DIGICITIES USE CASES AND LEARNINGS

James Allan

Project End

ERA-Net Smart Energy Systems

This project has received funding in the framework of the joint programming initiative ERA-Net Smart Energy Systems, with support from the European Union's Horizon 2020 research and innovation programme.

INTERNAL REFERENCE

Deliverable No.:	D 4.4
Deliverable Name:	
Lead Participant:	Empa
Work Package No.:	4
Task No. & Name:	T4.1, T4.2, T4.3
Document (File):	D4 Digicities usecase learnings and opportunities_v0.3.docx
Issue (Save) Date:	2026-04-10

DOCUMENT STATUS

Version	Date	Author(s),	Description
0.1	2025-10-22	**James Allan**	Initial Draft and contents
0.2	2025-10-31	James Allan	Inputs from digital services need-owners
0.3	2026-03-26	James Allan	Executive summary and corrections

DOCUMENT SENSITIVITY

☐	**Not Sensitive**	Contains only factual or background information; contains no new or additional analysis, recommendations or policy-relevant
☒	**Moderately Sensitive**	Contains some analysis or interpretation of results; contains no recommendations or policy-relevant statements
☐	**Sensitive**	Contains analysis or interpretation of results with policy-relevance and/or recommendations or policy-relevant statements
☐	**Highly Sensitive Confidential**	Contains significant analysis or interpretation of results with major policy-relevance or implications, contains extensive recommendations or policy-relevant statements, and/or contain policy-prescriptive statements. This sensitivity requires SB decision.

TABLE OF CONTENT

About ERA-Net Smart Energy Systems

ERA-Net Smart Energy Systems (ERA-Net SES) is a transnational joint programming platform of 30 national and regional funding partners for initiating co-creation and promoting energy system innovation. The network of owners and managers of national and regional public funding programs along the innovation chain provides a sustainable and service oriented joint programming platform to finance projects in thematic areas like Smart Power Grids, Regional and Local Energy Systems, Heating and Cooling Networks, Digital Energy and Smart Services, etc.

Co-creating with partners that help to understand the needs of relevant stakeholders, we team up with intermediaries to provide an innovation eco-system supporting consortia for research, innovation, technical development, piloting and demonstration activities. These co-operations pave the way towards implementation in real-life environments and market introduction.

Beyond that, ERA-Net SES provides a Knowledge Community, involving key demo projects and experts from all over Europe, to facilitate learning between projects and programs from the local level up to the European level.

www.eranet-smartenergysystems.eu

1 EXECUTIVE SUMMARY

This deliverable documents the use cases that guided the development of the Digicities platform, the methodology applied in each, and the learnings that emerged from implementation. It is accompanied by two companion deliverables: D1 Infrastructure and Implementation, which details the system architecture and platform modules, and D2 Data Structures and their Application, which describes the ontology design and its extension across use cases.

The Digicities platform was developed and validated through three use cases spanning operational energy management and strategic energy planning. The Austrian UC1 (Energy Dashboard) structured building-level energy consumption and generation data for a performance monitoring dashboard, with component types aligned to the Brick ontology's taxonomy of energy consumers and generators. The Swiss UC1 (Hydro Scheduling) applied model predictive control to optimise hydropower plant production and battery storage management for a solar self-consumption community, using LightGBM-based demand and inflow forecasting trained on data organised through the platform's semantic layers. The Swiss UC2 (Regional Energy Planning) connected municipal planning data and technology databases to the Sympheny regional energy planning API through the platform's scenario builder and API submission modules.

Each use case exercised a different subset of the platform's capabilities. The energy dashboard validated the ontology's ability to accommodate external vocabulary alignment with minimal structural overhead. The hydro scheduling use case demonstrated end-to-end data integration from semantic organisation through ML-driven forecasting to automated control, achieving estimated monthly savings of 16 kCHF (AEM) and 96–220 kCHF (SES) through peak demand reduction, alongside systematic battery sizing analysis for a self-consumption community. The energy planning use case demonstrated the complete workflow from data product creation through scenario construction to external API submission, while revealing that stakeholder interest focused primarily on data governance capabilities rather than computational efficiency gains.

