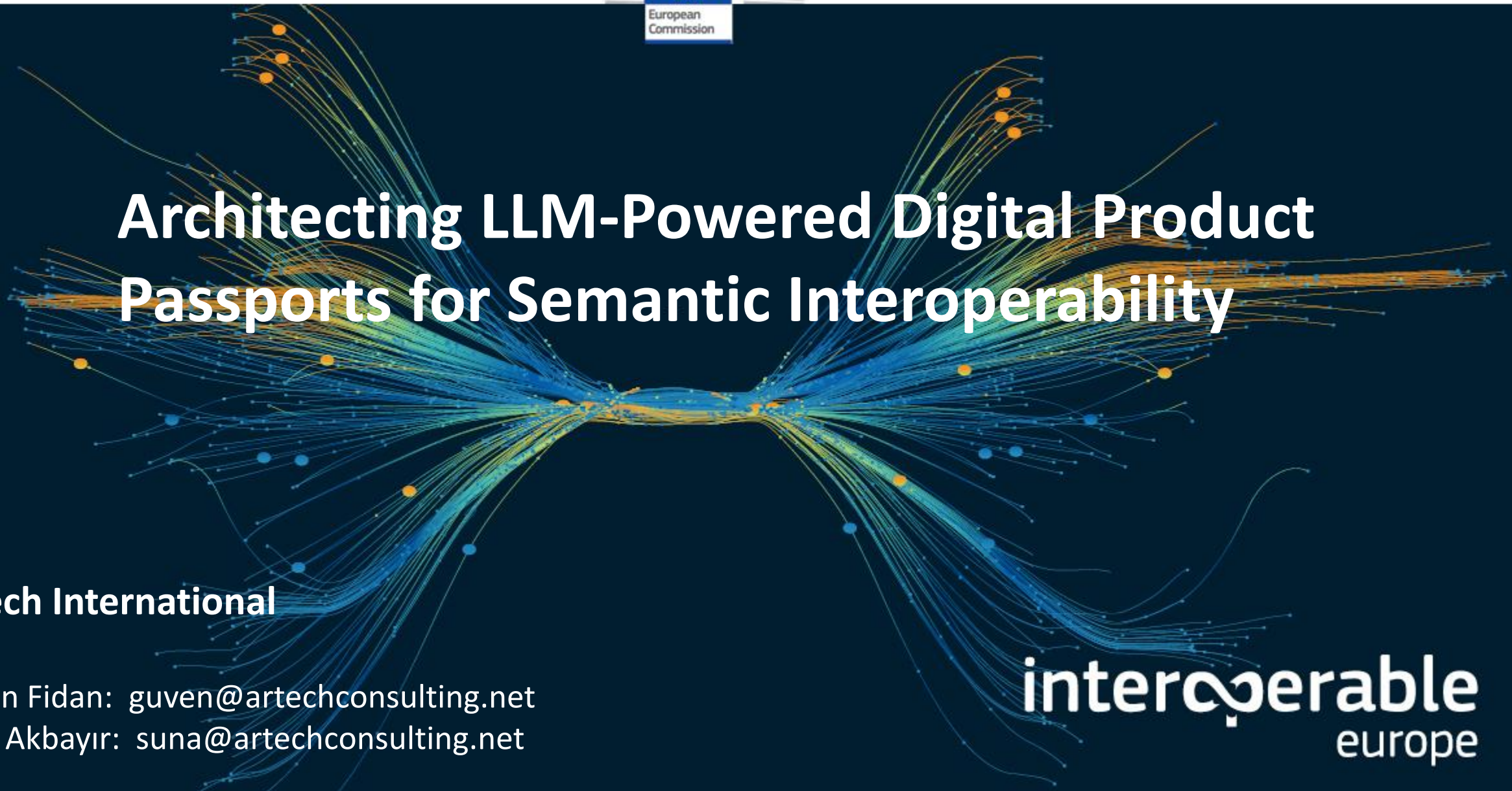


Architecting LLM-Powered Digital Product Passports for Semantic Interoperability

Artech International

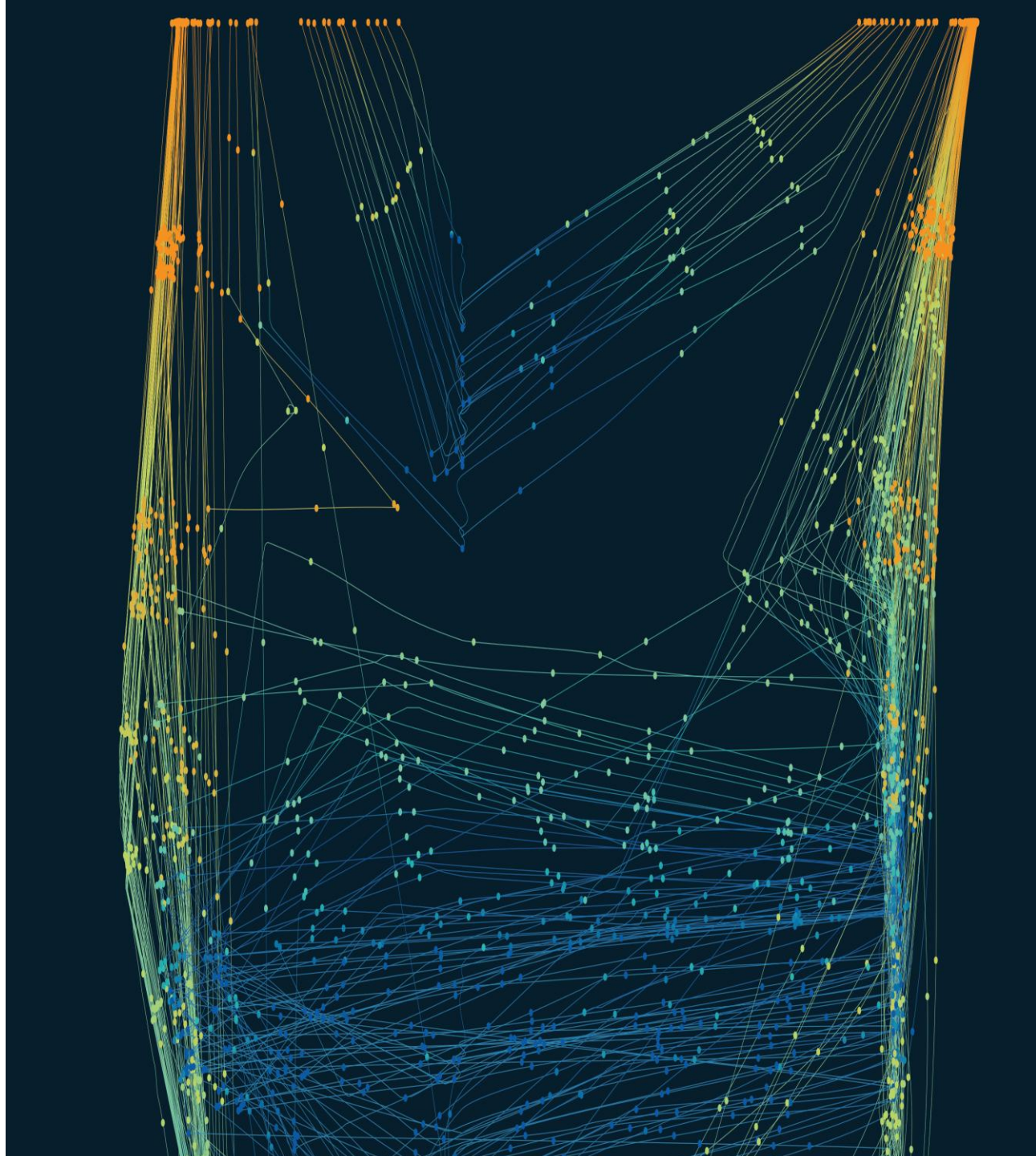
Guven Fidan: guven@artechconsulting.net
Suna Akbayır: suna@artechconsulting.net

1

Why DPPs Need Semantic Interoperability

EU mandates DPPs for batteries, textiles, electronics, with implementation starting in 2024 under the Eco design for Sustainable Products Regulation (ESPR) and by 2027 it is mandatory for Textile and Batteries.

