2020
DataPorts	Online reinforcement learning for self-adaptive business processes	Technology	This technique can be used to automatically take decisions whether or not to adapt a running process based on predictions generated from the above innovation. Underlying technology is reinforcement learning.	Medium		Economic	4		Technology	Architectural	new in 2020
DataPorts	Public catalogue of APIs for Smart Container related data sharing	Technology	Implementation of APIs to expose the smart container data via DataPorts Platform.	Medium		Economic	5	Catalogue	Transport/Logistic	Incremental	new in 2020
DataPorts	Data governance engine to protect the competitive advantages of the stakeholders	Technology	Credentials, rules and terms to be able to access the data via APIs.	Medium		None	4	Prototype	Technology	Incremental	new in 2020
DataPorts	Blockchain as a service	Technology	Software component for assigning permissions on datasets and data files. It will be available for all the partners who share	Significant	This component will ensure security	Economic	4	Prototype	Technology	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			on-chain and off-chain datasets in order to assign permissions.								
<b>DataPorts</b>	Visual Analytics Platform	Technology	Visual analytics platform for analysis and metrics, customized for the Thessaloniki port authority use case. Based on previous Analytics Platform.	Medium		Economic	4	Prototype	Technology	Incremental	new in 2020
<b>DataPorts</b>	DataPorts common data model	Technology	DataPorts common data model consisting in the metadata of the dataset.	Medium		Economic and Societal		Draft	Transport/Logistic	Architectural	new in 2020
<b>DeepHealth</b>											
<b>DeepHealth</b>	DeepHealth toolkit	technology	European open-source asset for health AI-based solutions: an European unified framework completely adapted to exploit underlying heterogeneous HPC, Big Data and cloud architectures assembled with Deep Learning and Computer Vision capabilities, ready to be integrated into any software platform that allows an easy and fast development and deployment of new applications for specific problems	Significant	Provide EU technology and toolkit in an ecosystem led by US and Asia. Increases productivity and facilitates daily work of data scientists since its unified DL&CV framework for health, modular and easy-to-use. Streamlines development of models and specific biomedical applications, maximizing their accuracy and minimizing time-to-market. With the flexibility of a multiplatform solution, DeepHealth toolkit makes transparent, optimized use of state-of-the-art HPC technologies	Economic and societal	TRL7-8	Supported by several prototypes and large scale pilots	Health, Others	Incremental	new in 2020
<b>DeepHealth</b>	European Deep Learning Library (EDDLL)	technology	Deep Learning library ready to run on Hybrid HPC + Big Data architectures. EDDL will provide all the functionalities of DNNs needed for medical image processing. Includes specific functions for running EDDL in distributed environments.	Significant	Same as in the toolkit	Economic and societal	TRL7-8	Supported by several prototypes and large scale pilots	Health, Others	Incremental	new in 2020
<b>DeepHealth</b>	European Computer Vision Library (ECVL)	technology	ECVL will provide the functionalities of image processing and computer vision needed to medical image processing,	Significant	Same as in the toolkit	Economic and societal	TRL7-8	Supported by several prototypes and	Health, Others	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			including data loading, adaptation and input to EDDL					large scale pilots			
DeepHealth	HPC infrastructure for DeepHealth	technology	Common framework consisting of a tailored SW platform to orchestrate the distributed and parallel execution on training and inference operations in heterogeneous HPC and cloud-based computing infrastructure featuring GPU, CPU, FPGAs	Medium	Allows portability fo workflows to any kind of HPC, cloud or hybrid infrastructure, enhances the usability, abstracts the parallel execution from the underlying infrastructure, making it transparent to the data scientist, and promotes a cloudified approach to HPC	Economic and societal	TRL5-6	Supported by several prototypes and large scale pilots	Health, Others	Incremental	expected in +2021
DeepHealth	7 improved platforms (biomedical and AI platforms)	technology, product	7 industrial and research platforms, provided by industrial and research partenrs enhanced integrating the DeepHealth libraries and adapted to solve different use-cases in the health domain	Medium	The creation of new biomedical applications to improve diagnosis, treatment and care of diseases	Economic and societal	TRL6-8	Prototypes and demonstrators developed in the project and tested with specific use-cases	Health	Incremental	expected in +2021
DeepHealth	Data-lake of medical images	Methods	Guidelines for the creation of data lakes of fully anonymized biomedical images, which purpose is to make available collections of images for the training of predictive models and future research in other projects	Medium	Availability of structured medical imaging data for AI purposes	Economic and societal	TRL8	Supported by its implementation in different use-cases, publication of the procedureç	Health	Architectural	expected in +2021
<b>EUHubs4Data</b>											
EUHubs4Data	Federated catalogue of data driven services and datasets	Platform, system, methods	One of the main outcomes of the project is the creation of a federated catalogue that brings together (and combine when possible) assets (services, datasets, trainings) coming from the DIHs composing the federation, and from other actors in the EUH4D ecosystem	Significant	The creation of this catalogue has impact in two dimensions: (i) at European level, it offers a unique reference in data driven innovation and experimentation, collecting novel solutions, services, etc ... from different stakeholders, providing a level of interconnection among them, and providing a one-stop-shop for all actors interesting in the adoption of data driven techniques, (ii) at regional level, increasing the	Economic		A first version of the catalogue has been published to be ready by the first open call for experiments in the project ( <a href="https://euhubs4">https://euhubs4</a>	All sectors	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
					offer of the respective DIHs and given access to the SMEs in their respective ecosystems to data driven solutions coming from Europe, but also increasing the usage of the services of the DIHs beyond their ecosystems			data.eu/). The catalogue will evolve during the project by: (i) enriching it with more assets coming from existing and new DIHs, (ii) refining the interface to improve its usability, (iii) providing different levels of interoperability, and (iv) creating local instances of the catalogue for each DIH to offer it at regional level			
<b>EUHubs4Data</b>	Model of interoperability and application to federated assets	Technology, methods, system	A model of interoperability for data driven services and datasets is being defined in the project to be applied to the federated catalogue (see above)	Significant	This model will increase the collaboration among DIHs in the federation, to combine services and to give access to datasets in a trusted data space. This model will also allow the project to implement its own data space, and to link with other data spaces, data platforms, and initiatives	Economic		An increasing level of interoperability will be provided during the project, and aligned with	All sectors	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
								each round of experiments			
<b>EUHubs4Data</b>	Organization and governance model of the federation of data driven innovation hubs	Model, methods, processes	A model that includes all aspects to be considered in the organization, management and governance of the European federation of data driven innovation hubs, to be adopted and applied during the project and which rule the evolution of the federation beyond the project, in a scalable, efficient and sustainable way.	Significant	The definition of this model is crucial to guarantee the sustainability of the federation beyond the project.	None		An outline of the main aspects to be considered in the model has been submitted as deliverable in Feb 2021. The model will be complemented and refined taking into account the rest of activities of the project. The objective is to end up with a complete federation handbook by the end of the project Dec 2023	All sectors	Incremental	new in 2020
<b>EW-Shopp</b>											
<b>EW-Shopp</b>	ABSTAT: knowledge graph summaries with ABstraction and STATistics	technology	ABSTAT is a tool to create semantic profiles of knowledge graphs. Semantic profiles describe structural properties of knowledge graphs and the vocabulary used therein. Semantic profiles are applied to support different tasks: data understanding and	Significant	ABSTAT is a first open source tool that allows users to generate knowledge graph profiles from a user interface. It is also the only tool that computes statistics about cardinality of schema patterns.	Economic and Societal	6	Main object of work for UNIMIB, aims at TRL 6 by December 2019	ICT	Disruptive	Already reported in 2018 but progress in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			exploration, data quality, vocabulary recommendation and analytics. ABSTAT is currently developed to profile RDF-based knowledge graphs but could be extended to profile knowledge graphs represented other data models (e.g., graph databases). ABSTAT is a free and open source software. Services on top of ABSTAT (deploy, customization, etc.) may be exploited.								
EW-Shopp	ASIA with Data Reconciliation and Extension Services	technology	ASIA is a semantic table annotation tool to generate knowledge graphs and enrich tabular data with third-party data sources. It interoperates with semantic reconciliation services, to reconcile values against shared vocabularies and systems of identifiers, e.g., Wikipedia, GeoNames, GFK Products, and data extension services, to fetch data from, e.g., GeoNames, ECMWF, GFK Products. On top of Grafterizer, ASIA supports end-to-end data cleansing, transformation and semantic enrichment at scale. ASIA is a free and open source software. Customization, e.g., deploy on big data architectures, require expertise that is marketed as additional services.	Significant	ASIA is the first tool that embed semantic table annotation functionalities in an end-to-end tool for data manipulation, semantic enrichment and generation of RDF data using a selected vocabulary/ontology. It is also the first tool to natively support the above mentioned functionalities for large volumes of tabular data.	Economic and Societal	4	A prototype of the solution is available for internal use within the consortium. Parts of the templates for container orchestration are published on the project GitHub page.	ICT	Architectural	Already reported in 2018 but progress in 2019
EW-Shopp	Big Data workflow tool based on container orchestration with DataGraft and Grafterizer	technology	The Big Data workflow tool provides clustered infrastructure for deployment, management and scaling of data workflows. The solution is based on Docker containers, a container orchestration system and a distributed file system. It will provide a high-level interface to handle and monitor data	Medium	The solution allows for easy deployment of big data processing workflows on existing hardware infrastructures. It has the potential to make big data processing more accessible to companies that lack technical knowledge about the topic.	Economic	3	a freely-accessible, open-source API with documentation used in business case pilots	ICT, Marketing, Commerce, CRM, IoT, BI	Architectural	Already reported in 2018 but progress in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			workflows, a processing engine for carrying out and scaling the workflow steps and an extensible library of composable steps, including an interface to ArangoDB, to be used as part of the data workflows.								
<b>EW-Shopp</b>	Weather data API	technology	The developed weather data API is a wrapper around the official European Centre for Medium-Range Weather Forecasts (ECMWF) data archive access API. It simplifies access and retrieval of weather data for non-meteorological purposes by simplifying the code for ease of use and limiting its scope to just the weather values likely to be useful in business applications (e.g. temperature or relative humidity) as opposed to those of greater interest to meteorology (e.g. albedo or temperature of lake surface). The API is only usable with a valid ECMWF API key which can only be obtained at ECMWF.	Medium	The solution eases the use of weather data in data analytics applications	Economic and Societal	8	prototypes being tested in a business environment	ICT, eCommerce	Disruptive	Already reported in 2018 but progress in 2019
<b>EW-Shopp</b>	Product Data Management Tools	technology	We developed two tools that use machine learning technology to automate different aspects of product data management. The first tool is the categorizer, which classifies products into existing (possibly hierarchical) taxonomies or categories based on their (non-structured) textual description. The second tool is the de-duplicator which identifies different textual descriptions of the same product and supports their reconciliation. Both of these are currently	Medium	The solution provides classification and feature extraction for product catalogs	Economic	7	a freely-accessible, open-source API with documentation used in business case pilots	ICT, Marketing, Commerce, CRM, IoT, BI	Disruptive	Already reported in 2018 but progress in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			highly labor-intensive tasks in online stores or comparison shopping engines.								
EW-Shopp	Event and Weather Analytics Toolset	technology	A set of tools and services built on top of the QMiner data analytics platform for building analytic models using events and weather data. These data features are joined with other data as external factors and their influence is modelled. An example of its application is modelling the effects of weather and major local events on sales in online stores. The tools support data ingestion, transformation, as well as model building and deployment (as a REST service).	Significant	The tool functionality can offer a significant competitive advantage to businesses in the e-commerce sector. By being specialized for its task it enables them to generate insights from their data faster than with general-purpose analytics platforms and the built-in prediction server allows rapid deployment to testing and production environments.	Economic and Societal	7	prototype tested in a lab environment	ICT, Marketing	Incremental	Already reported in 2018 but progress in 2019
EW-Shopp	Keyword Clustering Tool	technology	A tool for clustering of keywords and keyword phrases based on their semantics. The tool uses pre-computed word embeddings built using deep neural networks. For multi-word phrases the embeddings are aggregated using an approach based on Smooth Inverse Frequency (SIF) weighting. By using a metric in the embedding space the tool is able to find semantically related keywords for a given seed keyword. For example by using a category name (e.g. "Football") as the seed, the tool is able to find a set of keywords most likely belonging to this category (e.g. "football game", "referee", "FC Barcelona").	Medium	The solution helps dimensionality reduction in large-scale analysis of keyword-based marketing campaigns	Economic	5	Initial testing in the EW-Shopp ecosystem. Maturity will be reached at end of 2019	ICT, Marketing, Commerce, CRM, IoT, BI	Incremental	Already reported in 2018 but progress in 2019
EW-Shopp	Event Data Management for Data Analytics	methods	An ontology and a methodology to support event data exchange for event-based analytics.	Medium	The solution eases exchange of event data to support event-based analytics using and	Economic and Societal	4	Initial pilot service implemented	Marketing	Incremental	Already reported in 2018 but progress in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
					extending the widely adopted Schema.org reference ontology and JSON-LD			showing the adequance of the EWSHOPP approach and methodology			
<b>EW-Shopp</b>	Campaign BOOSTER	service	Digital marketing campaign impact strongly depends on the user behavior, which is directly affected by the weather conditions. In this scenario, JOT is developing a new service integrating weather forecast and marketing performance data enabling the implementation of 4 main functionalities: SCHEDULER, ALARM and IMPACT PREDICTOR (weather and event versions). They are based on keyword behaviour modelling (exploiting the campaign performance historical data base). This enables account managers to boost the impact of the campaign based on new insights generated by the correlation with the weather forecast and events and adapt the campaign bidding strategy accordingly, with the final goal of maximizing the number of impressions (as key impact factor).	Significant	The innovation created in EWSHOPP project will impact directly in the performance of the marketing campaigns, as it will allow the implementation of data driven campaign management including external factors not considered by the account managers so far. Due to the company business model, the better the performance indicators are, the more revenue is generated and more gross margin.	Economic and Societal	4	Applied as a methodology for the evaluation of all EW-Shopp business cases. Also implemented in BC3 as a supporting methodology	All possible sectors dealing with complex dataflows	Incremental	Already reported in 2018 but progress in 2019
<b>EW-Shopp</b>	EW-Shopp DataFlow Methodology	method / dataflow	This is a cross-domain, multi-actor methodology which enables to have a consistent workflow for data management and in particular multi-step dataflows. It concentrates on the logical steps (process) and 'actors' of a whole data management and sharing 'food chain' and provides a	Medium	Impact can vary depending on the data-sets at stake and actors (human or technical) involved. In some cases the impact could be significant in terms of rationalising existing complex flows.	None	n.a.	Initial pilot Performance Insights tool implemented and shown to selected partners to	Commerce and E-Commerce, Retailers, Vendors, Shoppers	Disruptive	Already reported in 2018 but progress in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			<p>general process for: 1) Data Integration; 2) Data Enrichment; 3) Data Analytics; 4) Data Visualisation. Additionally documentation can be enforced throughout all steps.</p> <p>The methodology is designed to allow/combine automatic, semi-automatic and human sub-flows and sub-processes as well as to support all phases of a dataflow starting from design up to implementation and running.</p> <p>Potential markets and exploitation scenarios are those where either an existing dataflow needs to be 'rationalised' (e.g. legacy dataflows) or ones where a new dataflow needs to be set-up.</p> <p>As it is the methodology is mostly logical and functional, i.e. not tied to any particular tool or technology. To this end this approach was chosen so as to keep it open to different uses. It could of course be formalised / adapted to- several existing dataflow tools.</p>					gather first impressions and feedbacks. Weather aware web site widget for shoppers implemented and tested in real live environment			
<b>EW-Shopp</b>	PERFORMANCE INSIGHTS: E-commerce Insights tool for shoppers, sellers and vendors	service	Performance Insights tool helps shoppers, sellers and vendors to use short term product search insights to support their smart decisions. The tool is build on top of years of user's uniquely integrated behavior data of searching products, weather data and different category events data. Shoppers will be able to use relevant purchase information	Significant	This solution will provide insights on the performance of products, brands and retailers based on the analysis of searches in a comparison shopping platform	Economic	3	Weather related prototype upgraded with Custom Event related-one, which has been build within a newly designed	Contact Centers, CRM and CIM systems offering services for different industries e.g.:	Disruptive	Already reported in 2018 but progress in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			and real time predictions to support their purchase decisions. Sellers will manage better pricing, category and campaign management and vendors get brand and SKU level purchase information in the shortest possible time to react and improve pricing, sales and campaign efficiency.					Custom Event Handling architecture and Custom Event API.	Telecommunications, Finance (Banking & Insurance), Commerce, Services & Tourism, Publishing, Energetics, Marketing & Research, ...		
<b>EW-Shopp</b>	Contact Center Workforce Optimization Manager	service	Digital marketing campaign impact strongly depends on the user behavior, which is directly affected by the weather conditions. In this scenario, Browsetel is developing a new service integrating weather data with existing Contact Center services KPI data, to get an innovative traffic prediction model. This will enable account managers to predict the adequate time to launch new campaigns and to better predict the needed Contact Center workforce.	Medium	The solution will influence on decreasing the probability of the Contact Center resources aggravation or saturation and increasing the probability of the call to be successfully served to the customer. Consequently, we expect improvements at "Call Response / Number of Agents / Customer Satisfaction" rate.	Economic	3	Initial pilot prototype developed. Full-fledged solution expected at the end of 2019	Marketing, Salesforce management (possibly also CM and SCM); Brand manufacturers /distributors/retailers in the segment of CE&Home Appliances.	Disruptive	Already reported in 2018 but progress in 2019
<b>EW-Shopp</b>	Weather and Event-aware BI Sales Strategy Advisor	service	Due to lack of systemized evaluation on how weather and events influence daily business, the service will enable to determine which type(s) of marketing activity(es) give(s) the most optimal results in given conditions; which type of event has more influence on demand/sales variables; how to manage activities on different customer	Medium	The solution will advise decision makers on how to exploit weather and events to boost product sales	Economic	3	Initial pilot prototype developed. Full-fledged solution expected at the end of 2019	Marketing, IoT, BI	Incremental	Already reported in 2018 but progress in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			touchpoints/channels and possibly how a price discount affects sales (on category level). Therefore thorough service we aim for optimization of Salesforce and resources within Marketing/Campaign and possibly also Category/Supply Chain Management.								
<b>EW-Shopp</b>	Event and Weather Scout Dashboard	service	Our dashboard helps businesses allocate their advertising budget based on quantitative historical analysis of customer flows vs seasonality/weather trends and internal events. We analyzed the number of people who visited a particular store in past periods of time and defined the impact of a particular event on people's behavior measuring the impact of marketing campaign/events that happened in the store. The results are visually presented on the dashboard in the form of easy understandable figures that give a clue to our clients about the effectiveness of past campaigns and weather temperature vs number of visitor relations.	Medium	The solution will provide Measurence Scout customers with the capability of understanding and evaluating the impact of events and weather on the visits in their stores	Economic	4				
<b>Examode</b>											
<b>Examode</b>	HistoGrapher	platform/framework	Histographer is a software platform for building holistic solution integrating multimodal histopathology data. This solution support histopathologists in making more informed decisions based on a larger amount of data (judging from similarity to other cases in their clinical practice or the identified likelihood in scientific literature). A	Medium		Societal	7	Demonstrator	Information Service activities Health and Wellness	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			<p>prioritization of cases is considered, so that the cases with higher severity and likelihood of specific diseases will be presented for confirmation first, and the cases with smaller likelihood – later, when there is enough time.</p> <p>Topic:Deep Tech Maturity:Business Ready Market creation potential:High</p>								
<b>Examode</b>	Platform for structuring diagnosis information based on ontologies and FHIR standard	platform/framework	<p>Platform for automatic structuring clinical text and concepts normalization to standard classifications and ontologies, like SNOMED CT.</p> <p>Topic:Deep Tech Maturity:Business Ready Market creation potential:High</p>	Medium		Societal	6	Demonstrator	Information Service activities Health and Wellness	Disruptive	new in 2020
<b>Examode</b>	VIRTUM HP - Cloud-based platform for whole slide image management and visual analysis	platform/framework	<p>This innovative solution is based on MicroscopeIT's proprietary Virtum Technology which includes, inter alia, pre-developed modules for storage, management, sharing, processing, and annotating image data, such as Whole Slide Images (WSIs) and other types of biomedical data.</p> <p>Virtum HP is a digital histopathology-dedicated cloud-based application with virtually unlimited storage and computational resources and can be used to build customized image-analysis applications with integrated AI models.</p> <p>Topic:Deep Tech</p>	Medium		Societal	7	Prototype	Information Service activities Health and Wellness	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			Maturity:Business Ready Market creation potential:High								
<b>ExtremeEarth</b>											
ExtremeEarth	Strabo 2	system	Distributed version of Strabon for scalable storing and querying of big linked geospatial data.	Medium	Strabo 2, the new distributed version of Strabon will be delivered as a demonstrator showcasing the scalable storing and querying of big linked geospatial data and enabling techniques for querying and analytics.	None	4	technology validated in lab	Technology	Incremental	new in 2020
ExtremeEarth	JedAI-spatial	system	Distributed version of JedAI for scalable interlinking of big linked geospatial data.	Medium	JedAI-spatial will be delivered as a demonstrator showcasing the scalable linking of geospatial data sources using geospatial relations.	None	4	technology validated in lab	Technology	Incremental	new in 2020
ExtremeEarth	GeoTriples-Spark	system	Distributed version of GeoTriples for scalable conversion of data into RDF.	Medium	GeoTriples-Spark, the Spark-based version of GeoTriples will be delivered as a demonstrator for converting data from various geospatial formats into RDF in a scalable manner.		4	technology validated in lab	Technology	Incremental	new in 2020
ExtremeEarth	Semagrow	system	Distributed and geospatial-enabled federated query processing, for scalable integration and processing of geospatial and relational linked data.	Significant	Extension, adaptation, and integration of the Semagrow federation engine and related technologies that facilitate the federation and unified access of geospatial data at very large scale in the Hops data platform. The resulting component is used to federate geospatial data sources in Strabo 2. SemaGrow is currently the first engine for federating big geospatial data sources.	None	8	system complete and qualified	Technology	Incremental	new in 2020
ExtremeEarth	Hops	technology	World's most scalable Hadoop platform with a distributed metadata architecture enabling high throughput, scalability and extreme geospatial analytics.	Medium	The new version of Hops supports federated data processing with support for discovering and sharing datasets between different platforms as well as managing EO data, including performance improvements for storing low-resolution satellite images and reduced cost for storage capacity.	Economic	9	actual system proven in operational environment	Technology	Incremental	new in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
ExtremeEarth	Hopsworks	technology	Data intensive AI platform, build on Hops, that provides users with a multi-tenant environment and the tools to develop and operate scalable end-to-end Deep Learning pipelines.	Significant	Hopsworks is the REST API and UI for Hops that integrates Hops, Spark, TensorFlow, and other services. It is extended to support arbitrary applications (not just data parallel Hadoop applications) with Kubernetes and Docker support. Application in languages like C, Julia, Perl are deployable on the new platform using both the UI and the updated REST API. We have also built out the model serving infrastructure for Deep Learning using Kubernetes to make it elastically scalable.	None	9	actual system proven in operational environment	Technology	Incremental	new in 2019
ExtremeEarth	CIRFA	service	Enhanced information extraction on sea ice characteristics from large scale datasets.	Significant	CIRFA has developed deep learning based sea ice analysis tools that also take into account the challenge of precisely labeling huge amount of sea ice data. Our tools combine collected labeled data with unlabeled data to improve the performance of sea ice classification. CIRFA also extended the system to distributed deep learning leveraging the computation power during the training stage.	Economic and Societal	9	actual system proven in operational environment	Technology	Incremental	new in 2020
ExtremeEarth	DeepArch	technology	Deep architecture tailored to the specific properties of Sentinel 2 data for large scale applications.	Significant	The deep architectures developed during the project are completely new and tailored to the specific properties of the Copernicus data.	None	7	system prototype demonstration in operational environment	Technology	Incremental	new in 2020
ExtremeEarth	TrainData	product	Publicly available large training database of multitemporal Sentinel 2 images for agricultural classification.	Medium	New approaches to extract reliable reference data from existing thematic products. The new version of the methods is defined to: (i) deal with different existing cartographic/thematic products, (ii) generate a very large database of labeled samples.	None	8	system complete and qualified	Technology	Incremental	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
ExtremeEarth	CuBOLD	service	Automatic detection of crops boundaries using long time series of Sentinel 2 images.	Medium	New methods for the automatic detection of crops boundaries using long time series of Sentinel 2 images at prototype level for local-scale applications. The methods will be improved and extended to large-scale applications and deployed in the food security case.	None	8	system complete and qualified	Technology	Incremental	expected in +2021
ExtremeEarth	EOBenchmark	system	Active learning tool for creating benchmarking data and/or semantic classification maps and statistical analytics. Input data Copernicus data and third Party data.	Medium	EOBenchmark is composed of several modules which offer functionalities such as ingestion and feature extraction from multispectral and SAR data, metadata extraction, semantic definition of the image content through machine learning and data mining methods, advanced querying of the image archives based on content, metadata and semantic categories, also for multitemporal datasets. EOBenchmark evolved from EOMiner (based on TerraSAR-X images) and is now able to generate large EO image benchmark datasets using different SAR and multispectral sensors (e.g., Copernicus and third-party sensors).	None	7	system prototype demonstration in operational environment	Technology	Incremental	expected in +2021
ExtremeEarth	EO DDN	technology	DDN architecture for charactering Copernicus data and extracting the relevant information.	Medium	EO DNN includes a hibird architecture. This architecture is based on LDA (latent Dirichlet allocation) and CNN (convolutional neural network).	None	4	technology validated in lab	Technology	Incremental	expected in +2021
ExtremeEarth	PCDSS	technology	Provision of data over intermittent and low band-width connections.	Significant	The PCDSS application requires better sea ice information to support the delivery of IMO Polar Code risk assessment information to ship users. ExtremeEarth outputs fit these requirements and will expand the PCDSS functionality and range of use.	None	6	technology demonstrated in relevant environment	Technology	Incremental	expected in +2021
ExtremeEarth	Semantic Catalogues	technology	Extend Copernicus data catalogue with semantics to provide annotation over EO	Significant	The updated semantics-based catalogues we developed in ExtremeEarth expose the	None		technology demonstrated in	Technology	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			data and expose knowledge hidden in satellite images.		knowledge hidden in Sentinel satellite images and related datasets, and allow users to ask sophisticated queries, which previously could not be answered, although the data to provide this knowledge is available in the Sentinel archive and related European datasets.			relevant environment			
<b>INFORE</b>											
INFORE	INFORE Streaming Extension	System	The INFORE Streaming Extension allows the graphical design of data processing workflows by the data analysts with zero programming required. The processing of the data workflows is executed over multiple available Big Data platforms.	Medium	The INFORE Streaming Extension dramatically reduces the time to program and to execute data processing workflows, as the only thing required is its graphical design.	Economic	5	Demonstrator & Released Open-Source	Business Intelligence	Disruptive	new in 2020
INFORE	INFORE Synopses Data Engine	System	The INFORE Synopses Data Engine can provide concise summaries of the massive, high-velocity input data streams it processes, using a Synopses-as-a-Service paradigm. These synopses help perform streaming data analytics and enhance the scalability of a system.	Medium	The INFORE Synopses Data Engine allows the execution of many complex data processing tasks in one data pass, thus significantly improving the performance and throughput of a system.	Economic	5	Demonstrator & Released Open-Source	Business Intelligence	Disruptive	new in 2020
<b>IoTwins</b>											
IoTwins	data ingestion workflows	Technology, service, process	Data ingestion workflows automatize the process of including datasets into the decision making process. Workflows make the data process easy-to-use, with no manual intervention and because of that, the data management become replicable and reproducible. Therefore, this methodology achieves significant benefits such as scalability (small and large datasets can be handled) and automatic data validation.	Medium	Reduce operation costs. Adhere reproducibility standards	Economic and Societal	TRL 6-8	validation phase	Crowd Management (sports + events), Urban Planning, Urban Mobility	Incremental	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			Automation of data management process allows the tool to be used in a wide variety of data science projects. In this way, the exploitable value relates to the generalization of the workflow to make it usable for data analysis projects, increasing the target audience.								
<b>IoTwins</b>	complex multi stage location simulation capability (PANDORA further development)	Technology/Service	<p>Pandora is a framework designed to create, execute and analyse agent-based models in high-performance computing environments. It has been programmed to allow the execution of large-scale agent-based simulations, and it is capable of dealing with thousands of agents developing complex actions.</p> <p>Regarding TB5, Pandora introduces several improvements related to natural path finding and agent synchronization, as well as agent configuration using the outcomes from the analysis on the previous section directly on the simulator configuration.</p> <p>The software is published under a Free Software License and the target audience is focused on the scientific community. By developing new features related to urban mobility IoTwins increases the potential adoption of Pandora by the urban planning research community.</p>	Significant	In particular, it provides the frameworks to simulate the mobility of around 100000 spectators of a sport event (TB5).	Economic and Societal	TRL 5-7	validation phase	Crowd Management (sports + events), Urban Planning, Urban Mobility	Disruptive	expected in +2021
<b>IoTwins</b>	ML models for profiling	Technology, service, process	Spectator Profiling from Data Analysis and Prediction: The main outcome are: the influx characterization (distribution of number of	Significant	Provide insights for spectator profiling in large amount of data. Improve user experience	Economic and Societal	TRL 5-7	validation phase	Crowd Management (sports +	Disruptive	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			visitors and their arrival times at different types of event), the spectator group sizes (distribution of number of people in groups that visit the venue), the length of stay prediction (the total time people will spend in the venue), the demographic attributes of visitors (particularly, gender and age group), and the origin of visitors (either local or tourist). The tool can be applied in a variety of sectors related to mass gathering						events), Urban Planning, Urban Mobility, Advertising, Social Studies, Public Management (Administrations/Cities), Retail		
<b>IoTwins</b>	Definition and use of ML models for mobility forecasting and predictive facility management	Technology, service, process	Using Machine learning models to understand human traffic flows and foresee behaviours and situations enables organizations to ensure a safe experience, optimize the operational management, maximize revenue and deliver a positive guest experience. The model can be applied to many different sectors that move or gather big amounts of people. But also, can be applied to smaller companies that will benefit of a process optimization through data and predictive modeling.	Significant	Reduce significantly the operations cost, improves safety and reduces the chances to suffer accidents due to accumulations or poor forecast, increases the chances to increase sales and improves user's satisfaction .	Economic and Societal	TRL 6, 8	Validation phase, currently proving on demonstrator, in order to get a very first comercial approach.	Crowd Management (sports + events), Urban Planning, Urban Mobility, Advertising, Social Studies, Public Management (Administration/Cities), Retail	Disruptive	expected in +2021
<b>IoTwins</b>	advanced cloud infrastructure management	Technology/Service	reuse skills and technology adopted for internal scientific communities to external ones, i.e. medicine, bioinformatics, biology, industry4.0 etc	Medium	provide high-value IT services to external communities	Societal	TRL 8	production IaaS/PaaS service	ICT service providers	Architectural	expected in +2021
<b>IoTwins</b>	advanced data management for	Technology/Service	reuse skills and technology adopted for internal scientific communities to external	Medium	provide high-value IT services to external communities	Societal	TRL 8	production data management services	ICT service providers	Architectural	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
	sensitive / personal data		ones, i.e. medicine, bioinformatics, biology, industry4.0 etc								
IoTwins	ML models for predictive maintainance (Data Centres)	Technology	apply ML techniques developed in other field to the datacenter predictive maintainance	Medium	improve data center availability, performance and efficiency	Economic and Societal	TRL 6	prototype applications	Data Centres owners	Incremental	expected in +2021
IoTwins	ML models for unsupervised anomaly detection	Technology	Detects anomalies in data streams, without need for human supervision.	Medium	Improve availability and decrease maintenance costs.	Economic	TRL 4	Prototype, demonstrator, real-world systems	manufacturing, cybersecurity and surveillance	Architectural	expected in +2021
IoTwins	ML models for time series forecasting	Technology, system	Provides predictions about future behaviors of a system, actions can be taken if the predicted future state is undesirable	Medium	Improve availability and decrease maintenance costs.	Economic	TRL 4	Prototype, demonstrator	manufacturing, cybersecurity and surveillance	Architectural	expected in +2021
IoTwins	Interpretability for ML	Technology, processes, methods	Allows the operator to interpret the predictions made by the machine learning system. The system provides an explanation that is interpretable by humans. Increases the confidence and trust of the operator in the system, which otherwise would be a black-box	Significant	Can have a wide range of applications and allows to speed up introduction of AI in real-world systems	Economic and Societal	TRL 4	Prototype, demonstrator	any sector	Radical	expected in +2021
IoTwins	ML models for predictive maintainance	Service & Technology	Systems simulation (introduction of defects in key components) ML Learning for time series analysis (TDA)	Medium		Economic and Societal	TRL 3 (2020 minimum of the various systems)	(parametric) Defects modeling is part of the Simulation-X commercial product. Times series analysis has been	wind turbines (and potential replicability in any sector where motor drives are required for movement of equipment (i.e.	Disruptive	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
								demonstrated at the laboratory scale only.	conveyer belts))		
IoTwins	Optimisation of the yaw drives control	Technology and processes.	<p>The market value of this innovation lies in the possibility for the customer to be able to operate his wind turbine with the lowest possible downtime due to unplanned service, to be able to identify operation conditions where wear and tear of yaw drives exceeds earnings from energy production and finally to extend the life of the yaw drives.</p> <p>Innovation is highly exploitable. It can be used directly on similar components in a wind turbine, such as pitch drives. With further efforts, innovation can be projected onto other wind turbine components such as gearbox, generator, etc. Furthermore, innovation can also be used in other industries where drives are used, such as conveyor belts.</p>	Significant	<p>It is a common perception in the wind industry, that reducing operating and service costs is the next step in the eternal pursuit of making wind energy even more competitive. Make it even more reliable, make it the cheapest and make it the preferred energy source in the world and hereby support the UN's Sustainable Development Goals to achieve a better and more sustainable future for all.</p>	Economic and Societal	TRL 4 – Technology validated in lab	<p>Estimation of life of yaw drives has already been proven on demonstrator while its final integration with the wind turbine data and application of digital twin related technologies is at prototype level.</p>	wind turbines (and potential replicability in any sector where motor drives are required for movement of equipment (i.e. conveyer belts))	Innovation will be applying new technology, and possible spin-off new processes and way off working, to current and future market of the wind industry.	expected in +2021
IoTwins	New multisensor Device	Product	<p>the market value of this innovation is represented by the possibility to sense in a sort of real time those variables that indicate the health status of the components of the wind turbine generator. This kind of continuous check can allow the user to prevent planning scheduled checks and reduce the cost of maintenance by tailoring on each WTG the maintenance planning</p>	Significant	<p>till today we rely on a scheduled maintenance and supposition based on the design factors and not on the field. this approach highlights the fact that meteorological condition or other not calculated variables are not included in a standard design, but having a system able to let you know what kind of possible anomalies happen it represents a strong innovation. In addition the possibility to open to modeling of system behaviors can</p>	Economic	TRL 4	<p>there are several prototypes manufactured but a field test is required in order to validate its maturity</p>	wind turbines (and potential replicability in any sector where motor drives are required for movement of equipment (i.e.	Disruptive	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
					revolutionize the way of how we design systems.				conveyer belts))		
IoTwins	Improved IoT Edge Architecture	Technology	<p>The three major improvements of the IoT Edge Architecture are its resource friendliness, the possibility to manage edge functionality together with the backend processes, and the possibility to use it on devices which are not always connected to the backend.</p> <p>By integrating resource restricted devices into an application management system for application on the cloud level, it becomes possible to manage functionality on the edge together and depending the functionality in the backend. In the backend ML training functionality can be deployed and started depending on the available data and quality of the currently used ML model. The resulting ML models and monitoring functionalities can be deployed to the edge depending on the training results of the backend processes.</p> <p>Because of its resource friendly operation mode, it becomes possible to use spare resources on existing filed hardware.</p> <p>Because of its ability to operate in settings where the connectzion to the backend is not always stabel it become possible to extend</p>	Significant	The integration of spare processing resources to monitor and control low voltage grids is on the one hand expected to enable the increased integration of functionalities like distributed renewable energy sources and energy communities. This is expected to reduced the dependmnce on centrally created energy and to reduce the amount of energy consumption because the new functionallities could be introduced. On the other hand it is expected that low voltage grids can be operated closer to their limits, which reduces the amount of necessary reconstruction works.	Economic and Societal	TRL 6 - Technology demonstrated in relevant environment (industrially relevant environment in the case of key enabling technologies)	Demonstrator	smart grids and building energy management	Incremental	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			the integrated management of edge devices to areas which poor connectivity and to reduce the usage of data connections to a minimum.								
IoTwins	Automatic training platform for optimized predictive smart grid operations	System, Processes, and Service	<p>The value of this innovation lies in the possibility to train measurement prediction values for measurement points for which measurements are only available time-delayed or on a non-sufficient resolution. Thereby the real measurements are taken to train a ML model on the cloud level and the resulting model is deployed to the edge device to predict the values and control the grid in accordance with the predicted values. A re-training is scheduled and executed once the quality of the predictions is no longer sufficient.</p>	Medium	Using the trainingsplatform it becomes possible to monitor and control the state of low voltage grids in which measurements are not directly available. This will reduce costs for sensor hardware, reduce transmission costs, and allow re-usage of low resolution smart meter measurements for higher resolution grid control.	Economic	TRL 4 – Technology validated in lab	Prototype	smart grids and building energy management	Incremental	expected in +2021
IoTwins	Scalable edge platforms	Technology	While software platforms for edge computing traditionally target industry PCs, there is an increasing trend towards a broader application of the underlying technologies to different kinds of devices. This includes small controllers (partially legacy devices) as well as powerful clusters for hyper-converged edge computing. As part of the IoTwins project, various technologies were evaluated and prototypes have been implemented that allow to better address the customers' needs and to open up new opportunities for edge computing.	Significant	There are several use cases at Siemens that require low-cost solutions for edge computing. In order to maintain competitiveness, pilot projects have been started that aim to address the technical challenges.	Economic	TRL 4 – Technology validated in lab	Prototype	Smart grids, factory automation, process industry	Incremental	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
IoTwins	Methodology for application of AI to edge platforms	System, Processes, and Service	Artificial intelligence helps to process the large amounts of data generated by field devices. However, running AI algorithms closed to the field is often a non-trivial task due to the diversity of hardware platforms used for edge computing. Moreover, the lifecycle of AI applications is more difficult as it includes additional steps - compared to pure cloud-based approaches - such as packaging, deployment, and monitoring. In the context of IoTwins, we address these issues and develop an integrated CI/CD pipeline with a focus on edge and IoT devices.	Medium		Economic	TRL 4 - Technology validated in lab	Demonstrator	Smart grids, factory automation, process industry	Incremental	expected in +2021
IoTwins	use of Digital Twins for production process optimisation through PaaS service	Services for improving production. Innovative business models.	The value of this innovation lies in the possibility of avoiding the current approach in finding the optimal combination of parameters that allow to control the process degradation that is lengthy, based on a try-and-catch strategy, which requires specific knowledge of the process, by replacing it with digital twins and neural networks that identify the optimal set of parameters based on the collected historical data	Significant	The benefits of the usage of digital twins lies into the possibility of testing the behaviour of configuration into a virtualized environment: this operation creates value because saves time and helps less qualified operators in the process of configuring the IoT machine	Economic	TRL 4 - Technology validated in lab	A demonstrator is the output of the TB12. The focus is validating innovative business models based on the IoTwins platform.	manufacturing sector (using numeric control machines)	Disruptive	expected in +2021
IoTwins	new knowledge on application of IoT technology and communication protocols to	Technology	Installation of sensors and plcs to machines to retrieve process and production data. With this innovation will be possible to obtain specific production information in relation to planned estimates	Significant	Accuracy of estimates	Economic	TRL 8	Prototype	manufacturing sector	Disruptive	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
	manufacturing operations										
<b>IoTwins</b>	adaptable ML models for predictive maintenance	Technology	Creation and validation of ML models to predict maintenance activities. This innovation has the purpose to decrease downtime due to unplanned maintenance and to increase production OEE value.	Significant	Increase of production level	Economic	TRL 3	Prototype	manufacturing sector	Disruptive	expected in +2021
<b>IoTwins</b>	Distributed control framework for manufacturing, bridging the gap between IT and OT.	technology, system, methods	The value of this innovation lies in the extension of IEC 61499 to provide a distributed real-time system for industrial control. The development of a generic approach is intended to consider currently available open source software and generic hardware to provide support for the real-time capability on distributed functions blocks. This effort extends the existing state of the art by providing the user a capability to schedule and execute hard real time function blocks on generic hardware.	Medium		Economic and Societal	TRL 3	Under development	manufacturing sector	Incremental	expected in +2021
<b>IoTwins</b>	ML models for anomaly detection (semi-supervised and supervised)	Technology	Identify anomalous situations in a target system (using historical, labeled data formatted as time-series); the technology is based on the usage of Deep Learning models (autoencoder neural networks)	Medium	Improve availability and decrease maintenance costs.	Economic	R4	Supported by demonstrator	manufacturing, large scale data centres	Incremental	new in 2020
<b>IoTwins</b>	Optimization service for hyperparameter fine-tuning	Technology	Automatically fine-tune the parameters which constitute a ML model (hyperparameter calibration); the fine-tuning is performed using Bayesian Optimization	Medium	Simplify the usage of other ML models	Economic	R4	Supported by demonstrator	manufacturing, large scale data centres	Incremental	new in 2019
<b>IoTwins</b>	methodology for translating industrial	Methods	Definition of a methodology and guidelines to help SME and industrial partners in the applications of AI-based solutions to specific	Medium	Support for industries and SMEs	Economic	n.a.	Under development	manufacturing, large scale data centres		expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
	requirements in Data Analysis and ML models and implementation of flexible and extensible templates of ML and Deep Learning models for industrial applications		domain problems (e.g., documentation of existing AI-based services, process to identify the most suitable ML model, definition of a pipeline of actions required to process the available data and apply ML models using those data, etc.)								
<b>IoTwins</b>	Data acquisition algorithm optimized for ball-bearings	Technology, methods	This data acquisition system will automate feature extraction for each part of the ball bearing, providing unique values to evaluate ball-bearing behavior	Medium	Automated measurement	Economic	TRL 8	Supported by demonstrator	Manufacturing sector	Incremental	expected in +2021
<b>IoTwins</b>	RUL estimation for ball-bearings based on novel machine learning algorithms	Technology, system, methods, product, service	A new algorithm based on HMM is able to estimate the ball-bearing remaining useful life (RUL) without a preliminary training. This is important as real-world ball-bearings are subject to different operating conditions that might affect the RUL estimation. In this case, the algorithm will learn during operation and will provide a health index that can be used to determine ball-bearing maintenance	Significant	Reduced machines unexpected downtimes. This is due to the RUL estimation that can help to plan the ball-bearing maintenance before unexpected failure occurs	Economic and Societal	TRL 4	Under development	Manufacturing sector	Disruptive	expected in +2021
<b>IoTwins</b>	Algorithms and procedures for data mining and simulating physical systems	Technology, methods	A new algorithm based on Topological Data Analysis (TDA) that allows extracting concise, complete and compact descriptors of time series, a sort of QR-code (fingerprint or identity card) that profits the invariance features of topology-based metrics. The	Significant	Allows real time data analysis and learning in the very low data limit. It can be applied in fault identification, anticipation, classification, in many domains.	Economic and Societal	TRL 4	First tests successfully carried out	manufacturing, transport, energy	Disruptive	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			resulting persistence images allows to compare data, to create regressions, ... and can be used for anticipating faults								
IoTwins	Algorithms and procedures for predicting spindle behaviour	Technology, methods, product, service	New algorithms based on machine learning for predicting the spindle behaviour are developed. The analyses and the services based on them enable the improvement of the component quality as well as the shortening of the cycle time for a resource-saving production.	Significant	Improved component quality, reduced cycletime, reduced downtimes	Economic	TRL 5	Under development	FILL machine tools (and potential replicability in any sector where spindle drives are for component production in terms of machining)	Disruptive	expected in +2021
IoTwins	Experience in Edge-/Cloud-Platform design for advanced service distribution	methods	The main value of this innovation lies in transfer of knowledge to other fields of application platforms, such as 5G Core Network virtualization.	Medium		Economic	n.a.	Under development	Industrial IoT and 5G network virtualization	Incremental	expected in +2021
IoTwins	integrated HPC/cloud infrastructure management	Technology, methods	An orchestration workflow letting HPC-based cloud resources interwork with edge resources and IoT devices for implementing our distributed digital twins.	Medium		Economic	TRL 4 - Technology validated in lab	Under development	HPC services providers	Incremental	expected in +2021
IoTwins	Datacenter digital twin	Technology/Service	A tool for: (i) estimating the operation cost reduction achievable by energy minimization; (ii) demonstrating the operation cost reduction by combining	Significant	We are not aware of similar tools	Economic	TRL 6 - Technology	Demonstrator	ICT providers	Incremental	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			energy-to-solution minimization and proactive maintenance; and (iii) developing digital twins to minimize cooling costs without reducing IT performance or increasing hardware failures.				demonstrated in relevant environment (industrially relevant environment in the case of key enabling technologies)				
<b>IoTwins</b>	Toolkit of edge node data services	Services	Repository of configurable services to choose from in order to build a data management solution on the Edge	Medium		Economic	TRL 4 – Technology validated in lab	Under development	Industrial IoT, Smart Cities, Smart transportation, Logistics	Architectural	expected in +2021
<b>IoTwins</b>	Methodology to design and deploy data	Methods	Guidelines and best-practices useful to practitioners to build, set-up, run and monitor data services in Edge nodes	Medium		Societal	n.a.	Under development	Industrial IoT, Smart Cities, Smart	Architectural	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
	management services in Edge environments								transportation, Logistics		
<b>IoTwins</b>	Application creation for Smart Manufacturing, ML for Smart Manufacturing, Physical modeling of machining	Service & Product	<p>1/ Collect data</p> <p>2/Analyse data (ML) for predictive maintenance &amp; quality control (surface quality, dimensional control)</p> <p>3/ Physics based simulation of the machining process itself; provides information on cutting forces, vibrations, surface quality. Simulation leads to the creation of a reduced model. Reduced model that can also be used in combination with data for monitoring.</p> <p>4/ Adequate Real-time advanced monitoring application</p>	Medium		Economic and Societal	TRL 3 (2020 minimum of the system)	<p>1/ Data collection system to be completed in the next 3 months</p> <p>2/Data analytics available in commercial product</p> <p>3/Machining Physics-bases Simulation tool 'ESI Machining' marketed today.</p> <p>4/ Adequate Real-time advanced monitoring application : specifications</p>	Machining Aero / Auto / Energy	Disruptive	new in 2020
<b>Kraken</b>											
<b>Kraken</b>	Personal local data wallet	Product	Several products exist in the market, even in the SSI arena. They however do not integrate with cloud backup facilities, nor they support the consent flow.	Significant	Change of paradigm for identity management providing whole control to the data user	Economic and Societal	7	Developing	Data sharing	Disruptive	expected in +2021
<b>Kraken</b>	Privacy protected cloud data wallet	Service	Based on CREDENTIAL project deliverables for privacy protected storage. It supports	Significant	Change of paradigm for identity management providing whole control to the data user	Economic and Societal	7	Developing	Data sharing	Disruptive	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			close integration and synchronization with a local data wallet.								
<b>Kraken</b>	SDK for VC and consented flow	Product	Data producers, including eIDAS identity providers, eIDAS nodes or generic ID brokers can leverage on this component to produce verifiable credentials It also provides data issuers with the capability to issue VCs signed with a qualified digital signature.	Significant	Change of paradigm for identity management providing whole control to the data user	Economic and Societal	7	Developing	Data sharing	Disruptive	expected in +2021
<b>Kraken</b>	SSI authentication components	Technology	SSI authentication leveraging on identifying credentials is core to usage of marketplace and access to services.	Significant	Change of paradigm for identity management providing whole control to the data user	Economic and Societal	7	Developing	identity	Disruptive	expected in +2021
<b>Kraken</b>	Integrated marketplace and underlying infrastructure	Product	Commercially active system serving customers globally containing publicly available wellbeing data streams. A generalised infrastructure of its technology stack will be replicated for the pilots and updated with consent management solutions for private data.	Medium		Economic and Societal	8	Improvement	Data sharing	Incremental	expected in +2021
<b>Kraken</b>	Enhanced MHMD platform	Product	Publicly available, fully GDPR compliant research data platform currently deployed in 4 EU hospitals and accessed by various research. It will be equipped with the Streamr commercial transactions functionalities.	Medium		Economic and Societal	8	Improvement	Data sharing	Incremental	expected in +2021
<b>Kraken</b>	Cryptographically secure privacy-preserving data analytics	Technology	Well-tested high-quality implementations of cryptographic libraries are a key enabler of the KRAKEN ambition. In particular, highly flexible analysis on encrypted data needs to be available and brought to a high technological readiness level.	Significant	Protecting privacy when processing data	Economic and Societal	6	Based on patent by SIC partner	Data sharing	Architectural	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
<b>Kraken</b>	Data analytics on authentic data	Technology	Analytics on authentic data allow to not only use a data market for self-claimed data, but also for such data that has previously been signed, e.g., by national authorities, and giving the recipient high authenticity guarantees while still maintaining the data subjects' privacy.	Significant	Protecting privacy when processing data	Economic and Societal	5	Based on patent by SIC partner	Data sharing	Architectural	expected in +2021
<b>LEXIS</b>											
<b>LEXIS</b>	Accounting and billing engine for HPC centers	Product	Adaptation of the Cyclops flexible billing engine for use in HPC context, including providing support to the HPC resource approval process, support for credits management, converged invoicing of HPC + Cloud resources in an on-demand, continuous manner. The innovation is close to be on the Market, but a BP is needed.			Economic	near to 8		HPC-BD and CLOUD	Incremental	expected in +2021
<b>LEXIS</b>	Data Infrastructure based on iRODS/EUDAT-B2SAFE & REST APIs	Technology-Service	In LEXIS, a secure and reliable Distributed Data Infrastructure (DDI), federating IT4I, LRZ and prospectively further data centres has been developed on the basis of iRODS and EUDAT-B2SAFE. We provide installation recipes for our redundant iRODS setup, for future members to join the LEXIS DDI or also for third parties to build similar data infrastructures. Data on the DDI is managed via HTTP REST APIs, facilitating e.g. data up-and download, listing (including metadata search), staging and rights management. The			Economic and Societal	near to 8		HPC-BD and CLOUD	Disruptive	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			code for these APIs, based on open-source toolkits, is provided as well.								
LEXIS	Improving Orchestration in distributed HPC infrastructures for HPC/Cloud/BigData workflow management	Technology-Service	LEXIS orchestration service is a key technological element enabling the full exploitation of the distributed HPC and Cloud resources (IT4I and LRZ, potentially including other future centers). To this end, the orchestrator has been developed around a set of technologies (YSTIA, HEAppE, OpenStack), whose integration allows to seamlessly execute users' workflows on the distributed resources, taking advantage of the DDI. Furthermore, LEXIS orchestrator, in its final form, will leverage on a dynamic mechanism to select the most suitable resources for the execution of the workflow tasks.			Economic and Societal	near to 8		HPC-BD and CLOUD	Incremental	expected in +2021
LEXIS	LEXIS Platform for SMEs, Industry & Society	Technology-Service	Specialized platform aimed at lowering the entry barriers to the worlds of supercomputing, cloud, and big data without sacrificing the execution performance is being created. Platform includes: dynamic data-aware orchestration of complex workflows; data sharing between Cloud and HPC resources and distributed data management with an appropriate data backend to the orchestration solutions; access to HPC/BigData/Cloud resources for SMEs and Industry, and; control of workflows and data through a web interface, with remote visualization services.		The Lexis platform is a specialized platform aimed at lowering the entry barriers to the worlds of supercomputing, cloud, and big data without sacrificing the execution performance is being created and will allow access to HPC infrastructure to who needs huge amount of resources, in particular SMEs at accessible costs.	Economic and Societal	near to 8		HPC-BD and CLOUD	Disruptive	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
LEXIS	Novel CFD numerical solver for design/analysis of Turbomachinery operating in Aeronautical engines	Product	Nowadays, thanks to GPUs introduction, SW application evolution is opening again to a new family of solvers with predictive capabilities more attractive for the aeronautical products. Lexis is porting the CFD solver "TRAF" from a solely CPU-based release to a more innovative GPU-enabled one. Research steps have been various, we can cite: adaptation/modification of data structures and re-writing of subroutines/functions, profiling aimed at identifying application bottlenecks, optimization of CPU-GPU data exchange/communication.			Economic	near to 8		HPC-BD and CLOUD	Incremental	expected in +2021
LEXIS	Real-time model of computation for high performance computational workflows under real-time deadlines	Technology-Service	In the context of answering to emergency situation, computation and simulation results have to be provided according to strict real-time deadlines to be of use for urgent decision making. The innovation is then to bring to the HPC/Cloud orchestration domain a model of computation with an ability to express and verify deadlines, while maintaining a view of a computational flow. This opens the ability to use such systems as a component of a Cyber-Physical System (CPS) and hence offers a strong potential as an innovation.			Economic and Societal	near to 8		HPC-BD and CLOUD	Incremental	expected in +2021
LEXIS	Smartgateway with embed data processing/validation features	Product	Smartgateways are growing with a speed close to IoT sensors, they need to be Low Power, energy autonomous, capable to connect into resilient sensors network and to		The Smartgateway is embedding native connectivity with twitting sensors, having onboard pre-processing and data validation before cloud transfer, based around innovative	Economic and Societal	9		HPC-BD and CLOUD	Incremental	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
	before cloud transfer		perform decentralised computing moving towards deep edge capability. LEXIS will contribute to solve this challenge embedding native connectivity with twitting sensors, having onboard pre-processing and data validation before cloud transfer, based around innovative low power chips leveraging on low-level firmware management for Low Power features optimisation. It is then designed with a multiplatform cloud protocol interpreter.		low power chips leveraging on low-level firmware management for Low Power features optimisation. It is then designed with a multiplatform cloud protocol interpreter.						
<b>Lynx</b>											
Lynx	Lynx Services Platform	product	The main outcome in Lynx is the Lynx Services Platform (LynxSP) including a set of 16 services for legal NLP based on a Legal Knowledge Graph that can be used standalone or orchestrated for certain requirements.	Medium	n/a	Economic and Societal	7-8	Demonstrator plus 3 Pilot prototypes are already in place	Compliance and LegalTech	Architectural	new in 2020
<b>MOSAICrOWN</b>											
MOSAICrOWN	Policy model and language	Technology	A formal model and a language to regulate access and usage of data to enforce privacy/anonymity requirements, protect company-confidential information, and enable only authorized visibility and operations throughout the whole data processing life-cycle	Significant	Balance of expressiveness/flexibility and simplicity/performance. Policies can refer to both source (raw) data or wrapped/sanitized (as well as derived) data, and can refer to data that are protected (sanitized/wrapped) before their release/processing	Economic and Societal		Supported by a prototype	cross sector	Disruptive	new in 2019
MOSAICrOWN	Policy engine	Technology	The policy engine is the tool responsible for parsing the MOSAICrOWN policy and checking whether a subject request is permitted	Significant	Policy engine is in charge of enforcing the specifications provided through the policy language	Economic and Societal	4	Prototype and demo	cross sector	Disruptive	new in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
MOSAICrOWN	Reference metadata model	Technology	The metadata model represents the basis for the definition of the model and language able to support the specification of data handling policies, taking into account different properties of data, and allowing data from all actors involved.	Significant	Metadata facilitate the portability of data and their regulations, enable controlled sharing and empowering data owners, allow control over data flow and usage	Economic and Societal		Prototype and demo	cross sector	Disruptive	new in 2019
MOSAICrOWN	Dynamic wrapping for collaborative computations	Technology	A solution for the enforcement of authorizations dynamically and automatically disabling/enabling visibility of data in collaborative computations.	Significant	Computations can see the controlled involvement (for protected data) of subjects not completely trusted to access data but offering computational capabilities at convenient prices, with clear economic benefits while not compromising on security.	Economic and Societal	4	Prototype and demo	cross sector	Disruptive	new in 2019
MOSAICrOWN	AONT with splitting and fragmentation	Technology	Application of an ALL-Or-Nothing Transform encryption mode, combined with fragmentation and slicing, for protecting data in storage	Significant	Ensures resilience to failures and strong protection even in case of key leakage	Economic and Societal	4	Prototype and demo	cross sector	Disruptive	new in 2019
MOSAICrOWN	UI-based data catalog software for the selection of data wrapping techniques	Technology	UI allows wrapping techniques to be applied to datasets, tables, individual columns. UI allows for dropdown selection of techniques. Based upon selected technique, UI kicks-off appropriate back-end scripts that load dataset and process it using selected technique	Medium		Economic and Societal	5	Prototype and demo	cross sector	Disruptive	new in 2019
MOSAICrOWN	Sanitization in Apache Spark	Technology	Anonymization of a collection of sensitive data with Mondrian algorithm with the Apache Spark framework, which permits to use the computational capacity of a distributed system	Significant	It offers solutions at scale, which work efficiently for any dataset size	Economic and Societal	4	Prototype and demo	cross sector	Disruptive	new in 2019
MOSAICrOWN	Membership Inference Service	Technology	A tool able to identify records used to train a target model, framed for privacy-preserving	Significant	Impact on supply chain insights, marketing budget, customer satisfaction, spending and sales optimization	Economic and Societal	5	Prototype and demo	cross sector	Disruptive	new in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			consumer analytics in cloud-based data markets								
<b>Musketeer</b>											
Musketeer	Client Connector	Technology	The client connector is the component required for a participant to join the data processing federation. It is the software application supporting participant in his data exchange, share and process, so to guarantee that he is sovereign of his data.	Medium	To improve the accuracy of the ML models thanks to the distributed knowledge provided by the federated participants	Economic	5	Prototype	sector independent		new in 2020
Musketeer	pycloudmassenger	service	System for privacy preserving federated machine learning.	Medium	To improve the accuracy of the ML models thanks to the distributed knowledge provided by the federated participants	Economic and Societal	5	Prototype	Agnostic	Disruptive	new in 2019
Musketeer	ML models supporting Prostate Cancer Diagnosis	Product/service	The ML models developed in the healthcare pilot of MUSKETEER, based on the privacy-preserving FML MUSKETEER system, will support the diagnosis of prostate cancer based on mp-MRI medical images, providing segmentation of the prostate and classification of findings. After certification under the medical devices regulations, these models can be integrated into 3Dnet radiology reporting system, increasing the speed and accuracy of the prostate cancer diagnosis.	Medium		Economic and Societal	6	Prototype	Health	Incremental	expected in +2021
Musketeer	Federated Learning	Process	Federated (Machine) Learning is a type of distributed machine learning to preserve data privacy. Data privacy and communication efficiency are two main motivations for its adoption in different scenarios as manufacturing	Medium		Economic	6 (final)	Prototype		Architectural	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
Musketeer	Federated Machine Learning	Service, processes	Development of an Industrial Data platform with scalable algorithms for federated and privacy-preserving machine learning techniques with detection and mitigation of adversarial attacks and a rewarding model capable to fairly monetize datasets according to the real data value.	Medium		Economic	6	supported by a prototype	Manufacturing	Incremental	expected in +2021
Musketeer	MMLL-POMs1,2,3	Software technology	MMLL (MUSKETEER Machine Learning Library). MMLL is a set of machine learning algorithms (Neural networks, K-Means, SVM, etc.) able to run under a set of privacy operation modes within MUSKETEER platform. The privacy operation modes (POMs) are different configurations of federated training. Each of them reaches different privacy level (1,2,3). The MMLL POMs 1,2,3 can be exploitable due to the interest of a wide variety of use cases in performing federated training processes.	Significant	The federated training process could overcome a lot of privacy constraints in a collaborative training process	Economic and Societal	TRL 6 – technology demonstrated in relevant environment	Medium-low. An open source code to be released in 2021. The innovation still needs an extra development stage for solving (among others) scalability issues that could appear in a production environment.	Almost all. Healthcare, manufacturing, public administration, financial, pharmaceutical, agrifood, energy, etc.	Radical	expected in +2021
<b>OpertusMundi</b>											
OpertusMundi	topio.market	system	Digital marketplace for proprietary geospatial asset trading	Significant	Topio.market is the first digital marketplace in Europe focused on trading proprietary geospatial data and services. With ~80% of data being geospatial, and an estimated market size of data suppliers ~100billion by 2025, topio.market aims to be the single place in Europe where geospatial data are published and traded	Economic	6	The platform is currently in development, with a closed beta starting in Q1 2022. An alpha system is already	Geospatial/ICT	Disruptive	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
								operational, with more than 100 data suppliers agreeing to participate in the beta phase			
<b>PIM-City</b>											
<b>PIM-City</b>	Data Portability Control (DPC)	system	Allow users to migrate their data to new platforms, in a privacy-preserving fashion.	Medium	Provide entities processing data necessary tools to comply with adopted legislative initiatives and in line with commonly agreed ethical and societal principles	Societal	6	Developing	Digital Marketing	Incremental	Expected in +2021
<b>PIM-City</b>	Data Valuation Tools from Market Place (DVTMP)	system	Allow data buyers and users to have an estimated value of audiences traded in major marketing platforms	Medium	Provide all stakeholders with necessary information in order to develop a sustainable and ethically acceptable environment	Economic and Societal	6	Developing	Digital Marketing	Incremental	Expected in +2021
<b>PIM-City</b>	Personal Data Safe (PDS)	system	Allow users to store personal data in a transparent, flexible and scalable way, allowing total control on their data	Medium		Economic and Societal	6	Prototype Ready	Digital Marketing	Incremental	Expected in +2021
<b>Pim-City</b>	Privacy Preserving Analytics (PPA)	system	Allow data buyers to query user data with strong and flexible privacy guarantees, in the form of a library or service	Medium	It will allow the integration of privacy preserving analytics in an easy way	Economic and Societal	6	Developing	Digital Marketing	Incremental	Expected in +2021
<b>PIM-City</b>	Data Provenance (DP)	system	Allows participants to control, verify and certify the original and provenance of the data	Medium	It will enable the current centralized data economy industry to shift the value of data assets to a future decentralized setting without loss of data ownership values	Economic and Societal		Developing	Digital Marketing	Incremental	Expected in +2021
<b>PIM-City</b>	Privacy Metrics (PM)	service	Allow users to information about data buyers participating PIMCity and about web services in general and invite data buyers and web	Medium	It will induce data buyers and web services to improve their online posture in terms of privacy, security and transparency	Economic and Societal	6	Developing	IT Security, Digital Marketing	Incremental	Expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			services to actively provide information about how they plan to use users' data.								
PIM-City	Personal Consent Manager (PCM)	system	Allow people to give their explicit consent for the use of their personal data. Allow companies and organizations to get users consent respecting users privacy choices.	Medium	It will enable for users to exercise one of their first rights: to opt-in and opt-out and in this way be a decision maker in the data ecosystem	Societal	6	Developing	Digital Marketing	Incremental	Expected in +2021
PIM-City	Data Trading Engine (DTE)	system	Allow users to get value from their data with their explicit consent using an API. With this system, companies and people will be able to interact directly and be part of the multi-billion market data economy. Community developers can also create tools to integrate with DTE and participate in this economy	Significant	Data economy have being increasing in an exponential way but people (real owners of the data) have been taken aside. With this, people will get benefits of the value that their data generates.	Economic and Societal	6	Developing	Digital Marketing	Incremental	Expected in +2021
PIM-City	Data Aggregation (DA)	system	Allow user to export data (that initially may contain personal information) in an aggregated and anonymized form to share/sell to be used by third parties without risking their privacy	Medium	Make available large and useful datasets for commercial or research use in data marketplaces. Metadata also will be available to describe properties of the dataset	Economic	6	Developing	Data management	Incremental	Expected in +2021
PIM-City	Data Knowledge Extraction (DKE)	system	Allows users to get a fitting initial user profile as well as updates for their profiles to keep them as accurate as possible. Enables the evaluation of the user's data with their explicit consent.	Medium	It will enable to translate the user's raw data into an accurate user profile that is necessary for the user to understand the value of their data.	Economic and Societal	6	Developing	Digital Marketing	Incremental	Expected in +2021
PIM-City	Personal Data Avatar	system	Allows users full control over their user profiles. They can access and control their profiles, refine information and decide what they want to share and what they want to keep private.	Medium	It will enable users to always be in control over their own data by giving them the possibility to change everything about their own profile as well as deciding what information will be shared.	Economic and Societal	6	Developing	Digital Marketing	Incremental	Expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
PIM-City	Transparency Tags	system	Allow users to get easy-to-understand information about data buyers participating PIMCity and about web services in general	Medium	It will assist users to better understand the nature of business around their data, and it will induce data buyers and web services to improve their online posture in terms of privacy, security and transparency	Economic and Societal	6	Developing	IT Security, Digital Marketing	Incremental	Expected in +2021
PIM-City	Dashboard	system	Provide intuitive visualizations of personal mobility/ location data to identify habits/ behavior of the user	Medium	Allow users to analyze their data to discover patterns, behavior and understand better the value of their data through comparing them with other users data	Economic and Societal	6	Developing	Data analytics	Incremental	Expected in +2021
PIM-City	Marketplace	product	Allow companies to get data from users with their explicit consent. With the marketplace, people and companies will be part of the multi-billion market data economy	Significant	Data economy have being increasing in an exponential way but people (real owners of the data) have been taken aside. With this, people will get benefits of the value that their data generates.	Economic and Societal	6	Developing	Digital Marketing, Big Data	Incremental	Expected in +2021
PIM-City	EasyPIMS	product	A fully fledged PIMS working in both B2B and B2C use cases	Medium	EasyPIMS will be a fully working PIMS that would allow individuals and companies to monetize their data in an open market.	Economic and Societal	7	Developing	Digital Marketing	Incremental	Expected in +2021
<b>Platoon</b>											
Platoon	PLATOON Reference Architecture + Interoperability layer / pipeline (Semantic Data Models + OpenAPIs)	Technology, Methods, Process	Facilitate the interoperability and exchange of heterogeneous data and representation of data amongst partners that form an ecosystem based on federated platforms/services	Significant	Basis for the creation of data spaces: Allow multiactor/multiplatform exchange of self described information	Economic	5 to 7	To be validated in large scale demonstrators	electricity (RES, smart grids and Smart buildings)	Architectural	new in 2020
Platoon	Data governance components (IDS connector, borker and vocabulary provider)	Technology	Allow multiparty exchange of data and services while maintaining security, privacy and sovereignty.	Significant	Basis for the creation of data spaces: guarantes the security, privacy and sovereignty of exchanged data and services	Economic and Societal	5 to 7	To be validated in large scale demonstrators	electricity (RES, smart grids and Smart buildings)	Architectural	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
<b>Platoon</b>	Data Analytics tools	Technology (potential patents)	Data analytics tools using AI and Digital twins which allow to extract the value from heterogenous self described data from different partners along the energy value chain	Significant	End goal of data spaces: extract value from data for different applications (O&M, smart grid management optimisation, energy efficiency optimisation...)	Economic and Societal	5 to 7	To be validated in large scale demonstrators	electricity (RES, smart grids and Smart buildings)	Incremental	expected in +2021
<b>QROWD</b>											
<b>QROWD</b>	<b>QROWD Platform</b>	Platform	Crowdsourcing-enabled big data integration and analytics services platform for urban mobility management (e.g. City governments). Allows the efficient management of the entire mobility related data value chain	Medium		Economic and Societal	7	Supported by demonstrator	ICT	Architectural	Already reported in 2018 but progress in 2019
<b>QROWD</b>	<b>Touristic Network services</b>	Service	Provide travel time and delays to reach several touristic destinations (e.g. ski resorts), so tourists can choose the best option(s).Ensures better tourist customer experience; Allows City tourism authorities to suggest touristic hotspots to visit based on current estimated travel time;Reduces Traffic congestion.	Medium		Economic and Societal	7	Supported by demonstrator; In integration with current services of industrial partner	Tourism, Transport	Incremental	already reported in 2018 but progress in 2019
<b>QROWD</b>	<b>On-Street parking</b>	Service	Inform drivers about available parking spot. Improves drivers' mobility experience; Improves efficiency in usage of parking facilities; Better Parking facilities ROI	Medium		Economic and Societal	7	Supported by demonstrator; In integration with current services of industrial partner	Transport	Incremental	already reported in 2018 but progress in 2019
<b>QROWD</b>	<b>Historical Analysis Reporter</b>	Service	Provides traffic reports to Mobility managers, to better understand specificities of traffic conditions in a concerned place, e.g. due to bad weather conditions, or man-made events. Improves road safety; Reduces traffic	Significant	Reduces traffic congestion	Economic and Societal	7	Supported by demonstrator; In integration with current services	Transport, Smart Cities	Architectural	already reported in 2018 but progress in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			congestion; Improves drivers' mobility experience.					of industrial partner			
QROWD	Road Event Reporter	Service	Enables traffic managers to have an overview of the traffic in real time and to be active on it, e.g. . report road closures, road works, accidents and other disruptions. Reduces traffic congestion as citizens are better informed and the city can better control traffic.	Significant	Reduces traffic congestion	Economic and Societal	7	Supported by demonstrator; In integration with current services of industrial partner	Transport, Mobility, Smart Cities	Architectural	already reported in 2018 but progress in 2019
QROWD	TomTom City	Service	Allows city's traffic/mobility managers/staff to better monitor current traffic conditions and on recent improvements/ alterations. It also compares the situation with similar cities (Traffic management tools; traffic patterns benchmarking)	Medium		Economic and Societal	7	Supported by demonstrator; In integration with current services of industrial partner	Transport Mobility, Smart Cities	Incremental	already reported in 2018 but progress in 2019
QROWD	Municipality Dashboard	Service	Dashboard for City staff showing the city's mobility data and the results of analytics. Facilitates Decision making processes	Medium		Economic and Societal	7	Supported by demonstrator; In integration with current services of industrial partner	Smart Cities	Incremental	already reported in 2018 but progress in 2019
QROWD	Citizen Dashboard	Service	Dashboard for citizens. It will be a subset of the Municipality Dashboard plus some extra services for citizens around mobility. Better-informed citizens	Medium		Economic and Societal	7	Supported by demonstrator; In integration with current services of industrial partner	Smart Cities	Incremental	already reported in 2018 but progress in 2019
QROWD	iLog	Service (Data collection as a service)	Mobile application to (i) Collect sensor data from engaged citizens in both active and passive way (ii) Put citizens "in-the-loop" for validation and curation of their own data (or	Medium		Economic and Societal	7	Supported by demonstrator; Spin-off created	Smart Cities, Citizen Engagement	Disruptive	already reported in 2018 but progress in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			collectively generated by them) . Allows City staff to get data and feedback from citizens about their mobility patterns. Enables /improves/promotes citizen engagement in data driven decision making related to their city								
QROWD	Modal Split	Service	Allows City Management to run travel surveys using citizen's mobile phones instead of cumbersome Telephone interviewing or paper surveys. . Allows shifting from expensive, limited, years old static data to up-to -date cost-efficient data driven decision making, simulations, what-if analysis and policy formulation. Its systematic and timely usage generates disruptive positive changes in traffic management	Significant	disruptive improvement in traffic management, congestion reduction and ultimately reducing traffic-related pollution levels	Economic and Societal	7	Supported by demonstrator	Transport, Mobility, Smart Cities	Disruptive	already reported in 2018 but progress in 2019
QROWD	Qrowd-DB	Analytics back-end	Back-end component that stores sensor streaming data (big data) and linked data. The i-Log (mobile app) requires it to collect data e.g. for the Modal Split calculation	Medium		Economic and Societal	7	Supported by prototype	ICT	Incremental	already reported in 2018 but progress in 2019
QROWD	Virtual City Explorer	Tool	Provides City management a better way to generate maps of points or items of interest in a city that can be found in their Google Street View virtual representation.It solves the problem of completing mobility infrastructure records at a fraction of the cost of sending Municipality workers to do surveying/inventories/counting.	Medium		Economic and Societal	7	Supported by demonstrator	Smart Cities	Disruptive	already reported in 2018 but progress in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
QROWD	QROWDSmith	Crowdsourcing platform	Allows city staff to define and run crowdsourcing tasks and citizen engagement challenges with gamification. Dramatic improvement in engaging citizens and data contributors to solve city issues	Medium		Economic and Societal	7	Supported by demonstrator	ICT	Disruptive	already reported in 2018 but progress in 2019
QROWD	Data Acquisition Framework	Framework	Supports city staff in the acquisition of data and associated methodology. Improves data management	Medium		Economic and Societal	7	Supported by prototype	ICT	Architectural	already reported in 2018 but progress in 2019
QROWD	Crowd Feedback-aware link prediction	Tool	Link prediction engine for city staff, capable of iterative improvement of generated datasets using data quality assessment from the crowd. Improve the performance of the analytics required to deliver the QROWD mobility services	Medium		Economic and Societal	7	Supported by prototype	ICT	Incremental	already reported in 2018 but progress in 2019
QROWD	Analytics with Crowd Feedback	Tool	In case a machine learning/analytics algorithm only performs poorly (e.g. due to insufficient training data) crowd/expert feedback is introduced by city staff to improve the learning results. Improve the performance of the analytics required to deliver the QROWD mobility services	Medium		Economic and Societal	7	Supported by prototype	ICT	Incremental	already reported in 2018 but progress in 2019
QROWD	Analytics on Data Streams	Tool	Allows city staff to perform analytics on non-static and static data that comes in as a continuous data stream. Improve the performance of the analytics required to deliver the QROWD mobility services	Medium		Economic and Societal	7	Supported by prototype	ICT	Incremental	already reported in 2018 but progress in 2019
QROWD	Spatio-temporal Analytics	Tool	Allows city staff to do analytics based on spatial and temporal relations like 'inside', 'near', 'after', 'before' and so on. Improve the performance of the analytics required to deliver the QROWD mobility services	Medium		Economic and Societal	7	Supported by prototype	ICT	Incremental	already reported in 2018 but progress in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
QROWD	Integrated Processing of Static and Stream Data	Method	In case static and stream data need to be processed for analytics purposes by city staff, those need to be integrated and provided through defined interfaces. Improve the performance of the analytics required to deliver the QROWD mobility services	Medium		Economic and Societal	7	Supported by prototype	ICT	Incremental	already reported in 2018 but progress in 2019
QROWD	Automated Personal Assistant (AIMEE)	Service	Allows users to automate processing of emails and email attachments	Medium		Economic and Societal	8	Supported by demonstrator; In integration with current services of industrial partner	Business services	Disruptive	new in 2019
<b>SmartDataLake</b>											
SmartDataLake	SmartDataLake Toolkit	System	The SmartDataLake toolkit offers components for data profiling, data virtualization, approximate query processing, storage tiering, similarity search, graph mining, and visual analytics.	Medium	SmartDataLake reduces the time and effort required for extracting useful insights from raw data.	Economic	5	Demonstrator	Business Intelligence	Disruptive	new in 2020
<b>Smashhit</b>											
Smashhit	smashHit platform	technology, system	provides basic functionalities for the user management, management of semantic models for consent and legal rights and offers context sensitivity support	Significant	The smashHit platform provides basic functionalities in the smashHit framework	Economic					
Smashhit	Contract Support Tool	technology, system, service	provides automatic GDPR compliant consent document generation, consent compliance verification and consent traceability	Significant	The smashHit solution shall accelerate the consent creation process to increase the efficiency in the data economy and to reduce the efforts needed for a consent creation and consent management. This component aims to allow a semi-automatic consent form creation process based on a law compliant data usage and processing rules	Economic					