DPPs (Digital Product Passports) aggregate lifecycle data: materials, production, emissions, recyclability



2

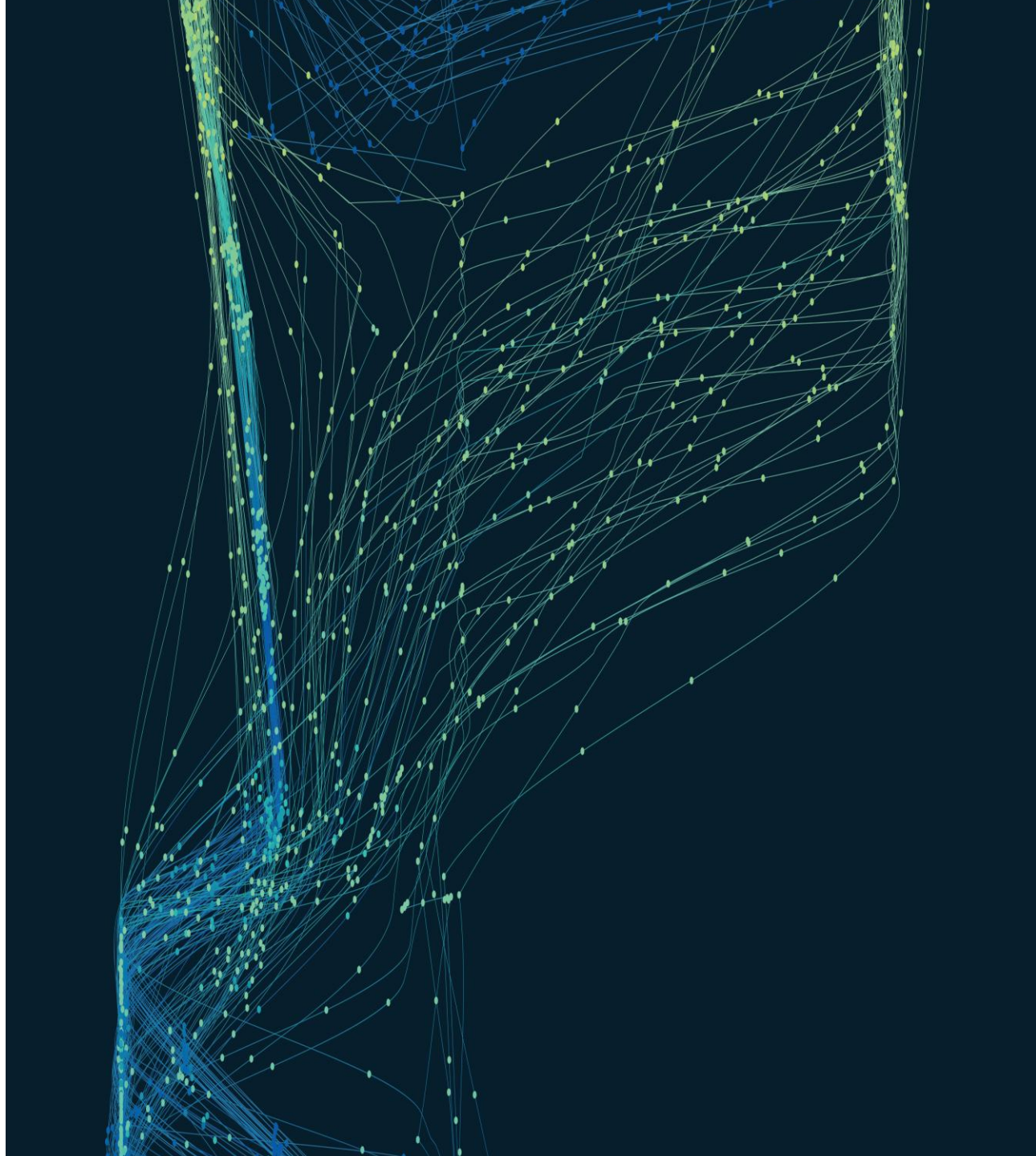
Challenge

Data comes from fragmented sources (ERP, PLM, certificates, suppliers). Semantic interoperability ensures machine-readable, standard-compliant, cross-domain data exchange.