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
Smashhit	Consent Certification	technology, system, service	provides functionalities for the certification of consent	Significant	This approach will enable to manage and control consent about data usage with two or more involved participants in the "data use traceability module".	Economic					
Smashhit	Data use traceability	technology, system, service	allows the tracing of the data usage and the authenticity verification of leaked (and possibly modified) data as well as possibly the point of data leakage	Significant	The module addresses the need of OEMs, Data Providers and Data Processors to get a solution which enables the identification of data leakages or misuses	Economic					
Smashhit	Security & Privacy Mechanisms and Metrics	technology, system, service, processes	assure security and privacy within the smashHit framework	Significant	This component focussed on how to combine privacy/security policies and their associated mechanisms in a complementary, if not integrated, way.	Economic					
<b>Synergy</b>											
Synergy	SYNERGY Data & AI Marketplace	Technology, Product, Service	The SYNERGY Data & AI Marketplace allows stakeholders in the electricity data value chain to share their data assets in a reliable and trustworthy manner. Access is controlled according to the preferences of the data asset providers who always decide whether they will proceed to negotiate the terms of a data asset sharing agreement (formalized through a contract) with any interested stakeholder. The full lifecycle of such data asset contracts is effectively managed, from their creation and negotiation to finalisation and (payment) settlement. As multiple parties may be involved in a data asset (e.g. in the case of analytics results), SYNERGY ensures that all involved stakeholders will be appropriately reimbursed according to the contract's provisions.	Significant	The SYNERGY Data & AI Marketplace brings a new mentality to the electricity data value chain that shall allow them to improve their operations, open new data-driven revenue channels and leverage novel business models innovations.	Economic	6	Under development	Energy	Disruptive	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
Synergy	SYNERGY Big Data Platform	Technology, Product, Service	The SYNERGY Big Data Platform allows energy data value chain stakeholders to attach value to their own data assets and unveil insights they could not previously acquire. Through its data collection, security and analytics functionalities, significant challenges (like secure data management, semantic interoperability, lack of analytics skills) can be effectively addressed.	Significant	The SYNERGY Big Data Platform facilitates the digital transformation of the electricity data value chain by allowing them to improve their operations and launch new services.	Economic	6	Under development	Energy	Disruptive	expected in +2021
Synergy	BM1- Data Intelligence as a Service for Advanced Electricity Grid Asset Management and Planning	Service	This business model is based on "Coordinated maintenance works and coordinated expansion planning strategies in the operation landscape of the electricity networks". This business model is based on several infrastructure planning and asset management optimization software tools providing value-added services on top of the SYNERGY core big data platform and AI Analytics Marketplace and will provide active and reflective, time-based probabilistic asset management techniques that will aim to also reduce the whole system cost including the cost passed to network operators.	Medium	N/A	Economic and Societal	N/A	Under development	Energy	Disruptive	expected in +2021
Synergy	BM2 - Data value as a service for Advanced Asset Management	Service	Real-time data streams from equipment items available in electricity system infrastructures need to feed existing Predictive Maintenance Software solutions utilized by a variety of stakeholders of the electricity data value chain, and more specifically, Network Operators, RES Plant Operators and Building/Facility Managers.	Medium	N/A	Economic	N/A	Under development	Smart cities, Energy	Disruptive	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			Having gone through such processes, the owned data of Beneficiaries can facilitate more effective daily maintenance-related works. I.e. asset management, decision-making maintenance scheduling.								
<b>Synergy</b>	BM3- Data Intelligence-driven Advanced Predictive Maintenance for RES Power Plants.	Service	Added value capabilities and technical features stemming from AI analytics are feeding in novel Advanced existing Predictive Maintenance Software solutions used by RES Plant Operators. More value is captured from the use of advanced analytics enhancing the daily maintenance-related works, i.e. asset management, decision-making of operation and maintenance scheduling. RES Plant Operators will be rewarded with up to date and near real-time information about real-time predictive needs, proactive diagnosis of equipment faults, probability of fault occurrence to proceed with the timely replacement of equipment reaching end-of-life or components with high probability to fail.	Medium	N/A	Economic	N/A	Under development	Energy	Disruptive	expected in +2021
<b>Synergy</b>	BM4- Dynamic Enhanced Energy Performance Certificates	Service	Aims at advancing the "Business-as-Usual" in the area of Energy Performance Certification by allowing ESCOs to access and process real-time data streams from a variety of buildings towards enabling the delivery of novel and advanced EPC services, characterised by the dynamic assessment of energy performance and enhanced with appropriate smartness indices as instructed	Medium	N/A	Economic and Societal	N/A	Under development	Smart cities, Energy	Disruptive	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			by recent EU regulations around the Smart Readiness of buildings.								
Synergy	BM5- Synergetic Energy Performance Contracting (design)	Service	Aims at transforming current business practices followed during renovation projects, that consider generic templates for the anticipated used of a building, by relying in real-life data streams in buildings so as to introduce in Energy Performance Simulations accurate profiles of occupancy and comfort and in this way enable better prediction of energy consumption in different building spaces, as well as, in the whole building, as part of the design process.	Medium	N/A	Economic and Societal	N/A	Under development	Smart cities, Energy	Disruptive	expected in +2021
Synergy	BM6- Synergetic Energy Performance Optimization (self-consumption)	Service	Aims at extending the offer of ESCOs to Building Managers/ Facility Managers/ Prosumers towards enabling further reduction of the performance gap (or even eliminating it) and allowing customers to enjoy additional energy cost savings and reduce dependency on the grid (reduction of network charges) through flexibility-based control of building energy systems (generation, storage, demand) for self-consumption maximization.	Medium	N/A	Economic and Societal	N/A	Under development	Smart cities, Energy	Disruptive	expected in +2021
Synergy	BM7- Intelligence-Driven Long-term Generation Planning and PPAs Advisory	Service	The further penetration of PPAs requires a better understanding on the side of RES operators on the profitability of previous PPAs, the volume they shall make available into such bilateral agreements and the correct timing of establishing them in relation to prices offered in wholesale	Medium	N/A	Economic	N/A	Under development	Energy	Disruptive	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			markets. This requires the execution of advanced analytics to reveal the best strategy a RES operator shall establish with regards to generation capacity building and price and duration negotiation in the frame of the agreement.								
<b>Synergy</b>	BM8- RES Power Plant Optimizer for GPPA maximization	Service	It builds on the previous innovation and the initial high-level advisory provided to RES Operators and focuses on the analysis of specific retailer's power needs, as the preliminary step towards the definition of the exact amount of power to be offered to them and the contractual parameters related to price and duration.	Medium	N/A	Economic	N/A	Under development	Energy	Disruptive	expected in +2021
<b>Synergy</b>	BM 9- Transparent GPPA Marketplace	Service	It will step on advanced analytics in order to equip RES Operators (beneficiary) and Retailers (customer) with the required forecasts and estimations that will allow them to place accurate bids in an online marketplace environment operated by the RES Operator and get engaged in transparent, long-term PPAs. The combination of the aforementioned analytics results will allow RES operators and retailers to get involved into a negotiation process regarding the establishment of long-term Green Power Purchase Agreements that will mutually benefit both sides, allowing RES operators to avoid losing significant revenues by non-sold energy, while increasing the competitiveness	Medium	N/A	Economic	N/A	Under development	Energy	Disruptive	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			of the offering of electricity retailers by introducing green energy tariffs (lower than the conventional ones) resulting from the favourable GPPAs established between the two involved stakeholders.								
Synergy	BM 10. RES Virtual Power Plant (VPP)	Service	Aims at minimizing the curtailment of RES and utilizing their excess power output for the provision of ancillary services to TSOs. This will be performed through an aggregation process performed by Aggregators over RES available at different locations of the network (Transmission Network) towards formulating ad-hoc Virtual Power Plants as a Virtual Battery for cost-effective RES market's uptake. We propose a paradigm change, focusing on the business activities at the power generation level to provide support to TSOs and avoid "energy curtailment" in PV or Wind farms.	Medium	N/A	Economic	N/A	Under development	Energy	Disruptive	expected in +2021
Synergy	BM 11 – Objective Dynamic Pricing of Electricity	Service	Aims at enabling Retailers to apply Dynamic Pricing Schemes over their portfolio to satisfy a variety of business goals by properly extracting and analysing the price-based flexibility (elasticity) of their consumer portfolio as a first step towards the definition of alternative pricing strategies. Significant value is expected to be generated for the retailer itself which will be able to setup an internal analytics environment for optimizing their portfolio performance in terms of energy efficiency	Medium	N/A	Economic	N/A	Under development	Energy	Disruptive	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
Synergy	BM 12 – Retailers as Non-Energy Service Providers	Service	Derived from the previous, in this case having access to large datasets coming from their clientele (prosumers, facility managers), will allow retailers to perform advanced personal analytics functions over each client's data and return back (under specific service provision contracts) novel services (security, health, well-being and comfort) that are highly valued and can create significantly increased profitability, especially when compared to traditional commodity sales business practices.	Medium	N/A	Economic	N/A	Under development	Energy	Disruptive	expected in +2021
Synergy	BM13- Flexibility and portfolio analytics (sales of insights)	Service	Aims at the development and promotion of a new service from Aggregators for the delivery of pre-trained analytics and insights to other stakeholders of the electricity data value chain, towards facilitating their understanding about flexibility (flexibility profiles, flexibility forecasts) and how flexibility could be utilized as part of the realization of their individual business objectives.	Medium	N/A	Economic	N/A	Under development	Energy	Disruptive	expected in +2021
Synergy	BM 14- Flexibility VPP configuration for ancillary services	Service	The involvement of aggregators into flexibility markets and trading requires the continuous optimization of their bidding activities in the respective markets in order to increase their revenues and profits while avoiding unnecessary penalties related to the inability to deliver the contracted flexibility in the frame of ancillary services to electricity grid operators.	Medium	N/A	Economic	N/A	Under development	Energy	Disruptive	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			To achieve this, aggregators need to reach out to cheap flexibility sources (small residential and tertiary prosumers) and exploit their unleashed and vast flexibility potential in a reliable manner (through accurate forecasting and clustering/ VPP formulation), by introducing them as a competitive flexibility source against traditional ones, in the respective markets.								
<b>Synergy</b>	BM 15-TSO-DSO Collaborative Network Management	Service	Aims at making the maximum use of the concept of flexibility when it comes to ensuring efficient operation of the electrical system and the highest quality in the power supply most cost-effectively (taking the most profit of the flexibility potential instead of purchasing additional capacity or activating market mechanisms in order to solve supply-demand imbalances). For tackling this, a coordination platform needs to be established where the TSOs and DSOs will apply advanced simulations to estimate at the local level and in each local DSO zone - and in a coordinated manner between all of them- the level of flexibility required for each of them, in advance.	Medium	N/A	Economic and Societal	N/A	Under development	Energy	Disruptive	expected in +2021
<b>Synergy</b>	BM16- Urban Planning crowdsourcing marketplace	Service	Aims at facilitating the realization of ambitious sustainability goals set by City Authorities by optimizing urban planning and transforming it into a crowdsourcing (collective data-driven) process where all city-relevant stakeholders are incentivized to	Medium	N/A	Economic and Societal	N/A	Under development	Smart cities, Energy	Disruptive	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			provide their data for achieving important sustainability targets.								
<b>Synergy</b>	BM17- Synergetic Energy as a Service Model (Retailer-ESCO & Retailer-Aggregator)	Service	To sustain growth, improve competitiveness and drive business value, the industry must transform and take advantage of business opportunities that arise from a low-carbon energy system. Such business opportunities require a major shift away from the traditional commodity sales business model and adoption of a more profitable business orientation that is based on EaaS (Energy as a Service) offering towards their clients, spanning advanced and personalized energy analytics for energy efficiency, flexibility trading, but also service for intelligent controls and (where possible) smart automation of consumer amenities.	Medium	N/A	Economic and Societal	N/A	Under development	Energy	Disruptive	expected in +2021
<b>Infinitech</b>											
<b>Infinitech</b>	Car data ingestion	Technology/System	This technology is tasked with the capture, homogenization, distribution and storage of the datasets that support the connected car pilot. It involves the design and implementation of the IoT agents that adapt the data available from the Vehicle, as well as deploying the necessary modules for data storage and distribution (FIWARE Orion, FIWARE Cygnus, etc)	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Data Protection Orchestrator (DPO)	Technology/System	It is an enabler for embedding and automating the assurance of security and privacy by design and by default in heterogeneous and complex business flow. It	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			orchestrates various privacy and security management functions (such as access control, encryption and anonymization). It will be used as part of the data governance Framework of the Project, and towards establishing the regulatory compliance tools in the project's sandboxes.								
<b>Infinittech</b>	Digital User Onboarding System (DUOS)	Technology/System	It is a solution for dealing with virtual identities in a mobile device. It provides remote user registration using eID or passport	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinittech</b>	EASIER-AI	Technology/System	EASIER-AI is an Hybrid (Cloud/Edge) platform that facilitates to develop, measure, monitor and deploy your AI models. The platform facilitates the data science tasks and it is focused on working on Hybrid Infrastructure and exploiting data generated by IoT. The platform synchronizes Cloud and Edge, keeping the Edge always up to date to run always the most accurate model.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinittech</b>	Driver profile classifier	Technology/System	This technology is aimed at the use of high-quality vehicle data allows insurance companies to offer customized products. The application of supervised machine learning techniques is proposed to classify drivers' profiles which generates a customized insurance premium. The resulting model is then deployed with TensorFlow serving and integrated as part of the cloud platform with a wrapper, achieving an accuracy of 85.7%.	Medium		Economic and Societal		Prototype/Demonstrator	Finance	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
Infinittech	Distributed near-real-time HPC processing and exchange of IoT streaming data	Technology/System	AI algorithms' optimization exploiting GPUs capabilities	Significant	Accelerates AI pipelines execution in the finance sector	Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
Infinittech	Botakis Chatbot Development Network	Technology/System	A tool for rapid development of chatbots applications, which will be used for the development of chatbots features in the INFINITECH pilots (i.e. notably the LIB, BOC and NBG led pilots).	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
Infinittech	AI-Engine-for-Psychometric-Profilng and Personalization	Technology/System	AI-driven engine to extract four categories of behavioural features, grouped according to the type of spending behaviour they capture: (i) overall spending behaviour, (ii) temporal spending behaviour, (iii) category-related spending behaviour, and (iv) customer category profile. The engine uses these features to predict the personality of customers.	Medium		Economic and Societal		Prototype/Demonstrator	Finance	Incremental	new in 2020
Infinittech	OpenSource AI/ML frameworks	Technology/System	These frameworks facilitate the development of AI/ML based tools, which shall be applied to Financial Crime and Fraud, e.g. on so called Instant Loans	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
Infinittech	Data Layer - REST API	Technology/System	A Data Layer to support Security Data Model with REST API based on a not relational database (MongoDB). Supports heterogeneous sources. Developed upon FLASK-Python3 framework and dockerized to be deployed on Kubernetes infrastructure.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
Infinitech	Terrier Information Retrieval Platform	Technology/System	Search Engine for BigData sets that offers integration with Spark for distributed processing	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
Infinitech	Anonymization tool	Technology/System	A tool that anonymizes data in order to preserve privacy. It also provides metrics that allow to measure the risk of the anonymized data and the impact of the anonymization process on the utility of the data. The tool will be used in pilot #11 and pilot #12. The component needs a specific configuration/development for each pilot in which it is used	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
Infinitech	Polyglot Database Management System	Technology/System	The LXS DBMS is a polystore database that provides access to different and heterogeneous datastores via a common interface. It allows for the data user to submit a query, whose scan operators can request data that are stored in external datastores, and combine their intermediate results with data coming from other sources, either LXS internal datastore or others. For instance, a JOIN operator might require to JOIN table A (resigned in LXS) and table B (resigned in a MongoDB or a Hadoop DataLake). At this phase, there is a support for a limited target datastores, for Proof-of-Concept of the prototype. Moreover, the user has to write queries for the target datastores in the specific dialect. What is more, JOIN operations are not efficient, as they require	Significant	Accelerates the development of data-intensive applications for the Digital Finance Sector	Economic		Prototype/Demonstrator	Finance	Disruptive	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			all data resigned in an external datastore to be retrieved in the query engine level.								
Infinittech	HTAP Database for the financial and insurance sector	Technology/System	An ultra-scalable SQL Database and real-time big data platform that revolutionize the business database management systems by introducing the next generation business database that can scale in any of the three Vs of Big Data (Volume, Velocity and Variety). In more technical details, it provides an ultra-scalable transactional management system that can scale out to 100s of nodes, which is typically a bottleneck in traditional database systems that provide transactional semantics, while on the same time, is full SQL compatible and ensures all ACID transactional properties. It additionally exposes an interface for direct access of its key-value storage engine, thus providing a dual access without downgrade transactional semantics. It offers OLTP and OLAP integration, thus providing support for HTAP that allows for analytical queries over operational data, which realizes the concept of real time business intelligence. Finally, it enables for the execution of polyglot query processing across different and heterogeneous data sources. Modules of the database will be used for implemented all building blocks of INFINITECH that are related to data management.	Significant	Alleviates Tedious ETL Functions in the Banking Sector Saving Significant Time and Costs, while Boosting Data Integration	Economic		Prototype/Demonstrator	Finance	Disruptive	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
<b>Infinitech</b>	Natural Language Processing for real-time, high-accuracy Credit Risk assessment	Technology/System	ReportBrain's NLP functionalities: -Provide real-time structured feeds on risk-assessment worthy information sourced from the news -Interlink entities, updates and maintains knowledge graphs in real-time with all the interlinking of entities -Use "visual" algorithms to collect and analyse the news in real-time in 65 languages 24/7 - Classify articles in real-time by their content (politics, business etc.) -Use its own models, identifies entities (organizations, persons and locations) in real-time	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>		Technology/System				Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Wenalyze Big data analytics platform	Technology/System	Platform that collect and process information from multiple open data sources regarding SMEs and apply cognitive algorithm to detect risk and changes in financial needs. The tool will be use in pilot 13	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Octopush geospatial enabling framework	Technology/System	Octopush is a geospatial enabling framework, developed by AgroApps, allowing the collection, pre-processing, post-processing and distribution of geospatial data products and services, either referred to remote sensing (satellite, drones) acquisitions or multidimensional data outputs from numerical simulations. Octopush allows users to have access through a centralized access point to decentralized services, while Octopush SDK enables IT developers, to easily adapt or	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			<p>expand the provided geospatial services.</p> <p>Octopush was created by AgroApps aiming to address the company/operations and services need for a modular system, independent from any third-party service provider (excluding those offering raw data like Copernicus, NASA etc.), a framework that will be easily adapted to the market needs and follow the service-oriented business model of the company. Octopush is the baseline framework that addresses the AgI companies need in geospatial information, either through the development of new services and data models or the adaptation of the existing ones. Some of the services that are currently offered by Octopush are: a significant number of Vegetation Indices derived from Optical and SAR imagery and crop specific biophysical parameters including leaf area index, chlorophyll content and above ground biomass; crop specific yield estimation; farm management information services such as irrigation scheduling and variable rate fertilization; weather-driven models of possible pests and diseases outbreaks; high resolution weather forecasts and specific agrometeorological parameters; crop damage assessment services.</p>								
<b>Infinitech</b>	AgroApps Weather	Technology/System	Weather Intelligence Engine, developed by AgroApps, is a numerical weather prediction	Significant	Enables Novel Agriinsurance Applications based on IoT Data (Weather Related)	Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
	Intelligence Engine		and atmospheric data assimilation processing chain, based on the WRF numerical weather prediction model. Weather Intelligence Engine is producing operationally all the needed weather data products (Near-real-time, medium-range weather forecasting, subseasonal to seasonal forecasts) by AgroApps offered services								
<b>Infinittech</b>	Sentiment Analysis Tool	Technology/System	Reportbrain Sentiment analysis tool uses application programming interface (API) calls to search existing news article index - Elastic Search Index. The results of the search are processed in Reportbrain's Sentiment Analyzer and the outcome is returned to the caller as a REST API response. This means that articles that are requested by an authorized caller are evaluated in real-time for sentiment and returned to the API caller. Sentiment evaluation describes sentiment as 0 for neutral, -1 for negative or +1 for positive.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinittech</b>	Partitioned and Distributed Transaction Graphs	Technology/System	Ethereum and bitcoin public blockchain transaction datasets.	Significant	Enables Fraud Detection in Cryptocurrencies transactions in an era where cryptoinvestments are booming and trusted is needed more than ever before	Economic and Societal		Prototype/Demonstrator	Finance	Disruptive	new in 2020
<b>Infinittech</b>	ALIDA: A Micro-service based platform for composition, deployment and	Technology/System	A micro-service based platform for composition, deployment, optimisation, execution and monitoring of big data analytics workflows (covering ingestion, preparation, analysis and visualization). It is designed and developed on top of the most	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
	execution of BDA applications		cutting-edge open source Big Data technologies and framework.								
<b>Infinitech</b>	Text Analysis Tool	Technology/System	Reportbrain text analysis tool generates insights from both structured, semi-structured and unstructured text data using natural language processing (NLP). Such insights include sentiment analysis, key phrases, language, and entities, among others. The Reportbrain Analytics Engine uses advanced parallel processing and combines complex NLP tasks in real-time to produce desired results.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Blockchain tokenization	Technology/System	Hyperledger Fabric blockchain support for tokens	Significant	Enhanced Hyperledger Fabric with Tokenization impacting significantly the vibrant Fabric Community of the Linux Foundation	Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Healthentia LifeSciences BigData Platform	Technology/System	BigData platform providing data sources aggregation and management, as well as tools for analytics and visualization. It will be used in the IoT-based Life Insurance pilot (ARI/RRD Lead pilot).	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Event-registry	Technology/System	Event Registry (ER) is real-time cross-lingual global media monitoring service for modelling global social dynamics (eventregistry.org) developed by the JSI. ER aggregates and analyses news content for over 120,000 news sources published globally in 100+ languages. Events mentioned in the news are identified and relevant information about them is automatically extracted and stored in a searchable form. The data can be accessed	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			directly on the platform or via the API. ER supports various analytics including deep analytics of the events and correlations between events, extracted entities and financial data extracted from the main financial indexes.								
<b>Infinitech</b>	Stream Story	Technology/System	Stream-Story multi-resolution modelling and explanation of (possibly real-time) streaming data: (1)Exploratory data mining - A system for the analysis of multivariate time series. It computes and visualizes a hierarchical Markov chain model which captures the qualitative behaviour of the systems' dynamics.; (2) Multi-scale representation - The hierarchical model allows users to interactively find suitable scales for interpreting the data; (3) Real-time monitoring Visualizes streaming data by mapping it to the hierarchical model. It can provide predictions and alarms for different behaviour.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Qminer	Technology/System	QMiner is an analytics platform for large-scale real-time streams containing structured and unstructured data. It is designed for scaling to millions of data points on high-end commodity hardware, providing efficient storage, retrieval and analytics mechanisms with real-time response.			Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	SSC - Super Stream Collider, a multiformat	Technology/System	The SSC enables distributed cloud-based high-performance processing of semantically linked streams i.e. it is an enabler for	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
	Data Management and Query System		semantic analytics. It will be used for analytics over semantically unified/interoperable streams (in WP5), as well as in the KYC and customer-centric services pilot (BOI-led pilot).								
<b>Infinittech</b>	The Global Engine with Neural Network Intelligence (GENNI)	Technology/System	AI Engine executing DL (deep neural networks) algorithms over semantically annotated streams. It will be used in the pilot for customer centric analytics (BOI led pilot).	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinittech</b>	Relational Database	Technology/System	This component consists in the central data repository of the platform. It enables transactional semantics and provides capabilities for query processing based on standard SQL statements. It can scale out on the runtime while continuing serving operational workloads. It can support analytical processing in combination with operational data modifications with the level of isolation to be snapshot isolation. That is, it enables real-time business analytics.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinittech</b>	Polyglot Query Processing	Technology/System	This component enables the query execution over more than one datastore in seamless manner. The data user can submit a single statement and let this component to execute the query by pushing it down to the target database. By doing this, it can process the data on premise, and retrieve only the results, thus it is convenient for cases where data cannot be loaded the platform and they	Significant	Accelerates Application Development in the Finance Sector	Economic		Prototype/Demonstrator	Finance	Disruptive	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			need to be accessed from an external datastore.								
<b>Infinitech</b>	Incremental Analytics	Technology/System	This component enables the query execution in an incremental fashion. The data user will be able to submit a continuous query to the datastore, which will be continuously and incrementally validated. This means that the initial results will be retrieved first, and as data arrives to the data repository, they will be validated against the submitted query, and if it validates the statement, it will be returned to the user.	Significant	Speeds up significant AI Analytics; Enables Real-Time Analytics in the Finance Sector	Economic		Prototype/Demonstrator	Finance	Disruptive	new in 2020
<b>Infinitech</b>	OneHotEncoder	Technology/System	Service to represent categorical variables as binary vectors	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Stream Processor	Technology/System	This component will provide streaming processing capabilities. The data user can declare continuous queries that will be executed over the data stream. It will also allow to combine streaming data with data at-rest and also enable the storage of data streams even when injected in very high rates.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Online Aggregates	Technology/System	This component allows for the execution of aggregate processing operators in an online manner. This way, the definition of the aggregate operations can be defined, and the result of the execution can be pre-calculated in an online manner, preserving data consistency and transactional semantics. When requested, the result of this execution can be retrieved with a GET operation,	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			removing the need to scan the whole dataset/								
<b>Infinitech</b>	Data Collection	Technology/System	This component provides the data ingestion mechanism that:a) enables the acquisition and retrieval of heterogeneous data from a variety of diverse data sources and data providers,b) facilitates the data annotation of the retrieved data by enabling the mapping between the data entities included in the retrieved data and the provided by the data provider data model,c) enables the design and execution of data cleaning operations towards the increase of the data quality of the retrieved data.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Anonymization tool	Technology/System	The anonymization tool modifies data in order to preserve privacy. It is especially indicated in those cases where a dataset contains personal data and it has to be outsourced or shared with a third party. The tool includes different anonymization algorithms that aim at avoiding the appearances of data combinations that could lead to a possible re-identification of the data subjects. It also includes a set of privacy and utility metrics that allow to measure the risk that remains after anonymizing the dataset, and the impact of the anonymization process on the quality of the data.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	DUOS (Digital User On-	Technology/System	Provides remote user registration using eID or passport, dealing with virtual identities in a mobile device. It uses various identity	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
	boarding Services)		proofing and verification services that link new user eID creation (virtual or derived eID) with government issued e-ID. It verifies electronic data stored on chip and machine readable zone. Provides Flexible Multi-factor authentication for different users or identities								
<b>Infinitech</b>	DPO (Data Protection Orchestrator)	Technology/System	The DPO embeds and automats the assurance of security and privacy by design and by default in complex business flows. It provides orchestration of Privacy Enhancing Technologies and related services using BPM tools in order to integrate privacy or data protection perspectives into business processes. It can orchestrate any kind of REST service.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Blockchain Reader	Technology/System	Fetches requested data from blockchain ledger. The specific component is part of a Blockchain chaincode. As the chaincode is tightly connected to the business operation that is performed on top of different business objects, different flavours of the chaincode exist depending on the business use cases (Consent Management, Know Your Customer / Know Your Business, Asset Tokenization).	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Blockchain Writer	Technology/System	Submits transactions on the blockchain ledger. The specific component is part of a Blockchain chaincode. As the chaincode is tightly connected to the business operation that is performed on top of different business objects, different flavours of the chaincode	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			exist depending on the business use cases (Consent Management, Know Your Customer / Know Your Business, Asset Tokenization).								
<b>Infinitech</b>	Smart Contract Executor	Technology/System	Executes smart contracts on the blockchain ledger. The specific component is part of a Blockchain chaincode. As the chaincode is tightly connected to the business operation that is performed on top of different business objects, different flavours of the chaincode exist depending on the business use cases (Consent Management, Know Your Customer / Know Your Business, Asset Tokenization).	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Blockchain Data Visualizer	Technology/System	Queries and displays information about blocks, transactions, chaincodes and transaction families, network name, status and nodes list, organizations list and peers list. The specific component is part of a Blockchain chaincode. As the chaincode is tightly connected to the business operation that is performed on top of different business objects, different flavours of the chaincode exist depending on the business use cases (Consent Management, Know Your Customer / Know Your Business, Asset Tokenization).	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Blockchain Authenticator	Technology/System	Grants access to specific channel(s) of the blockchain network. The specific component is part of a Blockchain chaincode. As the chaincode is tightly connected to the business operation that is performed on top of different business objects, different flavours of the chaincode exist depending on	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			the business use cases (Consent Management, Know Your Customer / Know Your Business, Asset Tokenization).								
<b>Infinitech</b>	Blockchain Encryptor	Technology/System	Encrypts the clients' sensitive data within the smart contract using AES256. The specific component is part of a Blockchain chaincode. As the chaincode is tightly connected to the business operation that is performed on top of different business objects, different flavours of the chaincode exist depending on the business use cases (Consent Management, Know Your Customer / Know Your Business, Asset Tokenization).	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Blockchain Decryptor	Technology/System	Decrypts data retrieved from the blockchain ledger.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Blockchain Transaction Dataset Preparation Component	Technology/System	This component will be responsible for retrieving raw transaction blocks from the Bitcoin and Ethereum blockchains and parsing the blocks in order to extract Bitcoin, Ethereum and major token transactions. After retrieving all the blocks up until now, this component will be run periodically (e.g. once a week or as needed) to retrieve newly generated blockchain blocks during the period.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Scalable Transaction Graph Analysis Component	Technology/System	This component will be responsible for taking massive bitcoin and ethereum public transaction data. Since transaction graph size massive and growing, it will use parallel algorithms to achieve scalability. It will	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			utilize graph and machine learning algorithms to analyse fraudulent transactions								
Infinittech	Semantic Streams Analyzer	Technology/System	Semantic Streams Analyzer Middleware-Engine - SeSAME, The SeSAME component is a data mashup builder for the financial sector that can be used as a data processing component for your data management application, it enhances the capacity to process financial and insurance data in the form of batches and provides a single output, it is ideal when multiple sources have different data formats, it is built to be compatible with the most common data formats in the financial and Insurance sector i.e. FIBO, FIGI and LKIF and additionally it uses INFINITECH Core Graph Data Model to enhance performance. The SeSAME component is designed as a dataflow/workflow execution framework connecting various data input/outputs through the concept of pipelines for creating the data mashup. Conceptually, each financial operator has input data or streams and SeSAME provides one output data or stream. The multiple inputs can be used simultaneously while a single output in RDF is provided. Only the final operator of a workflow can return a format other than RDF, if necessary by defining and transforming the data into the desired format. The data Operators can be of three	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			modes via APIs: API 1) a data acquisition operator is used to collect or receive data from data sources or gateways and can be pull-based or push-based. API 2) a stream processing operator defines stream processing functionalities in a declarative language, e.g., CQELS. API 3) a streaming operator streams the outputs of the final operator of a workflow to the consuming applications. In these three API modes operators of the data transformations and alignments can be done to produce a normalised RDF output format.								
<b>Infinitech</b>	Semantic Reasoner	Technology/System	Enhanced Distributed Reasoner over FinTech Ontologies - EnDoRFIN Semantic Reasoner, Enhanced Distributed Reasoner over FinTech Ontologies - EnDoRFIN Semantic Reasoner, The EnDoRFIN component is a tool for inferring knowledge from data streams, it uses some rules as conditions for defining logics conditions and as a result logical consequences are provided as outcomes. The inference rules are defined based on the most commonly used financial and insurance vocabularies i.e. FIBO, FIGI and LKIF and the way to process the rules is using APIs for defining the logical descriptions for the data applications it is introduced. This component allow the use of other languages but need to be upgraded to the target vocabulary and additionally the EnDoRFIN uses INFINITECH	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			Core Graph Data Model to ensure the inference is applicable to all the involved domains in FINTECHs.								
<b>Infinitech</b>	Ontology Mapping	Technology/System	INFINITECH GRAPH Data Model - Online Ontology Mapping Framework and Toolkit, The INFINITECH GRAPH Data Model is the set of online tools referring to the graphs, formats, vocabularies and ontologies used in the INFINITECH project. The INFINITECH GRAPH Data Model is provided in the form of a set of online accessible files, schemas and metadata model diagrams that represent the way the INFINITECH data can be organised and structured, it also contains the metadata in two different formats, json-ld and owl. The Ontology Mapping Ontologies section contains online machine-readable files both in OWL and JSON-LD format for online accessibility, both files are maintained and updated regularly to keep the latest version of the ontology files up to date using a versioning method..	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Semantic Annotator-Preprocessing	Technology/System	Semantic Annotator-Middleware Preprocessing Layer for FinTechs - SAMPLE-FIN, The INFINITECH SAMPLE-Fin is the support online tool for transforming datasets into RDF-compatible format, beside the online available tool, a set of documentation is provided providing the necessary steps to transform data sets from any data-exchange format i.e. CSV, XLS, etc into RDF. This tool is	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			provided as enabler for a semantic layer where enriched data can be processed more efficiently. INFINITECH SAMPLE-Fin is the mechanism INFINITECH uses for addressing cross domain and cross pilot Data Interoperability and Data Exchange and it also provides the pre-processing layer for the interoperability requirements in INFINITECH project.								
<b>Infinittech</b>	Smart Fleet (IoT Context Management and Historical data component)	Technology/System	A FIWARE-Based framework designed to capture, homogenise, process and distribute real time traffic and smart vehicle's information (it will also allow other related context information). It will implement Pub/Sub mechanisms and support Geolocation and Time series tools. Additional tools to build custom dashboards will be included.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinittech</b>	Fraud Detection Service Training	Technology/System	This component trains the corresponding Machine Learning model that will assign a driver's profile and helps to identify the driver's behaviour using the provided driving routes and vehicle's technical data.	Medium		Economic and Societal		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinittech</b>	Fraud Detection Service Execution	Technology/System	This component provides the drivers' profile and classification (behaviour) for the last given driving route. It refers to the last uses the model trained in component 3.6.1.	Medium		Economic and Societal		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinittech</b>	Pay As You Drive Service Training	Technology/System	This component trains the Machine Learning model that will classify the drivers' behaviour while driving, according to the	Medium		Economic and Societal		Prototype/Demonstrator	Finance	Disruptive	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			data collected from the driven cars (driving routes)								
<b>Infinittech</b>	Pay As You Drive Service Execution	Technology/System	This component will classify the drivers' behaviour while driving according to the data collected from their car and exploiting the Driver Classifier/Driving Profiling ML Models	Medium		Economic and Societal		Prototype/Demonstrator	Finance	Disruptive	new in 2020
<b>Infinittech</b>	Investment Recommendation Engine Training	Technology/System	This component trains the personalized Investment Recommendation engine, to provide a set of recommendations for the financial instruments categories suitable for each customer and his/her investment profile, based on Market Index & Financial Instruments Sentiment data and Customer Risk Profiling ML Models.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinittech</b>	Investment Recommendation Engine Execution	Technology/System	Financial Instruments Personalized Investment Recommendation engine suitable for each customer and his/her investment risk profile, based on Market Index & Financial Instruments Sentiment data.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinittech</b>	Recommender	Technology/System	Generates an actionable insight depending on the output from the remaining components within the Data Analytics layer, i. e. components P5b_Analytics_02 to P5b_Analytics_08.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinittech</b>	Cash Flow Prediction	Technology/System	ML/AI model used to indicate and predict the available working capital (operating cash flow) of the SME (AS-IS & near term future).Alerts/notifications (via BOC Middleware PushNotifications) to be pushed to the respective SME in case of potential	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			lack of liquidity and/or balance moving below a threshold.Cashflow Data/Insights to be provided to BOC Middleware Mobile/Web BPF.In order to provide valuable insights to the SME the data should be collected and streamed in real (or near real) time, since whenever a new object/entry appears the model should retrain and adapt.								
<b>Infinitech</b>	Budget Prediction	Technology/Sy stem	AI (ML) model is used to support the budget target setting for the various categories used by the respective SME. Doing so by providing budget predictions for each utilized category. The underlying model will take into consideration the cash flow analysis output, benchmark, macroeconomic and other available SME data (Business Plan).	Medium		Economic		Prototype/Demo nstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	KPI Engine	Technology/Sy stem	The KPI engine calculates KPIs with regards to the Financial Health and Performance of the respective SME.Doing so by taking into consideration the respective SME profile (e.g. maturity stage), accounting-wise optimal KPI values and how other similar SMEs perform.	Medium		Economic		Prototype/Demo nstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Transaction (Txn) Monitoring	Technology/Sy stem	A dynamic complex event processing (CEP) mechanism that monitors the transactions of the user. In case transaction amount or type deviates from normal behaviour the user will be informed of abnormal transactions in order to be safeguarded from double payment mistakes and potential fraud attempts.In addition, expense pattern are also analysed to identify potential savings	Medium		Economic		Prototype/Demo nstrator	Finance	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			for instance multiple subscription spending or high ATM fees.								
Infinittech	Transaction (Txn) Categorization	Technology/System	Smart transaction auto-classification which would also allow the user to manually override the given transaction category and define a new one (re-classify). The categorization performed will be based on the needs of the individual SME.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
Infinittech	Invoice Processing	Technology/System	Processes transaction data and ERP data in order to keep record of the invoices which have been partially or fully paid by the SME. Provides insights regarding the respective VAT amount payable and the optimization of cash flow by providing background info on invoices such as paying a invoice at the "right" time. In cases where available invoice data is limited the engine will utilize a simplified approach in order to derive the expected VAT amount to be paid by the SME at the next VAT due date. Doing so by utilizing the banks transaction info or/and past VAT payment amounts.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
Infinittech	KMeans	Technology/System	Batch BDA Service: Given a set of observations $(x_1, x_2, \dots, x_n)$ , where each observation is a $d$ -dimensional real vector, $k$ -means clustering aims to partition the $n$ observations into $k$ ( $k \leq n$ ) sets $S = \{S_1, S_2, \dots, S_k\}$ so as to minimize the within-cluster variance.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
Infinittech	Random Forest (Model)	Technology/System	Batch BDA Service: An ensemble learning method for classification, regression and	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			other tasks that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. Random decision forests correct for decision trees' habit of overfitting to their training set								
<b>Infinitech</b>	Random Forest (Predict)	Technology/System	Streaming random forests algorithm	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Client Contextual Information	Technology/System	The component generates and updates a relevant client contextual information related to the clients' data and behaviour	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Financial Fraud/Crime Risk Score	Technology/System	Using the clients' contextual information and transactions a risk score of a fraudulent request of an online instant loan is evaluated	Medium		Economic		Prototype/Demonstrator	Finance	Disruptive	new in 2020
<b>Infinitech</b>	Anomaly Analysis	Technology/System	Anomaly Analysis provides two main functionalities: Anomaly detection· Anomaly prediction for time series data.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Pattern Analysis	Technology/System	Pattern Analysis provides two main functionalities: Pattern matching· DiscoveryThe component will provide support for detection of complex patterns on data graphs	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>	Stream Story	Technology/System	Stream Story is a component for the analysis of multivariate time series. It computes and visualizes a hierarchical Markov chain model which captures the qualitative behaviour of the systems' dynamics, where system is described with a group of timeseries	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
Infinittech	Open API Gateway	Technology/System	This component provides the single point of entry for the added-value functionalities (such as the Machine Learning (ML) / Deep Learning (DL) analytics functionalities) of INFINITECH which are based on microservices. The specific component enables the discovery and invocation of the dynamically registered microservices, effectively handling the incoming requests towards these microservice instances.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
Infinittech	User Interface for Blockchain Transaction Reports and Visualization Component	Technology/System	This component will be responsible for providing user interaction with the Scalable Transaction Graph Analysis component within the bank and collect/manage user as well as annotated blacklisted blockchain addresses. It will utilize OpenAPIs (REST APIs) to submit queries consisting of customer blockchain addresses and blacklists to transaction graph analysis component and generate web based reports and visualization based received results	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
Infinittech	Visualization Preparation	Technology/System	Stream BDA Service: Service to prepare data to the visualization depending on the type of incoming data or the data you want to view	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
Infinittech	Real Time Visualization	Technology/System	Useful tool for displaying charts through web application	Medium				Prototype/Demonstrator	Finance	Incremental	new in 2020
Infinittech	INFINISTORE	Technology/System	This is a GENERIC DATA STORE implementation for the INFINITECH Project as a microservice on top of a noSQL DB (MongoDB) instance. It is fed by different	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			data ingestions servers and supports all other services.								
<b>Infinittech</b>	UI Risk Assessment based on VaR	Technology/System	Web application to monitor portfolio risk in real time, perform what-if analysis, providing also several statistics of the underlying financial assets.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinittech</b>	Pseudo-anonymization tool	Technology/System	A tool that pseudo-anonymize data in order to preserve privacy. The component needs a specific configuration/development for each pilot in which it is used.	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinittech</b>	Health insurance risk assessment service	Technology/System	Algorithm yielding risk based on user RWD and a pre-trained classifier. Implemented as a Python script	Medium		Economic and Societal		Prototype/Demonstrator	Finance	Disruptive	new in 2020
<b>Infinittech</b>	Health insurance fraud detection service	Technology/System	Algorithm detecting fraudulent behavior of insurance company customers. Implemented as a Python script	Medium		Economic and Societal		Prototype/Demonstrator	Finance	Disruptive	new in 2020
<b>Infinittech</b>	Well-being outlook classifiers	Technology/System	Classifiers to be used by the health insurance risk assessment algorithm. The current set includes Random Forest, Logistic Regression and Neural Network classifiers of varying complexity. The format of the classifier depends on its type.	Medium		Economic and Societal		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinittech</b>	Synthetic RWD for well-being analytics	Technology/System	Synthetic data of the same format as those collected in Pilot 12. The current version of the data span 1,000 people simulated for 116 weeks each	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinittech</b>	Open Banking Aggregator Solution	Technology/System	Crowdpolicy Open Banking Aggregator Solution is a modular architecture (UIs, connectors & APIs) platform so that it can be integrated into web / mobile banking	Medium		Economic and Societal		Prototype/Demonstrator	Finance	Incremental	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			applications, by the existing provider of the Bank in the form of API integration, but also as a separate application that could be made available to the users of Bank's online services Compatibility with best known market standards based on the European PSD2 Directive (Berlin Group, Open Banking UK, STET) Support for PISP & AISP services based on the PSD2 European Directive: - Payment Initiation Services - Account information Services								
<b>Infinitech</b>	BADP (Big data analytics platform)	Technology/System	Platform that collect and process information from multiple open data sources regarding SMEs and apply cognitive algorithm to detect risk and changes in financial needs. The tool will be use in pilot 13	Medium		Economic		Prototype/Demonstrator	Finance	Incremental	new in 2020
<b>Infinitech</b>										Incremental	new in 2020
<b>TT</b>											
<b>TT</b>	HORUS DASHBOARD-AUSOL/NORTE LITORAL	Software	Customized dashboard for AUSOL and NORTE LITORAL based on HORUS software. It has been developed for helping and assisting road operators into operations and maintenance works in a road infrastructure.			Economic			Transport (Highways)	Incremental	Delivered in 2018
<b>TT</b>	MAID - Animal intrusion detection system	Technology	Passive Infrared detector has turned out as a reliable technology to detect persons and animals within short ranges. Scare-out systems based on ultrasound generators can create effective, invisible and non-physical fences only audible by medium to big animals like cows, wild pigs and dogs. Both systems work together thanks to the			Economic			Transport (Highways)	Incremental	Delivered in 2018

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			deep learning models that decide the better reaction to each species of animal								
TT	Big Data - Descriptive Mobility Patterns	Knowledge	The results of the project can be exploited for the benefit of Cintra and Indra. The conditions of exploitation have not yet been determined but it is sure that the conclusions and the most important, the knowledge learnt along the project could be exploited for our benefit			Economic			Transport (Highways)		
TT	S2N - O-D Matrixes and user loyalty (Bluetooth / Wi-Fi)	Knowledge	Traffic real time data is being gathered and loaded into the corporate IoT Platform to predict traffic flows and segment users Capture public information that vehicles and drivers make available themselves: Bluetooth and WIFI signals Compliance with the GDPR The information we get by scanning the RF space at the BT and WIFI frequencies, gives us an anonymous heat map of user's preferences, which is crucial when talking about user experience.			Economic					
TT	Connected Cars Platform: Value added services on car data in the cloud.	Service	Descriptive and predictive analytics based on data coming from connected cars sensors.			Economic			Transport (Connected Cars)	Incremental	
TT	Improved Routing Service for Trucks	Component	Enhanced component to provide better and more detailed planning possibilities for truck routing and processes of Logistic Service Providers			Economic			Transport (Connected Trucks)	Incremental	
TT	Predictive & automated	Software	Predicting and ultimately enabling the prevention of failure of assets within the			Economic			Transport, (Railways)		

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
	diagnostic capabilities in the mainline rail industry		mainline rail industry by consuming large amounts of data from many different sources								
TT	Descriptive and predictive models for the infrastructure maintenance of railway high speed lines	Software models	These models will try to identify the relation between all the aspects of the railway exploitation in order to be able to improve the maintenance activities of the railway infrastructure focus on High Speed Lines. There are three main objectives at this pilot, which try to optimize the maintenance of these infrastructure elements, tracks, switch and crossing and slopes.			Economic			Transport, (Railways)		
TT	Port and Terminal Productivity Cockpit	Process	The Port & Terminal productivity cockpit is a web-base tool for supporting a common set of shared and specific metrics to evaluate the logistic process of a port terminal. This logistic process could be depicted as a hub process where the clients are inputs/outputs from different transportation models: vessel, train, and truck. One of the main performance indicators for a terminal is to provide the shortest possible time to leave or pick up containers at the lowest possible cost. This time-cost ratio is an overall indicator to measure performance that is transversal for the domain. The most widely used Key Performance Indicator (KPI) is the Truck/Train/Vessel Turnaround Time (T/VT): the time a specific means of transport spends in the terminal to fulfil an order. At this point	Medium		Economic	TRL6	Prototype	Maritime Ports and Terminals	Incremental	Delivered in 2018

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			turnaround time is known in isolation, i.e. in a specific terminal, and calculated after the order is processed. Our solution integrates Big Data from several stakeholders involved into the containers management in order to provide better insights and metrics about the overall efficiency at both port and terminal levels. The defined KPIs provide benchmarking capabilities (e.g. related to costs and performance) that may indicate different levels of competitiveness. A key feature of the Terminal productivity cockpit goal is to achieve an improvement in KPIs as the Truck Turnaround Time by predicting its trend using analytical models. These models have been developed using machine learning techniques and the whole set of data available. Therefore these models provide insights about current operations, thus helping stakeholders to understand how to improve the related KPIs.								
TT	Crane and Spreader Sensorization Infrastructure	Product	Sensorization of the Crane's Spreader to support the collection and transmission of signals in order to provide the necessary inputs for task such as operative monitoring or supporting maintenance tasks			Economic			Maritime Ports and Terminals	Incremental	
TT	Terminal Productivity Cockpit	System	The terminal productivity cockpit will exploit advanced data processing and predictive analytics to facilitate proactive decision-making and process adaptation. The productivity cockpit will leverage cutting-			Economic			Ports and Terminals	Incremental	

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			edge predictive business process monitoring solutions, i.e., real-time predictive big data analytics for terminal processes. Predictions are made for train delay based on real-time machine data.								
TT	Operation Management Predictive Optimization Module	Software module	This module will exploit predictive analytics with Passenger Flow data obtained in real time from airport and airline systems in order to facilitate proactive decision making in real time whenever there is any disruption over the initial plan			Economic			Transport, Airport management		
TT	Descriptive passenger behavior system	Software system	These models will identify the different type of passengers and their behaviors, to predict when they arrive to the airport, how they move across the airport terminal in order to dimension airport resources accordingly to the expected demand and optimize airport and airline procedures.			Economic			Transport, Airport management		
TT	Web page with Traffic related tweets for specific region and analysis of traffic cameras	Service	The asset is a web page, composed of twitter feeds related to traffic in a city or area.			Economic and Societal			Transport, Urban Mobility		
TT	Parking Reservation APP	Software	The application which is built on top of the parking management system, contains information on parking places. Drivers and logistic operators can reserve parking spaces through the app. Parking wardens can verify through the parking management system the validity of the reservation.			Economic and Societal			Transport, Urban Mobility		

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
TT	Urban Traffic Fluency Model	software model	The existing fluency model is remarkably improved by the use of floating car data (FCD) from the city buses.			Economic			Transport, Urban Mobility		
TT	Customized traffic incident alert	Service	Traffic incidents are automatically alerted to end users using Twitter as delivery channel			Economic and Societal			Transport, Urban Mobility		
TT	Traffic Model for freight transport in Valladolid City	Software model	A traffic model for particular areas in the city where freight transport has more impact to understand and gain insights about the patterns and attitudes of delivery areas in city centres			Economic			Transport, Urban Mobility		
TT	Dashboard for micro-simulation insights in Valladolid City Centre	Software model	Dashboard with micro-simulation traffic model results to show valuable insights from Data Analytics tasks done within TT project			Economic			Transport, Urban Mobility		
TT	Integrated Platform	System	The integrated platform of WP10 in which all the algorithms and datasets are successfully integrated and a ready to use system is designed			Economic			Dynamic Supply Networks, e-Commerce		
TT	Descriptive Analytics Service	Service	A descriptive analytics tool that will offer to the potential users: a) interaction with data through the selection of various criteria/dimensions, and b) Detailed analysis of data behavior in 3PL partners through a set of appropriate visuals, and deduction of patterns and trends.			Economic			Dynamic Supply Networks, e-Commerce		
TT	Forecasting Analytics Service	Service	A forecasting tool that will allow the potential user to periodically predict the behavior of different groups of data depending on the selection of various			Economic			Dynamic Supply Networks, e-Commerce		

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			criteria/dimensions, while presenting its forecasts through simple and convenient graphs								
TT	Data Mining Approach	Method	An artifact, which is a clustering-based approach that accomplishes behavioral segmentation and characterization of shoppers' orders, by examining the product categories included in each customer delivery			Economic			Dynamic Supply Networks, e-Commerce		
TT	Route Optimization Service	Service	Inventory routing in an omnichannel environment			Economic			Dynamic Supply Networks, e-Commerce		
TT	Customer Reviews Analysis Service	Service	Extract online consumer insights with respect to logistics and delivery processes			Economic			Dynamic Supply Networks, e-Commerce		
<b>Safe-DEED</b>											
Safe-DEED	Private Selective Aggregation Library	Software technology/process	It is a versatile protocol using homomorphic encryption with a broad range of privacy-preserving computation applications ranging from COVID-heatmaps to privacy-preserving questionnaires.	Medium		Economic and Societal		prototype	agnostic	Disruptive	new in 2020
Safe-DEED	Private Set Intersection Java Library	Software technology/process	A special-purpose secure multi-party computation library. It allows companies to compute the intersection of their customer-relationship-management database privately.	Medium		Economic		prototype	agnostic	Disruptive	new in 2020
Safe-DEED	Infineon-SafeDEED Demonstrator	Software Technology	The demonstrator offers an easy and secure platform for exchanging the order data needed for Lead Time Based Pricing while	Significant	The demonstrator integrates MPC technology into an existing order management platform (analogous functionality to an e-commerce store)	Economic		already supported by a	Supply Chain	Architectural	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			also keeping the data of the supplier and the customer private and secure. In order to achieve this, secure multi-party computing (MPC) is applied.		where existing customers of a manufacturer can purchase unallocated ATP from the manufacturer privately and securely. Enable the manufacturer or supplier to implement a Lead Time Based Pricing (LTBP) strategy to charge more money for an improved service.			Prototype / Demonstrator			
Safe-DEED	A Horizontal Patent Test Collection	Test Collection	A publicly available test collection to test de-anonymisation algorithms based on text features. Published in SIGIR 2019. <a href="https://dl.acm.org/doi/10.1145/3331184.3331346">https://dl.acm.org/doi/10.1145/3331184.3331346</a>	Medium		Economic		published	agnostic	Incremental	new in 2019
Safe-DEED	PrioPrivacy	Technology/method	A k-anonymisation tool called PrioPrivacy that considers the priorities of attributes given by the user and is capable of outperforming the state-of-the-art. Published in WI 2019. <a href="https://dl.acm.org/doi/10.1145/3358695.3360918">https://dl.acm.org/doi/10.1145/3358695.3360918</a>	Medium		Economic		demonstrator	agnostic	Incremental	new in 2019
Safe-DEED	De-Anonymisation Risk Analysis toolkit	Technology/method	A de-anonymisation risk analysis toolkit with 3 modules corresponding to different types of datasets (tabular, aggregated, invoices). Part of a publication in RCIS 2020. <a href="https://link.springer.com/chapter/10.1007/978-3-030-50316-1_32">https://link.springer.com/chapter/10.1007/978-3-030-50316-1_32</a>	Medium		Economic		demonstrator	agnostic	Incremental	new in 2020
Safe-DEED	Data Valuation Component (DVC)	Technology/method	The input to the DVC is a structured data set together with the context in which it will be used. The data valuation process then maps various aspects related to the data set to the value in the given context. These aspects are: data properties (size, age, accessibility etc.), data quality (completeness, domain validity,	Significant	The dual use of the platform will enable organisations to be more aware of the value of data they are holding. An economic valuation, such as that occurring in a market could put a euro-tag on such data. The biggest benefit in a market environment would be the promotion of	Economic and Societal	4	Prototype and Demonstrator	All	Disruptive	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			format validity, timeliness), context (systems & economics, legal & terms, business, technical). The output of the DVC are 3 scores (contextual, quality, aggregated value) that reflect the value of data from each of these perspectives. The DVC can be used both as a standalone application for internal valuation of an organisation's own datasets, as well as deployed as part of data markets, thus increasing the trust and transparency of data exchanges.		transparent data valuation methods in the context of data exchanges.						
<b>Safe-DEED</b>	Data-Driven Business Modell Innovation Toolset	Service	A toolset of three methods (Data Map, Data-Service Cards, Data-Driven Business Canvas) to support the value creation process (innovation process) based on data. The tools are integrated into an existing commercial platform for business model innovation	Medium		Economic	7	published	agnostic	Incremental	new in 2020
<b>TRUST</b>											
<b>Trust</b>	TRUSTS platform	System	Platform for secure sharing and trading of data assets	Medium	FOAK with decentral architecture, utilizing IDS Framework	Economic	7	MVP in 2021, prototype in 2022	n/a	Disruptive	expected in +2021
<b>Trust</b>	TRUSTS Knowledge Dissemination	Other	Project-related knowledge dissemination activities	Medium	Elevated societal awareness	Societal		n/a	all	Incremental	new in 2020
<b>Trust</b>	TRUSTS R&D Reports	Other	Publication of project reports	Medium	Open access sharing of conceptual R&D project results	Economic and Societal		n/a	all	Incremental	new in 2020
<b>EDI INCUBATOR</b>											
<b>Bettair (EDI)</b>											