3

Key Challenges in DPP Creation

- Heterogeneous formats (PDF, JSON, XML, RDF)
- Manual ontology mapping is slow & error-prone
- Regulatory compliance varies across domains
- Language, schema, terminology inconsistency
- Data security and proprietary information protection



Modern LLMs: Capabilities & Fit (2025)

Latest Models (2025)

GPT-4.5, GPT-4o, GPT-o3-mini: Advanced reasoning, multimodal capabilities

Claude 3.7/3.5 Sonnet: Enhanced context understanding, reduced hallucinations

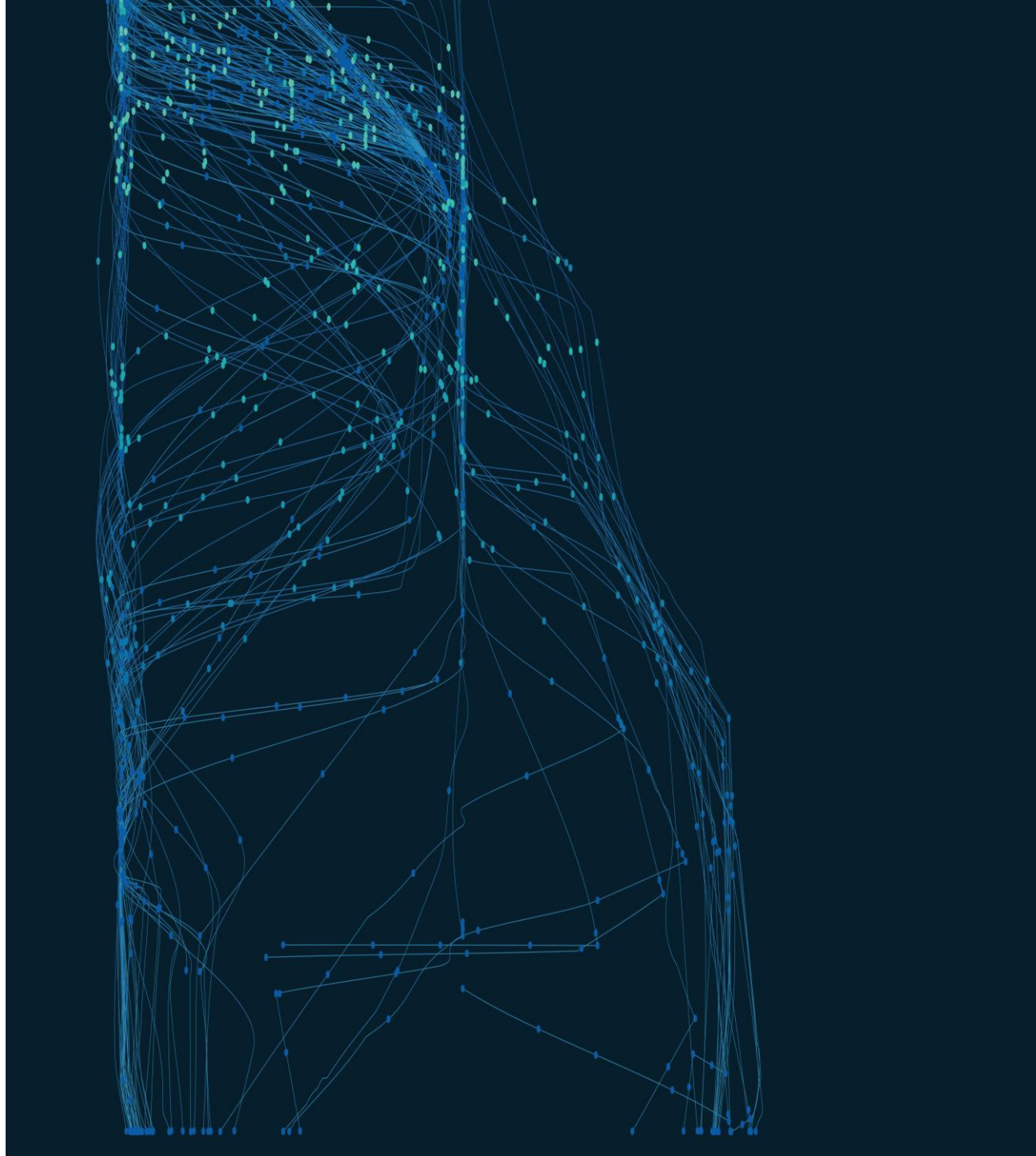
4

Gemini 2.0 Pro/Flash: 2M token context window, real-time processing

Grok-3, Mistral Large 2: Competitive open alternatives

Key Capabilities

- 200K-2M token context windows for full-document understanding
- Advanced multimodal processing (text, images, charts, tables)
- Zero-shot schema mapping and terminology normalization
- Fine-tunable for domain-specific terminology (battery composition, textile labeling)
- Improved factuality with reduced hallucinations