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
Bettair	Prediction of Pollution Episodes	Technology, service	We can offer a new service in our Platform: forecasting of air pollution	Significant	If permit to be more attractive and more competitive in public procurements	Economic and Societal	8	It is implemented, ready to the next phase, production	Air pollution	Architectural	expected in +2021
EDI											
EDI	Bettair Cities - Prediction of Pollution Episodes	Technology, service	We can offer a new service in our Platform: forecasting of air pollution	Significant	If permit to be more attractive and more competitive in public procurements	Economic and Societal	8	It is implemented, ready to the next phase, production	Air pollution	Architectural	expected in +2021
EDI	Aindo - Synthetic data generator	Technology	We developed a synthetic data generator, based on Artificial Intelligence algorithms, to enable companies to share their data in a secure and confidential manner. This technology fosters open innovation and research projects between institutions ensuring GDPR compliance.	Significant			8	We are in the process of submitting an application for a patent.	cross-industry	Disruptive	
EDI	Builtrix - Building clustering algorithm	Software	This innovation is able to provide a meaningful benchmarking of buildings based on their consumption data and meta-data. This innovation enables asset managers to keep tracking their energy efficiency in a fair comparison with other buildings in Builtrix database.	Significant	The impact of this innovative feature on Builtrix business was relatively high, because it was built based on a strong demand of Builtrix clients. Furthermore, it increased the competitive advantage of Builtrix in the energy analytics market.	Economic and Societal	TRL 8	Commercial MVP	Energy	Incremental	new in 2019
EDI	Energus - GEM-ANALYTICS	Product	The Genius Energy Manager (GEM) Analytics is a software platform that includes Artificial Intelligence algorithms performing automated energy analytics based on the data collected by sensors and installed in customers' buildings. The GEM-Analytics	Significant	GEM-Analytics has opened a new market for the company, allowing us to attract and manage more industrial and tertiary customers. The impact on the company is radical: unlike the other products, which are supplied on a project basis, GEM-Analytics is a fully automated	Economic and Societal	8	The software developed in R&D starting from June 2020 entered the company	Energy	Disruptive	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			platform operates on top of a distributed computing infrastructure (based on Fog Computing). The GEM-Analytics algorithms are tailored to the specific characteristics and needs of each customer, and are supported by historical knowledge about previous measurements and analysis carried out for the same and (possibly) other customers. The platform is modular and is composed of a series of artifacts that solve specific energy problems and that can be combined with each other. The goal is to optimize the performance of machinery and buildings to save energy and money. The target of the product are energy-intensive industries, commercial buildings and shops.		software for energy management, and allows us to scale much faster, thus significantly increasing our revenues.			portfolio and is now offered to our customers in Software-as-a-Service (SaaS) mode with an annual fee. In February 2021, we received orders for more than 120,000 euros in annual fees			
EDI	CyC - CUSTOMER DATA	SERVICE	Thanks to EDI we propose, developed and test a solution to serve new insights to customer care professionals through the combination of AI and Business Intelligence solutions to the envisaged markets (service companies)+	Medium		Economic		Our solution was build and tested through a prototype	Service companies (utilities, banking, assurance, telcos)	Incremental	new in 2019
EDI	CyC - APPLIED AI FOR...	SERVICE	Thank to the approach tested on EDI, we were able to apply the technology on completely different scenarios with the same components, architecture and concepts: AI must uncover interesting insights and the data, and context, should be offered to the final user through data analysis solutions.	Significant	Though they have been small initiatives, since EDI we have been able to apply some technologies, concept and approaches on several scenarios including health, automotive and public administration.	Economic		Already supported by demonstrator on different scenarios.	Public Administration, health and automotive	Architectural	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
EDI	Feelingstreams - chat automated analysis	product	large corporations want to improve chatbot conversations. We help them to understand where the bot can't understand the customers. Customers are more happy to talk with a chatbot	medium	better customer experience	Economic	7	industrial prototype	finance	Disruptive	new in 2020
EDI	Maxfone - Total Wall	Service	Total Wall is an innovative web based solution for the real-time promotion, management and analysis of every type of digital or hybrid events. The platform allows the user to monitor the images posted on Twitter and Instagram and to embed live streaming videos from Facebook, LinkedIn, Youtube, or video calls from Zoom, Webex and GoToMeeting. The interface can be viewed on full screens or any display. In addition, it is possible to activate a call to action in order to boost the engagement level during the event and foster the posting of new contents on social media. The Wall can be shared directly as a web link. Finally, a refined technological architecture analyzes in real-time the performances of the event by measuring interactions and behaviors and studying the image content (AI integration).	Significant	Before the pandemic emergency the impact foreseen was medium. Nowadays, since the event industry has been obliged to switch to a full digital format, the impact is valued as significant.	Economic and Societal	9	Before the pandemic our data provider (IAAF) expressed interest in our solution, however the spreading of the Covid_19 has canceled the event itself. During 2020 the service has been acquired by other customers and applied for different events (national events), ranging from sports to entertainment, to conventions. The starting project has been boosted on the analysis level	Mainly the event industry (from sports to entertainment, to conventions and exhibitions). And secondly th retail industry, editorial industry, advertising industry and any kind on industry where the service can be applied as a call-to-action tool.	Disruptive	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
								through the introduction of an artificial intelligence model for business intelligence operations (information on audience' traits) and content validation (image recognition). In addition was implemented an audio analysis feature. The service is fully developed and currently marketed in Italy, but we are planning to also distribute it internationally.			
EDI	SMAP - Meter data interpolation algorithm	Methods (software)	This project resulted in the demonstration of a method for estimating missing energy consumption from smart meter data. This innovation may have applications in assessing accuracy of energy consumption	Medium		Economic	6	industrial prototype	finance	Incremental	new in 2019

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
			and estimating energy consumption where smart meter data is otherwise unavailable.								
EDI	uVantage	technology, methods	uVantage delivers behavioral risk analytics to the global finance community, lenders, transaction processing service providers, asset managers and loan issuing institutions to provide the best available fraud and loan repayment prediction, by analyzing unstructured and structured information and alternative data with machine learning technology.	Significant	Results include an increase in lending opportunities, a reduction in fraud and an increase in credit access even in situations where typical credit scores are unavailable.	Economic	7	industrial prototype	finance	Disruptive	new in 2020
EDI	InnovaHealth - Prediction of Recovery Days	Technology	We developed a prediction system through algorithms that estimates a number of days of recovery and a final state of recovery, based on a set of anonymized data such as pathology or socioeconomic parameters.	Significant	Predicting those patients who are most likely to have a worse recovery allows health professionals to be more intensive in the healthcare provided, thus improving healthcare	Economic and Societal	6	Prototype	Healthcare	Disruptive	new in 2020
EDI	Orbem - Genus Branta	Method	The Genus Branta combines artificial intelligence with magnetic resonance imaging to analyze and classify eggs. The Branta measures the following features prior to incubation: fertility status, egg position, egg volume, internal egg quality, presence/absence of double yolks,	Significant	Instead of wasting billions of eggs, these can be removed prior to incubation and used as a novel food source instead.	Economic and Societal	4	Prototype in development	Agriculture	Disruptive	new in 2020
EDI	adiutaByte - automatic Depot optimization	technology + service	find best locations for new depots based on historical events and data	Significant	less CO2 emissions + more flexibility	Economic	TRL 6/7	prototype	Fleetplanning & Logistics	Architectural	new in 2020
EDI	adiutaByte - automatic Depot allocation	technology + service	find best allocation of vehicles/ employees to depots based on historical events and data	Significant	less CO2 emissions + less waiting time for customers	Economic and Societal	TRL 6/7	prototype	Fleetplanning & Logistics	Architectural	new in 2020
EDI	adiutaByte- automatic shift-	service	define best fitting shift times for employees based on given the depot allocation solution	Significant	less waiting time for customer + less idle time for employees	Societal	TRL 6/7	prototype	Fleetplanning & Logistics	Architectural	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
	time optimization										
EDI	Advanced Infrastructure - PV Outturn Visualisation	Methods	A way of displaying PV outturn using population a proxy for power flow analysis	Medium	Incremental innovation allowing more intuitive visualisation of PV outturn from an otherwise difficult to interpret dataset	Economic	5	In the process of commercial exploitation by the company that invented it	Energy	Incremental	expected in +2021
EDI	Amplify Analytics - CMAB for Smart Bidding	technology, system, product, service, processes	Application of Contextual Multi-Arm Bandit methods to keyword bidding in paid search	Significant	Accounts for a dynamic contextual landscape in marketing and in a way that expects changing conditions while optimizing for budget allocation and learning from human expert knowledge.	Economic	TRL6	A working MVP has been developed already. Operational testing remains to be done.	Advertising	Architectural	new in 2020
EDI	Amplify Analytics - Human-in-the-loop	technology, system, product, service, processes	Intervention by a human subject-matter-expert, which can guide the learning of the CMAB agent	Significant	Allows human expert knowledge to be incorporated into the bidding output of the model and allows the model to learn the behaviours of campaign managers in identifying optimal keyword bids.	Economic	TRL6	A working MVP has been developed already. Operational testing remains to be done.	Advertising	Architectural	new in 2020
EDI	Amplify Analytics - Privacy Preservation in Smart Bidding	technology, system, product, service, processes	Contextual Bandit learning on Personal data at Client's side in applications that involve user data	Significant	Compliance with the strict GDPR guidelines, yet leveraging the power of Bandits	Economic and Societal	TRL2	Technology concept and/or application formulated	Advertising	Disruptive	expected in +2021
EDI	Amplify Analytics - Custom Analytics on Smart Bidding performance	technology, system, product, service, processes	Tiered pricing for custom insights presented on a dashboard for smart bidding	Medium	Differentiates from other solutions in the market, in that it combines bespoke analytics with a product offering	Economic	TRL3	Technology concept and/or application formulated	Advertising	Incremental	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
EDI	Creation Labs - Computer Graphics Simulation Engine	Technology	Computer graphics simulation engine for autonomous vehicles. Allows you to simulate different scenarios and setup scenarios	Medium			6	Prototype	Technology	Incremental	new in 2020
EDI	Creation Labs - GAN-SIM	Technology	AI domain adaptation technology that makes computer rendered images look more realistic	Significant	If you can train an AI in a simulation engine it significantly reduces the cost of acquiring data	Economic	7	Prototype	Technology	Disruptive	new in 2020
EDI	Divizend Maximizer	Service	The first fully digital platform for private investors to reclaim foreign withholding tax on dividends to be much more cost-efficient than traditional methods (e.g. tax advisors), to save time and in the local language. Worldwide market size of unreclaimed withholding taxes: \$18B per year.	Significant	Until now, no such application exists for private investors, and often it was not feasible or profitable to reclaim withholding taxes. After many proposals by the EU to fix this issue in a cross-border way, we are the first FinTech to tackle this problem. This way, cross-border investments are encouraged again and the goal of a European Capital Markets Union comes closer.	Economic and Societal	9	Fully developed MVP, currently testing with real customers. International expansion in the coming months.	Financial Services	Disruptive	expected in +2021
EDI	Divizend Screener	Service	The world's first interactive dividend screener, which allows finding attractive dividend stocks from the 20 most important countries for securities. High value add for customers, as with the Screener, they have access to high quality stock and dividends data (powered by our SecDB) for the very first time.	Significant	Until now, normal retail investors have a hard time if they want to make informed decisions about their dividend stocks, as such information is not easily available. With its innovative filter methods, the Screener solves that problem.	Economic and Societal	7	Fully developed MVP, currently testing with real customers.	Financial Services	Disruptive	expected in +2021
EDI	Divizend - Securities Aggregation API (SecAPI)	Technology/Service	The common, unified, country-independent interface for accessing the APIs of banks and other financial institutions, most importantly for aggregating securities accounts.	Significant	Access to securities accounts is not standardized yet, in contrast to PSD2 for payments, thus startups and other companies have a high entry barrier if they want to offer personalized stock-related services to their customers. By offering	Economic and Societal	8	The SecAPI is already one of the underlying components of the Maximizer. New supported	Financial Services	Radical	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
					the SecAPI to other market players in the future, Divizend will solve this problem.			bank APIs are added constantly. Currently, we have covered most of the German and Austrian market.			
EDI	Divizend - Securities Database (SecDB)	Technology/Service	The single point of truth for all international information about companies, stocks and dividends (e.g. date, amount, history). For this database, we ensure highest data quality and the most recent data through processes like automatic scraping, data curation, cross-validation with other data sources and manual validation.	Significant	Most stock- and company-related information is either behind a strict paywall of large data provider companies (e.g. FactSet) and possibly even only available to institutions, or only scattered on a public website, where it is infeasible to process it in an automated way. As the SecDB is based on graph database Neo4j, it is highly flexible, provides best data quality and allows extending the data model arbitrarily. By offering SecDB access to other market players in the future, they will be able to provide their users much more accurate analyses etc.	Economic and Societal	7	The SecDB is the basis for the SecAPI and the Divizend Screener. It is constantly expanded with the most recent stock, dividend and company information, as well as stocks from new countries.	Financial Services	Radical	expected in +2021
EDI	Fractal Mind - Contextual Forecasting	technology, methods, product, service, methods, processes	Contextual justification of demand drivers in product forecasting to sustain more accurate and context aware forecasting.	Significant	Standard forecasting platforms enable for a preview of a product's performance but lack detailed justification regarding the drivers of demand, for each product in any given moment in time. This informs decision makers regarding strategic decisions, towards more efficient processes.	Economic	TRL 6	A working MVP has been developed already. Operational testing remains to be done.	Retail, Health, Employee Engagement	Architectural	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
EDI	Fractal Mind - Product Performance Simulation	technology, methods, product, service, methods, processes	Simulation of variable shock in product forecast and evaluation of cross effect within product categories, useful to increase leverage in negotiation and assess impact of business decisions in the competition.	Significant	Simulating product variable variations like price or marketing investment and evaluating impact of said decisions in product performance and the overall competitor market, enables decision makers to have more leverage during negotiation processes.	Economic	TRL 6	A working MVP has been developed already. Operational testing remains to be done.	Retail, Health, Employee Engagement	Architectural	new in 2020
EDI	Fractal Mind - Simulation based Gamified Engagements	technology, methods, product, service, methods, processes	Generation of gamified engagements from simulated forecast scenarios, connecting business analysis to operational decisions with consequence.	Significant	Transforming simulated scenarios into actionable engagements enables operational managers to have a direct consequence of business analysis efforts, turning a complex forecast into a list of actionable behaviors, tailored with game mechanics to increase engagement.	Economic and Societal	TRL 6	A working MVP has been developed already. Operational testing remains to be done.	Retail, Health, Employee Engagement	Architectural	expected in +2021
EDI	Smart Planner - Grama	Product	Smart Planner is a tool to identify the best combination of locations to place electric vehicles charging points.	Medium	Large corporations often need to determine the best location for their long-term infrastructure investments. When using information of existing infrastructure, demographics, business related information and natural geography features you can save millions by improving the outcome while planning new infrastructure. Considering the EVs charging operators market in Europe it is foreseen to grow from a 600M€ of investment in new charging stations in 2020 to 80B€ in 2030, with the number of charging points growing from 18k in 2020 to 2.9M in 2030 . If with our solution we may contribute with a marginal optimization of 1% of the foreseen investment we may provide to our customer benefits of 800M€ in 2030 alone.	Economic	TRL6	MVP	Energy, Evs	Architectural	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
EDI	ID Ward	Technology	ID Ward is a decentralised solution to GDPR compliance and the deprecation of cookies. It aggregates personal data on the device and anonymises it using decentralised machine learning, ID Ward allows publishers and marketers to achieve privacy-compliant marketing in cookieless browsers and operating systems, protecting user data and marketing revenues.	Significant	ID Ward allows privacy-preserving audience targeting without using cookies. This saves publishers and advertisers over 60% of their revenues by allowing them to optimise their marketing spend, and reduces legal risks and liabilities related to the collection and management of personal data.	Economic and Societal	8	The technology is available for public use in beta mode.	Digital marketing	Disruptive	expected in +2021
EDI	iERP.ai	technology, methods	iERP.ai delivers innovative sales forecasting solution for retail industry. By implementing industry specific external factors, this solution is delivering much higher precision of forecasts for retail businesses	Significant	Having precise sales forecastings, retail businesses can reduce tier costs, environmental impact and waste of products	Economic	7	Prototype	Retail	Disruptive	new in 2020
EDI	IN2 - Automatic Video Chaptering Solution	product	We have developed a new product for content owners to better exploit their content by improving the metadata they have available for their videos. Using our product they can bypass the costly process of manually annotating videos and enrich the data about their content automatically. The product can be used for new videos as well as for large archives.	Significant	Because of this automatic enrichment, content owners can re-use the video in OTT services, license it out and generate new revenue streams from their archive. Moreover, staff that were previously tasked with manual video annotation can now focus on more creative tasks for which they are better suited.	Economic	7	Demonstrator with content owner	Media / Content	Disruptive	new in 2020
EDI	Instaon - Keyword processing	technology + service	Instaon filters and finds a large volume of relevant keywords for any business to ensure high ROI for search advertising.	Significant	less CO2 emissions + more flexibility	Economic	TRL 6/7	prototype	Fleetplanning & Logistics	Architectural	new in 2020
EDI	Instaon - Keyword categorization	technology	Instaon's technology can categorize any keyword into a highly detailed category for further use.	Significant	less CO2 emissions + less waiting time for customers	Economic and Societal	TRL 6/7	prototype	Fleetplanning & Logistics	Architectural	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
EDI	Instaon - Ad generation	technology + service	Thanks to the keyword categorization & text generation technology, Instaon can automatically create relevant ad creatives	Significant	less waiting time for customer + less idle time for employees	Societal	TRL 6/7	prototype	Fleetplanning & Logistics	Architectural	new in 2020
EDI	Open source energy forecasting platform	Methods	Will open up the energy forecasting industry to be much more transparent	Significant	Providing a completely unique service on an existing market. We empower the user to achieve their goals to because an AI company instead of providing blackbox that competes with those interests.	Economic	7	Prototype	Energy	Disruptive	expected in +2021
EDI	Automated hyperparameter tuning using Python and DASK	Product	Allows for new business models based on data-as-a-service and compute-as-a-service in existing market	Medium			7	Prototype	Energy	Incremental	expected in +2021
EDI	Automatic rolling backtesting of energy forecasts	Product	Allows to backtest in a rolling horizon fashion to give an estimate of energy forecast performance	Medium			8	Prototype	Energy	Incremental	expected in +2021
EDI	Forecast deployment and service monitoring	Product	Deployment and monitoring in energy forecasting platform	Medium			8	Prototype	Energy	Incremental	expected in +2021
EDI	Followup of forecasting quality	Product	Transparent quality followup and evaluation	Medium			8	Prototype	Energy	Incremental	expected in +2021
EDI	Rebase Energy (Greenlyics) - Action recognition for large scale media content	Technology	By using action recognition AI, we are able to understand actions in the videos at scale	Significant	Disruptive impact	Economic and Societal		We have trained our action recognition AI on large scale open-source dataset and got accurate results	Media and Entertainment	Disruptive	new in 2020

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
EDI	Rebase Energy (Greenlytics) - Measuring gender diversity (Total speaking time ratio of female/male) by using AI	Technology, system and methods	By combining different type of AI such as gender recognition, action recognition, object tracking and object recognition, we are able to measure speaking time of each individual and the ratio of total speaking time of females/males	Significant	Disruptive impact on media industries. They are not able to measure gender diversity accurately and the process is manual (people are counting speaking times on a couple of videos instead of accurately measuring the real ratio), expensive and not scalable. However our solution can handle thousands of hours of videos cost-efficiently.	Economic and Societal		We have demonstrated our capabilities. We have prepared a simple API for customers to upload video and get the gender speaking time and ratio immediately.	Media and Entertainment	Disruptive	new in 2020
EDI	Vodéna - Foxtail	Technology	Automated Machine learning system for production, management, and deployment of predictive models on industrial scale. Foxtail enables companies to automatically create and maintain a large number of machine learning models in a reasonable timeframe with a minimal engagement of the end user.	Medium		Economic	7	Minimal Viable Product is developed	IT	Incremental	new in 2020
EDI	WearHealth	Technology, service	We can offer the new service that will help doctors to identify workers at risk of not improving their CVD risk factors	Significant	We will be able to expand to other market sectors. / A better and more personalized treatment can be offered to CVD risk workers.	Economic and Societal	7	MVP ready	Health	Architectural	expected in +2021
EDI	HOPU - Best price calculation	Service	Based on the water consumption data from previous years of a supplier company, and its billing model, this service can calculate the best price for each type of billing and each calibre diameter. These prices will equalize the company's benefits and expenses since, due to its characteristics, it should not obtain services.	Significant	This service aims to make water fairer's price in the different consumption ranges, rewarding responsible water consumption and thus contributing to global warming and urban heat islands.	Economic and Societal	7	A prototype has been created that works with real data supplied by the data provider, yielding valid results.	Water Supply	Disruptive	expected in +2021

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
EDI	starkdata - Medication Consumption Forecasts	Algorithm	A machine learning models that predicts sales by product and region and correlates it with current socks in pharmacies and predicts shortages of medications	Medium	By using machine learning models to forecast sales consumptions, it is possible to correlate those outputs with information from the pharmacies and detect possible medication shortages	Economic and Societal		models tested with a 40Gb dataset having an accuracy of ~95%	Pharma Industry	Disruptive	new in 2020
EDI	starkdata - Mercury SaaS platform	Product	A SaaS platform with insights, forecasts, and data analysis of the performance of pharma companies	Medium	With medication forecasts, pharma companies can manage their stock inventory, and medication production with with higher accuracy.	Economic	TRL 5	MVP developed	Pharma Industry	Disruptive	new in 2020
<b>Rebase - Energy (EDI)</b>											
Rebase - Energy	Open source energy forecasting platform	Methods	Will open upp the energy forecasting industry to be much more transparent	Significant	Providing a completely unique service on an existing market. We empower the user to achieve their goals to becuase an AI company instead of providing blackbox that competes with those interests.	Economic	7	Prototype	Energy	Disruptive	expected in +2021
Rebase - Energy	Automated hyperparameter tuning using Python and DASK	Product	Allows for new business models based on data-as-a-service and compute-as-a-service in existing market	Medium			7	Prototype	Energy	Incremental	expected in +2021
Rebase - Energy	Automatic rolling backtesting of energy forecasts	Product	Allows to backtest in a rolling horison fashion to give an estimate of energy forecast performance	Medium			8	Prototype	Energy	Incremental	expected in +2021
Rebase - Energy	Forecast deployment and service monitoring	Product	Deployment and monitoring in energy forecasting platform	Medium			8	Prototype	Energy	Incremental	expected in +2021
Rebase - Energy	Followup of forecasting quality	Product	Transparent quality followup and evaluation	Medium			8	Prototype	Energy	Incremental	expected in +2021
<b>Verchable (EDI)</b>											

Project	Innovation name	Type of innovation	Brief description	Impact	Impact description	Impact type	TRL	Maturity level	Sectors	Type of innovation	Timing
Verchable	Action recognition for large scale media content	Technology	By using action recognition AI, we are able to understand actions in the videos at scale	Significant	Disruptive impact	Economic and Societal		We have trained our action recognition AI on large scale open-source dataset and got accurate results	Media and Entertainment	Disruptive	new in 2020
Verchable	Measuring gender diversity (Total speaking time ratio of female/male) by using AI	Technology, system and methods	By combining different type of AI such as gender recognition, action recognition, object tracking and object recognition, we are able to measure speaking time of each individual and the ratio of total speaking time of females/males	Significant	Disruptive impact on media industries. They are not able to measure gender diversity accurately and the process is manual (people are counting speaking times on a couple of videos instead of accurately measuring the real ratio), expensive and not scalable. However our solution can handle thousands of hours of videos cost-efficiently.	Economic and Societal		We have demonstrated our capabilities. We have prepared a simple API for customers to upload video and get the gender speaking time and ratio immediately.	Media and Entertainment	Disruptive	new in 2020

Table 15 Answers from project to the "Innovations to Market Questionnaire"

## 5.2 New economic viable services of high societal value (KPI II.7)

Projects were asked to quantify, list, describe new economically viable services of high societal value developed or resulting from the project during 2019-2020 and also to forecast (2021-2024). They were also asked to describe why services qualified.

**146 new economically viable services of high societal value** were developed during 2019-2020 (67 in 2019 and 79 in 2020) as result of the projects who replied to the questionnaire (40 projects). Considering those answers, during the period 2019-2020, 32,5% (13 projects) contributed to this KPI for 2019 and 37,5% (15 project) for 2020. **Table 16** provides details per project.

Project	#	Description of the services of high societal value (2019)	Description of the services of high societal value (2020)
BigMedilytics	8	<ul style="list-style-type: none"> <li>• Solution for tracking efficiency of emergency departments in hospitals</li> <li>• Solutions for treating prostate cancer patients more effectively</li> <li>• Predicting inpatient treatments of patient with known heart failure with health insurance data</li> <li>• Search engines enabling to exploit healthcare data to improve individual diagnostic quality and corresponding care</li> <li>• A knowledge graph that can be explored and traversed</li> </ul>	More advanced versions of those from 2019
AEGIS	1	. The main technological result of the AEGIS project is an online platform providing data analytics as a service. It reduces the barrier for organisations to start using big data analytics, but requires from them to have data analysts – the target users of the AEGIS platform.	N/A
DMS	1	Acceleration programme focused on trainings and support services in 5 categories. Within the programme, more than 50 services are provided to the startups. The programme as a whole could be considered a viable set of services of high societal value.	Acceleration programme focused on trainings and support services in 5 categories. Within the programme, more than 50 services are provided to the startups. The programme as a whole could be considered a viable set of services of high societal value.
EW-Shop	0	In the project several business services have been developed, but with limited societal value (except for job creation and improvement of data-driven technologies adopted in SMEs).	N/A
CLASS	0	N/A	The development and execution of big-data analytics in a distributed computing environment composed of edge and cloud resources requires a redesign of big data system stacks, ranging from the programming model to the data-analytics runtimes. This innovation addresses this need by providing a SW architecture capable of distributing big-data analytics along the compute continuum while providing real-time guarantees on the end-to-end response time of complex big-data analytics. This innovation was selected by the European Commission's (EC) Innovation Radar among the most highly innovative EU projects. CLASS' innovation was assessed as having a "High" level of Market Creation Potential. Only innovations that are showing multiple signals of market creation potential are assigned a value under this indicator system
INFORE	6	see Pilot descriptions	see Pilot descriptions
Infinitech	>7	N/A	Indicative List: BFM Services for SMEs Fraud detection in Blockchain Trusted Transactions in SEPA/EEA Personalized Usage Based Insurance (Vehicles, Healthcare)

BD4OPEM	0	0	The marketplace is under development
SmasHit	N/A	N/A	Services in automotive/insurance/smart cities areas are currently under specification, development of early prototype just started
Cross-CPP	6	N/A	Meteologix: Various Weather Services, including: The UltraHD "City-Scale" Weather Model The Smart Building Forecast The Virtual Radar The Live Storm-Cell and Moving Vehicle Intersection Siemens: Real Time Weather Warning & Navigation Service Volkswagen: E-Mobility Service More information: <a href="https://www.cross-cpp.eu/services">https://www.cross-cpp.eu/services</a>
Examode	5	N/A	All developed innovative solutions support healthcare practitioners to make more informed decision making based on large amount of data (judging from similarity to other cases in their clinical practice or the identified likelihood in scientific literature). A prioritization of cases is considered, so that the cases with higher severity and likelihood of specific diseases will be presented for confirmation first, and the cases with smaller likelihood – later, when there is enough time. The developed services for automatic transformation of textual unstructured clinical information in structured format support the work of health practitioners, saving time for manual coding of the information and allow secondary use of clinical data for training complex model for clinical research, that can enable improvement of healthcare.
Synergy	7	N/A	Those services marked as "Economic and Societal" on the "BDV PPP_Significant innovations to Market MR2019-2020_SYNERGY" spreadsheet. They qualify since they imply a reduction in energy costs for the customers and/or provide customer empowerment tools with regards to the use of the energy.
INFORE	0	N/A	N/A
Boost 4.0	32	Research findings have also been incorporated in so-called technology transfer projects that it's OWL offers to SMEs. So the benefits and learnings of BOOST 4.0 can be disseminated within our cluster community. FDT, FPT, DSS, Fusion Isostatic Press Data Analyses Glazing Data Analyses	Proposition of new approach for identifying process parameters to our ESI customers and predictive maintenance algorithm on Cognitive Value Chain Workload consolidation methods and framework for edge computing devices. Methods for deployment and orchestration of virtualized computation modules Edge computing platform (NERVE)
DataPorts	0	0	0
ExtremeEarth	0	N/A	N/A
DeepHealth	0	N/A	N/A
Safe-Deed	1	0	The Data Valuation Component (DVC) may end up being instrumental in the establishment of trusted and transparent data markets. This could end up promoting transparent and fair data exchanges and data markets.
EDI	36	<a href="https://edincubator.eu/startups-2019/">https://edincubator.eu/startups-2019/</a>	<a href="https://edincubator.eu/startups-2020/">https://edincubator.eu/startups-2020/</a>
CloudButton	2	IBM-PyWren: a framework that allows to process large data sets at massive scale with PyWren over IBM Cloud Functions.	Lithops: a framework to massively scale the execution of Python code and its dependencies on serverless computing platforms and monitor the results.
TT	5	a) Improved Routing Service (Connected Trucks): Enhanced component to provide better and more detailed planning possibilities for truck routing and processes of Logistic Service Providers; Category: Software; Main features are described as follows: • Calculation of alternative routes according to travel	N/A

		times and POI analyses • Data analyses of different data sources and comparison to real LSP data • Provision of recommendations for planners and drivers b) Descriptive Analytics Service (Dynamic Supply Networks in e-commerce): it provides interaction with data through the selection of various criteria/dimensions, and as well it detailed analysis of data behaviour in 3PL partners through a set of appropriate visuals, and deduction of patterns and trends. c) Forecasting Analytics Service (Dynamic Supply Networks in e-commerce): A forecasting tool that will allow the potential user to periodically predict the behaviour of different groups of data depending on the selection of various criteria/dimensions, while presenting its forecasts through simple and convenient graphs d) Route Optimization Service (Dynamic Supply Networks in e-commerce): Inventory routing in an omnichannel environment to facilitate the postcode assignment to hubs. e) Customer Reviews Analysis Service (Dynamic Supply Networks in e-commerce): Extract online consumer insights with respect to logistics and delivery processes to perform a sentiment analysis over the reviews.	
SmartDataLake	6	see Pilot descriptions above	see Pilot descriptions above
Lynx	1	A portal for citizens that provides free access to legislation of European countries.	A portal for citizens that provides free access to legislation of European countries.
Data Pitch	8	3 startup projects that help tackle climate change, 2 startups that work in smart transport to help reduce air pollution and traffic collisions, 3 startups working to improve healthcare	N/A
QROWD	1	The combination of modal split and i-Log enables the implementation of agile travel surveys that also increase citizen engagement.	N/A
DataBench	2	Beta Toolbox and benchmarks	Release Toolbox and benchmarks

**Table 16:** List and description of services of high societal value per project (delivered 2019-2020)

Additionally, projects have forecasted (2021-2024) additional **64** new economically viable services of high societal value as result of their projects in the upcoming 4 years (2021-2024). 47,5% of the projects active in 2019-2020 are planning to contribute. **Table 17** provides additional details.

Project	#	Description of the services of high societal value (2021 - 2024)
BigMedilytics	10	Same as above
Everest	1	Accelerator of earth observations will enable faster predictions of catastrophic events. Such services can be offered for example to companies for weather-based services and local authorities to reduce their impact.
More	3	A data management service, a forecasting service and a diagnosis tool for renewable energy sources, that will increase the effectiveness of renewable energy sources and precipitate the green transition
DMS	1	Open Material, Some DMS services are available for the general public and thus for the startups and SMEs
CLASS	same as 2020	same as 2020
INFORE	5	see Pilot descriptions above
Infinitech	>14	Indicative List: Fraud detection in Blockchain Trusted Transactions in SEPA/EEA Personalized Usage Based Insurance (Vehicles, Healthcare) BFM Services for SMEs Personalized Insurance Products Recommendations
BD4OPEM	1	The marketplace will provide a number of solutions of high societal value
SmashHit	0	Services in automotive/insurance/smart cities areas are currently under specification, development of early prototype just started

Cross-CPP	6	Meteologix: Various Weather Services, including: The UltraHD "City-Scale" Weather Model The Smart Building Forecast The Virtual Radar The Live Storm-Cell and Moving Vehicle Intersection Siemens: Real Time Weather Warning & Navigation Service Volkswagen: E-Mobility Service More information: <a href="https://www.cross-cpp.eu/services">https://www.cross-cpp.eu/services</a>
Examode	2	Additional services for support decision making of healthcare practitioners.
Synergy	1	It is expected that at least one of the selected apps to be developed under the RFP Open Call, will be focused in providing high societal value.
DataPorts	1	Improvement of operational ports activities via data driven solutions.
ExtremeEarth	2	Polar TEP Service: We will develop a series of services integrated in the TEP that utilize EO data to produce ice maps for the Polar Regions. The Polar Regions are important globally. They play an important role in regulating and driving the global climate, but are experiencing significant change. An associated effect is growing global interest in the Polar Regions both politically and economically. These new economic opportunities drive increased attention and traffic, which in turn raise widespread concern about impact on this delicate and pristine environment. New information is vitally important in order to better predict and mitigate the resulting global economic and environmental consequences. It is also vitally important when developing regional and national policy, and monitoring their effectiveness. Sea ice information is used by operators involved in a broad range of activities in, or related to, ice-covered waters in the Polar Regions. Food Security TEP Services: We will develop a series of services integrated in the TEP that utilize EO data to address irrigation for farms in the Danube catchment area. All societal and economic activities depend on the availability of water, with the pressure on the water system only ever increasing with population growth and climate change. The Food Security Use Case with its focus on water availability thus has a wide range of societal and environmental implications. Water is one of our most precious commodities, and one that in many regions of the world is scarce and leads to many political conflicts, sometimes even to wars. Using the freshwater we have available in the most efficient, sustainable way and making sure everyone's needs are met will be one of the most pressing challenges of dealing with climate change in the next decades.
DeepHealth	5	On one hand, services provided around the DeepHealth toolkit, the major outcome of the project. Services around the OSS result will be deployed according to a sustainability plan that will increase the productivity and facilitate the daily work of data scientists working in the health domain. On the other hand, specific services provided by enhanced biomedical platforms and the specific applications developed in the project by industrial and research partners for the diagnosis, monitoring, and treatment of different pathologies, with a great impact to increase early diagnosis and improving treatments, extend the knowledge about diseases and pathologies and at the end, to save direct and indirect healthcare costs.
CloudButton	1	Full Python data analytics stack supported in the cloud enabling true big data analytics democratization. We expect to leverage the scalability of the cloud to support a full stack of Python toolkits (JupyterLab, Dask, numpy, Pandas...).
PimCity	3	The aim of Pimcity is to implement a PIMS development kit (PDK) to commoditize the complexity of creating PIMS. This lowers the barriers for companies and SME to enter the web data market. At the same time, the PDK democratize the development and adoption of PIMS. This is a ground-breaking change that will finally help the data economy accelerate and become transparent. Pimcity offers innovative solutions for hard technical problems like privacy-preserving data analytics, easy to use consent manager, instruments to increase awareness in users, tools to transparently trade data, methods to port data to different platforms, etc. Furthermore, Pimcity will deliver the fully fledged EasyPIMS, that has been bootstrapped thanks to data already available among consortium partners, involving all stakeholders in its design. It would stop the current arms race between users and services. From a legal and ethical standpoint, due to the absence of a harmonised and tailored legislative initiative addressing PIM-City legal and ethical challenges, the consortium has carried out its activity considering the normative production developed within the EU in the last years in such context. In particular, such an activity has taken into account regulatory context and the ethical debate and the most recent jurisprudence of the European Court of Justice. As a result, legislative gaps have been filled considering ethical, societal and economic implications the foreseen solutions might have on end-users and, in particular, consumers. Pimcity will run one of the largest demonstrators in PIMS, involving SME that are strongly committed and large companies that are eager to enter the data market, in a transparent way.
Reach	/	Difficult to estimate. 30 Startups will reach the experimentation oriented phase here, while the rest will only pitch and may not be successful afterwards
SmartDataLake	6	see Pilot descriptions above
Lynx	1	A portal for citizens that provides free access to legislation of European countries.

**Table 17::** List and description of services of high societal value per project (additional expected 2021-2024 by projects active in 2019-2020)

### 5.3 New Systems and Technologies (KPI II.17)

Projects have reported **337 new systems and technologies** developed during 2019-2020 (141 in 2019 and 196 in 2020). Many of them are reported already as part of the KPI “Significant innovations to Market”. Others are additional technological components (not directly related to an exploitable value in the market).

**Table 18** provides details per project I

n

2019. **Table 19** provides information per project in 2020

Project	#	List and description systems and technologies developed 2019
<b>BigMedilytics</b>	9	Multiparty Computation (MPC): a method/tool to analyse sensitive data from multiple parties (who do not wish to share data with each other) without revealing the data itself, but only revealing the outcome of the (data) analysis, i.e. the coefficients of the learned model. • A LASSO based regression that allows for prediction of comorbidities of hospital admissions out of > 4.000 features in a Health Insurance database. • GDM system • GDM mobile App • GDM medical portal • GDM analytics • Questionnaire • A lung-cancer pilot is under development what intends to contribute in the reduction of costs in hospitals, but also in a reduction of toxicities • Location based analytics for hospitals - asset management and workflow optimization
<b>AEGIS</b>	1	The AEGIS platform provides a multi-tenant data management and processing services for big data. The multi-tenancy behaviour allows different users and services to securely and privately access and process their data. The AEGIS platform enables users to share their data with other users on the platform and allow access for specific services. Also, users can use different data processing services that are supported by the platform to process and visualise their data. AEGIS is built on top of Hopworks and Hops. It provides an integrated support for different data parallel processing services such as Spark, Flink, and MapReduce, as well as a scalable messaging bus with Kafka, and interactive notebooks with Zeppelin and Jupyter. Under the hood, the data is mainly stored in the AEGIS data store; however, the AEGIS data store APIs are kept hidden from users. Instead, the AEGIS platform provides a Project/Dataset service to allow users to upload/download, explore, and do analysis on their data in a secure way without interacting with the AEGIS data store directly.
<b>DataMarketServices</b>	5	Mimirium Surveys - a SaaS platform for internal surveys for companies that need more engaging way for interaction combined with 100% anonymity of the votes and rewarding points system. One more hardware product to the lineup. So 2 products became 3 since DMS. Based on feedback from existing partners, we have refined our MVP Shifted towards Analytics Urban design tool based on data services
<b>EW-Shop</b>	9	- ABSTAT: knowledge graph summaries with ABstraction and STATistics - ASIA with Data Reconciliation and Extension Services - Big Data workflow tool based on container orchestration with DataGraft and Grafterizer - Weather data API - Product Data Management Tools - Event and Weather Analytics Toolset - Keyword Clustering Tool - Event Data Management for Data Analytics - Campaign BOOSTER - EW-Shopp DataFlow Methodology - PERFORMANCE INSIGHTS: E-commerce Insights tool for shoppers, sellers and vendors - Contact Center Workforce Optimization Manager - Weather and Event-aware BI Sales Strategy Advisor - Event and Weather Scout Dashboard
<b>CLASS</b>	1	CLAS SW architecture ecosystem for distributing big-data workloads along the compute continuum while providing real-time guarantees
<b>INFORE</b>	3	- Forecasting price swings - Systemic risk analysis - Decision-Support for Investment Opportunities
<b>MosaiCrown</b>	8	See Table 11 on Innovations to Market Questionnaire
<b>Cross-CPP</b>	4	Cross CPP Big Data Marketplace Cross CPP Big Data Analytics Toolbox CIDM - Common Industrial Data Model Cross Sectorial Services
<b>Boost 4.0</b>	26	IDS Connector Blockchain Applied in the supply chain of a manufacturing industry Generic Data Model for injection moulding machines Boost 4.0 Advanced Mapping Tool Quality Dashboard Spare Parts Augmented Forecasting
<b>ExtremeEarth</b>	N/A	See Table 11 on Innovations to Market Questionnaire
<b>Safe-Deed</b>	2	See Table 11 on Innovations to Market Questionnaire
<b>IoTwins</b>	1	See Table 11 on Innovations to Market Questionnaire
<b>EDI</b>	18	18 new data-driven solutions More details about each of the solutions can be encountered at <a href="https://edincubator.eu/startups-2019/">https://edincubator.eu/startups-2019/</a>

Project	#	List and description systems and technologies developed 2019
<b>CloudButton</b>	1	Lithops: a framework to massively scale the execution of Python code and its dependencies on serverless computing platforms and monitor the results.
<b>Fandango</b>	2	Multilingual text analytics - Service of multilingual text analytics to be used by journalists, media professionals and fact checkers to detect misleading messages or disinformation claims Spatio-temporal analytics - Analyzing news posts and finding duplicates or near duplicate posts in the past or referring to other geographic/physical locations or in other contexts.
<b>TT</b>	11	All these systems and technologies are embedded in "Innovations to Market": - Process - Port and Terminal Productivity Cockpit - Crane and Spreader Sensorization Infrastructure - Service - Emission Reduction System - Simulation model - Traffic micro-simulation models in city centres - Model - Predictive High-Speed Network Maintenance - Predictive & automated diagnostic capabilities in the mainline rail industry - Dashboard to improve probability for trucks to arrive at planned time - Dashboard to predict car breakdown - Dashboard to predict and prevent failures in railways - Dashboard to measure cockpit efficiency and analyse truck turnaround time - MAID - Animal intrusion detection system
<b>Musketeer</b>	3	pycloudmessenger, MMLL, client connector.
<b>SmartDataLake</b>	1	SmartDataLake Toolkit
<b>SmartDataLake</b>	2	Application for • Financial market analysis and trading • Market analysis and price hedging for producers of agricultural goods
<b>Lynx</b>	18	The Lynx Services Platform (LySP) that offers a cloud-based platform of orchestrated (mainly NLP) services on top of a Legal Knowledge Graph, see: <a href="https://lynx-project.eu/doc/api/">https://lynx-project.eu/doc/api/</a>
<b>Data Pitch</b>	N/A	See Table 11 on Innovations to Market Questionnaire
<b>QROWD</b>	8	QROWD Platform

**Table 18:** List and description of systems and technologies delivered in 2019 per project

Project	#	List and description systems and technologies developed 2020
<b>BigMedilytics</b>	N/A	See Table 14 on year 2019
<b>DataMarketServices</b>	10	Mobile & Cloud Solution Application of technology to sports new product to predict the consumption of energy for clients digital products Data-as-a-service energy data analysis platform mobile app Cloud service Market research, and customer data development Not applied for IPS, web based platform Cloud service, no IPR requests raised
<b>CLASS</b>	1	CLAS SW architecture ecosystem for distributing big-data workloads along the compute continuum while providing real-time guarantees. (the innovation is the same as 2019 but evolved)
<b>INFORE</b>	3	- Forecasting price swings - Systemic risk analysis - Decision-Support for Investment Opportunities
<b>Infinitech</b>	>20	Overall 20 technologies and proof of concept concerning BigData, AI and Blockchain Technologies in digital finance, covering wealth management, customer centric analytics, fraud detection, AML, cybersecurity, usage based insurance, personalized insurance recommendations, agro-insurance
<b>Kraken</b>	8	Personal local data wallet Privacy protected cloud data wallet SDK for VC and consented flow SSI authentication components Integrated marketplace and underlying infrastructure Enhanced MHMD platform Cryptographically secure privacy-preserving data analytics Data analytics on authentic data
<b>MosaiCrown</b>	8	Descriptions are in the excel file.
<b>SmasHit</b>	N/A	System is under specification, development of early prototype just started
<b>Cross-CPP</b>	4	See Table 11 on Innovations to Market Questionnaire
<b>Examode</b>	3	VIRTUM HP - Cloud-based platform for whole slide image management and visual analysis HistoGrapher ( Innovation Market - <a href="https://marketplace.big-data-value.eu/node/411">https://marketplace.big-data-value.eu/node/411</a> ) Demonstrator for semantic knowledge management of diagnosis related medical data using advanced text analytical technologies and knowledge graphs Platform for structuring diagnosis information based on ontologies and FHIR standard
<b>INFORE</b>	2	INFORE Streaming Extension INFORE Synopses Data Engine
<b>Boost 4.0</b>	20	From Sheep to Shop with a single source of truth IDS connector for Obelisk Cognitive AI & Robot Control FIWARE EIDS CIM REST Connector Teralab AI marketplace Casting Process Data Analysis Models Edge Computing Platform Nerve

<b>DataPorts</b>	20	See Table 11 on Innovations to Market Questionnaire
<b>EUHubs4Data</b>	1	EUH4D federated catalogue of data driven services ( <a href="https://euhubs4data.eu/services/">https://euhubs4data.eu/services/</a> ) and datasets ( <a href="https://euhubs4data.eu/datasets/">https://euhubs4data.eu/datasets/</a> ).
<b>ExtremeEarth</b>	N/A	See Table 11 on Innovations to Market Questionnaire
<b>DeepHealth</b>	18	This includes: 1) the DeepHealth toolkit and its four components (EDDL, ECVL, back-end and front-end); 2) 7 enhanced biomedical and AI platforms that integrate the toolkit and are being adapted to different use-cases developed by everis, Philips, Thales, CEA, Wings, Unit and CRS4; 3) 6 DeepHealth HPC infrastructure components: enhanced COMPS distributed programming framework, the StreamFlow framework, customized PCIe card, FPGA kernels for the libraries, runtime management software and an hybrid cloud solution.
<b>Safe-Deed</b>	6	Private Selective Aggregation Library - Private Set Intersection Java Library - Infineon-SafeDEED Demonstrator - De-Anonymisation Risk Analysis toolkit - Data Valuation Component (DVC) - Data-Driven Business Model Innovation Toolset. More info in Table 11.
<b>IoTwins</b>	3	• ML models for predictive maintenance • ML models for anomaly detection (semi-supervised and supervised) • Application creation for Smart Manufacturing ML for Smart Manufacturing, Physical modeling of machining See Table 11 Innovations to Market” questionnaire for more information
<b>EDI</b>	18	18 new data-driven solutions More details about each of the solutions can be encountered at <a href="https://edincubator.eu/startups-2020/">https://edincubator.eu/startups-2020/</a>
<b>CloudButton</b>	2	Faasm: a novel runtime that can execute application components that are expressed using the universal WebAssembly representation in a serverless cloud context. Infinispan AnchoredKeys: an alternative implementation of the scaling algorithm of the Infinispan data grid which avoids costly state transfer when adding new nodes so that throughput is maintained.
<b>PimCity</b>	15	description embedded in "innovation to Market" file about the followings technologies under development: Data Portability Control (DPC) Data Valuation Tools from Market Place (DVTMP) Personal Data Safe (PDS) Privacy Preserving Analytics (PPA) Data Provenance (DP) Privacy Metrics (PM) Personal Consent Manager (PCM) Data Trading Engine (DTE) Data Aggregation (DA) Data Knowledge Extraction (DKE) Personal Data Avatar Transparency Tags Dashboard Marketplace EasyPIMS
<b>Fandango</b>	2	Copy-move detection on audio-visual content - The innovatio is focused in the detection of the intentional manipulation of images and videos to add or remove or in any way modify the visual content. Source credibility - Profiling of the sources of the news posts and apply graph analytics to detect paths and nodes that tend to produce such fake news and defuse them widely on the public web.
<b>OpertusMundi</b>	1	See Table 11 on Innovations to Market Questionnaire
<b>SmartDataLake</b>	1	SmartDataLake Toolkit
<b>SmartDataLake</b>	2	Application for • Financial market analysis and trading • Market analysis and price hedging for producers of agricultural goods
<b>Lynx</b>	18	The Lynx Services Platform (LySP) that offers a cloud-based platform of orchestrated (mainly NLP) services on top of a Legal Knowledge Graph, see: <a href="https://lynx-project.eu/doc/api/">https://lynx-project.eu/doc/api/</a>
<b>BodyPass</b>	2	Blockchain network for exchanging 3D personal data extracted from medical images and body scanners Full-body mapping and automatic registration of medical (CT) images for geometrical and tissue queries

**Table 19:** List and description of systems and technologies delivered in 2020 per project

## 6 SUPPORT MAJOR SECTORS AND MAJOR DOMAINS BY BIG DATA TECHNOLOGIES AND APPLICATIONS (II.13)

This KPI measures the number of sectors and major domains supported by Big Data technology and applications developed in cPPP projects. As described in KPI I.4 the BDV cPPP projects address a great variety

of sectors<sup>19</sup>, **10% of the projects are fully cross-sectorial** (their outcomes can be used in any sector or application domain) and **77,5% of the projects are working in more than 1 sector or application domain** (this explain why the total is superior to 100% in Table 20). In particular the BDV cPPP projects address a great variety of sectors<sup>20</sup> as shown in Table 20.

Sector/application domain	Projects addressing this sector/application domain (% over active projects in 2019-2020)	Innovations to market (delivered in 2019-20, % over the total)
Cross-sectorial	10%	23%
Public services	25%	4%
Smart Cities	32,5%	4%
Transport and logistics	47,5%	12%
Mobility	40%	12%
Retail	30%	-
Business Services	35%	-
Environment	27,5%	-
Health and healthcare	47,5%	10%
Manufacturing	35%	9%
Media	30%	2%
Finance	42,5%	27%
Telecommunication	20%	-
Energy	37,5%	3%
Bio-Economy	15%	1%
Water and natural resources	17,5%	
Earth Observation	27,5%	
Others	20%	

Table 20: Support major sectors and domains

The BDV cPPP lighthouse projects<sup>21</sup> active in 2019-2020 focused on the Bio economy (Agriculture, Fisheries, and Forestry) (**DataBio project**), Transport, mobility and logistics (**Transforming Transport project**), Health and Healthcare (**BigMedilytics project**) and Manufacturing (**BOOST4.0**), with a total of **4 major sectors**

19 Grouped with a good level of alignment to the NACE registry These categories are part of the information in the BDV CPPP Marketplace that will be used for promoting all exploitable solutions coming out of the cPPP Projects (if needed new categories can be added)

20 Grouped with a good level of alignment to the NACE registry These categories are part of the information in the BDV CPPP Marketplace that will be used for promoting all exploitable solutions coming out of the cPPP Projects (if needed new categories can be added)

21 Large-scale data-driven innovation and demonstration projects that aim at creating superior visibility, awareness and impact in specific relevant economic sectors

**supported** by Lighthouse projects and therefore widely supported with multiple use cases, scenarios and solutions.

Considering the whole project portfolio, the number of sectors supported is larger than **17**, with a solid distribution of use cases, experiments, solutions, and outreach activities among different sectors.

The BDV CPPP Marketplace as a supporting tool will help to further promote the marketable and exploitable results of the cPPP Projects to create a higher impact of the cPPP and contribute to its sustainability.

In addition to the projects, BDVA's Task Force 7<sup>22</sup> (application) bring together European stakeholders to identify and act on the needs (technology, skills, etc.) of different industrial sectors as well as the areas of interest applicable to different industrial sectors such as language technologies, HPC, etc. During 2019-2020 twelve different application domains were represented in BDVA, organised in different subgroups.

In particular:

- TF7.SG2 Telecom (Lead: Robert Seidl, Nokia)
- TF7.SG3 Healthcare (Lead: Tiblets Demewez, Philips)
- TF7.SG4 Media (Lead: Luk Overmeire, VRT) (acting)
- TF7.SG5 Earth Observation & Geospatial (Lead: Florin Serban, Terrasigna)
- TF7.SG6 Smart Manufacturing Industry (Lead: Davide Dalle Carbonare, Engineering. Co-Lead: Sergio Gusmeroli, Polimi)
- TF7.SG7 Mobility and logistics (Lead: Vivian Kiousi, Intrasoft Intl.)
- TF7.SG8 Smart Governance and Smart Cities (Lead: Roberto Di Bernardo, Engineering; Co-lead: Vega Rodríguez, ITA Innova)
- TF7.SG9 Agriculture (Lead: Nuria de Lama Sánchez, ATOS)
- TF7.SG10 Finance (Lead: Vittorio Monferrino, GFT Italy)
- TF7.SG11 Automotive (Leads: Ohiana Otaegui, Vicomtech and Bernhard Peischl, AVL)
- TF7.SG12 Energy (Lead: Philippe Calvez, ENGIE)
- TF7.SG13 Security (Leads: Nizar Touleimat, CEA and Vito Morreale, Engineering)

## 7 EXPERIMENTATION (II.11, II.12, II.14)

### 7.1 II.11: Large Scale experiments conducted in cPPP projects and i-Spaces involving closed data

Projects reported **1065 large scale experiments** (aggregated number of experiments conducted by all projects in 2019-2020), of which the majority involve closed (private) data (with the project Exa-mode reporting more than 1000 experiments).

Projects also provided criteria to for experiments to qualify as “large-scale” with following criteria being mentioned:

- Amount of data
- Speed of data
- Quality of data
- TRL
- Number of users
- Large number of data owners

---

<sup>22</sup> <http://bdva.eu/task-force-7>

- Large geographical coverage
- Use large data sets and involvement of real-life operations and processes
- A combination of data volume, data complexity and velocity.
- Complex datasets integrating data in some case from different countries

**Table 21** provides details in the information provided by each of the projects contributing to this KPI.

Project	Number of large scale experiments (2019)	Number of large scale experiments (2020)	Experiments involving closed data	Description of criteria to consider large scale
<b>BigMedilytics</b>	14	14	12	Amount of data, Speed of data, Quality of data, TRL
<b>INFORE</b>	2	2	1	Cross-border, 3 or more partners involved, 200 or more users
<b>Infinittech</b>	N/A	4	N/A	- Cross-Border - TRL>=6 - Value Chain / Finance Supply Chain Participants >=3
<b>Cross-CPP</b>	8	12	8	large scale experiments involving large amounts of data, coming from various Cyber physical products, from many different geographical regions and involving many different data owners.
<b>Examode</b>	1000+	1000+	the majority of experiments are performed on closed data	The experiments are "large-scale", because: * This is a study on clinical data for several thousands of people * The experiments are performed cross-border, because the data are collected from patients in Italy and Netherlands, but processed by project partners from different countries. * Some of the tools used for experiments have TRL 6-7
<b>Boost 4.0</b>	14	14	14	These are considered large-scale because of the number of companies involved and the investment made to conduct them
<b>DeepHealth</b>	1	5	1	We consider an experiment as large-scale based on the computing requirements to train Deep Learning models. Despite the datasets used so far are not larger than 1TB, the iterative process of training during hundreds of epochs and performing data augmentations on-the-fly convert the experiments in high-demanding of computing power and memory. Furthermore, these experiments are only possible to be run on computers with GPUs, otherwise the training process is not feasible due to the required time.
<b>Safe-Deed</b>	3	6	3	Privat Data: We consider a trial as large scale if we use large data sets (e.g. complete CRM databases) and involvement of real-life operations and processes.  Industrial Data: The experiment qualifies as "large-scale" because the QSD is based on large order data from a high number of companies.
<b>IoTwins</b>	1	1	1	The experiment involved 7.000 computing nodes
<b>Fandango</b>	4	4	4	-
<b>TT</b>	13	-	13	A combination of data volume, data complexity and velocity. Number of users is also a criterium for large scale experiment in some cases.
<b>SmartDataLake</b>	3	3	1	Cross-border, 3 or more partners involved, 200 or more users

<b>QROWD</b>	1	0	1	Geographical impact (Whole Trento region) and number of users (250)
--------------	---	---	---	---

**Table 21:** Number of large-scale experiments

BDVA i-Spaces reported **24 large scale experiments, 14 of them involving closed data**. In the same way than in the projects main criteria to qualify as large scale relate to number of companies, size, and impact. 3 BDVA i-Spaces provided input to this KPI and in particular (**Table 22**):

BDVA i-Space	Number of large scale experiments (2019)	Number of large scale experiments (2020)	Experiments involving private data (2019 + 2020)	Description of criteria to consider large scale
<b>Know-Center GmbH</b>	4	8	6	minimum of 3 companies involved
<b>ITI - instituto tecnológico de informática</b>	2	4	6	Number of companies involved, more than 3. Intensive use of GPUs.
<b>Universidad Politécnica de Madrid</b>	3	3	2	The main experiment involving closed data, run for CRTM (Consortio Regional de Transportes de Madrid) where transport card validations data from the last 5 years on a segment of the population (senior citizens) up to March 2020 (before the pandemic) was used. This results in approximately 25 millions of validations.

**Table 22:** Number of large-scale experiments in BDVA i-Spaces

## 7.2 II.12: Uptake of BDV use cases and experiments

This KPI intends to assess the year over year increase of the number of Big Data Value use cases and experiments supported in cPPP projects and BDVA i-Spaces.

Projects reported **1251 use cases or/and experiments** conducted during 2019-2020 with a contribution from 40 different projects (many experiments performed along the 2 years). **Table 23** provide details per project of the experiments and use cases developed during the two years. It is important to notice that the large-scale experiments reported in II.11 are normally a subset of this list (as normal rule).

BDVA i-Spaces have reported additional **194 experiments** with 3 i-Spaces contributing to this KPI. Details can be found in Table 24. In total the cPPP has developed **1445 experiments**.

project	Number of data experiments/use cases 2019 (number)	Number of data experiments/use cases 2020 (number)	Please provide some qualitative data to support your answer (i.e list, brief description,..)
<b>BigMedilytics</b>	15	15	<ul style="list-style-type: none"> <li>• Involved in 5 pilots in BigMedilytics: Population Health; Prostate cancer; Stroke workflows; Sepsis workflows; Asset management</li> <li>• Analysis of EHRs for detecting comorbidities clusters: this experiment will find using ML techniques how a combination of several relevant diseases could led to a higher risk of hospitalization and mortality.</li> <li>• initial pilot experiments to prepare for evaluation campaign assessing value of search in radiologists workflow. We deployed the prototype of the search engine at the radiology department of MUW/AKH and conducted initial trials to finalize deployment and understand the use case better.</li> <li>• The first use case aims to relate the patient characteristics with the reason and frequency of teleconsultation, which is a service established by the HUMPH hospital in Madrid with a view of reducing the unscheduled hospital visits.</li> <li>• After developing a framework defining the relevant clinical items, we have deployed an infrastructure to run the image analysis</li> <li>• Prepared data usage access and data selection with Curie and IBM for breast cancer pilot.</li> </ul>
<b>Daphne</b>	0	5	Prototypical MLIR-based compiler and runtime, initial experiments with fused operator pipelines.
<b>AEGIS</b>	5	0	<p>The work was performed in 3 demonstrators:</p> <ol style="list-style-type: none"> <li>1) The Automotive Demonstrator: safe driving indicator and regional driving safety risk estimator scenarios</li> <li>2) The Smart Home and Assisted Living Demonstrator:               <ol style="list-style-type: none"> <li>a) provision of anonymised group-based or even personalised notification and recommendation services to the at-risk individuals, as well as to the at-risk individuals' informal carers;</li> <li>b) provision of an increased indoor comfort and welfare service to the at-risk individuals, empowered by smart home automation functionalities.</li> </ol> </li> <li>3) The Insurance Demonstrator : Personalised early warning services for asset protection</li> </ol>
<b>EW-Shop</b>	5	N/A	<p>The five business use cases in the project all include many data experiments that yield to the final development of business data-driven services. In addition all technologies have been developed with a substantial experimental basis. See success stories for more relevant experiments. The complete list of experiment-intensive developments is:</p> <ul style="list-style-type: none"> <li>- ABSTAT: knowledge graph summaries with ABstraction and STATistics</li> <li>- Big Data workflow tool based on container orchestration with DataGraft and Grafterizer</li> <li>- Product Data Management Tools</li> <li>- Event and Weather Analytics Toolset</li> <li>- Keyword Clustering Tool</li> <li>- Event Data Management for Data Analytics</li> <li>- Campaign BOOSTER</li> </ul>

			<ul style="list-style-type: none"> <li>- PERFORMANCE INSIGHTS: E-commerce Insights tool for shoppers, sellers and vendors</li> <li>- Contact Center Workforce Optimization Manager</li> <li>- Weather and Event-aware BI Sales Strategy Advisor</li> <li>- Event and Weather Scout Dashboard</li> </ul>
<b>INFORE</b>	2	2	see Pilot descriptions
<b>Infinitech</b>	N/A	8	<p>P01 Invoices Processing Platform for a more Sustainable Banking Industry</p> <p>P02 Real-time risk assessment in Investment Banking</p> <p>P03 Collaborative Customer-centric Data Analytics for Financial Services</p> <p>P04 Personalized Portfolio Management (“Why Private Banking cannot be for everyone?”)</p> <p>P05b Business Financial Management (BFM) tools delivering a Smart Business Advise</p> <p>P06 Personalized Closed-Loop Investment Portfolio Management for Retail Customers</p> <p>P07 Avoiding Financial Crime</p> <p>P08 Platform for Anti Money Laundering Supervision (PAMLS)</p> <p>P09 Analyzing Blockchain Transaction Graphs for Fraudulent Activities</p> <p>P10 Real-time cybersecurity analytics on Financial Transactions’ BigData</p> <p>P11 Personalized insurance products based on IoT connected vehicles</p> <p>P12 Real World Data for Novel Health-Insurance products</p> <p>P13 Alternative/automated insurance risk selection - product recommendation for SME</p> <p>P14 Big Data and IoT for the Agricultural Insurance Industry</p> <p>P15 Open Inter-banking Pilot</p>
<b>BD4OPEM</b>	0	0	In 2020, they have been defined but not conducted yet.
<b>Kraken</b>	0	4	Sell and buy data through the KRAKEN marketplace. Data monetization. Data Analytics as a Service,
<b>MosaiCrown</b>	3	3	The project includes three Use Cases for deploying and validating the developed techniques and tools. Each Use Case represents a real-world problem where one of the industrial partners (EISI, SAP, MasterCard) participating in the project is involved, and for which there are open privacy/confidentiality problems (considered by MOSAICrOWN) that cannot be addressed with current technology. The Use Cases also cover diverse application domains and provide scenarios able to promote the MOSAICrOWN’s innovations. Each Use Case focuses on deployment of specific MOSAICrOWN solutions, as well as exploitation in specific scenarios.
<b>Cross-CPP</b>	8	12	as above mentioned, experiments (CPP Marketplace, Analytics toolbox) with data coming involving large amounts of data, coming from various Cyber physical products, from many different geographical regions and involving many different data owners were performed. Those experiments were used to validate the robustness and scalability of the Cross-CPP solution.
<b>Examode</b>	1000+/ 1 use case	1000+/ 4 use cases	<ul style="list-style-type: none"> <li>* Experiments for WSI annotations</li> <li>* Experiment for Processing Megapixel Images with Deep Attention-Sampling Models</li> <li>* Experiment for compound figures separation and association to related text</li> <li>* Experiments with Web-oriented search system that allows to search, access, and explore nanopublications on the Web.</li> <li>* Experiment with Multi-resolution convolutional neural networks for semantic segmentation in histopathology whole-slide images.</li> <li>* Experiments with a deep learning system for weakly supervised object detection in digital pathology whole slide images</li> <li>* Experiments for large field-of-view semantic segmentation</li> </ul>

			<ul style="list-style-type: none"> <li>* Experiment with the invertible mapping of the color information in histopathology Whole Slide Image patches</li> <li>* Experiments for automatic extraction of concepts from clinical reports</li> <li>* Experiments for heterogenous data integration and transformation in Knowledge Graphs</li> <li>* Experiments for data enrichment with external sources information</li> <li>* Experiments for KG Completion</li> </ul>
<b>INFORE</b>	3	3	<ul style="list-style-type: none"> <li>(1) Large-scale simulations involving drug administration over cancer cells and forecasting of promising simulations</li> <li>(2) Predicting Price Swings, Systemic Risk and Forecasting Investment Opportunities</li> <li>(3) Increase Maritime Situational Awareness by detecting and classifying possible threatening maritime activity</li> </ul>
<b>Boost 4.0</b>	14	14	<p>TRIAL 1 – NEMAK: Will work to develop a big data based support application called Assistant Systems.</p> <p>TRIAL 2 – FILL: Will work to improve their engineering process with machine data obtained during the operation phase into the design of new machines.</p> <p>TRIAL 3 – VOLKSWAGEN AUTOEUROPA: the main goal is to fully integrate the material flow, from the reception at the unloading dock to the point of fit.</p> <p>TRIAL 4 – GF: the ambition is to exploit the opportunities provided by BOOST 4.0 enablers for a zero-defect factory for milling spindles for GF machines.</p> <p>TRIAL 5 – FIAT – Connect automated guided vehicles (AGV) with the laser machine in order to further automatize the production process.</p> <p>TRIAL 6 – PHILIPS CONSUMER LINE – Improve the control system for the moulding process providing more data to the operator and using data analysis to predict and prevent the defective parts production.</p> <p>TRIAL 7 – GESTAMP – Improve the navigation information of the AGV to optimize the logistics in the warehouse</p> <p>TRIAL 8 – TRIMEK – Develop a new high resolution metrology and visualization system to achieve zero-defects manufacturing.</p> <p>TRIAL 9 – WHIRLPOOL – Update their forecasting tool for spare parts distribution and manufacturing incorporating big data analysis tools to process and predict the needs based on the data produce across all the product lifecycle.</p> <p>TRIAL 10 – BENTELER – Develop a machine health monitoring system that will allow to improve the predictive maintenance plan and prevent the defective production.</p>
<b>ExtremeEarth</b>	28	41	<ul style="list-style-type: none"> <li>1. Experiments to test and evaluate the new distributed versions of linked geospatial data tools.</li> <li>2. Experiments with EO data to extract information to assist in irrigation recommendation.</li> <li>3. Experiments with EO and non-EO data to create Deep Learning models for crop type detection.</li> <li>4. Experiments with EO and non-EO data to create Deep Learning models for ice type detection.</li> <li>5. Experiments with big data to test and evaluate the capabilities of the Hops platform.</li> <li>6. Experiments with big data to test and evaluate the new features of Hopsworks.</li> <li>7. Experiments with big data to test and evaluate the integration of Hopsworks with CREODIAS platform.</li> <li>8. Experiments with big data to test and evaluate the new services implemented on the Polar and Food Security TEPs.</li> </ul>
<b>DeepHealth</b>	1	5	The same commented in the previous question.
<b>Safe-Deed</b>	5	19	<p>2019:</p> <p>MPC: 2 experiment/use cases</p> <p>De-anonymisation risk analysis: 4 experiments/use cases</p> <p>Data valuation: 1 experiment/use case</p> <p>2020</p> <p>IFAG: 3</p>

			MPC: 2 experiment/use cases De-anonymisation risk analysis: 4 experiments/use cases Data valuation: 1 experiment/use case
<b>IoTwinS</b>	12	12	The use-cases considered in the IoTwinS project are the following: <ul style="list-style-type: none"> <li>• TB01 – Wind-farm predictive maintenance system</li> <li>• TB02 – Machine tool spindle predictive behaviour</li> <li>• TB03 – Crankshaft manufacturing system predictive maintenance test-bed</li> <li>• TB04 – Predictive maintenance and production optimization for closure manufacturing (beverage and spirits caps)</li> <li>• TB05 – NouCampNou facility management and maintenance</li> <li>• TB06 – IT facility predictive maintenance and operation optimization</li> <li>• TB07 – Facility power quality management through smart grids</li> <li>• TB08 – Pre-industrial prototype on smart manufacturing</li> <li>• TB09 – Reuse of the closure manufacturing use-case in different plants</li> <li>• TB10 – Deployment of Examon infrastructure in INFN and BSC facilities</li> <li>• TB11 – FCB use-case in smaller sport facilities</li> <li>• TB12 – Use-case on business models</li> </ul>
<b>EDI</b>	18	18	Full details about each of the experiments generated in 2019 ( <a href="https://edincubator.eu/startups-2019/">https://edincubator.eu/startups-2019/</a> ) and 2020 ( <a href="https://edincubator.eu/startups-2020/">https://edincubator.eu/startups-2020/</a> ) are given
<b>CloudButton</b>	2	2	Metabolomics experiment 1: General Annotation. This experiment is representative of a normal use-case on METASPACE, which makes it suitable for head-to-head comparisons. There is often limited time available on the higher-spec PC used for initial data capture as it is a shared resource, so usually the analysis will be performed from scientists' or students' lower-spec laptops.  Metabolomics experiment 2: Interactive reprocessing. This is representative of a new type of functionality that we currently don't support in METASPACE because it's uneconomical with the serverful approach. While looking for specific compounds, scientists tend to have relatively short lists of molecules of interest, and iteratively try different adducts or modifiers until they find the data they're interested in.  Metabolomics experiment 3: "Stress test", aims to ensure that the limits of are similar to METASPACE, this is one of the larger datasets that has been processed.  Geospatial experiment 3: Water Consumption. It consists in comparing water use estimates obtained from two different databases over extended regions of irrigated crop fields and then map differences in the water use footprint of irrigated arable lands in representative large areas of Peninsular Spain.
<b>Fandango</b>	4	4	The project foreseen four use case in four different domains: immigration, european policies, climate change, plus Covid 19 (we added this during the last year)
<b>TT</b>	48	-	<ol style="list-style-type: none"> <li>1. Smart highways pilot: Understand better the road traffic and mobility patterns</li> <li>2. Smart highways pilot: Optimize highway operations</li> <li>3. Smart highways pilot: Guarantee safer roads and make a better use of these roads</li> <li>4. Connected cars pilot: Breakdown estimator system.</li> <li>5. Connected cars pilot: Emission reduction system.</li> </ol>

6. Connected cars pilot: Traffic jam detector system.
7. Connected trucks pilot: Using Big Data approach on traffic and movement data specifically related to truck fleets.
8. Connected trucks pilot: Generation of truck specific patterns.
9. Connected trucks pilot: Provision of truck specific traffic information (actual and predictive).
10. Connected trucks pilot: Event related information on specific POIs (ramps, ports, toll stations).
11. Connected trucks pilot: Analyses on infrastructure, events, lane-specific traffic conditions for trucks (satellite images).
12. Rail Initial pilot use case 1: Overhead line equipment (OLE)
13. Rail Initial pilot use case 2: Point Machines
14. Rail Initial pilot use case 3: Track Circuits
15. Rail Initial pilot use case 4: Train Track Interface
16. Rail Replication pilot Use Case 1: Prediction of the degradation of point machines
17. Rail Replication pilot Use Case 2: Prediction of track profile degradation
18. Rail Replication Pilot Use Case 3: Use predictive models to optimise railway operation in the Rail Traffic Management System
19. Valencia port pilot: Yard Crane Scheduling Optimization
20. Valencia port pilot: Predictive Maintenance
21. Valencia port pilot: Predictive Web Cockpit
22. DUISPORT pilot: Terminal Productivity Cockpit
23. DUISPORT pilot: Predictive Maintenance System
24. Athens airport pilot: Operation Management Predictive Optimization
25. Athens airport pilot: Descriptive passenger behaviour system
26. Athens airport pilot: Retail model and the consumption ratio per passenger at a flight level
27. Malpensa airport pilot: Improve allocation of STA based on historical operation data analysis
28. Malpensa airport pilot: Track and predict Aircraft taxi in and out times
29. Malpensa airport pilot: Track and predict boarding process
30. Malpensa airport pilot: Track and predict Aircraft turn-around timing process
31. Malpensa airport pilot: Track and predict Aircraft sequence before departure
32. Malpensa airport pilot: Feed turnaround prediction data into A-CDM platform
33. Malpensa airport pilot: Feed enhanced ETA into Airport A-CDM platform based on historical operation data analysis
34. Malpensa airport pilot: Monitor performance of the Airport A-CDM platform estimations
35. Malpensa airport pilot: Coordinate prioritization with ATC
36. Malpensa airport pilot: Assess and predict time for passenger to reach the gate
37. Malpensa airport pilot: Track and predict Aircraft de-icing time and sequence before departure
38. Valladolid urban mobility pilot: Generating a traffic model for particular areas in the city where freight transport has more impact.
39. Valladolid urban mobility pilot: Analyse different freight delivery scenarios and make best decision according to the results.
40. Valladolid urban mobility pilot: Creation of a planning tool for delivery fleets and to inform drivers about the optimal route plan
41. Tampere urban mobility pilot: Provision of tools for urban traffic management for diagnosis of traffic status and for selection of alternative solutions for mitigating the impact of roadworks and other events. Addition of new data sources for improved situational awareness, such as social media and traffic cameras.
42. Tampere urban mobility pilot: Provision of tools for drivers and travellers regarding traffic status.

			<p>43. Tampere urban mobility pilot: Provision of tools to improve the access of goods delivery vehicles to parking places.</p> <p>44. Dynamic supply networks pilot: Identification of delivery patterns and problematic issues and forecasting</p> <p>45. Dynamic supply networks pilot: End to end information sharing</p> <p>46. Dynamic supply networks pilot: Shared micro-hubs (in terms of space) where customers collect their online order.</p> <p>47. Dynamic supply networks pilot: Inventory routing in an omnichannel environment</p> <p>48. Dynamic supply networks pilot: Online consumer insights regarding logistics and delivery processes</p>
<b>SmartDataLake</b>	3	3	(1) Business-to-Business Portfolio Recommendation (2) Seasonal Pattern Assessment for the Production, Delivery and Pricing of Goods (3) Intelligent Advisors for Innovation, Procurement and Investment Decisions
<b>Lynx</b>	3	3	<p>Use Case 1: QA on labour law (Spanish labour law)</p> <p>Use Case 2: Recommender on geothermal energy (Dutch and European geothermal energy regulations and permits)</p> <p>Use Case 3: contract analysis and management (Austrian case law and above)</p>
<b>Data Pitch</b>	29	not applicable	Data Pitch supported 29 startups to work with on business and/or societal challenges using shared data provided by corporations and public sector organisations.
<b>QROWD</b>	7	0	<p>1-3) QROWDLabs 5,6, and 7: Small-scale experiments in preparation for Official QROWD-Lab experiment for modal split data</p> <p>4) QROWDLab official: Large-scale use case to collect modal split data</p> <p>5) Experiment to collect bike rack locations and images from in-place volunteers</p> <p>6) Experiment to collect bike rack locations and images from Crowd-workers</p> <p>7) Experiment to fusion geospatial data from different sources (5 + 6 + existing static sources)</p>
<b>BodyPass</b>	0	3	<p>Case 1: analysis of CT scans and body scans with obesity patients in two hospitals (Spain &amp; Italy)</p> <p>Case 2: use of 3D scan databases for designing apparel</p> <p>Case 3: use of 3D scan databases to advise online purchase of clothing</p>

**Table 23:** Experiments performed by projects during 2019-2020

BDVA i-Space	Number of data experiments/use cases 2019 (number)	Number of data experiments/use cases 2020 (number)	Please provide some qualitative data to support your answer (i.e list, brief description,...)
Know-Center GmbH	60	70	25 (2019) /35 (2020) projects running within the center; all of them involving several AI and data experiments.
ITI - INSTITUTO TECNOLÓGICO DE INFORMÁTICA	20	38	Most of the experiments are related to Manufacturing or Health Sector, driven both by companies or by research centres/Universities.
Universidad Politécnica de Madrid	3	3	Cities: The i-Space is providing support to several cities (Madrid, Zaragoza, Santiago de Compostela, A Coruña) for their open data publishing strategies. More specifically, in the case of Madrid, a platform is being created as a spin-off of the initial living labs activity, with the purpose of allowing the usage of open data, with a special focus on air quality and mobility.  Geographic Data: Data from the National Geographic Institute on geolocated resources in Spain, published as Linked Data.

**Table 24:** Experiments performed by BDVA i-Spaces

### 7.3 II.14: Amount of data that has been made available for experimentation (cPPP projects and i-Spaces)

This KPI measures the number of Exabytes of data made available in cPPP projects and i-Spaces (including closed/private data). This includes data made available for experimentation (e.g. Data incubators and i-Spaces) and data made available internally in the projects for experimentation purposes (e.g. Lighthouses and other projects).

20 projects provided data for this KPI reporting a total of **0,001669 Exabytes (1,669 Petabytes)** for 2019-2020.

**Table 25** offers details per project about the amount and type of data used (source, providers, etc). It is important to notice that some of the projects are not only providing internal access to diverse data sets from different sources, but also are improving and creating new valuable datasets (e.g of ExtremeEarth). Additional projects reported experimentation started in 2021 and therefore they don't contribute to this KPI in 2019-2020.

Project	Amount of data made available in the project in 2019 (including closed/private data)	Amount of data made available in the project in 2020(including closed/private data)	Additional information about the data (source, providers, etc)
<b>BigMedilytics</b>	>5 TB	>5TB	a) The main data source are structured text files describing features from EHRs. These files have been exported from the current Healthcare Information Systems of the Valencian Region.  b) EHR data of more than 5 million patients  c) Some pilots are not about data volume, but focus more on using complex data

			<p>d) Imaging data (computed tomography) and connected information (report information)</p> <p>e) Text and images</p> <p>f) Processing biomedical articles with Natural Language techniques resulted in a 6.5 Gbytes file that is stored in a graph database (Neo4j).</p> <p>g) Medical images in multiple modalities supported by hospital record data from hundreds of patients.</p> <p>h) IoT data from real-time locating systems</p>
<b>AEGIS</b>	10 GB	0	<p>automotive sensor data collected by VIF</p> <p>smart home sensor data collected by Hypertech and Konkat</p> <p>private data of the insurance company HDIA</p> <p>weather data</p> <p>social media data</p> <p>open data from European Data Portal</p>
<b>More</b>	0	10 TBs	The data are private closed data from wind turbines and solar parks. Access to them has been provided by the project partners Engie Laborelec and INACCESS
<b>EW-Shop</b>	0.05	N/A	approx. 60 petabyte of data were used/shared, with dominant weight of weather data
<b>INFORE</b>	8	8	For the SPRING use Case and Pilot: Mainly stock market data from various exchanges.
<b>Infinitech</b>	N/A	100 TB	The 15 Pilots of the project have they own set of data. Summing the data managed by the first wave od demonstrators will manage 10-100 terabytes. Pilots will process a subset of the data expected in productions. In some cases data are samples and in other sythetic. Quantities are expected exceed these number by 1 or 2 orders of magnitude.
<b>Cross-CPP</b>	big amount of datapackages coming from CPPs in the CPP Marketplace available, exact amount of GB hard to estimate	> 12 Mio. datapackages coming from CPPs in the CPP Marketplace available, exact amount of GB hard to estimate	data coming from Cyper Physical Products of different industrial sectors (cars, smart buildings). Volkswagen provides car data, Siemens provides smart buildings data.
<b>Examode</b>	3 Terabytes	18 Terabytes	<p>Anonymized annotated data WSI + clinical reports, Catania hospital, for Breast, Lung, Colon and Cervix, 16.4 Terabytes</p> <p>Anonymized annotated data WSI + clinical reports, Annotated training data and models for WSI, RABBOUDUMC, for CeliacDisease and Uterine Cervix, 1.5 Terabyte</p> <p>Annotated text descriptions with SNOMED CT codes for diseases, procedures and findings, Ontotext, 4.4GB</p>
<b>Boost 4.0</b>	40Tb	40Tb	The data providers are basically the pilots. As there are many pilots, data types and characterization is very diverse. There is structured and unstructured data, technical drawings, sensor data flows, GPS devices information, etc.
<b>ExtremeEarth</b>	1 TB	4 TB	<p>1. Earth Observation data (sentinel-1/2/3) provided by CREODIAS.</p> <p>2. INVEKOS dataset for Austria, provided as open data from the country.</p> <p>3. LUCAS 2018 dataset provided by Copernicus Land Monitoring Service.</p>

			<p>4. EU-Hydro River Network dataset provided by Copernicus Land Monitoring Service.</p> <p>5. Corine Land Cover 2018 dataset provided by Copernicus Land Monitoring Service.</p> <p>6. Ice maps dataset archive from 2018, provided as open data by our partner METNO.</p> <p>7. Global Administrative Divisions dataset (GADM) provided online as open data.</p>
<b>DeepHealth</b>	6 TB	6 TB	<p>Around 6 terabytes taking into account all the datasets of all the use cases.</p> <p>Mainly biomedical images, depending on the use case image type can be in X-Ray, Computerized Tomography, Magnetic Resonance Image, or RGB photos of the skin. Additionally, UC13 is based on EEGs and UC1 in several parameters (some are vital signs) measured during so long periods of time.</p>
<b>Safe-Deed</b>	were in the order of Gigabytes	Data were in the order of Gigabytes	<p>The QSD is based on order data from a high number of companies.</p> <p>It allows for research while at the same time keeping the data confidential ensuring the privacy of the data.</p>
<b>MediaFutures</b>	0	1	<p>We have made available a data catalogue for applicants at <a href="https://mediafutureseu.github.io/datasets">https://mediafutureseu.github.io/datasets</a></p>
<b>IoTwins</b>	10 Terabytes	76 Terabytes	<p>Data come from on-field sensors provided by the partners running the 12 test-beds. Data type is very heterogeneous: velocity, temperature, load, forces, vibrations to name a few.</p>
<b>EDI</b>	20 terabytes	30 terabytes	<p>In 2019:</p> <p>DATA PROVIDER DIMENSION</p> <p>Cartiere del Garda - Lecta Group 1GB</p> <p>EDP &lt;8GB</p> <p>EMASESA TBD</p> <p>ENERGA 165KB</p> <p>IAAF 40GB</p> <p>Ibermutua TBD</p> <p>MIGROS TBD</p> <p>SONAE TBD</p> <p>Technical University of Munich 0.001GB</p> <p>Telia Eesti AS TBD</p> <p>TUB 100GB</p> <p>UBIMET 1GB</p> <p>VPS 1-50GB</p> <p>VRT TBD</p> <p>YKT 16GB</p>

			<p>In 2020:</p> <p>DATA PROVIDER DIMENSION</p> <p>RACC 200 MB</p> <p>JOT Giga</p> <p>HMR 40 GB</p> <p>EDP N/A</p> <p>IBERMUTUA Tera</p> <p>MIGROS 5 GB</p> <p>VRT N/A</p> <p>EMASESA 100 GB</p> <p>UBIMET 210 MB</p>
<b>CloudButton</b>	100 TB	0	EMBL made available METASPACE public raw data (>100TB).
<b>TT</b>	According to the data received from the pilots, 0.000051393 exabytes (51.39TB) are currently being used or produced and a total amount of 0.00015595 exabytes (155.95TB) have been used during the project. This amount of data is expected to grow as the project progresses.	-	The description of all the datasets used by the project is publicly available in the following link: <a href="http://data.transformingtransport.eu/">http://data.transformingtransport.eu/</a>
<b>SmartDataLake</b>	10	10	For the SPRING use Case and Pilot: Mainly stock market data from various exchanges.
<b>Data Pitch</b>	Unknown	not applicable	<p>The Data Pitch datasets catalogue is here for the second cohort of Data Pitch startups, which took part in the 2019 accelerator: <a href="https://campaign.dawex.com/en/lp/datapitch-datasets/2018/">https://campaign.dawex.com/en/lp/datapitch-datasets/2018/</a></p> <p>Masai Mobility datasets were not matched with a startup to share data however. In addition, 14 startups joined the accelerator that had found their own data providers who shared data.</p> <p>Further information about data usage across both cohorts of Data Pitch can be found here: <a href="https://datapitch.eu/wp-content/uploads/2020/01/D2.5_v_0.7.pdf">https://datapitch.eu/wp-content/uploads/2020/01/D2.5_v_0.7.pdf</a></p>
<b>QROWD</b>	1000 Terabytes	0	<p>Geospatial data provided by TomTom (closed)</p> <p>Floating car positions for Trento provided by TomTom (closed)</p> <p>Heterogeneous datasets provided by the Municipality of Trento (open)</p> <p>Sensor data collected from participants in the Trento Use Case (closed)</p> <p>Images and coordinates of mobility infrastructure (bike racks)</p>
<b>BodyPass</b>	1 Tera	1 Tera	3D body scans

**Table 25:** Data made available in the projects in 2019-2020 – details per project

Additionally, the BDVA labelled i-Spaces also contribute to this KPI (further information here: <http://www.bdva.eu/node/1172>).3 i-Spaces contributed to this KPI reporting additional **0.04 Petabytes** of data for experimentation. In particular:

i-Spaces	Amount of data made available in the i-Space in 2019 (including closed/private data)	Amount of data made available in the i-Space in 2020 (including closed/private data)	Additional information about the data (source, providers, etc)
Know-Center GmbH	not available	not available	not available
ITI - Instituto Tecnológico De Informática	15TB	25TB	<p>Datasets available in ITI Data Space can be published by partner organizations both publicly for all iSpace users and privately for members of their organization or other organizations with a data sharing agreement established.</p> <p>When a user registers an experiment request, they can select one or more datasets on which they have access permissions to be provisioned together with the computing resources for their experiment.</p> <p>The Data Space has datasets from various public and private entities. These datasets are of different types raw data, images, time series, json messages, Structured Data exported from databases, csv and nii MRI files</p> <p>The total size of these datasets is approximately 25 TB</p>
Universidad Politécnica Madrid	not available	not available	not available

**Table 26:** Data made available in BDVA i-Spaces

Combining reported information from projects and i-Spaces (members of BDVA) the number of Exabytes of data made available in cPPP projects and i-Spaces is **0,001709 Exabytes** (1,709 Petabytes).

## 8 SRIA UPDATE AND IMPLEMENTATION (II.9, II.10)

### 8.1 KPI II.9: Ensure efficiency, transparency and openness of the cPPP's consultation process

During 2019-2020 there was no updates of the BDA cPPP SRIA as the programme was ending. Efforts have been put to contribute to the AI, Data and Robotics PPP SRIDA, to the EuroHPC JU roadmap documents, and whitepapers produced by the association.

### 8.2 KPI II.10: Ensure that technology progress is in line with multi-annual roadmap of SRIA

To determine how well the SRIA technical priorities and challenges are covered by ongoing research and innovation activities of the running BDV CPPP projects, the BDVA (supported by BDVe) performed a systematic collection of data. To this end, a common template was shared among the BDV CPPP projects, to collect information about the contributions to the different technical areas of the SRIA.

At the end of the reporting period, 38 projects have provided the detailed mapping of their contributions and technologies to the SRIA technical priorities and challenges.

### SRIA Technical Priorities

The BDV SRIA<sup>23</sup> defines the overall goals, main technical and non-technical priorities, and a research and innovation roadmap for the European Public Private Partnership (PPP) on Big Data Value. The SRIA has been developed by the Big Data Value Association (BDVA), an industry-led organisation representing large businesses, small and medium-sized enterprises (SMEs), and research organisations in Europe.

Five *technical priority* areas have been identified for research and innovation:

**Data Management** as a foundation for higher-level big data techniques;

Optimized **Data Processing Architectures** for data-at-rest and data-in-motion;

**Data Analytics** to improve the understanding of data;

Mechanisms ensuring **Data Protection** and anonymization, to enable the vast amounts of data which are not (and never can be) open data to be incorporated into the data value chain;

Advanced **Data Visualization and User Interaction**.

### Results and Analysis of SRIA Coverage

Based on the data reported by the BDV CPPP projects, **Table 28** shows the consolidated and aggregated results over time showing only the main priorities. **Table 28** shows the same table but split in topics per technical priority<sup>24</sup>. The table depicts the number of technical contributions and big data technologies per SRIA technical priority and challenge, both as totals (first column) and per different phases / years of the PPP. In the rest of the columns, not only including the results of the year 2019, 2020 and ongoing and planned for 2021+, but results from the launched of the PPP (2017 and 2018 results correspond to the reports of those 2 years). As part of SRIA Version 4.0, four new technical challenges were introduced and are included in the table. As a result, the 2017 column does not report coverage of these challenges, as they were not part of the previous SRIA.

As can be seen on a high-level of analysis (totals in the table), each of the SRIA priorities is covered, as are all technical challenges.

The major focus of technical contributions lies in the "Data Analytics" priority, followed by "Data Management" and "Data Processing Architectures". All priorities are covered substantially and with a large number of contributions.

"**Data Protection**" shows a clear increase during the years 2019 and 2020 due mainly to the contribution of the ICT-13 projects or Data Platforms. "**Data Management**" has also been influenced by the ICT-13 with a solid increase in contributions during 2019 and 2020. "**Data Processing Architectures**" doubles its number of contributions in 2021+ thanks to the contribution of the recently started projects part of the ICT-51 topic. "Data Analytics", the priority with the highest number of contributions, keeps quite a steady line along the years, also showing an increase for the 2021+ contributions. "**Data Visualisation and user Interaction**" receives less contributions than the others but in a steady manner. In 2019 this priority was the year with the largest number of contributions, and similar numbers as expected in 2021+. Interesting to highlight that projects have reported quite a large number of contributions for 2021+ under the category "**others**", some specific contributions under the others category include e.g. Data asset supported through a Data & AI Marketplace (in which datasets, AI models, analytics results can be shared/traded in a trustworthy manner such as in the Synergy project), multi-party data contracts, Self-Sovereign Identity tool, VC issuing and signing tool, mapping data properties to data value, aggregation of data value dimensions, test collection for de-anonymisation algorithm, and legal compliance with existing binding and non-binding regulatory initiatives touching upon data processing

---

23 S. Zillner, E. Curry, A. Metzger, R. Seidl (Eds.), "European big data value strategic research and innovation agenda (SRIA)," Version 4.0, November, 2017; [http://www.bdva.eu/sites/default/files/BDVA\\_SRIA\\_v4\\_Ed1.1.pdf](http://www.bdva.eu/sites/default/files/BDVA_SRIA_v4_Ed1.1.pdf)

24 The questionnaire separated the last column "Ongoing in 2021 and planned 2021+" into 2 questions: "Ongoing (not delivered but initiated in 2019 or earlier)" and "Planned (for 2021+)". As some projects reply to one and not the other, and due to the ambiguity in between the 2 questions this report merged them into 1 considering contributions delivered in 2021 or later.