Architecture Overview (Hybrid Approach)

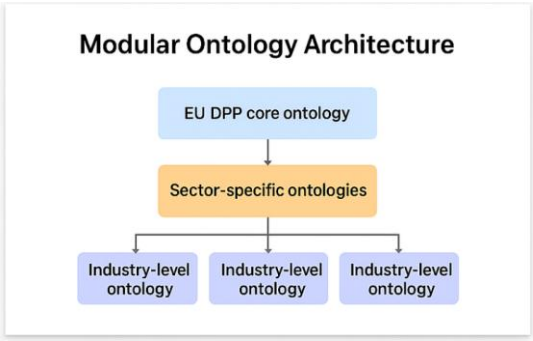
Data Sources: Supplier docs, certifications, ERP/PLM exports etc.
Preprocessing: OCR, language detection, metadata tagging
LLM Layer:

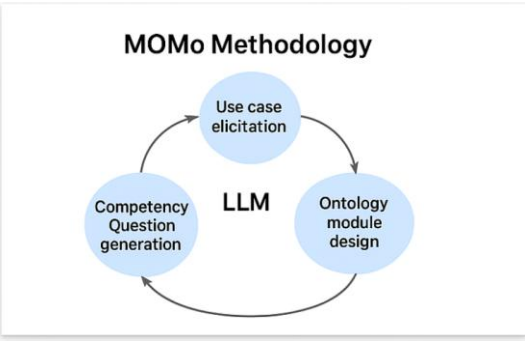
- Entity extraction, schema mapping, terminology normalization
- Summarization (for user-friendly DPP sections)
- Zero-knowledge proofs for proprietary data protection

Knowledge Graph Integration: Grounded concepts via ontologies (e.g., EUDPP, ISO/GS1 terms)
SHACL + Rule Validation: Enforce field constraints, logic, compliance
Output: JSON-LD DPP, PDF summary, digital twin integration, decentralized storage options
Dataspace Integration:

- DPPs operate across different dataspaces, enabling semantic interoperability.
- LLMs automate data mapping and integration between dataspaces.
- Semantic models can be generated in shared RDF/Turtle format.
- Data sharing is facilitated through integration with Dataspace Support Centers (DSSC).

LLM-Powered DPP Core Ontology: Enabling Semantic Interoperability





Modular Ontology Architecture (CIRPASS-2):

- **DPP Core Ontology**: Formal specification of cross-sectoral concepts
- **Sector-Specific Ontologies**: Battery, Textile, Electronics, etc.
- **Industry-Level Ontologies**: Low-level data structures aligned to sector UC

Ontologies linked via semantic anchors (e.g., Schema.org, UNTP, GS1)

Ontology Methodology & Requirements:

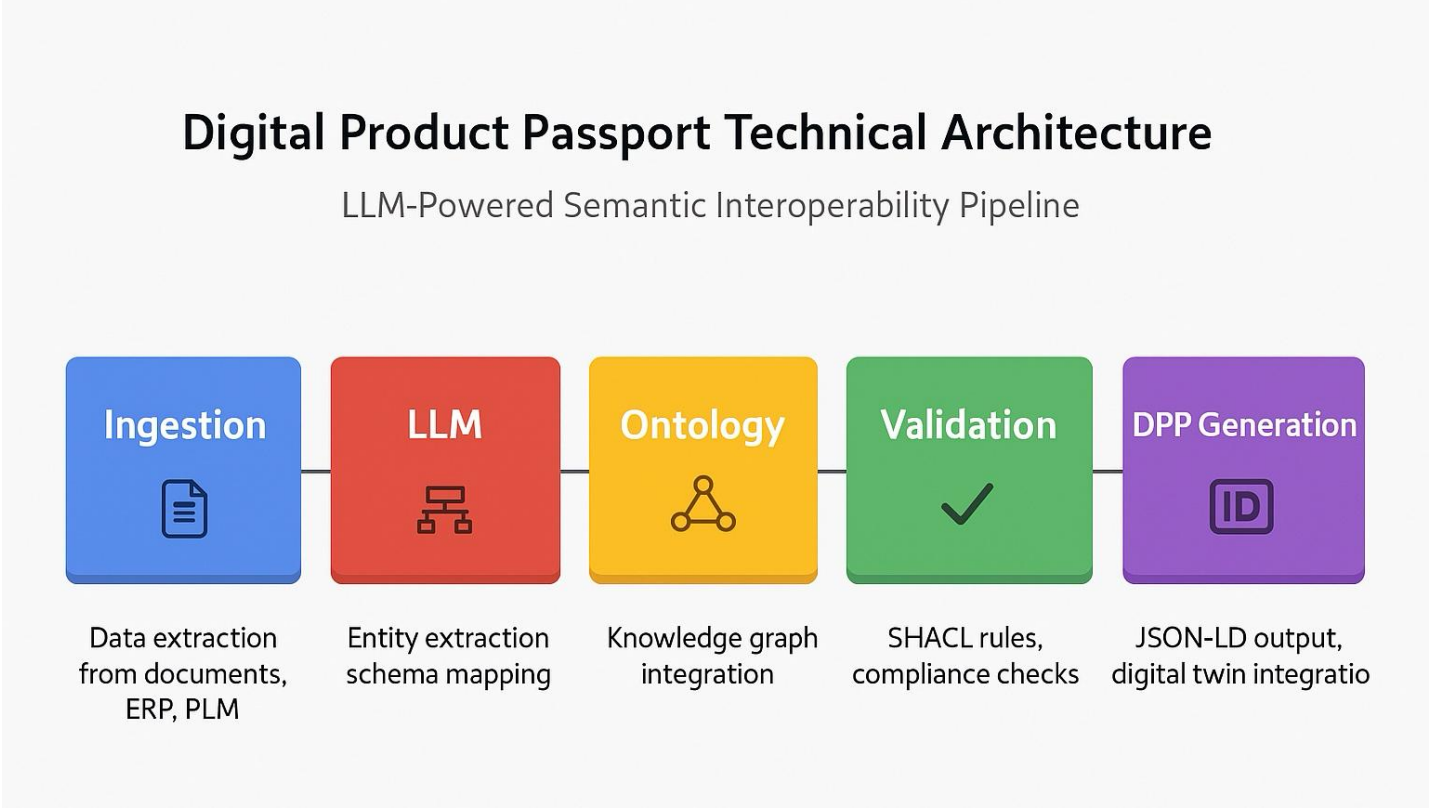
- **MOMo (Modular Ontology Modeling)** for scalable, domain-aligned module creation
- **Competency Questions (CQs)** define functional requirements (e.g., traceability, compliance, substances of concern)
- **Non-Functional Requirements (NFRs):** OWL 2 DL, JSON-LD support, FAIR principles, extensibility, multilingual support

LLM-Augmented Pipeline:

- **Input Normalization**: Extract from PLM/ERP/certificates using OCR, NER
- **Ontology-Driven Structuring**: Align terms, generate RDF/OWL modules, support CQ mapping
- **Output**: DPP in JSON-LD, SHACL-validated, ready for the DPP Vocabulary Hub
- **Decentralized Storage + Zero-Knowledge Proofs**: Secure proprietary data across sectors