Technical priorities and challenges	Number of contributions/technologies					
	TOTAL	Delivered in 2017	Delivered in 2018	Delivered in 2019	Delivered in 2020	Ongoing in 2021 and planned 2021+
Priority "Data Management"	547	44	122	66	145	170
Priority: "Data Processing Architectures"	510	9	126	58	117	200
Priority: "Data Analytics"	697	11	183	120	177	206
Priority: "Data Protection"	196	15	33	16	47	85
Priority: "Data Visualisation and User Interaction"	247	5	56	72	53	61
Other	21	0	0	0	0	21

**Table 27.** Coverage of SRIA Technical Priorities and Challenges (Number of Contributions/Technologies)

Looking at the totals over time (2017 to 2021+) we can observe that:

"**Data Analytics**" Is the priority with the **largest number of contributions** and In particular the subtopics of "**Data Analytics and AI**" (with clear Increase In contributions by the projects In the last 2 years), "**predictive and prescriptive analytics**", and "**analytics frameworks and processing**" are the ones with the larger number of contributions. The subtopics of "**Advanced business analytics and intelligence**" and "**Semantic and knowledge-based analysis**" also received a substantial number of contributions in a consisten way along the years. "**Content validation**" Is the subtopic receiving the lowest level of contributoins, although 2019 was the year with the largest number of contributoins In this area. **HPDA** Is still In Its Infancy. ICT-51 projects launched at the end of 2020 and beginning of 2021 are planned to contribute to this subtopic and additonal research Is expected as part of Horizon Europe.

"**Data Management**" Is the priority with the second largest number of contributions and In particular In the subtopics of "**Semantic Annotation of unstructured and semi- structured data**", "**Data-as-a-service**", and "**Integration of data and business processes**". Important to highlight the **fast growth In contributions In the area of "Distributed trust infrastructures for data management" In the last 2 years**, which Indicates the level of Interest In this area.

"**Data Processing Architectures**" has received a similar number of contributions that the "Data Management" priority. Important to highlight the large number of contributions expected for 2021+ in view of the ICT-51 projects launched at the end of 2020-2021. **Scalability** and **Heterogeinity** are the subtopics that have received the largest number of contributions, Followed by **Performance**, **Processing of data-In-motion and data-at-rest** and **Decentralisation**. In terms of trends In general all subtopics have Increased number of contributions from 2019 to 2021+.

Under the priority "**Data Protection**" the topics "**Generic and easy to use data protection approaches**" and "**Robust Data privacy (incl. multi-party computation)**" are the ones that have received the largest number of contributions with an Increasing trend In general. Finally the priority "**Data Visualisation and User Interaction**" has received over the years a steady number of contributions being "**Collaborative, intuitive and interactive visual interfaces**".

Technical priorities and challenges	Number of contributions/technologies					
	TOTAL	Delivered in 2017	Delivered in 2018	Delivered in 2019	Delivered in 2020	Ongoing in 2021 and planned 2021+
<b>Priority "Data Management"</b>	<b>547</b>	<b>44</b>	<b>122</b>	<b>66</b>	<b>145</b>	<b>170</b>
Semantic Annotation of unstructured and semi- structured data	107	9	31	13	31	23
Semantic interoperability	67	11	13	6	20	17
Data quality	65	7	15	12	15	16
Data Management lifecycle and data governance	71	9	11	6	20	25
Integration of data and business processes	90	3	44	9	21	13
Data-as-a service	93	4	4	11	24	50
Distributed trust infrastructures for data management	53	1	4	9	14	25
Other (specify)	1	0	0	0	0	1
<b>Priority: "Data Processing Architectures"</b>	<b>510</b>	<b>9</b>	<b>126</b>	<b>58</b>	<b>117</b>	<b>200</b>
Heterogeneity	104	4	33	12	21	34
Scalability	114	4	28	11	23	48
Processing of data-in-motion and data-at-rest	76	0	19	8	18	31
Decentralization	72	0	12	11	20	29
Performance	80	1	18	7	19	35
Novel architectures for enabling new types of big data workloads -- new in SRIA4.0	40		13	4	9	14
Introduction of new hardware capabilities -- new in SRIA4.0	22		1	5	7	9
Other (specify)	2	0	2	0	0	0
<b>Priority: "Data Analytics"</b>	<b>697</b>	<b>11</b>	<b>183</b>	<b>120</b>	<b>177</b>	<b>206</b>
Semantic and knowledge-based analysis	88	5	21	17	22	23
Content validation	36	2	6	7	18	3
Analytics frameworks & processing	135	4	25	19	31	56
Advanced business analytics and intelligence	105	0	53	19	19	14
Predictive and prescriptive analytics	140	0	53	28	28	31
High Performance Data Analytics (HPDA) -- new in SRIA 4.0	46		13	13	9	11
Data analytics and Artificial Intelligence -- new in SRIA 4.0	143		12	17	48	66
Other (specify)	4	0	0	0	2	2
<b>Priority: "Data Protection"</b>	<b>196</b>	<b>15</b>	<b>33</b>	<b>16</b>	<b>47</b>	<b>85</b>
Generic and easy to use data protection approaches	73	0	9	10	21	33
Robust Data privacy (incl. multi-party computation)	67	1	21	4	12	29
Risk based approaches	16	0	3	2	5	6
Other (specify)	40	14	0	0	9	17
<b>Priority: "Data Visualisation and User Interaction"</b>	<b>247</b>	<b>5</b>	<b>56</b>	<b>72</b>	<b>53</b>	<b>61</b>
Visual data discovery	61	3	10	14	16	18
Interactive visual analytics of multiple scale data	59	0	10	22	17	10
Collaborative, intuitive and interactive visual interfaces	79	2	31	18	8	20
Interactive visual data exploration and querying in a multi-device context	45	0	5	18	12	10
Other (specify)	3	0	0	0	0	3
<b>Other (specify)</b>	<b>21</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>21</b>

**Table 28.** Coverage of SRIA Technical Priorities and Challenges (Number of Contributions/Technologies) - With Subtopics

### Achievement of Specific KPI:

For what concerns SRIA coverage, the relevant specific KPI is "Ensure that technology progress is in line with multi-annual roadmap of SRIA" (KPI II-10).

As agreed by the respective BDVA task force, this KPI is measured as "% of research priorities covered compared to overall scope of research priorities defined in SRIA".

Based on the above table, the KPI is quantified as follows:

- Delivered in 2019: 100%
- Delivered In 2020 100%
- Ongoing 2021 and planned for 2021+: 100%

### Non-technical priorities:

In addition to the technical priorities' implementation is important to assess the non-technical aspects of the SRIA and assess status of implementation.

- **Skills**, is fully covered by the BDVe project and BDVA TF9 and complemented by many activities and contributions from most of the projects as detailed in section 2 of this annex.
- **Societal and Ethical** aspects of Big Data are covered by e-Sides project and BDVA TF5.
- **Policy** aspects are mainly covered by BDVA TF5.
- **Business** in particular regarding the support to SMEs, it is covered by BDVA TF8 and BDVe project and it has been reported under section 3 of this annex.

## 9 TECHNICAL PROJECTS (II.4, II.15, II.16)

### 9.1 KPI II.4: Privacy and Security KPI

The analysis of KPI II.4 “Enable advanced privacy and security respecting solutions for data access, processing and analysis” is based on two sources of information. On the one hand, we used the reports from projects about the coverage of the SRIA technical priorities (see section 8 of this report). On the other hand, we collected from the 40 projects that replied to the questionnaire specific information on three specific sub-KPIs:

- II.4a) Number of patents filed by cPPP projects that enable advanced privacy and security
- II.4b) Number of publications by cPPP projects that describe advanced privacy and security
- II.4c) Number of OSS contributions / products / SW components resulting from cPPP projects that enable advanced privacy and security

Based on the data reported by the BDV CPPP projects. Table 29 shows the consolidated and aggregated results for the three specific sub-KPIs.

The table also shows results reported in 2018 to show the evolution.

II.4 a) Number of patents filed that enable advanced privacy and security respecting solutions for data access, processing and analysis			II.4 b) Number of publications that describe advanced privacy and security respecting solutions for data access, processing and analysis			II.4 c) Number of OSS contributions / SW / Products that enable advanced privacy and security respecting solutions for data access, processing and analysis		
Year 2018	Years 2019-2020	2021-2024 (forecasted)	Year 2018	Years 2019-2020	2021-2024+ (forecasted)	Year 2018	Years 2019-2020	2021-2024 (forecasted)
2	10	49	61	173	189	24	158	115

**Table 29.** Sub-KPIs for KPI II.4 “Enable advanced privacy and security respecting solutions for data access, processing and analysis”

From the PPP projects, 3 technical projects were fully dedicated to advance the state of the art in privacy-preserving Big Data technologies: SPECIAL, SODA and MHMD and ended on 31/12/2019. Specific contributions from SPECIAL, SODA and MHMD can be found in **Table 30** However many other projects contribute to these KPIs as expected during 2019 and 2020 from projects belonging to H2020-ICT-13-2018-19 call for proposals, focused on scaling privacy preserving technologies and close to deployment industrial and data platforms.

Project	II.4 a) Number of patents filed that enable advanced privacy and security respecting solutions for data access, processing and analysis			II.4 b) Number of publications that describe advanced privacy and security respecting solutions for data access, processing and analysis			II.4 c) Number of OSS contributions / SW / Products that enable advanced privacy and security respecting solutions for data access, processing and analysis		
	2019	2020	2021 - 2024	2019	2020	2021 - 2024	2019	2020	2021 - 2021
SODA	1	N/A	N/A	10	N/A	N/A	5	N/A	N/A
SPECIAL	1	N/A	N/A	8	N/A	N/A	3	N/A	N/A
MHMD	1	N/A	N/A	5	N/A	N/A	1	N/A	N/A
Kraken	0	0	1	0	3	15	0	0	3
MosaiCrown	0	0	2	9	10	30	1	2	3
Safe-Deed	0	0	0	2	4	3	2	5	0
Pim-City	0	0	1	0	12	10	0	6	2
Data Ports	0	0	0	0	3	4	0	0	3
Musketeers	1	3	3	7	14	TBD	2	5	N/A
Opertus Mundi	0	0	0	0	4	22	0	18	24
SmashHit	N/A	0	0	N/A	0	2	N/A	0	3

9.2 **Table 30.** Sub-KPIs for KPI II.4 “Enable advanced privacy and security respecting solutions for data access, processing and analysis” for ICT-18-2016 projects (RIA) and H2020-ICT-13-2018-19 “Supporting the emergence of data markets and the data economy”

Moreover, in the healthcare sector, DeepHealth OSS contributions are directly linked with enabling advanced privacy and security respecting solutions for data access, processing and analysis.

3 projects in the retail sector have reported patents filed that enable advanced privacy and security respecting solutions for data access, processing and analysis: EW-Shopp, PimCity SmartDataLake.

Finally a large list of projects have contributed delivering OSS, software or products that enable advanced privacy and security respecting solutions for data access, processing and analysis during 2019-2020. In particular AEGIS, Infinitech, EW-Shopp, MosaiCrown, INFORE, IoTwins. Some examples include:

- AEGIS: the AEGIS platform
- Infinitech: Asset Tokenization in Blockchain Infrastructures, Open Source Service for Analyzing Blockchain Graphs for Fraudulent Activities, Secure Personal Data Marketplace based on Homographic Encryption and Tokenization
- EW-Shopp: ASIA - Weather data API - Event and Weather Analytics Toolset
- MosaiCrown: a policy engine responsible for parsing the MOSAICrOWN policy and checking whether a subject request is permitted; a tool for applying an All-Or-Nothing Transform encryption mode, combined with fragmentation and slicing, for protecting data in storage; a tool for anonymizing a collection of sensitive data with Mondrian algorithm with the Apache Spark framework

- INFORE: provided an OSS in the field of market data analysis, especially regarding stock market data  
 loTwins: provided two OSS products. An edge node manger based on the Apache Mesos and Docker tools (Mesos4Edge) and OPC UA PubSub, which uses a publish-subscribe communication model instead of a client-server approach

### 9.3 KPI II.15: Availability of metrics for measuring the quality, diversity and value of data assets

This KPI refers to the number of metrics developed by the cPPP that allow for assessing quality, diversity and value of data assets. From the running projects in 2019-2020, 42,5% of the projects that participated to the questionnaire confirmed they assessed quality, diversity and value of data assets.

Project	Metrics used to quantify quality, diversity and value of data assets
<b>BigMedilytics</b>	a) Time of care, visits to emergency room, visits to specialist, hospitalization, mortality, costs of health and labor.  b) Search speed in large-scale PACS archives
<b>Daphne</b>	end-to-end throughput, latency, resource utilization, memory consumption, energy, ML application accuracy.
<b>Everest</b>	Number of computations per unit of time. The project just started so we didn't measure any improvement yet.
<b>More</b>	number of data values processed per second
<b>EW-Shop</b>	The number of processed data (size).
<b>INFORE</b>	Velocity, Number of queries per second, Data load
<b>Infinitech</b>	- Number of Data Consumed per sector - Data Ingestion Rates - Time required for (real-time) data analytics including execution of ML/AI algorithms
<b>Kraken</b>	- Annual increase in the number of data provider organisations in the personal and industrial data platforms. - Annual increase in the number of data user/buyer organisations using industrial data platforms. - Annual increase in number of users (data subjects) in the personal data platforms. - Annual increase in volume of business (turnover) channelled through the platforms.  In all cases considering the previous EU-funded projects (MD-Paedigree, CARDIOPROOF, MHMD) and Streamr project.
<b>INFORE</b>	Query execution time and throughput in terms of the handled rate of input data streams.
<b>Safe-Deed</b>	GMiMC: A symmetric encryption scheme designed for efficient and secure multi-party computation.
<b>CloudButton</b>	We can use the total processing time of our experiments but also the amount of data processed by time unit. We were capable of launching thousands of cloud functions processing data from object storage, reaching aggregate throughputs of around 60 GB/s.
<b>SmartDataLake</b>	Query execution time and scalability with respect to the size of input data, in particular in approximate query answering and similarity joins.
<b>SmartDataLake</b>	Velocity, Number of queries per second, Data load
<b>Lynx</b>	Number of characters.
<b>Data Pitch</b>	We track the number of data providers sharing data with startups through our programme. The intention will be to change the mentality of data providers in working collaboratively, resulting in an increase in the amount of data being shared.

**Table 31.** Metrics used to quantify quality, diversity and value of data assets

These results show the intense usage of metrics to measure quality, diversity and value of data assets in projects and some projects have developed specific metrics and methods to ensure quality, diversity and value in the data.

### 9.4 KPI II.16: Increase the speed of data throughput compared to 2014

47,5% of the projects answering the questionnaire, reported they expect the project to improve data throughput. In particular INFORE, Kraken, SmartDataLakes and Infinitech reported expected outcomes in this area.

Some projects such Safe-Deed, Cloudbutton, and Lynx has measured improvement of 400-500%. Other such as INFORE has measured improvements over 70%, Kraken experienced an annual increase in the number of data provider organisations in the personal and industrial data platforms of 20%, number of data user/buyer organisations using industrial data platforms of 30%, users (data subjects) in the personal data platforms of 50% and finally an increase in volume of business (turnover) channelled through the platforms of 20%.

Table 32 provides details per project in relation to the metrics used and the specific levels of improvement.

Project	Metrics used for speed of data throughput	improvement (in percent) measured in the project
<b>BigMedilytics</b>	a) Time of care, visits to emergency room, visits to specialist, hospitalization, mortality, costs of health and labor.  b) Search speed in large-scale PACS archives	Up to 20% for both.
<b>Daphne</b>	end-to-end throughput, latency, resource utilization, memory consumption, energy, ML application accuracy.	0 (start Dec/2020)
<b>More</b>	number of data values processed per second	There have not been any results in the reference period.
<b>EW-Shop</b>	The number of processed data (size).	Not really measured. Estimated in around +30%
<b>INFORE</b>	Velocity  Number of queries per second  Data load	50%
<b>Infinitech</b>	- Number of Data Consumed per sector  - Data Ingestion Rates  -Time required for (real-time) data analytics including execution of ML/AI algorithms	-30% in the time need for data analysis (targeted)
<b>Kraken</b>	- Annual increase in the number of data provider organisations in the personal and industrial data platforms.  - Annual increase in the number of data user/buyer organisations using industrial data platforms.  - Annual increase in number of users (data subjects) in the personal data platforms.  - Annual increase in volume of business (turnover) channelled through the platforms.  In all cases considering the previous EU-funded projects (MD-Paedigree, CARDIOPROOF, MHMD) and Streamr project.	20%, 30%, 50%, 20% respectively.
<b>INFORE</b>	Query execution time and throughput in terms of the handled rate of input data streams.	70%
<b>Boost 4.0</b>	As this is not a main objective of BOOST 4.0 not all the pilots have defined any metrics, and not even all the pilots are measuring it. Some of the metrics that have been defined and is being measured or will be measured are:  - Lot-size-2 engineering lead time  - Processed sensor measurements per second	Depending on the pilot, from a 10 to a 40%

	<ul style="list-style-type: none"> <li>- Number of different data sources processed per second</li> <li>- Metrics defined by 3GPP in TS28.554</li> <li>- Collection, storage and data processing times</li> <li>- Number of samples per second</li> </ul>	
<b>Safe-Deed</b>	GMiMC: A symmetric encryption scheme designed for efficient and secure multi-party computation.	We observe improvements in throughput by a factor of more than 4 and simultaneously a 5-fold reduction of preprocessing effort, albeit at the cost of higher latency.
<b>CloudButton</b>	We can use the total processing time of our experiments but also the amount of data processed by time unit. We were capable of launching thousands of cloud functions processing data from object storage, reaching aggregate throughputs of around 60 GB/s.	With Lithops framework, we measured an improvement of 4x in the total processing time in the metabolomics experiments compared to a Spark cluster.
<b>Reach</b>	F6S is not the right partner in the project to answer this question	F6S is not the right partner in the project to answer this question
<b>SmartDataLake</b>	Query execution time and scalability with respect to the size of input data, in particular in approximate query answering and similarity joins.	50%
<b>SmartDataLake</b>	Velocity  Number of queries per second  Data load	60%
<b>Lynx</b>	Number of characters.	We do not know the state of affairs in 2014 but from project start until now it differs from service to services, for some it was between: 50% and 500%
<b>Data Pitch</b>	We track the number of data providers sharing data with startups through our programme. The intention will be to change the mentality of data providers in working collaboratively, resulting in an increase in the amount of data being shared.	We didn't collect percentage improvements, but did capture qualitative feedback from data providers that taking part in Data Pitch had changed their work processes and made them more willing to do open innovation projects and share data.

**Table 32.** Increase the speed of data throughput – details per project

## 10 MACRO-ECONOMIC KPIS (KPI II.1, II.2, II.3, II.5)

The monitoring of the macro-economic KPIs (KPI II.1, II.2, II.3 and II.5) is based on input from the European Data Market (EDM) study<sup>25</sup> as they are presented in the most recent report<sup>26</sup> by IDC (data March 2020, publication July 2020). Since 2013 this study has provided annual measurements of the European data market and economy based on indicators by Member State, industry and company size. This is all open data, published as an open dataset presenting the historical series of the indicators from 2013 to 2019 plus forecasts to 2025 under 3 alternative scenarios. The KPIs presented in this chapter are extracted from this study and cover the number of European Data Professionals, the number of European Data Companies (that is suppliers of data technologies and services), and the value of the revenues of data companies in the EU and three other countries (US, Japan and Brazil). In this monitoring report we present the 2013 baseline, figures for the period 2017-2019. All multi-annual comparisons are based on the 2013 figures.

25 SMART 2016/0063 – Study “Update of the European Data Market Monitoring Tool”, IDC and Lisbon Council

26 [The European data market monitoring tool - Publications Office of the EU \(europa.eu\)](https://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&language=en&plugin=1), Final Study Report (Deliverable D2.9), IDC and the Lisbon Council, July 2020

The IDC report provided estimates for 2020 which are now unreliable because of the COVID-19 pandemic impacts on the economy. The new set of data with updated 2020 and 2021 estimates will be produced in the fall of 2021 under a new release of the study (European Data Market Study 2021–2023). If feasible this data will be Included in an updated release of the Monitoring report.

The data presented is for the EU28 (EU27+UK) since Brexit did not formally happen until January 2020. For EU27 data the reader may refer to the EDM study dataset.

The IDC report provides forecast data for 2025 based on 3 alternative scenarios<sup>27</sup>:

- The Baseline scenario is characterised by a healthy growth of data innovation, a moderate concentration of power by dominant data owners with a data governance model protecting personal data rights, and an uneven but rather wide distribution of data innovation benefits in the society. This is considered the most likely scenario.
- The High Growth scenario is characterised by a high level of data innovation, low data power concentration, an open and transparent data governance model with high data sharing, and a wide distribution of the benefits of data innovation in the society.
- The Challenge scenario is characterised by a low level of data innovation, a moderate level of data power concentration due to digital markets fragmentation, and an uneven distribution of data innovation benefits in the society.

Based on estimates produced at the end of 2020 during the economic crisis driven by the COVID-19 pandemic, the study foresees a strong rebound of growth at the end of 2021, so the forecast scenarios remain broadly valid, but the forecast values are likely to be lower than the current ones, since the rebound will start from a lower market value than previously expected. Updated scenarios will be produced In October 2021.<sup>28</sup>

## 10.1 KPI II.1: Market share of the European Union in the global Big Data Market

KPI II.1 assesses development of the market share of the European Union in the global Big Data Market. As an indicator, we will compare the total revenues of EU Data Companies with other economies, i.e. the US, Japan, and Brazil, as they are used as a benchmark in the IDC report<sup>29</sup>.

The EU share of the total revenues in these economies the 2013 baseline was 27,7%. This share has remained substantially stable since then, with a slight decrease to 27.02% in 2019. However, the EU data companies' revenues grew faster than those from Japan and Brazil second only to the US companies growth. This is remarkable considering the dominance of global US platforms In the data market. In absolute terms, the amount of the total revenues of US Data Companies in 2019 was more than twice the revenues of EU28 Data Companies in the same year (€185 bn vs. €84 bn).

Table 33 presents a detailed comparison for the years 2018-2019 and shows that the EU Data Companies revenues compound annual growth for 2013-2019 (9.8%) outperform Japan (6.1%) and Brazil (9%) in terms of relative and absolute growth in revenues. The United States show strong growth in revenues of 11.3% in relative terms In the same period but starting from a much higher baseline and therefore by far the highest value of revenues in absolute terms.

World Region	Revenues (€B)		Share		CAGR 2013-2019
	2019	2018	2019 % of total	2018 % of total	
EU27+UK	83,545	77,297	27.02%	27.76%	9.8%
USA	184,873	163,993	59.78%	58.89%	11.3%
Brazil	7,905	7,373	2.56%	2.65%	9.0%

27 The EDM monitoring tool, ibidem, D.2.9 Final study report, chapter 2.1, pages 28-29

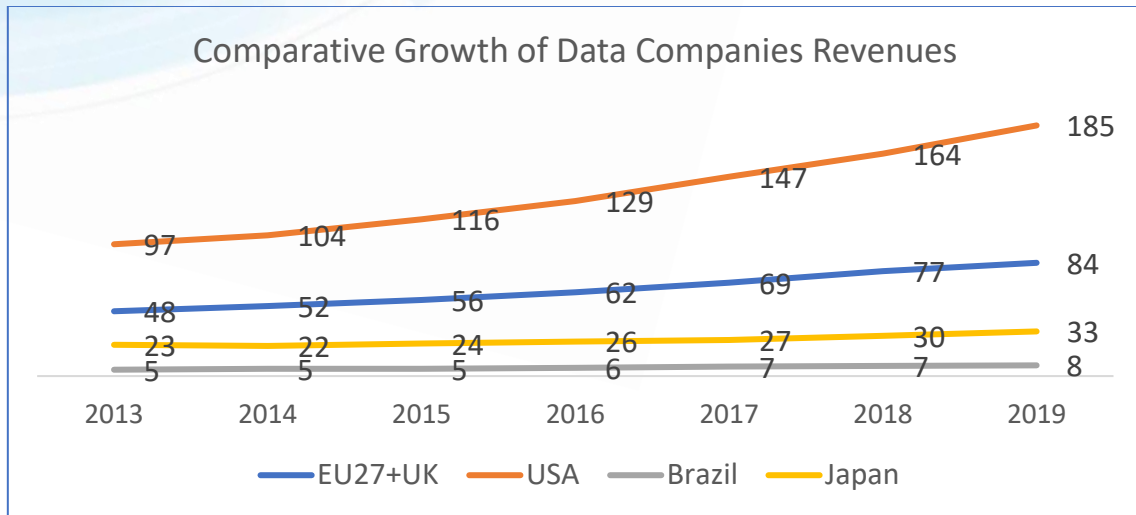
28 The EDM monitoring tool, ibidem, D.2.9 Final study report, The EU27 Data Economy post-Covid, page 66

29 The EDM monitoring tool, ibidem D.2.9 Final study report, chapter 2.5, pages 41-47

Japan	32,929	29,799	10.65%	10.70%	6.1%
Total	309,252	278,462	100.00%	100.0%	10.2%

**Table 33:** Comparison of data companies revenues by country

Finally, Figure 5 shows the increase of the total revenues of Data Companies revenues of the selected economies since 2013. The US and Europe show the highest growth rate and value compared to 2013, with the US clearly being ahead of the rest of the world in terms of the size of the industry.



**Figure 5:** Data Companies Revenues Growth, EU27+UK versus other economies, 2013-2019

**60% of the cPPP projects** active in 2019-2020<sup>30</sup> reported contribution to increase revenue share of EU companies against total of revenue of EU, US, Japan, Brazil. While some projects provide statements in potential impact in market share, others can already share some evidence in real examples: e.g giving cross-sectorial industries access to the great spectrum of sensor data coming from high volume products from various industrial sectors (SmashHit) With the increasing number of connected sensors and actuators within such mass products (so called cyber physical products), this number will rise exponentially in short-term. This enormous amount of data offered by various data sharing platforms represent a massive information resource to create new value, allowing the improvement of existing services or the establishment of diverse new innovative services. Other projects, e.g More, state that the successful completion of the project will increase the competitiveness of most stakeholders in the Wind and Solar energy industries. Moreover, it will create new jobs in the data management industry with specialized tools for the renewable energy industry.

Some projects contribute to the quick adoption of new technologies, making their user industries more competitive in a world-wide data economy. For example, MosaiCrown provides effective data protection solutions, that can encourage institutions and entrepreneurs to enter the data sharing market by enabling them to become trustworthy companies. SmashHit acknowledge the emergence in the last years of many data sharing platform approaches for different purposes (like B2B data exchange, data marketplaces or data processing services), giving cross-sectorial industries access to the great spectrum of sensor data coming from high volume products from various industrial sectors (vehicles, smart home devices, smart cities data etc.). This enormous amount of data offered by various data sharing platforms represent:

<sup>30</sup> Based on number of respondents.

- a massive information resource to create new value, allowing the improvement of existing services or the establishment of diverse new innovative services, by combining data streams from various sources
- a major big data-driven business potential, not only for the manufacturers of Cyber Physical Products (CPP), but in particular also for cross-sectorial industries and various organisations with interdisciplinary applications

Other ways projects are supporting this KPI:

- Incubators: Supporting EU data driven companies to build innovative solutions that can be scaled internationally, organising networking event with investors and mapping funding opportunities
- Accelerating adoption (mainly in lighthouse projects)
- Developing innovative technologies (software and hardware) that can be European companies more competitive
- adopting data protection approach that can make European data industries more competitive (SmaHit).
- Enabling industries to exploit their big data efficiently and therefore increase their market share and services provided to their customers
- Creating marketplaces that would enable radical innovations by numerous cross sectorial service providers (e.g. Cross-CPP)

Table 34 provides details of the qualitative contribution of projects to the Market share of the European Union in the global Big Data Market.

Project	Qualitative statements about how the project is contributing to Market share of the European Union in the global Big Data Market
<b>BigMedilytics</b>	<ul style="list-style-type: none"> <li>• We're targeting solutions in the Oncology, Population Health and Hospital Workflow Optimization domains.</li> <li>• New products exploiting large-scale healthcare data</li> </ul>
<b>Daphne</b>	system infrastructure as basis for potential startups and products in existing organizations
<b>Everest</b>	The solutions developed in the project are expected to create new business for the companies, especially due to a better exploitation of the underlying hardware technologies.
<b>More</b>	The successful completion of the project will increase the competitiveness of most stakeholders in the Wind and Solar energy industries. Moreover, it will create new jobs in the data management industry with specialized tools for the renewable energy industry.
<b>DataMarketServices</b>	<p>DMS organises plenty of networking events and meetings with investors to increase the number of clients and to get private funding for the startups. In addition, we also provide an extensive analysis of public funding opportunities for startups to help them increase their revenue. Last year, we organized the DMS Investors' Week, bringing together top EU investors and the best DMS startups. We continue to</p> <p>Furthermore, we provide participating startups with trainings that may directly</p>

	contribute to increasing their revenue, such as those focused on marketing and on sales strategies.
<b>EW-Shop</b>	All business products have created competitive advantages against competitors, which are often from US
<b>INFORE</b>	Project results will be exploited, following the exploitation plan.
<b>Infinitech</b>	Many partners of the Infinitech consortium are global companies (ATOS, IBM, FUJITSU, GFT... ) with branch offices all over the world. It is expected a revenue increase worldwide.
<b>Kraken</b>	SMEs are involved in the project.
<b>SmashHit</b>	<p>In the last years many data sharing platform approaches for different purposes (like B2B data exchange, data marketplaces or data processing services) emerged, giving cross-sectorial industries access to the great spectrum of sensor data coming from high volume products from various industrial sectors (vehicles, smart home devices, smart cities data etc.). With the increasing number of connected sensors and actuators within such mass products (so called cyber physical products), this number will rise exponentially in short-term. This enormous amount of data offered by various data sharing platforms represent:</p> <ul style="list-style-type: none"> <li>• a massive information resource to create new value, allowing the improvement of existing services or the establishment of diverse new innovative services, by combining data streams from various sources</li> <li>• a major big data-driven business potential, not only for the manufacturers of Cyber Physical Products (CPP), but in particular also for cross-sectorial industries and various organisations with interdisciplinary applications</li> </ul> <p>The majority of these data sharing platform approaches offer a scalable, secure and trusted sharing of data. However, they all lack in the provision of solutions for consent creation, management and observation between two or more involved actors that aim to share Personally Identifiable Information (PII) data. To fully exploit data sharing approaches in practice, they have to consider national and international laws like the European General Data Protection Regulation (GDPR). Among others, such regulations define that data owners can decide to whom and for which purposes their data can be used; they define how data providers have to handle data and they define how data have to be processed. A data sharing platform that consider such regulations needs an additional layer of consent creation linked to a contract which would guarantee the compliance with all requirements of the contracting parties and national and international laws. This gap needs to be filled and is where smashHit aims to bring an added value.</p>
<b>Cross-CPP</b>	<ul style="list-style-type: none"> <li>• Key motivation of Cross-CPP project is to give cross-sectorial industries access to the great spectrum of sensor data coming from high volume products from various industrial sectors (vehicles, smart home devices, etc.). With the increasing number of connected sensors and actuators within such mass products, this number will rise in short-term. This enormous amount of data continuously generated by mass products will represent: <ul style="list-style-type: none"> <li>o a NEW information resource to create new value, allowing the improvement of existing services or the establishment of diverse new cross-sectorial services, by combining data streams from various sources</li> <li>o a major big data-driven business potential, not only for the manufacturers of Cyber Physical Products (CPP), but in particular also for cross-sectorial industries and various organisations with interdisciplinary applications</li> </ul> </li> </ul> <p>The Cross-CPP marketplace will represent a disruptive innovation, enabling radical innovations by numerous cross sectorial Service providers. This would help to increase revenue of EU companies against total revenue of US, Japan, and Brazil.</p>
<b>Boost 4.0</b>	<p>One of the main objectives of the BOOST 4.0 project is to accelerate the adoption, by Industry 4.0, of data and advanced analytics platforms, which will have a direct input on the revenue share of EU companies.</p> <p>In general, the creation of the BOOST 4.0 connected smart Factory 4.0 model will</p>

	<p>help companies deploy big data solutions that will give them many competitive advantages, such as increase the operational efficiency, E2E manufacturing planning and deliver, improved smart product costumer experience, and also foster new business models, like outcome-based and product customization.</p> <p>A few examples of the specific solutions that will prove this impact and that are being deployed within the BOOST 4.0 pilots are:</p> <ul style="list-style-type: none"> <li>- Using Big Data analysis to simulate the manufacturing process and predicting machine breakdowns will improve the productivity and will reduce the energy consumption and the waste production.</li> <li>- Understanding better the manufacturing process will also help to optimize the general operation of the production line, increasing productivity.</li> <li>- Providing machines with IoT technologies and “smart” capabilities gives them a competitive advantages and creating new business opportunities.</li> <li>- By deploying better AGV (automated guided vehicles) technologies based on big data analysis will improve the internal logistics and the productivity.</li> <li>- Improving the information shared among the supply chain and monitoring the delivery network allow the prediction of delays.</li> </ul>
<b>DataPorts</b>	<p>Through the data-driven platform, DataPorts plan to contribute EU companies belonging to seaports’ ecosystems to increase their revenues by participating as Data or Service Providers. European seaports will have an advantage since new revenue streams will be created.</p>
<b>EUHubs4Data</b>	<p>One of main objectives of EUHubs4Data is to facilitate, mainly to the SMEs and start-ups, the access to cross border and cross sector data, as well as to big data services that support them in the development of new products and/or services or the improvement of existent ones. In addition to the creation of the federated catalogue and its promotion among these actors, the project will run a data driven innovation programme, which includes federated service provisioning at regional level and the launch of open calls for supporting experiments, where SMEs and start-ups will be able to use and validate, in collaboration with the DIHs, the cross border and sector data as well as the federated services. We expect to involve from 60 to 80 SMEs / start-ups / web entrepreneurs directly involved in the data driven innovation programme. We also expect for those companies to increase their revenues by +30% per year on average (KPI to be reviewed in view of the Covid crisis). We will support the exploitation of results coming from the data driven innovation programme, ensuring the most appropriate measures to speed up the access to market. Additionally, and as part of the evaluation plan in teh project, we will carry out a survey among engaged SMEs monitoring their baseline revenues before and after the programme.</p>
<b>ExtremeEarth</b>	<p>ExtremeEarth provides state-of-the-art technologies to manage and analyse Earth Observation data using Artificial Intelligence techniques. Equipping EU companies with an arsenal of such tools varying from cloud infrastructure and AI platforms, to Deep Learning models and linked data analysis, aims to increase the revenue share against other countries.</p>
<b>DeepHealth</b>	<p>DeepHealth strongly contributes to the development of a sustainable European AI for Health imaging value chain. Seven software platforms will benefit from DeepHealth as they will integrate the core technology libraries and, therefore, will take advantage of the DeepHealth toolkit to train models. The expected increase of productivity thanks to that will make the software platforms more competitive, so the industrial partners will increase their incomes. Specially, thanks to the fact that the great effort for improving industrial solutions falls on the use of the most recent advances in technology. By using the DeepHealth toolkit, companies offering end-user solutions do not have to spend too many efforts to maintain up to date their software platforms.</p> <p>Additionally, the DeepHealth toolkit can be integrated in any third party solution</p>

	and make it part of their business model to provide services based on it. DeepHealth will change the way we produce and use software, enabling new players and software-based services (SSBS) such as Big Data providers, to use the toolkit and support more precise health diagnostics, monitoring, research findings, etc.
<b>MediaFutures</b>	We are supporting startups in H2020 countries
<b>IoTwinS</b>	The application of the developed solutions in IoTwinS will allow easier access to AI and ML learning services and computational resources by companies and entities, as well as the offer of dedicated service by the partners involved, at first instance, in European countries. This will ultimately impact on revenue shares EU vs other countries.
<b>TT</b>	Advanced services developed and tested in TT will allow EU companies to commercialise cutting-edge services and products allowing to increase competitiveness against competitors and gaining market share. For instance, INDRA within the scope of Airports pilot, they are managing to replicate the outcomes achieved in the Retailing services business case trying to sell to other Airport Managers worldwide. The international presence of Indra could be a catalyst to speed up the market uptake of the solutions by other end-users. The same situation occurs with CINTRA and their international concessions of tolls. The positive impacts and achievements obtained in their operational efficiency measurements provides a competitive advantage to be replicate and implemented within other similar tolls around the world. These two cases (among others) states the contribution of TT to gain market share compared to companies and institutions of EU, US, Japan and Brazil.
<b>OpertusMundi</b>	The project will establish a startup company for the commercialization of topio.market.
<b>SmartDataLake</b>	The project develops Big Data analytics tools that can support EU companies, especially SMEs, in reducing time and effort and improving quality of their data analysis pipelines, thus generating more value from their data and becoming more competitive.
<b>SmartDataLake</b>	Project results will be exploited, following the exploitation plan.
<b>Data Pitch</b>	Data Pitch supported 47 European data driven companies to build innovative solutions that can be scaled internationally. These companies are solving real world challenges set by large corporations, creating competition for already established organisations within their respective markets.
<b>DataBench</b>	By using the results of DataBench and the tools made available companies can maximise their economic return from BDA benchmarking.

**Table 34:** Qualitative contribution of projects to the Market share of the European Union in the global Big Data Market

## KPI II.2: Growth of the number of Data Companies in the EU

Data Companies in terms of the European Data Market study are defined as organisations whose main activity is the production and delivery of digital data-related products, services, and technologies, including selling data. These companies represent the emerging data industry, that is the supply side of the data market.<sup>31</sup>

Regarding the number of Data Companies (KPI II.2), according to the most recent report<sup>32</sup> they increased to 290,000 by 2019, compared to 283,390 in 2018, with a growth rate of 2.3%. It should be noticed that, due to the high concentration of the ICT industry in the UK, almost half of them are based there.

The majority of Data Companies (285,750 in 2019, 98,5% of total) are SMEs under 250 employees, reflecting the structure of the European industry, but they accounted for 71.4% of revenues. The number of medium/large data companies reached 4,250 in 2019 reflecting a much faster growth than SMEs (+19.7%

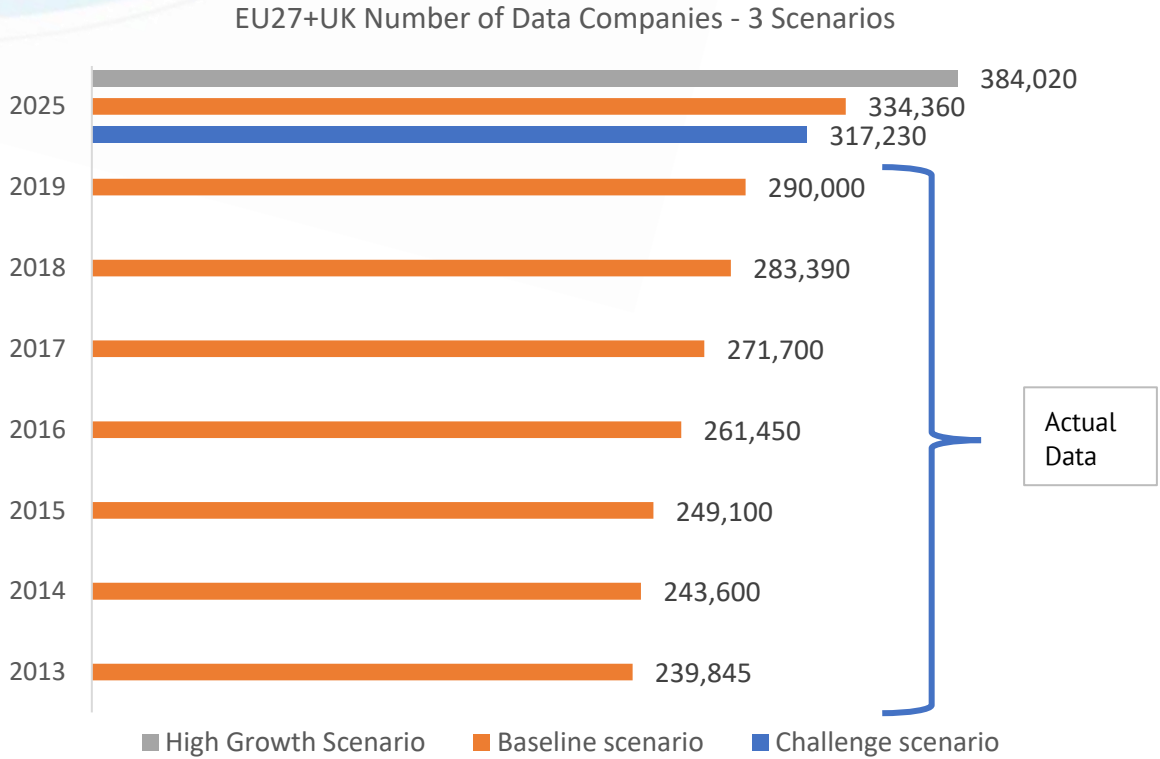
<sup>31</sup> The EDM monitoring tool, ibidem , chapter 2.3 pages 33-34

<sup>32</sup> The EDM monitoring tool, ibidem

over 2018 in number and +10.1% in revenues). The increase of the number of medium and large companies is very important for the competitiveness of the EU data industry.

In terms of future expectations, the report forecasts a total number of 334 360 Data Companies in the Baseline Scenario in 2025 representing a Compound Annual Growth Rate (CAGR) of 2.4% in the period 2020-2025. Under the favourable High Growth scenario data companies could increase to 377 720 (CAGR 4.8% on 2020); in the Challenge scenario the number could be lower at 317 230 (CAGR 1.5% on 2020). Figure 6 summarizes these numbers. The share of SMEs is expected to be a fraction lower in all three scenarios at roughly 98.4%, reflecting the structure of the EU industry, even though data companies of more than 250 employees are expected to grow twice as fast as smaller companies.

As anticipated above, these forecasts are likely to be revised slightly downwards at the next update of the study in October 2021, because of the effects of the economic crisis driven by the COVID-19 pandemic, even though a strong rebound growth is expected in the period 2022-2025.



**Figure 6:** Number of Data Companies in the EU27+UK, actual and forecast to 2025

**BDVA i-Spaces** and **Data Incubators** (ICT 14-b projects, so DataPitch and EDI) are in particular designed to contribute to this KPI as they support start-ups and entrepreneurs from early ideas to technical and business development till go-to-market<sup>33</sup>.

<sup>33</sup> Further information can be found in section 2.1 of this report.

85% of the BDV cPPP projects active in 2019-2020<sup>34</sup> reported contribution from their project to increase the number of European Companies offering data technology, applications. Projects contribute in different ways that can be summarized as follows:

- Creating tools that will stimulate the creation of new companies (e.g EW-Shopp, BigMedilytics or Cross-CPP)
- Creating new companies or enabling startups to be created as results of the project (e.g Daphne, Everest)
- Supporting EU data driven companies to each build innovative solutions that to solve data related challenges (data incubators such as DataPitch and EDI)
- Creating marketplaces and platforms for data analysis and data sharing (Smashit, AEGIS)
- Providing services on data protection and privacy (MosaiCrown, Kraken)
- Lowering the threshold to create new business in a particular sector:
  - More in the energy sector, providing sensor data from wind turbines and solar parks;
  - TT project some concrete targets have been set up for assessing how much SW and HW products has been produced in the project that might increase the product offering of the company and create new companies through spin-offs focused on the exploitation of new technology assets or services based on data. The answers provided in question 38 and 39 related on Large Scale Experimentations on Big Data demonstrate that End-users, Technology Providers, Industrial Partners and Research Centres has contributed to extend product systems capabilities by making use of data analytics (descriptive, predictive, etc.) and machine learning capabilities to enhance the current market positioning of the services portfolio of consortium partners.
  - Boost4.0 provides the EU industry with a model to take advantage of the big-data technologies, the project itself is a contribution to increase the number of European Companies offering data technology and applications.
  - Through DataPorts, European Companies related to seaports will have the opportunity to provide their data technologies and applications (also data and service sharing) and benefit from that offering.