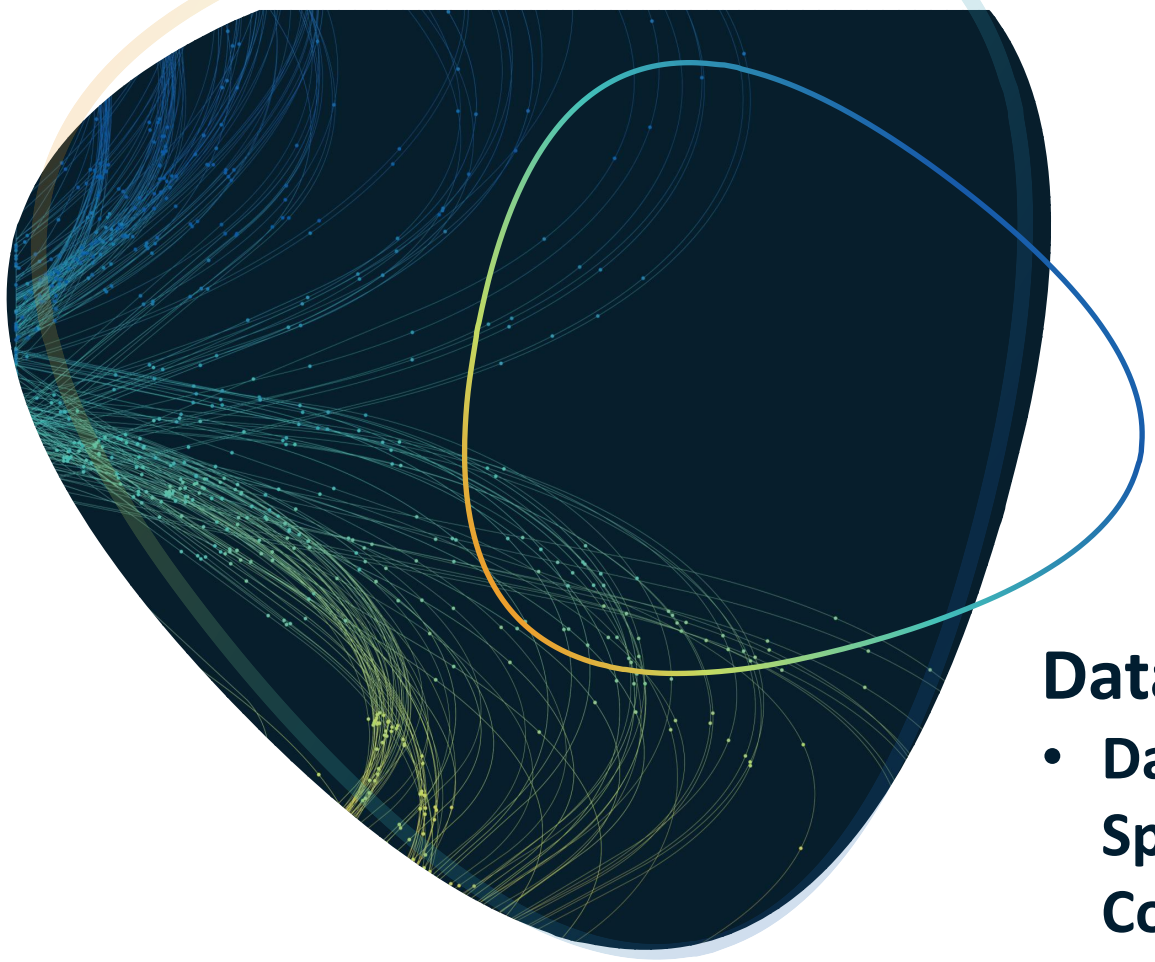
Standards Alignment & Future Integration:

- Compliant with **ESPR**, **Standardization Request 5423**, **EPCIS 2.0**, **CEN/CLC/JTC24**
- Prepares for **UNTP DPP compatibility**, **PROV-O/DQV-based data quality**, and **semantic data sharing via DSSC/IDS protocols**



Enhanced Features (2025)

- Retrieval-Augmented Generation (RAG) for LLM accuracy
- Validation feedback loop to correct hallucinated fields
- Integration with EPCIS 2.0 (ISO/IEC 19987:2024) event tracking
- Decentralized data storage options for enhanced security
- Zero-knowledge proofs for proprietary information protection
- GraphRAG frameworks for combining structured data with generative capabilities



Technical Architecture Diagram(Cont.)