Table 35 provides details per project to their **contribution to the development of the number of Data Companies in the EU.**

Project	Qualitative statements about the contribution to the Development of the number of Data Companies
<b>BigMedilytics</b>	<ul style="list-style-type: none"> <li>• It supports the involved companies and might aid additional companies serving as platforms for healthcare IT solutions.</li> <li>• Working with VTT to improve their algorithms</li> <li>• enabling My mhealth with big data and AI technologies</li> </ul>
<b>Daphne</b>	system infrastructure as basis for potential startups and products in existing organizations

<sup>34</sup> Based on number of respondents.

<b>Everest</b>	There is possibility to open new start-ups based on the technologies of the project.
<b>AEGIS</b>	We provided a platform enabling online data analyses and offering data analysis platform as a service
<b>More</b>	The project will create new tools for managing sensor data from wind turbines and solar parks. These tools will create a new industry that crossed the renewable energy and the data management markets. The project will also put emphasis on the exploitation of results by creating a stakeholder cluster that will help MORE tailor its solution to the market needs. These actions will provide the suitable environment for new companies to be created that will offer solutions to the renewable energy industry. ModelarDB, our partner is already once such company.
<b>DataMarketServices</b>	Our Project offers data-skills courses to the startups' employees. It helps them improve and gain data skills to offer data technology and applications.
<b>EW-Shop</b>	These services are products that consists on data technology: <ul style="list-style-type: none"> <li>- Campaign BOOSTER</li> <li>- PERFORMANCE INSIGHTS: E-commerce Insights tool for shoppers, sellers and vendors</li> <li>- Contact Center Workforce Optimization Manager</li> <li>- Weather and Event-aware BI Sales Strategy Advisor</li> <li>- Event and Weather Scout Dashboard</li> </ul>
<b>INFORE</b>	Project results will be exploited, following the exploitation plan.
<b>Infinitech</b>	Fintech companies composes the Consortium and have plans to expand their services which will produce a cascading effect on the value chain.
<b>BD4OPEM</b>	Within the project, some SMEs and startups are contributing to increase the number of European Companies offering data technology
<b>Kraken</b>	European companies could leverage the assets KRAKEN is developing for assuring security and privacy of personal data at the end of the project.
<b>MosaiCrown</b>	Providing effective data protection solutions, MOSAICrOWN can encourage institutions and entrepreneurs to enter the data sharing market by enabling them to become trustworthy companies.
<b>SmasHit</b>	The innovative inter-data-platform framework solution will have several positive effects on the European Data Economy, as e.g.: <ul style="list-style-type: none"> <li>• To overcome existing barriers, as e.g. the complicated and time-consuming consent/contract processes</li> <li>• Increase trust for all stakeholders in the data market, due to strong data traceability mechanism</li> <li>• Increase willingness of Data Providers (OEM) to provide CPP data for 3rd party use</li> <li>• Open new opportunities for value-creation by creation of innovative services using data from multiple platforms</li> </ul>
<b>Cross-CPP</b>	As mentioned before the Cross CPP project offers cross domain CPP data in a uniform format via big data marketplace. This marketplace is equipped with big data analytics toolbox thus, offering European companies to provide new innovative services.
<b>Synergy</b>	The project includes an RFP open call process to select SMEs and startups that will develop 8 applications to run in the SYNERGY big data platform.
<b>INFORE</b>	The INFORE project makes it simple for data analysts to execute complex data processing workflows, as the only thing required is their graphical design.

<b>Boost 4.0</b>	As BOOST 4.0 aims to provide the EU industry with a model to take advantage of the big-data technologies, the project itself is a contribution to increase the number of European Companies offering data technology and applications.
<b>DataPorts</b>	Through DataPorts, European Companies related to seaports will have the opportunity to provide their data technologies and applications (also data and service sharing) and benefit from that offering.
<b>EUHubs4Data</b>	One of main objectives of EUHubs4Data is to facilitate, mainly to the SMEs and start-ups, the access to cross border and cross sector data, as well as to big data services that support them in the development of new products and/or services or the improvement of existent ones. In addition to the creation of the federated catalogue and its promotion among these actors, the project will run a data driven innovation programme, which includes federated service provisioning at regional level and the launch of open calls for supporting experiments, where SMEs and start-ups will be able to use and validate, in collaboration with the DIHs, the cross border and sector data as well as the federated services.
<b>ExtremeEarth</b>	<p>The Extreme Earth technologies and development have a strong scientific and basic research background and will serve the further development and increased application of the EO, big data and linked data technologies. We provide Deep Learning models, training datasets and a semantic catalogue over EO data for the Polar Regions and Food Security domains. These technologies can be adopted by European Companies to develop new applications in the respective domains.</p>
<b>DeepHealth</b>	As commented in the question before, DeepHealth main outcome, the DeepHealth toolkit can be easily used to develop and deploy new applications in the medical imaging domain, i.e. it provides a tool for innovators to create new applications. Most of AI-based medical imaging innovations come from SMEs and start-ups. DeepHealth will highly impact on supporting European SMEs and start-ups constrained by their market accessibility potential to tools as well as difficulties to access leading AI R&D knowledge.
<b>MediaFutures</b>	We are supporting startups in H2020 countries
<b>IoTwinns</b>	The application of the developed solutions in IoTwinns will allow easier access to AI and ML learning services and computational resources by companies and entities, as well as the offer of dedicated service by the partners involved. It will thus allow to increase the number of European companies offering data-based technologies, applications and services.
<b>EDI</b>	We have incubated over 100 solutions produced by data-driven companies across Europe
<b>CloudButton</b>	Some of the partners are considering to create spin-off companies to leverage the results of the CloudButton project.
<b>PimCity</b>	The project aims at developing technologies that can be used by companies to enter data markets. As such we foresee new opportunities for European companies to provide new solutions and services.
<b>Reach</b>	30 data-driven startups will reach the experimentation oriented phase, before which they will conduct experiments with their data providers.
<b>TT</b>	TT has set up concrete targets for assessing how much SW and HW products has been produced in the project that might increase the product offering of the company and create new companies through spin-offs focused on the exploitation of new technology assets or services based on data. The answers provided in question 38 and 39 related on Large Scale Experimentations on Big Data demonstrate that End-users, Technology Providers, Industrial Partners and Research Centres has contributed to extend product systems capabilities by making use of data analytics (descriptive, predictive, etc.) and machine learning capabilities to enhance the current market positioning of the services portfolio of consortium partners. By M18 mid-term review, it has been reported a total of 4 new services produced with BD capabilities and 16 services which have been improved by integrating BD capabilities.
<b>OpertusMundi</b>	The project will establish a startup company for the commercialization of topio.market.

<b>SmartDataLake</b>	Most tools developed by SmartDataLake are made publicly available as open source software, making it easier for EU companies, especially SMEs, to develop and offer services for data analytics over data lakes.
<b>SmartDataLake</b>	Project results will be exploited, following the exploitation plan.
<b>Data Pitch</b>	Data Pitch supported early stage, data technology startups.
<b>QROWD</b>	One spin-off was created out of a subset of the innovations in the project.
<b>DataBench</b>	BDA benchmarking tools help companies prove their applications.

**Table 35:** Qualitative contribution of projects to the Development of the number of Data Companies in the EU

In addition, **27% of the BDVA members** reported their organisation run or support a programme that is specifically targeted at **supporting start-ups or entrepreneurs in the Big Data area**. Some examples include:

- TNO and the NL AI Coalition is a national programme that aims to support companies including startups
- Know-Center, who is partner of several incubation programmes and 2 national funded DIHs as well as in the EU-funded project EUH4D
- ITI Data Space has strong agreements with the main incubators in the region, like LANZADERA (MERCADONA) or Go Hub (GLOBAL OMNIUM). ITI Data Space is a reference partner for those incubators and we help start-ups to use and take profit of Data Space services for testing before investing. The relationship is as a service provider.

ITI Data Space is working with several start-ups and SMEs focused on Data Analysis. Those start-ups trust ITI Data Space and use it for accelerating the development of applications and the accuracy of the results. This, helps them to reach the market in better conditions, reducing the level of mortality of companies. We also are working with not so small and young companies but with consolidated companies that want to accelerate a business Unit in Data Analysis and Artificial Intelligence. We are providing best practices, training, feasibility analysis, proof of concept, validation of products and services, business models.

- IBM Research with a start up accelerator program in Israel
- UPM runs a programme to support technological start-ups, including Big Data area, providing several services such as coaching in entrepreneurship, empowering favourable environment to establish strategic alliances with the UPM and that combine its own activity with the carrying out of joint R&D activities with the UPM and with other entities of the ecosystem, technology acceleration through activities that facilitate the integration of technologies, development of spin-offs generated by the UPM or by other entities in partnership with the UPM Incubator more quickly

### KPI II.3: Growth of Data Companies revenues in the EU

The revenues of Data Companies in the European Union (KPI II.3), according to the IDC report<sup>35</sup>, reached 84 €B in 2019 compared to 77 €B in the year before, with a growth of 8.1%.

The value of revenues by SMEs in 2019 amounts to €59.6 €B (71.4% of the total revenues), an absolute growth of 4 €B on the year before. The growth rate of revenues increases in proportion to company size, with the revenues of large companies over 500 employees growing at 13.4% in 2019 over 2018.

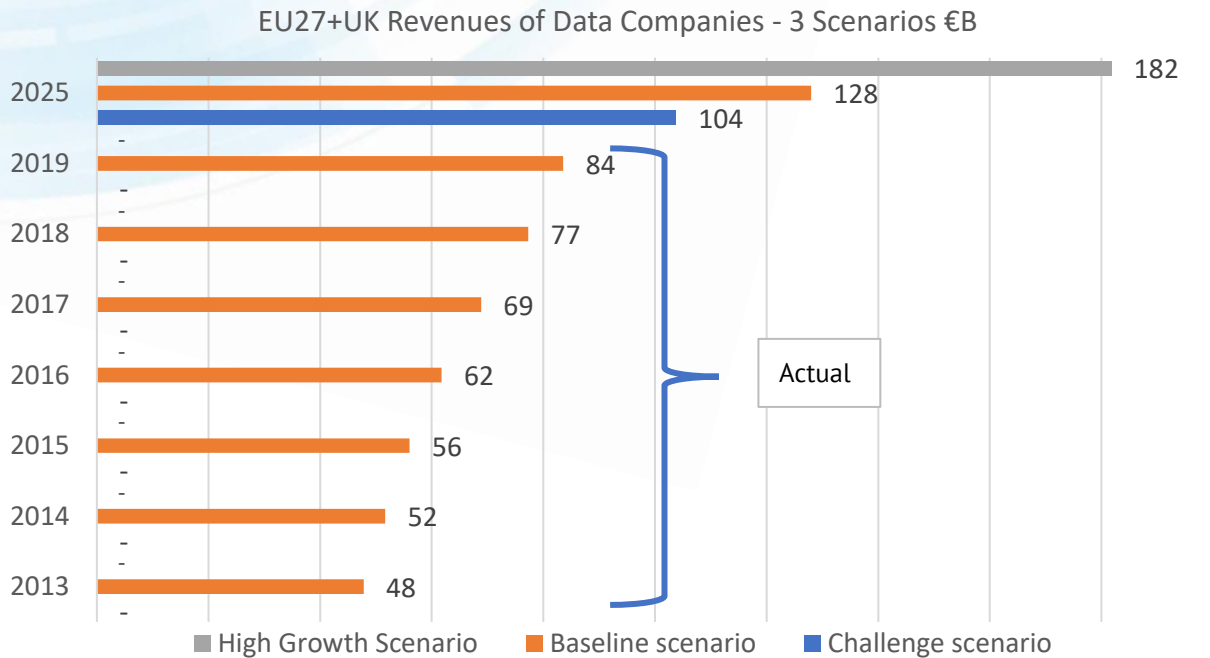
With respect to future growth IDC projects an increase in total revenues for the Baseline Scenario to 128 €B in 2025 representing a Compound Annual Growth Rate 2025/2020 of 7.4%. For the 2025 High Growth scenario the revenues could instead go up to 182 €B representing a CAGR on 2020 of 13.8%; in the Challenge scenario the revenues would be lower, reaching €104 bn, representing a CAGR 2025/2020 of 3.7%. The revenue share of SMEs remains quite stable at about 63% for all three scenarios with the High Growth Scenario resulting in the lowest share (63,5%)<sup>36</sup> (because of faster growth of large companies).

<sup>35</sup> The EDM monitoring tool, ibidem, chapter 2.3 pages 36-37

<sup>36</sup> The EDM monitoring tool, ibidem

**Figure 7** depicts the values from 2013 to 2019 and the projected three growth scenarios as discussed in this paragraph.

As anticipated above, these forecasts are likely to be revised slightly downwards at the next update of the study in October 2021, because of the effects of the economic crisis driven by the COVID-19 pandemic, even though a strong rebound growth is expected in the period 2022-2025.



**Figure 7:** Revenues of Data Companies in the European Union in billion EUR

**67,5% of the cPPP projects** active in 2019-2020<sup>37</sup> reported contribution (or plan to contribute) to the revenue generated by European Data Companies. The projects contribution to this KPIs is mainly by:

- **Providing data solutions to hospitals** for Oncology, Population Health and Hospital Workflow solutions (**BigMedilytics**)
- allow EU data suppliers to trade their geospatial assets across EU in a single, homogeneous, and fair digital marketplace (**Opertus Mundi**)
- **Boost4.0** in Manufacturing helping the industry to exploit for its benefit the huge amount of data that is already generating or that can be generated, to improve their revenues and reduce their expenses;
- **Offering direct support and getting funding for data start-ups to increase revenues.**
  - An independent impact assessment of Data Pitch forecast that annual total revenue generated by Data Pitch supported startups would grow by 73% per year until 2022. See

<sup>37</sup> Based on number of respondents.

full impact assessment here: <https://datapitch.eu/wp-content/uploads/2020/01/London-Economics-Data-Pitch-evaluation-FINAL-PDF.pdf>

- In EDI some of the companies who participated in the incubator have already reported important profits.
- Make **data processing easier and cheaper for companies** (e.g EW-Shopp for small e-commerce and retails companies) or **offering technologies or solutions that will make easier for European companies to do more big data analysis thus increasing revenue** e.g BigMedilytics in healthcare,
- **Through the commercialisation of new services with marketable value** to be developed, demonstrated and later exploited by the project partners (e.g TT in the transport, logistics and mobility market, DataBio in the bioeconomy sector)
- **Creating opportunities for common exploitation** based on joint Big Data technology pipelines (e.g DataBio in the bioeconomy) or developing **simplicity in some business ecosystems** (e.g Lynx will allow law firms, standardization bodies and companies to more efficiently deal with compliance issues).

Table 36 provides additional details on how the project contribute to the development of the revenue of Data Companies in the EU.

Project	Qualitative statements about how your project is contributing to the development of the revenue of Data Companies in the EU
<b>BigMedilytics</b>	<ul style="list-style-type: none"> <li>• Providing data solutions to hospitals for Oncology, Population Health and Hospital Workflow solutions.</li> <li>• It evaluates, establishes and helps deploying new products.</li> <li>• By providing the MPC solution, it will be easier for european companies to do more big data analysis thus increasing revenue.</li> </ul>
<b>Daphne</b>	system infrastructure as basis for potential startups and products in existing organizations
<b>Everest</b>	Thanks to the job profiles and solutions, the data companies can increase their services.
<b>More</b>	The project will create new tools for managing sensor data from wind turbines and solar parks. These tools will create a new industry that crossed the renewable energy and the data management markets. The project will also put emphasis on the exploitation of results by creating a stakeholder cluster that will help MORE tailor its solution to the market needs. These will allow MORE's outcomes to be easily exploited by the industry to increase its competitiveness and will be also contribute to the creation of new companies.
<b>DataMarketServices</b>	Although it is difficult to quantify, we have introduced the startups to investors, and we have helped them acquire new customers, thanks to our training and mentoring sessions in acceleration. In addition, we have delivered a customized funding opportunities report to the startups, to encourage them to apply for additional funding sources.
<b>EW-Shop</b>	The five business cases have been developed by European data companies.
<b>INFORE</b>	Project results will be exploited, following the exploitation plan.
<b>Infinitech</b>	It is possible and plausible that the new services created will also have cascading effects on data companies and related businesses.
<b>Kraken</b>	The plan will be included in the KRAKEN exploitation plan.
<b>MosaiCrown</b>	The technologies and solutions developed within MOSAICrOWN facilitate the portability of data and create new market opportunities, helping the volume of business in EU to increase economic benefits

<b>SmasHit</b>	<p>The innovative inter-data-platform framework solution will have several positive effects on the European Data Economy, as e.g.:</p> <ul style="list-style-type: none"> <li>• To overcome existing barriers, as e.g. the complicated and time-consuming consent/contract processes</li> <li>• Increase trust for all stakeholders in the data market, due to strong data traceability mechanism</li> <li>• Increase willingness of Data Providers (OEM) to provide CPP data for 3rd party use</li> <li>• Open new opportunities for value-creation by creation of innovative services using data from multiple platforms</li> </ul>
<b>Cross-CPP</b>	This project aims to provide cross domain CPP data to the service providers, they can be Data Companies providing different services to the end customer. The project also provides a big data analytics toolbox which can be helpful especially for SMEs who do not have big data expertise.
<b>Boost 4.0</b>	As answered in question 44, Boost 4.0 is about helping the industry to exploit for its benefit the huge amount of data that is already generating or that can be generated, to improve their revenues and reduce their expenses. That means more directly more revenue for the tech providers.
<b>DataPorts</b>	European Data Companies will be able to offer their data services through DataPorts' data-driven platform.
<b>Safe-Deed</b>	due to successful exploitation of the developed services in Safe-DEED by the consortium partners
<b>MediaFutures</b>	We are supporting startups in H2020 countries
<b>EDI</b>	Some of the companies who participated in the incubator have already reported important profits.
<b>TT</b>	Through the commercialisation of new services with marketable value to be developed, demonstrated and later exploited by the project partners. Indeed, TT has been the opportunity for consortium partners to explore and investigate which impacts occasioned the development and integration of BD capabilities in their service portfolio. The experimentations performance in each pilot involving real end-users which foster the trust and reliability of this solutions when end-users are planning to extend their services or products integration BD capabilities.
<b>OpertusMundi</b>	The project will allow EU data suppliers to trade their geospatial assets across EU in a single, homogeneous, and fair digital marketplace.
<b>SmartDataLake</b>	The project develops Big Data analytics tools that can support EU companies, especially SMEs, in reducing time and effort and improving quality of their data analysis pipelines, thus generating more value from their data and becoming more competitive.
<b>SmartDataLake</b>	Project results will be exploited, following the exploitation plan.
<b>Lynx</b>	By implementing Lynx services into existing products and services of partners.
<b>Data Pitch</b>	An independent impact assessment of Data Pitch forecast that annual total revenue generated by Data Pitch supported startups would grow by 73% per year until 2022. See full impact assessment here: <a href="https://datapitch.eu/wp-content/uploads/2020/01/London-Economics-Data-Pitch-evaluation-FINAL-PDF.pdf">https://datapitch.eu/wp-content/uploads/2020/01/London-Economics-Data-Pitch-evaluation-FINAL-PDF.pdf</a>
<b>QROWD</b>	Our results help European Data companies to integrate the different forms of Crowdsourcing into their data analytics processes.
<b>DataBench</b>	By using the results of DataBench and the tools made available companies can maximise their economic return from BDA benchmarking.

Table 36: Qualitative contribution of projects to the development of the revenue of Data Companies in the EU

## KPI II.5: Increase of Data Professionals in the European Union

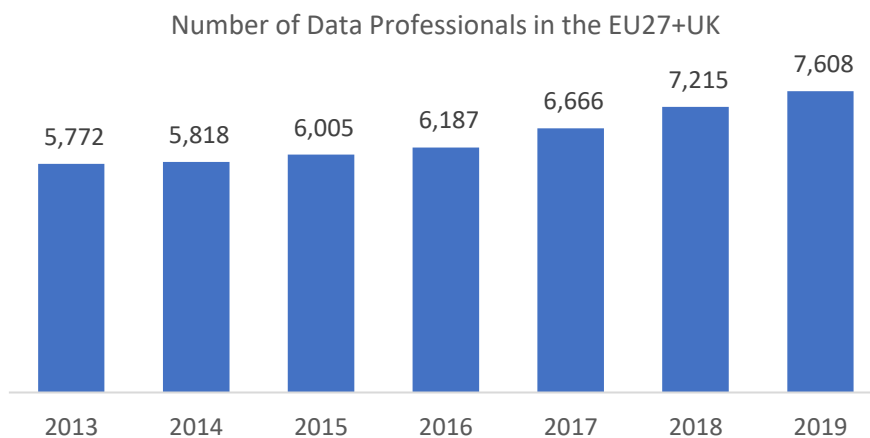
Data Professionals according to the EDM Monitoring Tool collect, store, manage, and/or analyse, interpret and visualise data as their primary or as a relevant part of their job activities: they should be able to work with massive databases and with emerging data base technology, as well as with structured and unstructured data, to elaborate and visualise data in order to use them for the interpretation of facts and as factors for decision-making processes.<sup>38</sup>

The baseline for Data Professionals in the EU27+UK in 2013 (KPI II.5) amounts to 5,77 million. The number of Data Professionals increased to a total of 7.6 million by 2019, resulting in a total absolute growth of 1,836 thousand professionals since 2013. Over 394 thousand positions were added in 2019 corresponding to 5.5% growth on the year before.<sup>39</sup>

According to the Baseline scenario, Data Professionals could become 11.3 million by 2025, representing an absolute growth of 3.7 million additional jobs from 2019, or 5.5 million from 2013. The compound annual growth rate (CAGR) between 2013 and 2025 in the baseline scenario is 5.8%. Based on the baseline scenario, the number of data professionals is expected to nearly double between 2013 and 2025.

Depending on framework conditions and the dynamics of the data market, the number of data professionals could reach 13,1 million under the High Growth or be lower (but still relevant) under the Challenge scenario at 10,2 million.

As anticipated above, these 2025 forecasts are likely to be revised slightly downwards at the next update of the study in October 2021, because of the effects of the economic crisis driven by the COVID-19 pandemic, even though a strong rebound growth is expected in the period 2022-2025.



**Figure 8:** Data Professionals in EU27+UK

80% of the cPPP projects active in 2018<sup>40</sup> reported contribution from their project to increase the number of Data Workers in Europe. Projects contribute to this KPI in different ways:

- **New organisations created as result of the project hiring new data professionals** to exploit project results (e.g Lynx, BigMedilytics) or new jobs for data workers created by the project (e.g PimCity)
- **Supporting emerging startups to grow** (e.g companies in DataPitch incubator created more than 70 data related jobs in their companies)

<sup>38</sup> The EDM monitoring tool, ibidem, chapter 2.2 pages 29-32

<sup>39</sup> The EDM monitoring tool, ibidem, chapter 2.2 pages 29-32

<sup>40</sup> Based on number of respondents.

- **More data-driven services** that will **require new data-workers** (e.g EW-Shopp, BigMedlitytics, Boot4.0, Lynx) or unlocking value of data services
- **Creation of new job profiles** e.g TT project, EDI incubator, Typhon project
- **Adoption of data solutions** in public sector and users in different sectors will increase the amount of data workers (e.g QROWD, BigMedilytics, TT, Boost4.0, Synergy)
- **Education and training** (e.g Examode, EDI incubator)

Table 37 provides additional details on how the project contribute to increase the number of Data workers in the EU.

Project	Qualitative statements about how they project is contributing to increase the number of data workers in Europe (referring to 2018 activities)
<b>BigMedilytics</b>	<ul style="list-style-type: none"> <li>• The following activities developed during 2019 &amp; 2020 would contribute to increase the number of data workers in Europe: Development of AI algorithms, Cluster installation, Anonymization process, Secure channel set up, Data curation and treatment</li> <li>• Currently in the health domain, there is many potential data to be analysed. Our pilot shows relevant results using such data, fact that encourages the employment of data workers.</li> <li>• All new employees will be data workers, since they will be involved in the curation, analysis, and management of complex medical imaging data and connected information</li> <li>• Education and training</li> <li>• New positions in the different knowledge-driven tasks required to create and make knowledge graphs available are required.</li> <li>• The Lung-cancer pilot application we are developing requires data analytics, machine learning, knowledge engineering, and software development skills</li> <li>• To have validated and integrated hospital data for projects like Breast cancer pilot new data preparation specialists are needed</li> <li>• there is a rapid increase of data scientist job recruitments in Europe</li> <li>• We discussed with key stakeholders the rules for the use of big data technologies in health care and summarized the results in infographics. This work supports the work of data workers in Europa.</li> </ul>
<b>Everest</b>	EVEREST plans to increase the number of data workers as it will open more possibilities to integrate heterogeneous data sources in complex applications with semi-automatic implementation onto the computational resources. Such data workers will be centred on data assimilation, data reorganization, and application development.
<b>More</b>	We expect that the project will help create more jobs, equal to 1% of the wind and solar industries as described in the previous question. Nevertheless, MORE was active only for 3 months in 2020 so this estimation is based on our expectance of succesful completion and not on the results on 2019 and 2020.
<b>DataMarketServices</b>	DMS fulfilled its expected impact of delivering training on standards and legal issues related to data. Programme monitoring activities showed that engagement with Data Skills courses had doubled in the span of two years.

	<ul style="list-style-type: none"> <li>• Fundamentals of AI for Business</li> <li>• Fundamentals of Data Science</li> <li>• Risk Analysis, policy compliance and elements of GDPR</li> </ul> <p>Services provided by: UNIVERSITY OF SOUTHAMPTON   KINGS COLLEGE LONDON</p> <p>These courses help the companies gain data skills, thus contributing to increase the number of data workers in Europe. Also, among the reported job creation by the startups, most of the hired profiles are data workers.</p> <p>These services are addressed to any companies looking to increase their staff skills, companies that want to create a new data- related department and companies looking to offer benefits to their employees.</p>
<b>INFORE</b>	Increase in the number of data workers mainly refers to the analysis of stock- and commodity markets.
<b>Infinitech</b>	Project is about DATA Management and AI therefore
<b>BD4OPEM</b>	As a project is related to Big Data, it needs data professionals to develop the activities performed within the project.
<b>Kraken</b>	<p>Personnel working in the project will acquire/improve skills and expertise in a set of technologies related to data-sharing platforms:</p> <ul style="list-style-type: none"> <li>- Development of tools aimed to integrate SSI solution in data sharing business flow compatible with different specific blockchain implementations for identity management</li> <li>- Design of data protection techniques based on advanced crypto tools (P/F/HE, FE, MPC, ABE...) coupled with privacy preserving (AI/ML) analytics</li> <li>- Data models design</li> </ul>
<b>MosaiCrown</b>	MOSAICrOWN enables data sharing and collaborative analytics in multi-owner scenarios in a privacy-preserving way, ensuring proper protection of private/sensitive/confidential information. The ability to have free movement and sharing of data helps the EU to increase the size of its data workforce, which is key to driving the data economy.
<b>SmasHit</b>	As mentioned above the project will develop an innovative inter-data-platform framework solution will have several positive effects on the European Data Economy.
<b>Cross-CPP</b>	<p>As mentioned above the project deals with cross domain CPP data and this data in nature is big data. Thus, such data require data workers who can not only analyse but also manage it in cross industrial data model.</p> <p>The most important project outcome should be the third party services based on CPP data. These services can be created in cross domains with different objectives in mind, such as the CPP data can be used for weather forecast and at the same time can be used for energy management in buildings etc. Thus, new service providers can take advantage of vast amount of data gathered from cross domain CPPs and can enhance their services. For the analytics toolbox, specialists developing further analytics tools are needed.</p>
<b>Examode</b>	The data processed in the ExaMode project need to be annotated by medical experts. Some medical experts were trained to use developed project tools for Whole Slide Image annotation, and for main processes for data management - like data cleaning, data preparation, data transformation, etc.
<b>Synergy</b>	On top of the previous comment, SYNERGY is establishing a data-driven mentality, long-term analytics skills, business models and procedures within

	the whole European energy value chain, that will have an impact in the increase of the number of data workers in Europe.
<b>Boost 4.0</b>	<p>In 2019 and 2020 the partners of the Boost 4.0 Consortium have created around 150 new jobs, most of which are data professionals (data scientists, data analysts, data modeling expert, data engineer, etc.). But the expectations until 2024 are even higher (as mentioned in the previous question, around 600 additional jobs are estimated).</p> <p>Beyond the project Consortium, the project has dedicated a lot of effort to contribute to increase the number of data workers by promoting through all the means that could be reached the benefits of Big Data for industry, showing the main results of the project.</p>
<b>DataPorts</b>	<p>As part of DataPorts work it is expected to develop an Automatic Predictions Engine component. The main goal of this component is to automatically generate a machine learning model using the data available in the DataPorts platform for training. Port stakeholders could select the specific Key Performance Indicator (KPI) to predict. Time Series Forecasting techniques analysis have been carried out as they address the most common prediction problems in the Ports domain.</p> <p>Furthermore, a process-based analytics component is currently being implemented to facilitate proactive management of port processes, it includes the following components: Ensemble Predictive Process Monitoring, Prescriptive Process Monitoring, and Explainable Predictive Process Monitoring.</p> <p>All these activities lead to solutions that need the involvement of data professionals in Europe.</p>
<b>ExtremeEarth</b>	The researchers employed in the ExtremeEarth project (including Ph.D. students and postdocs) will be data professionals with a lot of importance experience by the end of the project.
<b>DeepHealth</b>	As explained in the question above, DeepHealth will support the creation of new data-related jobs and contribute to increase the number of data workers, providing services and support to DeepHealth resulting products and developments.
<b>Safe-Deed</b>	Due to the integration and education of Master- and PhD Students within the project.
<b>IoTwinns</b>	<p>One of the observations that were made during the first phase of the project (e.g., during the collection of requirements from test-beds) is the fact that many SMEs currently lack data analysts and experts in the areas of big data and artificial intelligence.</p> <p>The IoTwinns project directly address these issues, as it aims to fill the gap between SMEs and data analysis. The first step consists in defining the architecture to collect the data from the IoT level to the cloud and providing support for the deployment of such infrastructure on the SMEs and data centres premises. Then, a set of diverse data analytics and artificial intelligence services have been developed, together with a comprehensive documentation to support the needs of the test-beds (e.g., predictive maintenance).</p> <p>However, SMEs and data centres will need data analysts beyond the duration of the IoTwinns project, in order to create useful “knowledge” out of the collected data.</p>
<b>EDI</b>	We have definitely helped companies to incorporate new data scientists in their companies. Besides, not only technical people around data solutions have been incorporated. We have also incorporated business people who have to put in place data-centred business models.
<b>CloudButton</b>	<ol style="list-style-type: none"> <li>1. Python is the largest programming community in the world and one of the most used in data science.</li> <li>2. Any existing Python data scientist has the potential of leveraging the huge</li> </ol>

	resources of the cloud (compute, storage) thanks to Lithops, a toolkit developed in the context of the CloudButton project.  3. CloudButton democratizes Big Data in the cloud running unmodified Python code with existing Python notebook technologies.
<b>PimCity</b>	Partners in the project have hired personnel that is currently working on data protection and privacy preserving solutions. Once the project is over, these people will result in data experts, with a deep understanding of the problems and technical solutions in such space.
<b>Reach</b>	The project has only started in September 2020.  When it comes to next 4 years, we will engage over 100 startups, 15 corporations, over 5 DIHs around data processing. Assume that minimum of 1 person per organization will actively engage around Data, you can say that we will contribute to the increase of 120 data focused people
<b>TT</b>	Qualitative: Thanks to TT data workers will gain relevance in the transport sector increasing the number and types of data-science related profiles working for  Quantitative: currently, TT has not direct metric that allow us to assess the number of data workers produced. Nonetheless, TT has planned to include this metric among the project KPIs in the second Period. The results achieved will directly come from the answers provided by partners in dedicated questionnaires that TT is sharing to feed the calculations of Project KPI levels.
<b>SmartDataLake</b>	Increase in the number of data workers mainly refers to the analysis of (agricultural) stock- and commodity markets.
<b>Lynx</b>	Its two-fold, on the one hand there are people working on the data and the legal knowledge graph and on the other hand as more and more organisations will integrate Lynx services into their own applications.
<b>Data Pitch</b>	In 2019, the second cohort of Data Pitch startups reported creating 70 new jobs.
<b>QROWD</b>	We provide a framework for collection and analysis of data remotely by crowdworkers (that then become data workers in full right), together with an analysis of the ethical aspects of such models.

Table 37: Qualitative contribution of projects to the increase of Data Workers in the EU

## 11 CONTRIBUTIONS TO ENVIRONMENTAL CHALLENGES (II.6, II.19, II.20, II.21)

### 11.1 Contribution to the reduction of energy use (KPI II.6)

This KPI intends to measure energy saved in big data analytics using a specified benchmark through solutions provided by specific cPPP projects compared to baseline at beginning of H2020.

Projects were asked: 1) whether they were monitoring Energy Efficiency in their project; 2) what metrics were they using to measure this parameter (if monitoring); and 3) which improvement (in percent) have they measured in their project.

5 projects (DataBench, TT, IoTwins, Everest and Daphne) reported they are monitoring Energy Efficiency in their projects. Only TT project provides specific results in terms of % of energy reduction, and it will be reported in II.19.

Projects provided the following input regarding the parameters used to measure energy efficiency:

**Daphne:** energy per operation workload + data transferred

**Everest:** Energy consumption. The project just started so we didn't measure any improvement yet.

**IoTwins:** Annual Energy Production, cost of energy, Active Energy Delivered, Active Energy Incoming, Apparent Power, Power Factor, Reactive Power, Voltage, Effective Power

**TT:** There is a specific "Assessment Category" to evaluate Energy Consumption (EC) which measures the energy required by the pilot's activities in both test scenarios (before and after Big Data technologies had been implemented), therefore potential savings signify improvements in efficiency.

The metrics used in TT are implicit in the units in which the KPIs are calculated. The indicators are adapted to assess the specific activities performed by the pilots. In addition, it is important to avoid external factors, such as seasonality or macroeconomic issues that may influence the volume of the activity. In road transport, energy is measured as kWh/100 km, while for airports energy is referred to daily values (kWh/day) and also applicable at passenger level.

Also, the scale to which the energy monitoring is applied varies depending on the specific pilot. Since the monitoring of the energy consumption in the global transportation system (eg. Airport energy consumption) would blur the impact achieved, the monitoring unit is scaled to the specific pilot objectives (eg. Energy consumption in the airport security area).

**DataBench:** one of the benchmarking communities involved does monitoring

## 11.2 Contribution to Energy use reduction (KPI II.19)

17,5% of the projects running in 2019-2020 have reported they contribute to the reduction of energy use (DataBench, TT, IoTwins, Boost 4.0, Synergy, Everest, Daphne). Synergy provides several measurements planned for the coming years of the project (see table 32).

**Table 38** provides details of the contribution from cPPP projects to energy use reduction.

Project	Qualitative and quantitative statements (2018 and planned)
<b>Daphne</b>	System infrastructure for integrated data analysis pipelines for improved performance, resource utilization, and energy efficiency.
<b>Everest</b>	FPGA acceleration is known to provide significant benefits in terms of energy use
<b>Synergy</b>	2019 and 2020: N/A Planned: - Energy Savings for consumers: 30-35% - Energy cost savings for energy consumers: 40% - Self-consumption achievement in local energy communities and buildings levels: 50-80% - Reduction of Levelized Cost of Energy for RES Operators: 30% on average - Reduction of O&M Costs for Network Operators, RES Operators and Facility Managers: 30% on average - Reduction of Network Losses: 22% - Network and Assets Down-time Reduction: 40-50% - Value of Lost Load gains: €26/kWh
<b>Boost 4.0</b>	Significant optimization of the manufacturing processes have been achieved in most of the lighthouse pilots that have been implemented in the Project. This optimization implies in most cases a reduction of the energy use, although it has not been monitored, so quantitative statements cannot be provided.
<b>IoTwins</b>	Testbed #2: Predicting and optimizing the spindle behavior leads to a reduction in cycle-time – the components can be processed in shorter time with equal or even better quality. The energy use per component and the energy use of postprocessing – due to better quality – decrease

	<p>Testbed #5: The uses cases that are being validated in this testbed aim to analyze the crowd flows and better understand how attendees interact with our sports facilities. When we make an attendance prediction and profile the expected behavior of the different archetyped attendees we get a powerful tool to adjust our provisioning to the expected demand for merchandising, food &amp; beverages, energy, etc</p> <p>Tested #7: The testbed has the goal to implement the IoTwins architecture in order to enable monitoring and control of low voltage grids. In the long run this will enable the energy grid operators to add functionalities like “energy communities” or “virtual power plants” even in legacy grids. This in turn will reduce the dependency on centrally created power based on fossil energy sources and increase the awareness about the energy consumption at the end user.</p>
TT	<p>As it has been mentioned, there is a specific Assessment Category to evaluate Energy Consumption (EC) which measures the energy required by the different pilots.</p> <p>All pilots have reported their final impact values by June of 2019. Results show that for the specific monitored items improvements in energy use efficiency: Highways achieved 33% of improvement, Connected Vehicles around 6%, Airports 28%, Urban Mobility 55%, whereas E-Commerce improved its efficiency by 46%.</p> <p>In general, percentage of KPI improvements have revealed a growing tendency throughout the different reporting periods within the project, overtaking most of the predicted target values for energy efficiency, meaning that the initially planned targets have been satisfyingly addressed.</p>
DataBench	provides a way to measure efficacy of Benchmarking techniques for various sectors including energy

**Table 38:** Contribution from cPPP projects to energy use reduction

### 11.3 Contribution to the reduction of CO<sub>2</sub> (KPI II.19)

27,5% of the projects replying to the questionnaire have reported contribution to the reduction of CO<sub>2</sub> and in particular BigMedilytics, Daphne, Everest, More, BD4OPEM, Synergy, Boost 4.0, TT, Lynx, QROWD, DataBench.

It is important to highlight that mostly all the projects contributing to this KPI are **linked to the transport and mobility sector or the energy sector** (maritime, aviation, terrestrial transport and mobility in general). Even projects focused on other sectors such as healthcare reductions of CO<sub>2</sub> relate to the reduction of needs or mobility.

Only one project, TT, provides evidence in measurements in CO<sub>2</sub> reductions reporting Improvements concerning CO<sub>2</sub> and other emissions such as NOX and PM, savings reach up to 37%.

Table 39 provides details per project in relation to their contribution to the reduction of CO<sub>2</sub>.

Project	Qualitative and quantitative statements (2019-2020 and planned)
BigMedilytics	Some pilots involve remote monitoring of patients. So they need to visit hospitals less often. That results in less travel and less usage of resources at the hospital.
Everest	FPGA-based data centers will require less energy and so CO <sub>2</sub> emission will be reduced.

<b>More</b>	None yet, but we plan to increase the efficiency of renewable energy sources. This will be done by diagnosing problems and malfunctions earlier than it is now possible.
<b>BD4OPEM</b>	<p>The project is developing some services which contribute to the reduction of CO2 emission.</p> <p>S3.2 Impact study PV, EV &amp; new loads</p> <p>S5.1 Flexibility Forecast</p> <p>S5.2 Flexibility aggregated services for BRPs</p> <p>S5.3 Flexibility aggregated services for DSOs</p> <p>S5.4 EV to Grid</p> <p>S6.1 Energy management at household or at community level</p> <p>S6.2 Energy forecasting</p>
<b>Synergy</b>	<p>2019 and 2020: N/A</p> <p>Planned:</p> <ul style="list-style-type: none"> <li>- Significant GHG and air pollutants reduction (CO2, CO, NOx, Dangerous Nanoparticles, Non-combustible Hydro-carbons) through the contribution to the avoidance of 217 TWh of RES curtailment in 2050 at EU level</li> </ul> <p>and the facilitation of the target to increase RES integration in the EU Energy Mix to over 90% by 2050</p> <ul style="list-style-type: none"> <li>- Contribution to 1bn tons annual CO2 emissions reduction by 2050 due to renewables integration at EU level</li> </ul>
<b>Boost 4.0</b>	Significant optimization of the manufacturing processes have been achieved in most of the lighthouse pilots that have been implemented in the Project. This optimization implies in some cases a reduction of CO2 emissions, although it has not been monitored, so quantitative statements cannot be provided.
<b>TT</b>	<p>In the case of vehicles, carbon dioxide emissions highly depend on energy consumption, therefore reductions are expected in proportion to energy consumption reduction. When energy savings are based on electricity supply, the CO2 savings depend on the generation mix operating in that specific moment. Final results show that for the specific monitored items improvements in energy consumption per unit of activity range between 6% and 73%, considering among others, contributions belonging to cars, and then standing for CO2 emissions, and contributions belonging to savings in airport energy consumption, which comes from electric supply: Improvement of 33% in the case of Highways, 6% for Connected Vehicles, 28% in Airports, 55% in Urban Mobility and 73% for the E-Commerce pilot.</p> <p>So, taking the average of all KPIs Improvements concerning CO2 and other emissions such as NOX and PM, savings reach up to 37% after the last report made by June of 2019.</p> <p>As said before, besides CO2 there are other emissions that have also being assessed, such as PM and NOX ones. Achievements show positive conclusions too, reaching up to 32% of emissions reduction, considering the average value of all this emission-related KPIs improvements, taken from those pilots addressing this kind of indicators (WP4-Highways, WP5-Connected Vehicles and WP9-Urban Mobility).</p> <p>Again, percentage of KPI improvements have shown a growing tendency among the</p>

	different reporting periods within the project, overtaking most of the predicted target values for energy efficiency. This leads to encouraging prospects for future similar actions on the transportation sector.
<b>Lynx</b>	By providing services for compliance in regards of internationalisation of SMEs we support reduction of travels.
<b>QROWD</b>	<ul style="list-style-type: none"> <li>- By enabling frequent measuring of modal split, data-driven policies that reduce CO2 can be designed and implemented.</li> <li>- Increase of the quality of mobility infrastructure data</li> <li>- Increase citizen engagement towards green mobility campaigns</li> </ul>
<b>DataBench</b>	as above provides a way to measure efficacy of Benchmarking techniques for various sectors including energy

**Table 39:** Contribution from cPPP projects to CO<sub>2</sub> reduction

## 11.4 Contribution to the reduction of waste (KPI II.20)

10% of the projects that replied to the questionnaire have reported contributions to the reduction of waste (DataBench, TT, Boost4.0, BigMedilytics).

In particular:

**DataBench:** provides a way to measure efficacy of Benchmarking techniques for various sectors including energy

**TT:** In the assessment methodology defined in WP2 there are no specific KPIs that will directly monitor the reduction of waste. However, it can be stated that the project contributes to decrease waste through savings in operational activities and improved management of assets. This will happen especially in those pilots that are implementing predictive maintenance activities (WP4 – Highways, WP6 – Railways, WP7 – Ports) that should extend the life of assets, parts and consumables. Final results concerning these pilots show a percentage of improvement higher than 32%, which can fairly well demonstrate a relative high level achievement of reduction waste itself at this final stage of the project.

**Boost 4.0:** Significant optimization of the manufacturing processes have been achieved in most of the lighthouse pilots that have been implemented in the Project. This optimization implies in some cases a reduction of waste, although it has not been monitored, so quantitative statements cannot be provided.

**BigMedilytics:** Pilot focuses on reducing the number of assets a hospital needs to own by optimizing workflows. Plan to completely replace the current paper based system for tracking readings by digital solutions.

## 11.5 Contribution to the reduction in the use of material resources (KPI II.21)

15% of the projects running in 2019-2020 and that have participated to the questionnaire have reported contribution to the reduction in the use of material resources. 3 running lighthouses projects (TT, BOOST4.0 and BigMedilytics) contribute to this KPI in addition to DataBench, IoTwins, DataPorts, Daphne. No quantitative data were provided by the responders, however the qualitative answers can be found in the following table 34.

Project	Qualitative and quantitative statements (2018 and planned)
<b>Daphne</b>	Use cases in the context of semiconductor manufacturing and vehicle development - improvements have the potential to reduce needed material resources.