Dataspace Integration Layer:

- **Data exchange aligned with International Data Spaces (IDS) protocols, Eclipse Dataspace Components (EDC) and DSSC Blueprint**
- **LLM-assisted creation and alignment of semantic models.**
- **Data translation and normalization between heterogeneous dataspaces.**
- **Data management aligned with FAIR principles (Findable, Accessible, Interoperable, Reusable).**

Example: Textile DPP Pipeline

Input: Fabric spec sheets, certificates, supplier declarations

LLM Processing:

- Parses fiber ratios, maps terms ("polyamide 6" → "nylon-6")
- Extracts sustainability metrics aligned with ESPR requirements
- Identifies substances of concern per latest REACH regulations

Outputs:

- Standard DPP JSON-LD + sustainability summary
- SHACL checks: mandatory fields, unit validation, vocabulary alignment
- Final DPP linked to QR code for consumer access
- Decentralized storage with selective disclosure capabilities

Example: Battery Passport Pipeline

Input: Material origin docs, BMS data, recycling events

LLM + KG Processing:

- Extract chemistry, emissions, lifecycle events
- EPCIS 2.0 event log parsing (manufacture → use → recycle)
- Carbon footprint calculation (mandatory from February 2025)

Enhanced Features:

- DPP includes provenance + SHACL-validated metrics
- DPP stored in public registry for compliance + second market
- Zero-knowledge proofs for sensitive supply chain data
- Real-time updates via IoT integration

Standards Integration

EUDPP-ESPR: Defines required fields, units, and ontologies (implementation from 2ESPR) **CEN/CLC/JTC24:** CEN/CLC/JTC24 is developing deliverables for the DPP framework and system

EPCIS 2.0 (ISO/IEC 19987:2024): Event tracking across lifecycle with enhanced interoperability

ISO/PWI 25534-1: ISO and UNECE joint working group for DPP Standards

SHACL: Data shape validation

GS1 Digital Link: URI access to product-level DPPs

ISO/IEC 19988 (CBV): Controlled vocabularies for semantic alignment

CEN/CLC/JTC24: Harmonized standards for DPP system (expected December 2025)

Dataspace Standards:

- Protocols defined by DSSC (BDVA, Fiware, IDSA and Gaia-X).
- GAIA-X federation services for trusted data infrastructure.
- Semantic vocabulary definitions using RDF/Turtle format.
- Shared ontologies for cross-dataspace semantic interoperability.

Deployment Considerations

Model Options:

- On-prem (e.g., Llama 3.1, Mistral Large 2) for sensitive data
- API-based (e.g., GPT-4.5, Claude 3.7) for quick prototyping
- Hybrid approaches with selective data processing

Security & Trust:

- Auditability: Prompt/output logging, field-source traceability
- Hallucination mitigation: Prompt constraints, RAG, rule validation
- Zero-knowledge proofs for proprietary information protection
- Decentralized storage options for enhanced data security

Versioning: Keep model + prompt history per DPP

Traditional vs LLM-Augmented DPP Pipelines

Feature	Traditional	LLM-Augmented
Data Mapping	Manual ETL	Dynamic extraction + mapping
Text Generation	Templates/manual	Automatic + explainable
Validation	Schema-only	SHACL + rules + feedback loop
Scalability	Costly per schema	Flexible across products
Transparency	Deterministic	Requires traceability layer
Security	Limited options	Zero-knowledge proofs + decentralized storage
Interoperability	Format-specific	Semantic understanding across formats

Closing

- LLMs accelerate DPP creation by enabling semantic automation across heterogeneous data sources.
- Hybrid architecture ensures compliance, auditability, and trust while protecting proprietary information.
- Standards alignment (CEN/CLC/JTC24 , ISO/IEC 19987:2024, EPCIS 2.0) is critical for the successful implementation of DPPs starting in 2025-2026.
- Beyond accelerating DPP creation, LLMs are also critical for enabling semantic interoperability across diverse dataspaces.
- Dataspace integration enhances the scalability and efficiency of the DPP ecosystem, allowing secure and meaningful sharing of product data throughout the supply chain.
- A hybrid architecture addresses key requirements for trust, compliance, and security in inter-dataspace data exchange.

Architecting LLM-Powered Digital Product Passports for Semantic Interoperability

SEMIC & BDVA Workshop

10.04.2025

Artech International
Rue de la Presse 4,
1000 Brussels

Suna Akbayır: suna@artechconsulting.net
Guven Fidan: guven@artechconsulting.net

interoperable europe