<b>Boost 4.0</b>	The reduction of the defective products is one of the main contributions of BOOST 4.0 to the reduction of material resources, but not the only one. Also the better planification of the material needed, in case of smart maintenance, where elements (machines, vehicles...) are not subject to a fixed maintenance plan, where parts are changed based on fixed criteria but to smart maintenance where the conditions of the element will indicate when the part needs to be changed,
<b>DataPorts</b>	Reduction of paperwork in ports, reduction of paperwork for port logistics and truck companies.
<b>IoTwins</b>	<p>Testbed #1: less material used during new installation thanks to the project outcomes and less material used thanks the correct maintenance based on Models and data</p> <p>Testbed #2: by increasing the quality of the components, the scrap rate is reduced, and material resources are saved as well as due to reduced postprocessing steps.</p> <p>Testbed #7: by enabling the smart grid operators to closely monitor and to control the state of the low voltage grids it becomes possible to operate closer to the limits of the grid. This in turn should reduce the necessity to upgrade existing power grids. Also, by detecting faults affecting transmission lines as well as equipment in an early phase, maintenance costs, which include spare parts, can be reduced.</p>
<b>TT</b>	Again, in the assessment methodology defined in WP2 there are just a few KPIs that will directly monitor the reduction of material resources consumption even though through improved maintenance and operation, savings in that sense are probable. The defined KPIs in this matter are mainly linked to savings in fuel consumption and in predictive maintenance of assets, whose qualitative achievements have been already presented.
<b>DataBench</b>	provides a way to measure efficacy of Benchmarking techniques for various sectors including materials

Table 40: Contribution from cPPP projects to waste reduction

## 12 STANDARDISATION ACTIVITIES (III.2)

In relation to the specific KPI, **over 30% of the projects** running in 2019-20 reported they perform **activities leading to data/Big data Standardisation** (some with a level of uncertainty).

<b>Project</b>	<b>Activities performed by the project during 2018 leading to standardisation</b>
<b>BigDataOcean</b>	Participation in relevant working group.

<b>DataBio</b>	<p>Contributions to:</p> <ul style="list-style-type: none"> <li>• OGC Testbed 13 – Exploitation Platform Thread</li> <li>• OGC Testbed 14 – EOC Exploitation Platform Thread</li> <li>• OGC DCAT Best Practice Paper OGC 18-001</li> <li>• OGC Big Data Domain Working Group</li> </ul>
<b>euBusinessGraph</b>	Development of an ontology for company information.
<b>FashionBrain</b>	We have standardized the aggregation of open-ended tasks on crowdsourcing platforms using AI models.
<b>SPECIAL</b>	WU have been involved in W3C Open Digital Rights Language (ODRL) working group which released two specifications in February 2018 <a href="https://www.w3.org/TR/odrl-model/">https://www.w3.org/TR/odrl-model/</a> and <a href="https://www.w3.org/TR/odrl-vocab/">https://www.w3.org/TR/odrl-vocab/</a> . On the 17th and 18th of April 2018 the consortium hosted a W3C Workshop on Privacy and Linked Data at WU Wien, Austria <a href="https://www.w3.org/2018/vocabws/">https://www.w3.org/2018/vocabws/</a> . The setup of a W3C Data Privacy Vocabularies and Controls Community Group (DPVCG) in order to address the needs of the broader business community <a href="https://www.w3.org/community/dpvcg/">https://www.w3.org/community/dpvcg/</a> , which was launched the day the GDPR came into effect.
<b>BigMedilytics</b>	<ul style="list-style-type: none"> <li>• Identification of clustering releasing recommendations to relevant scientific Societies</li> <li>• research publication about decentral AI execution in healthcare</li> <li>• We use DICOM and HL7 FHIR and may suggest and enhancement to HL7 FHIR to support medical analytics.</li> <li>• We are defining clinical items to characterise clinical outcome</li> </ul>
<b>BODYPASS</b>	Protocols for secure exchange and aggregation of 3D data
<b>Boost 4.0</b>	BOOST 4.0 has a tight bond with the International Data Space Association as one of the main outcomes expected for the project will be a full operational implementation of the IDS ecosystem. To achieve that, we are working closely with the IDSA on the definition and the development of the IDS components. Launch of W3C Web Standardisation for Graph Data Workshop and the agreement to launch a W3C Business Group. Contribution to NIST Quality Information Framework QIF) standard elevation to ISO standard. Start discussion for Pro-STEP collaboration. Proof-of-concept on the application of AI in network security in ETSI ISG ENI.
<b>TYPHON</b>	The following technologies under development will be proposed as new or extensions to existing standards: TyphonML Modelling Language, TyphonDL Hybrid Polystore Deployment Language, TyphonQL Query Language
<b>Cross-CPP</b>	In 2018, Cross CPP has started working towards standardisation of data model for various cross industrial data which is known as CIDM – Cross Industrial Data Model.

**Table 41:** Activities performed by projects leading to standardisation activities

**3 projects have reported contribution to European Standardization Bodies (ESBs) activities.** In particular:

BDV cPPP Monitoring report 2019-20 – Annex 2 (Detailed explanation of KPIs)

**DataBio** is involved in **6 Working Items** in European Standardization Bodies (ESBs) that are related to the topics of the PPP. In particular:

- OGC 18-001: GeoDCAT-AP Best Practice Report
- OGC 18-050: OGC Testbed 14, ADES & EMS Results and Best Practices ER
- OGC 17-023: OGC Testbed 13, EP Application Package Engineering Report.
- OGC 17-024: OGC Testbed 13, Application Deployment and Execution Service Engineering Report.
- OGC 17-003: GeoJSON(-LD) Metadata for EO Datasets
- OGC 17-084: GeoJSON(-LD) Metadata for EO Collections

It is also involved in **5 of Pre-Normative Research Files** under consultation in European Standardization Bodies (ESBs) that are related to the topics of the cPPP. In particular:

- OGC 18-001: GeoDCAT-AP Best Practice Report
- OGC 17-023: OGC Testbed 13, EP Application Package Engineering Report.
- OGC 17-024: OGC Testbed 13, Application Deployment and Execution Service Engineering Report.
- OGC 17-003: GeoJSON(-LD) Metadata for EO Datasets
- OGC 17-084: GeoJSON(-LD) Metadata for EO Collections

Boost 4.0 is involved in **3 Working Items** in European Standardization Bodies (ESBs) that are related to the topics of the PPP. In particular:

- ETSI GS ZSM002 (ETSI ZSM<sup>41</sup> framework reference architecture)
- ETSI GS ZSM005 (ETSI ZSM means of Automation)
- ETSI GS PDL002

Finally, Lynx project reports being involved in **2 Working Items** in European Standardization Bodies (ESBs) that are related to the topics of the PPP but does not provide additional information in the specific topics.

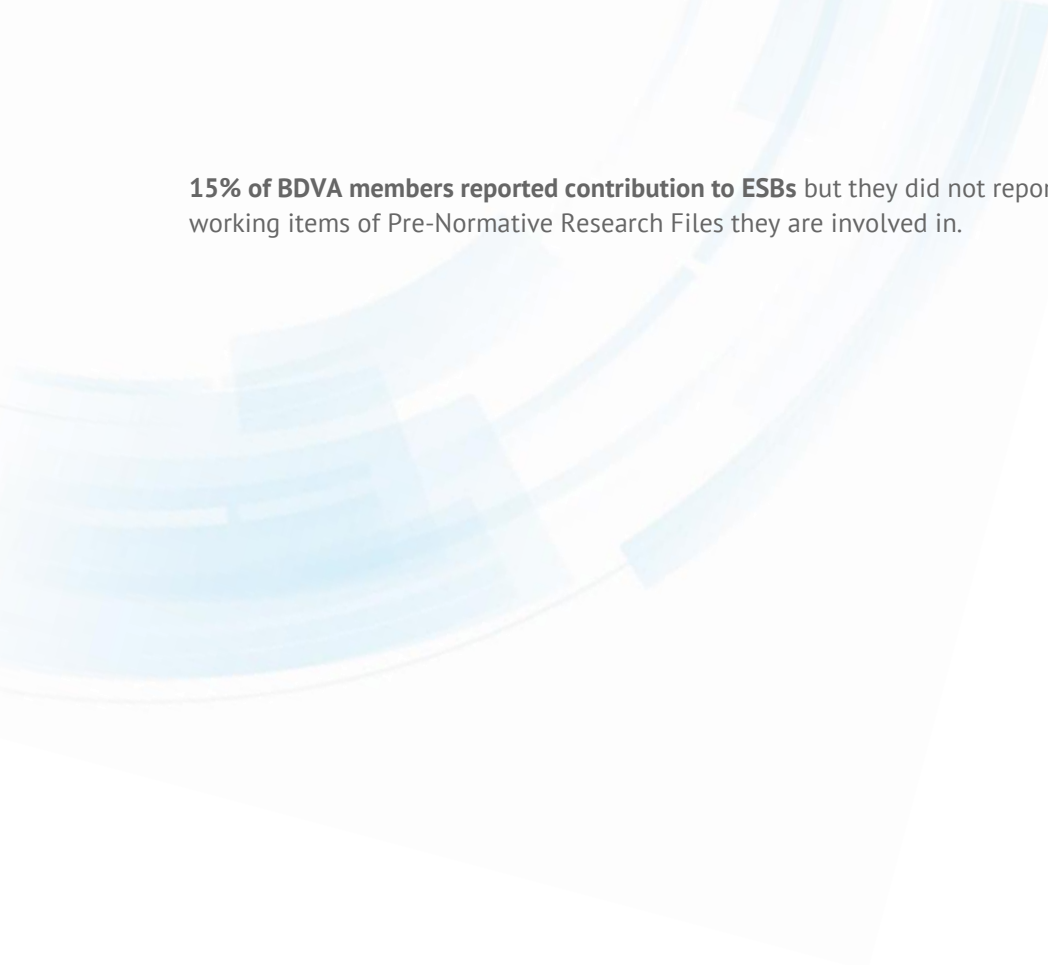
## **20% of BDVA members reported its organisations perform activities leading to data/Big data Standardisation.**

Some examples include:

- **Institute of Advanced Studies Kőszeg** is involved in standardization of health care patient journey data related to National Heart Registry and payer's patient journey data in cancer diseases.
- **Huawei Technologies Düsseldorf GmbH** is involved and contributing to the new ISO/IEC JTC 1 SC 42 on AI
- **Inria** contributes to TEI (stdfSpec) and in ISO on TC37/SC4, ISO 2473 (Language resource management - Lexical markup framework (LMF)). They are also involved in WG of IETF, IRTF, OpenMP, MPI forum, W3C.
- **Fraunhofer-Gesellschaft** drives to standardise primarily data architectures, interoperability methods at this institute. Other institutes will likely have additional efforts but there are few means of obtaining such statistics.
- **Prizztech Robocoast** participates in ISO/IEC JTC1/SC42 Artificial Intelligence, currently at WG1 and WG3 as members of Finnish Standardization Association SR315 group.
- **Siemens AG** contributes to IEC, DIN DKE and other consensus-base standardisation bodies.

---

41 Zero-touch network and Service Management



**15% of BDVA members reported contribution to ESBs** but they did not report specific data about the specific working items of Pre-Normative Research Files they are involved in.

# 13 QUESTIONNAIRES TO COLLECT INFORMATION

## 13.1 Questionnaire for PPP projects running in 2018

(Questionnaire fill in through an online tool)

### Section 0: Master Data

1. Name of your project:
2. H2020 Call topic addressed by your project:
3. Starting date and duration (in months) of your project:
4. Project coordinator (organization):
5. Please indicate here the name and contact details of a contact person from your project who can be reached in case of questions
6. Do you believe that your participation as a project in a larger programme (Big Data Value PPP Programme) brings any value for your project?
7. Please list and describe outcomes of your collaboration with other Big Data Value PPP projects or your participation in the Big Data Value PPP programme (e.g data shared with other projects, etc). We are particularly interested in capturing success stories that can justify the existence of a programme (and not only individual projects). Please focus on 2019 and 2020 outcomes

### Section I: Success Stories

8. List and describe **success stories** of your project for the years 2019 and 2020 (open question)

(Describe relevant project outcomes and achievements here. Add links to media, applications, etc if needed. We appreciate having easy to read / concrete examples that help the reader to understand why this is a success story)

### Section 2: job creation, job profiles and skills development

(Contributes to KPI I-2, II-5 and II-8)

9. Does your project contribute to job creation? (Job creation forecasted within a reasonable timeframe in the future, 3 years after H2020 ends)  
(Yes/No)
10. If yes, can you provide an estimation of the impact of your project in job creation forecasted within a reasonable timeframe in the future (or at least until 2024)? Please provide quantitative and qualitative answer if possible
  - Estimation of jobs created until 2024 (estimate a number)
  - Qualitative explanation of how the project will impact job creation (max. 2000 char)
11. Does your project contribute to increase the number of data workers in Europe? (Yes/No?)
12. If yes, please provide some qualitative statements about how your project is contributing to increase the number of data workers in Europe? (Please refer as much as possible to activities developed during year 2019 and 2020) (Open question. Qualitative answer)
13. Does your project contribute to the creation of new job profiles? (Yes/No?).
14. If yes, how many complete new job profiles have been established by your project in (or by end of) 2019 and 2020? List and describe the new job profiles created indicating skills required per profile
  - Number of new job profiles established by end of 2019 and 2020 (number only)

- List and describe the new job profiles created indicating skills required per profile (open text. 10.000 char)
15. How many complete new job profiles do you estimate will be created by the end of the project?
  16. List and describe the new job profiles created indicating skills required per profile (open text. 5.000 char)
  17. Does your project contribute to the creation/development of new skills? (Yes/No).
  18. If yes, explain how your project has contributed to the development/creation of new skills during 2019 and 2020, and list the new skills developed. (open question)
  19. Number of Master students involved in your project in 2019 (number FTEs)
  20. Number of PhD students involved in your project in 2019 (number FTEs)
  21. Number of Master students involved in your project in 2020 (number FTEs)
  22. Number of PhD students involved in your project in 2020 (number FTEs)
  23. Please indicate in which way your project has contributed and will contribute to the Skills Agenda for Europe<sup>42</sup>:
    - i. During 2019
    - ii. During 2020
    - iii. Forecasted 2021-2024
  24. Number of training activities and programmes (such as tutorials, webinars, etc.) organized by your project in 2018 and number of people benefited by those activities (totals). (total number of activities, estimation of people outreached by training activities, description of those activities (name, channel, attendees):
    - a. Number of training activities (total number 2019)
    - b. Number of training activities (total number 2020)
    - c. Number of people who benefit of those activities (Estimation of a total number 2019)
    - d. Number of people who benefit of those activities (Estimation of a total number 2020)
  25. List and describe the training activities (name, channel, content, audience) (Open question, max 5.000 characters)
  26. Number of European training programs involving 3 different disciplines established by your project by the end of 2019 and 2020 (include number of programmes, list them and indicate participants per programme). Separate the number and the description
    - a. Number of training activities involving 3 different disciplines in 2019
    - b. Number of training activities involving 3 different disciplines in 2020
    - c. Number of participants/beneficiaries of the activities in 2019
    - d. Number of participants/beneficiaries of the activities in 2020
  27. List and describe the training activities (name, channel, content, audience) (Open question, max 5.000 characters)

### Section 3: Innovations and technical results

---

<sup>42</sup> <http://ec.europa.eu/social/main.jsp?catId=1223>

(Contributes to I-4, II-4, II-7, II-10, II-13, II-15, II-16, II-17)

### **Innovations to market**

The European Commission has requested all the cPPPs to measure what is called “Number of significant innovations to market”. This KPI concerns all developed items that have a marketable value, including products, processes, instruments, methods, technologies, spin offs, patents, ... It should involve all items directly linked to the cPPP projects as developed foreground, as well as any items beyond the scope of the project that is linked to the foreground.

We are asking the projects to list and briefly described all developed foreground, tangible and intangible assets, that have a marketable or at least an exploitable value, including (not only) products, processes, instruments, methods, technologies, spin-offs, patents,....

For the MR2019-2020 input to this question is given through a different excel questionnaire to facilitate your input and the assessment. You can find the file here. If you cannot access the file ask your project coordinator or contact [operations@core.bdva.eu](mailto:operations@core.bdva.eu).

Further explanations about this question and KPI are given in the abovementioned file.

28. Can you confirm you understood this question is replied via a separate excel questionnaire and that you have access to it? (Yes/NO)

### **European provision of big data value creation systems and technologies**

*The Big Data Value Contractual agreement also incorporates a KPI (specific to this Big Data programme) called: “Increased competitive European provision of big data value creation systems and technologies”. For this we are asking the project to provide information about the “Number of systems and technologies developed in the relevant sector in cPPP projects (beyond state of the art)”*

*Theoretically this input should be a subset of your answers to the previous question.*

29. Number of systems and technologies developed in the relevant sector in the project during 2019 and 2020 (number and short description) (if your answer to this question is embedded in “Innovations to Market” questionnaire please indicate so in the description and just referred to the short name)

### **Support major sectors and major domains by Big Data technologies and applications**

30. List the sectors and major domains supported by Big Data technology and applications developed in your project (the following domains listed and more can be added as needed: Health and Healthcare, Bio-economy, Manufacturing, Transport and logistics, Mobility, Telecom, Earth Observation, Media, Retail, Energy, Finances, Public Services, Water and natural resources, Business services, Smart Cities, other)

### **Enable advanced privacy and security respecting solutions for data access, processing and analysis**

31. Number of patents filed by your project that enable advanced privacy and security respecting solutions for data access, processing and analysis:
- a. Year 2019
  - b. Year 2020
  - c. Forecasted 2021-2024

32. Number of publications by your project that describe advanced privacy and security respecting solutions for data access, processing and analysis:
  - a. Year 2019
  - b. Year 2020
  - c. Forecasted 2021-2024
33. Number of OSS contributions/ products/ SW components resulting from your project that enable advanced privacy and security respecting solutions for data access, processing and analysis (number and brief description):
  - a. Year 2019
  - b. Year 2020
  - c. Forecasted 2021-2024
34. Please provide here a short description of the abovementioned OSS contributions/ products/ SW components if applicable

**New economic viable services of high societal value**

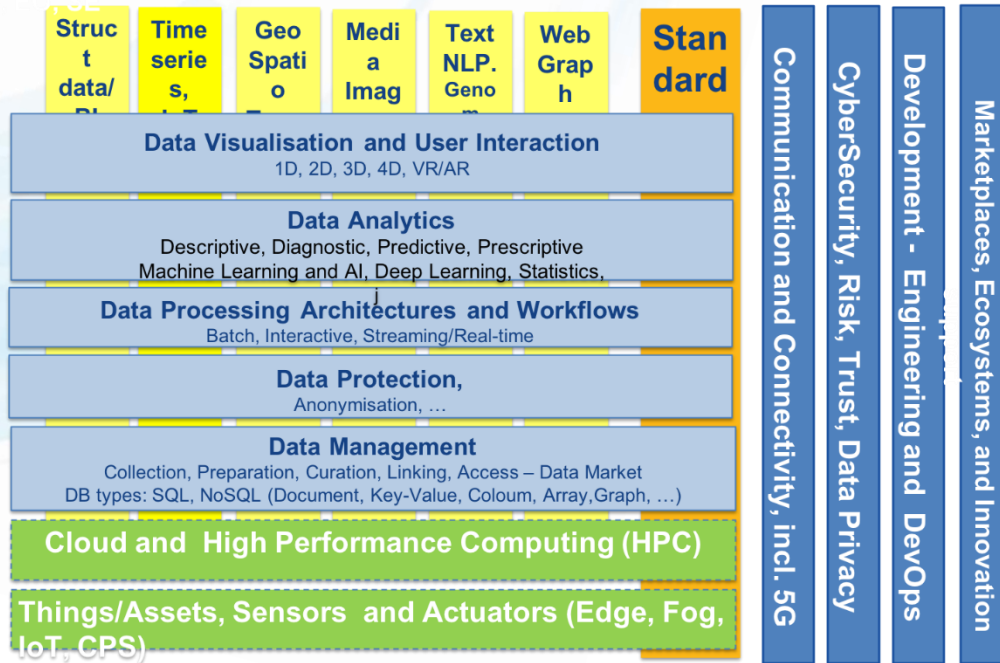
35. Please indicated below the number and provide a short description of new economic viable services of high Societal Value developed by your project.
  - a. Year 2019:
    - i. Number
    - ii. Description per service (describe why something qualifies)
  - b. Year 2020:
    - i. Number
    - ii. Description per service (describe why something qualifies)
    - iii.
  - c. Forecasted 2021-2024:
    - i. Number
    - ii. Description per service (describe why something qualifies)

(Note: Criteria for “high societal value” should be up to the project. Project should describe why something qualifies)

**Technology progress is in line with multi-annual roadmap of SRIA**

*With this KPI we intend to measure % of research priorities covered compared to overall scope of research priorities defined in SRIA (differentiate running, upcoming and not covered yet).*

*To answer this question, you need to use as reference the BDVA SRIA v.4.*



Link to SRIAv4: [http://www.bdva.eu/sites/default/files/BDVA\\_SRIA\\_v4\\_Ed1.1.pdf](http://www.bdva.eu/sites/default/files/BDVA_SRIA_v4_Ed1.1.pdf) Chapter 3 (overview in chapter 2.3, reference model)

A separate excel file is used to collect information with the appropriate level of details. You can find the excel file here.

Use this question just to confirm you understood you have to submit a separate file and you have access to it.

36. Can you confirm you understood this question is replied via a separate excel questionnaire and that you have access to it? (YES/NO)

#### Availability of metrics for measuring the quality, diversity and value of data assets

37. Are you assessing quality, diversity and value of data assets? (YES/NO?)

38. If yes, what metrics are you using to quantify them? (Open Answer)

#### Increase the speed of data throughput compared to 2014

39. Does your project expect to improve data throughput? (YES/NO?)

40. If yes, what metrics are you using to measure this? (Open Answer)

41. If yes, what improvement (in percent) did you measure in your project? (Numerical Percentage Value).

#### Section 4: Experiments in Big Data

(Contributes to II-11, II-12 and II-14)

#### Large Scale experiments conducted in cPPP projects and i-Spaces involving closed data

42. Number of large-scale experiments conducted in your project in 2019 (number)

43. Number of large-scale experiments conducted in your project in 2020 (number)

44. How many of those involved closed data in 2019? (number)

45. How many of those involved closed data in 2020? (number)

46. Please describe your criteria for an experiment to qualify as “large-scale” (e.g. cross-border / Geographical impact, Number of companies involved, Investment, Higher TRLs, number of users, etc) If possible, list the most relevant experiments and explain outcomes (open question)

#### **Uptake of BDV use cases and experiments**

47. Number of data experiments/use cases of any kind or size conducted in 2019 and 2020 (number)  
48. Please provide some qualitative data to support your answer (i.e list, brief description,..) (open answer)

#### **Amount of data that has been made available for experimentation (cPPP projects and i-Spaces)**

49. Number of Exabytes of data made available in your project in 2019 and 2020 (including closed/private data)

(This includes data made available for experimentation (e.g. Data incubators and i-Spaces) and data made available internally in the projects for experimentation purposes, so closed data (e.g. Lighthouses and other projects).

50. Add an approximative number and indicate the units (Terabytes, etc)  
a. for 2019  
b. for 2020  
51. Please add some qualitative information to support your answer (e.g. overview type of data, providers, etc)

#### **Section 5: Contribution to Macro-Economics KPIs**

(Contributes to II-1, II-2 and II-3)

52. Does your project contribute (has contributed or plan to contribute) to increase revenue share of EU companies against total of revenue of EU, US, Japan, Brazil? (YES/NO?)  
53. If yes, please provide some qualitative statements about how your project is contributing to this. (Open Answer)  
54. Does your project contribute (has contributed or plan to contribute) to increase the number of European Companies offering data technology, applications? (YES/NO?)  
55. If yes, please provide some qualitative statements about how your project is contributing to this. (Open Answer)  
56. Does your project contribute (has contributed or plan to contribute) to the revenue generated by European Data Companies? (YES/NO?)  
57. If yes, please provide some qualitative statements about how your project is contributing to this. (Open Answer)

#### **Section 6: Contributions to environmental challenges**

(Contributes to II-6, II-19, II-20, II-21)

##### **Energy saved in big data analytics using a specified benchmark through solutions provided by specific cPPP projects compared to baseline at beginning of H2020**

58. Are you monitoring Energy Efficiency in your project? (YES/NO?)  
59. If yes, what metrics are you using to measure this parameter? (Open Answer)  
60. If yes, what improvement (in percent) have you measured in your project? (Numerical Percentage Value).

### **Contribution to the reduction of energy use and CO2 emissions**

61. Does your project contribute to the reduction of energy use? (YES/NO?)
62. IF yes, please provide some qualitative and quantitative statements (2019, 2020 and planned)
63. Does your project contribute to the reduction of CO2 emission? (YES/NO?)
64. IF yes, please provide some qualitative and quantitative statements (2019, 2020 and planned)

### **Contribution to the reduction of waste**

65. Does your project contribute to the reduction of waste? (YES/NO?)
66. IF yes, please provide some qualitative and quantitative statements (2019 and 2020 and planned)

### **Contribution to the reduction in the use of material resources**

67. Does your project contribute to the reduction of material resources? (YES/NO?)
68. IF yes, please provide some qualitative and quantitative statements (2019 and 2020 and planned)

### **Section 7: Mobilisation of stakeholders, outreach, success stories**

(Contributes to section 2.2 of the monitoring report of the contractual PPPs)

69. Number of dissemination events, seminars, conferences organised by your project in 2019 and 2020 (number)
70. Estimate the number of people outreached by your dissemination and engagement activities in 2019 and 2020:
  - Total
  - In events (conferences, seminars, trainings)
  - Online (webinars, online trainings)
  - Via mass media (online or physical)
  - Others
71. List and describe all the activities performed in 2019 and 2020 to mobilise and outreach to stakeholders in your project (i.e. dissemination and communication activities of any kind, events, workshops, etc). Include in all cases the outreach in numbers of participants and in geographical outreach. (open question)

### **Section 8: Contribution to Standards**

72. Does your project perform any activities leading to data/Big data Standardisation? (Please include also AI related activities) (yes/no?)
73. If yes, list and describe activities performed by your project during 2019 and 2020 leading to standardisation
74. Does your project contribute (through one or several project partners) to European Standardization Bodies (ESBs) activities? (yes/no?)

If yes,

75. Indicate the number of Working Items in European Standardization Bodies (ESBs) that where your project is involved in and that are related to the topics of the PPP

(A Working Item is a registered topic in one of the ESBs that has been accepted as a matter of standardization activities).

76. List the IDs of the Working Items (list)
77. Indicate the number of **Pre-Normative Research Files** under consultation in European Standardization Bodies (ESBs) that your project is involved in and that are related to the topics of the cPPP. (A Pre-Normative Research File is a draft standard under consultation in one of the ESBs.) (number)

### **Section 9: Contribution to Big Data Ecosystem**

78. Does your project run or support a programme that is specifically targeted at supporting start-ups or entrepreneurs in the Big Data area and AI? (yes/no?)
79. If yes, please describe! (You will be contacted for further details, if necessary)
80. Is your project engaged in contributing to any National activities related to Big Data Value? (yes/no?)
81. If yes, provide the country, details of the activity, and an overview about your participation! Focus on 2019 and 2020 activities Some short statements are sufficient. You will be contacted for further details separately.
82. Is your project present in EU13 Member States? (yes/no?)
83. If yes, provide country(ies), and overview of activities in that country.
84. Is your project linked to another European PPP, ETP or European industrial Association?
85. If yes, provide PPP/ETP/International Association names and briefly describe common/linked activities.
86. Does participation in a larger programme (Big Data Value cPPP) bring any value to your project? (Yes/No)
87. Please list and describe outcomes of your collaboration with other BDV CPPP projects or your participation in the BDV CPPP programme (e.g data shared with other projects, etc). We are particularly interested in capture success stories that can justify the existence of programme (and not only individual projects). Please focus on 2019 and 2020 outcomes (Open question)

### **Section 10: Additional comments**

88. Provide any additional information you may find relevant (if needed)
89. Provide comments to this survey (if needed)

## **13.2 Technical priorities table for projects - Questionnaire**

Survey in excel distributed to all projects running in 2019 and 2020.

It relates to the Q28 of the MR2019-2020 questionnaire for projects.

*Only input in the cells that are relevant for your project.*

*In "focus" mark with "x" main focus of your project and "o" challenge is addressed in a vertical way given by the main focus.*

*In the columns "number of" indicate a number (e.g 0, 1, 2, et) when applies.*

*You can use the column comments to specify more details if you consider so important.*

(Add the name of your project here)		Delivered in 2018	Delivered in 2019	Delivered in 2020	Ongoing (not delivered but initiated in 2019 or earlier)	Planned (for 2021+)	
Technical priorities and challenges	Focus	Number of contributions/technologies (provided in 2018)	Number of contributions and technologies	Number of contributions and technologies	Number of contributions/technologies	Number of contributions/technologies	Comments
<b>Data Management</b>							
Semantic Annotation of unstructured and semi-structured data							
Semantic interoperability							
Data quality							
Data lifecycle management and data governance							
Integration of data and business processes							
Data-as-a-service							
Distributed trust infrastructures for data management							
Other (specify)							
<b>Data Processing Architectures</b>							
Heterogeneity							
Scalability							
Processing of data-in-motion and data-at-rest							
Decentralization							
Performance							
Novel architectures for enabling new types of big data workloads (Hybrid Big Data and HPC architecture)							
Introduction of new hardware capabilities							
Other (specify)							
<b>Data Analytics</b>							
Semantic and knowledge-based analysis							
Content validation							
Analytics frameworks & processing							
Advanced business analytics and intelligence							
Predictive and prescriptive analytics							
High Performance Data Analytics (HPDA)							
Data analytics and Artificial Intelligence							
Other (specify)							
<b>Data Protection</b>							
Generic and easy to use data protection approaches							
Robust Data privacy (incl. multi-party computation)							
Risk based approaches							
Other (specify)							
<b>Data Visualisation and User Interaction</b>							
Visual data discovery							
Interactive visual analytics of multiple scale data							
Collaborative, intuitive and interactive visual interfaces							
Interactive visual data exploration and querying in a multi-device context							
Other (specify)							
<b>Other (specify)</b>							

### 13.3 Significant Innovations to Market – Questionnaire

Survey in excel distributed to all projects running in 2019 and 2020.

It relates to the Q21 of the MR2019-2020 questionnaire for projects.

The excel file requested the project list, describe and characterised innovations to market including the following information per innovation:

- Innovation name (short name if exists)
- Type of innovation (technology, system, methods, product, service, methods, instruments, processes, spin-off, patent, other)
- Brief description (Including the marketable/exploitable value)
- Impact (Significant/Medium)
- Impact description (if significant please justify that assessment)
- Impact type (Economic/Societal)
- TRL (if applies)
- Maturity level (e.g whether already supported by a prototype or demonstrator, filed or grant patent, developing or published standard, or spin-off created)
- Sectors
- Type of innovation (details in readme)
- Timing (Already reported in 2018 but progress in 2019 / New in 2019 / New in 2020/ Expected in +2021)

The following guidelines were provided to support classification of innovations:

*The European Commission has requested all the cPPPs to measure what is called “Number of significant innovations to market”. This KPI concerns all developed items that have a marketable value, including products, processes, instruments, methods, technologies, spin offs, patents, ... It should involve all items directly linked to the cPPP projects as developed foreground, as well as any items beyond the scope of the project that is linked to the foreground.*

We are asking the projects to list and briefly described all developed foreground, tangible and intangible assets, that have a marketable or at least an exploitable value, including (not only) products, processes, instruments, methods, technologies, spin-offs, patents,....

Note 1: It should involve all items directly linked to the cPPP projects as developed foreground, as well as any new foreground beyond the scope of the project that is linked to the project results (most likely through IP). Information on innovations ready to be taken to the market would be an asset given that the EU seeks opportunities to stimulate more interest by investors. Also, other activities could be considered to carry value added, especially at lower technology readiness levels, such as advancement of a certain technology, clustering, and standardisation. However, the measurability must contain some exploitable value. As an example, a (part) contribution to an industrial standard which has an intrinsic value as a method on the market would count as an innovation.

Types of Innovation:

- **Incremental Innovation:** Improving existing technology to increase customer value (features, design changes, etc) within an existing market
- **Disruptive Innovation:** Applying a new technology or processes in an existing market
- **Architectural Innovation:** taking existing skills and technology and applying it in a different market
- **Radical Innovation:** New technology that creates new industries

Significant impact: High impact regardless of the type of innovation (e.g. and incremental innovation can have a major impact if scaled reducing CO2 emissions). Justify this significant impact. Rest goes to Medium

## 13.4 Investments questionnaire

Distributed to all for-profit companies (large and SMEs) part of projects active in 2019 and 2020 and to all for-profit members.

		2014	2015	2016	2017	2018	2019	2020	2021-2024	Additional info requested (see answer for general information, Reasoning text)	IPM F Classification
6	Regional name of the following figures	EUR									For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP.
7	Revenue	EUR								For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP.	This number is overall a public number published by companies
8	Employment	FTE								For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP.	This number is overall a public number published by companies
9	R&D expenses	EUR								For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP.	This number is overall a public number published by companies
10	Adjusted of the total amount of R&D expenses that are related to the cPPP and other related to the cPPP	EUR								For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP.	This number is overall a public number published by companies
11	Adjusted of the total amount of R&D expenses that are related to the cPPP and other related to the cPPP	EUR								For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP.	This number is overall a public number published by companies
12	Adjusted of the total amount of R&D expenses that are related to the cPPP and other related to the cPPP	EUR								For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP.	This number is overall a public number published by companies
13	Adjusted of the total amount of R&D expenses that are related to the cPPP and other related to the cPPP	EUR								For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP.	This number is overall a public number published by companies
14	Adjusted of the total amount of R&D expenses that are related to the cPPP and other related to the cPPP	EUR								For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP.	This number is overall a public number published by companies
15	Adjusted of the total amount of R&D expenses that are related to the cPPP and other related to the cPPP	EUR								For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP.	This number is overall a public number published by companies
16	Adjusted of the total amount of R&D expenses that are related to the cPPP and other related to the cPPP	EUR								For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP.	This number is overall a public number published by companies
17	Adjusted of the total amount of R&D expenses that are related to the cPPP and other related to the cPPP	EUR								For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP.	This number is overall a public number published by companies
18	Adjusted of the total amount of R&D expenses that are related to the cPPP and other related to the cPPP	EUR								For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP. For SMEs: Specify how you measure regional GDP.	This number is overall a public number published by companies

Table 42: Investment questionnaire sent to for-profit companies

	2018
EC: EU contribution received for projects under the cPPP	Extracted from Corda
<b>Direct leverage (A = A1 + A2)</b>	
A1: Direct contributions from the industry to the cPPP projects, as predetermined in the projects (co-funding)	Extracted from Corda
A2: Additional investment or resources spent by your company for the execution of cPPP projects, beyond the direct contributions under A1 (these are typically due to shortfalls in fixed overheads for a project, infrastructure costs not covered by direct funding, consumables and maintenance costs, equipment costs, additional personnel costs, other costs)	EUR
<b>Indirect leverage (B = B1 + B2)</b>	
B1: Additional investment or resources spent as a follow-up to the activities and results related to cPPP projects (e.g. to exploit or scale-up results) and that are not supporting directly the execution of these projects. This can include additional funding mobilised thanks to the cPPP projects such as venture capital, other public funding, etc.	EUR
B2: Additional investments or resources spent that support the objectives of the cPPP and that have been triggered/encouraged by the existence of the cPPP, but that are not following up activities or results of the cPPP projects as such (hence, not counted in A1, A2 and B1).	EUR

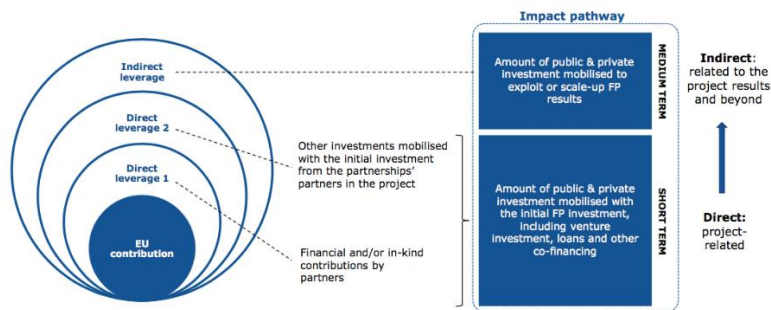


Figure 9: Table and illustration of the subset relationship of the expenses requested in (11) to (14)

## 13.5 Questionnaire for BDVA members

Distributed to all BDVA members. Responses collected online.

### Section 0: Master Data

1. Name of your Organisation\*
2. Membership type\*
3. My organisation is a University or Non-for-Profit (Research/Innovation) Organization (Yes/No)\*
4. My organization is an SME according to EC criteria (1. less than 250 employees, 2. annual turnover less than 50 MEUR or annual balance sheet less than 43 MEUR)\*

In case we have any questions to your answers we would like to ask you to provide your name and Email address\*

5. In case we have any questions to your answers we would like to ask you to provide your name and Email address:

### Section I: Success Stories

6. Provide any kind of success story of your organization during 2019 or 2020 in relation to BDVA and the Big Data PPP e.g. where the PPP clearly has proven its value or where the Big Data PPP triggered investments inside your organization, etc. that we are able to publish in the Monitoring Report (open question)\*

(Add links to media, applications, etc if needed. We appreciate having easy to read / concrete examples, text that we can “quote” that help the reader to understand why this is a success story)

### Section 2: job creation, job profiles and skills development

(Contributes to KPI I-2, II-5 and II-8)

7. Has your participation in the BDVA/BDV CPPP contributed directly or indirectly to job creation? (Yes/No)\*
8. If yes, can you provide an estimation of the impact in job creation during 2019 and 2020 (in your organisation or organisations linked to you)?
  - Estimation of jobs created by end of 2019 (estimate a number)
  - Qualitative explanation (e.g. profiles, etc)
  - Estimation of jobs created by end of 2020 (estimate a number)
  - Qualitative explanation (e.g. profiles, etc)
9. Does your organisation contribute to the creation of new job profiles in Big Data? (Yes/No?)\*
10. If yes, how many complete new job profiles (in big data/AI) have been established by your organisation in 2019 and 2020? Please list them:
  - Number of new job profiles established by end of 2019 (number only)
  - List and describe the new job profiles created indicating skills required per profile (open text. 5.000 char)
  - Number of new job profiles established by end of 2020 (number only)
  - List and describe the new job profiles created indicating skills required per profile (open text. 5.000 char)

11. Does your organisation contribute to the creation/development of new skills (scope of Big Data Value)? (Yes/No)\*
12. If yes, explain how your organisation has contributed to the development/creation of new skills during 2019 and 2020, and list the new skills developed (scope of Big Data Value) (open question)
13. Please indicate in which way your organisation has contributed and will contribute to the Skills Agenda for Europe<sup>43</sup>:
  - i. During 2019. \_\_\_\_\_
  - ii. During 2020. \_\_\_\_\_
  - iii. Forecasted 2021+. \_\_\_\_\_
14. Number of training programs in 2019 and 2020 (such as tutorials, webinars, etc.) conducted by your organisation that have arisen from cPPP activities and number of people benefited by those activities (totals)\* (Please don't include training programmes part of BDV CPPP projects and therefore reported as a PPP project activity. Ok if they are linked to topic and developed through other projects and programmes)
  - a. Number of training activities (total number 2019)
  - b. Number of people who benefit of those activities (Estimation of a total number 2019)
  - c. List and describe the training activities (name, channel, content, audience)
  - d. Number of training activities (total number 2020)
  - e. Number of people who benefit of those activities (Estimation of a total number 2020)
  - f. List and describe the training activities (name, channel, content, audience)
15. Number of European training programs in 2019 and 2020 involving 3 different disciplines conducted by your organisation that have arisen from cPPP activities in 2019 and/or 2020 (include number of programmes, list them and indicate participants per programme)\* (Please don't include programmes part of BDV CPPP projects)
  - a. Total number of programmes involving 3 different disciplines in 2019
  - b. List the programmes and number of participants per programme
  - c. Total number of programmes involving 3 different disciplines in 2020
  - d. List the programmes and number of participants per programme
  - e. Any other comments

### Section 3: Innovations and technical results

(Contributes to I-4, II-4, II-7, II-10, II-13, II-15, II-16, II-17)

#### Only for PPP projects

### Section 4: Experiments in Big Data

(Contributes to II-11, II-12 and II-14)

#### (ONLY for projects or BDVA labelled i-Spaces)

#### Large Scale experiments conducted in i-Spaces involving closed data

16. Number of large-scale experiments conducted in your project in 2019 and 2020 (number)
17. How many of those involved closed data? (number)

---

<sup>43</sup> <http://ec.europa.eu/social/main.jsp?catId=1223>

18. Please describe your criteria for an experiment to qualify as “large-scale” (e.g. cross-border / Geographical impact, Number of companies involved, Investment, Higher TRLs, number of users, etc) If possible list the most relevant experiments and explain outcomes (open question)

#### **Uptake of BDV use cases and experiments**

19. Number of data experiments/use cases of any kind or size conducted in 2019 and 2020 (number)
20. Please provide some qualitative data to support your answer (i.e list, brief description,..) (open answer)

#### **Amount of data that has been made available for experimentation (i-Spaces)**

21. Number of Exabytes of data made available in your i-Space in 2019 and 2020 (including closed/private data)

(This includes data made available for experimentation (e.g. Data incubators and i-Spaces) and data made available internally in the projects for experimentation purposes, so closed data (e.g. Lighthouses and other projects). PLEASE ADD an approximative number and indicate the units (Terabytes, etc). DON't add explanations here

22. Please add some qualitative information to support your answer (e.g. overview type of data, providers, etc)

#### **Section 5: Contribution to Macro-Economics KPIs**

(Contributes to II-1, II-2 and II-3)

Only for PPP projects (no need for members to input)

#### **Section 6: Contributions to environmental challenges**

(Contributes to II-6, II-19, II-20, II-21)

Only for PPP projects (no need for members to input)

#### **Section 7: Mobilisation of stakeholders, outreach, success stories**

(Contributes to section 2.2 of the monitoring report of the contractual PPPs)

Only for PPP projects and TF3/TF4

#### **Section 8: Contribution to Standards**

23. Does your organisation perform any activities leading to data/Big data Standardisation? (yes/no?)\*

24. If yes, list and describe activities performed by your organisation during 2019 and 2020 leading to standardisation

25. Does your organisation contribute to European Standardization Bodies (ESBs) activities? (yes/no?)

If yes

26. Indicate the number of **Working Items** in European Standardization Bodies (ESBs) where your organisation is involved in and that are related to the topics of the PPP

(A Working Item is a registered topic in one of the ESBs that has been accepted as a matter of standardization activities).

27. List the IDs of the Working Items (list)

28. Indicate the number of **Pre-Normative Research Files** under consultation in European Standardization Bodies (ESBs) that your organisation is involved in and that are related to the topics of the cPPP. (A Pre-Normative Research File is a draft standard under consultation in one of the ESBs.) (number)
29. List the prEN numbers of these Research Files
30. Does your organisation contribute to International Standardization Bodies (ISBs) activities? (yes/no?)\*

If yes,

31. Please describe your activities linked to topics of the BDV CPPP (Big Data/AI)

### **Section 9: Contribution to Big Data Ecosystem**

32. Does your organisation run or support a programme that is specifically targeted at supporting start-ups or entrepreneurs in the Big Data area? (yes/no?)\*
33. If yes, please describe! (You will be contacted for further details, if necessary)
34. Does your organisation participate in National activities related to Big Data Value? (yes/no?)\*
35. If yes, provide the country, details of the activity, and an overview about your participation! Some short statements are sufficient. You will be contacted for further details separately.
36. Does your organization participate in other European PPPs, ETPs or European industrial Associations?\*
37. If yes, provide PPP/ETP/International Association names and briefly describe activities and link with Big Data PPP

### **Section 10: Additional comments**

38. Provide any additional information you may find relevant (if needed)
39. Provide comments to this survey (if needed).



[www.bdva.eu](http://www.bdva.eu)