The stakeholder assessment identified a characteristic adoption pattern: digital service providers see immediate value in reduced integration workloads, standardised API data quality, and scalable multi-scenario handling, while energy planners and municipal stakeholders view migration costs and workflow disruption as primary barriers, requiring concrete case studies and quantified time savings before commitment. The technology potential analysis identifies the flexible ontology, semantic data products, decentralised workspace architecture, and service developer benefits as the platform's core value propositions, with semantic reasoning over ontology hierarchies and alignment with European Data Space initiatives (GAIA-X, IDSA) as priority areas for future development.

The consortium is considering releasing the full software stack as open source. While no formal business model for the platform software was developed during the project, this approach would enable community-driven refinement of the tooling and data ingestion methods, broaden the potential user base, and accelerate the accumulation of case studies that stakeholders identified as prerequisites for adoption. The commercial opportunity would then centre on the data products, integration services, and domain expertise built on top of the platform rather than on the software itself.

2 DIGICITIES USECASES

This section describes each of the Digicities use cases that guided platform development, the methodology and the learnings from implementation. For detailed descriptions of the system architecture and data vocabulary developed in the Digicities project, please see D1 Infrastructure and Implementation and D2 Data Structures and their Application that are published alongside this report.

2.1 Energy Dashboard (Austria)

2.1.1 Methodology

The Austrian use case focused on structuring energy consumption and generation data for a building-level performance dashboard at the AIT campus. The source data comprised historical time series from thermal and electric energy consumers (heating circuits, ventilation systems, lighting, appliances) and generators (solar thermal collectors, photovoltaic installations) within the building, which had already been conceptually mapped to the Brick ontology's distinction between energy consumers and generators.

The resulting knowledge graph provided the semantic structure for a dashboard application that could query energy flows by type (thermal/electric) and role (consumer/generator), aggregate consumption and generation across the building, and display historical performance trends. The simplicity of the attribute structure meant that the dashboard could be implemented with minimal ontology complexity while still benefiting from the semantic layer's ability to classify, filter, and aggregate data by component type.

2.1.2 Digicities learnings for energy dashboard use cases

The energy dashboard use case validated three aspects of the platform's design. First, the ontology extension mechanism accommodated Brick-aligned component types without requiring adoption of the Brick ontology in its entirety. The component names were mapped using equivalent or close matches, demonstrating that vocabulary alignment can be achieved at the component naming level while the underlying attribute structure follows the Digicities pattern. This is a substantially lighter-weight approach than importing Brick as a sub-ontology, avoiding the structural dependencies and maintenance burdens that full ontology adoption entails. The primary goal of Digicities is to handle the requirements of the service, nothing else.

Second, the use case confirmed that the platform can accommodate monitoring applications with minimal attribute complexity. Where the Swiss use cases required complex attribute structures spanning multiple types (Physical, Cost, ClassObject, CustomPhysicalRatio), the dashboard required only Annotations and Historic time series. The core ontology's type-based attribute classification handled this minimal case without imposing unnecessary complexity, confirming that the framework scales down as well as up.

Third, the use case highlighted a practical consideration for multi-national deployment: attribute naming conventions differ across linguistic contexts. The source data used German-language labels (Wärme Solar, Strom Photovoltaik Anlage) while the ontology component types used English names. The current approach is functional but raises questions about multilingual support that would need to be addressed for broader European deployment, particularly in the context of cross-border data exchange within Energy Data Spaces.

2.2 Hydro Scheduling (Switzerland)

2.2.1 Methodology

The hydro scheduling use case applied model predictive control (MPC) to optimise renewable energy resource scheduling for two small hydropower plants and evaluated battery sizing for a 10-household solar self-consumption community.

The data requirements were specified in an ingestion template containing five component types: HydroTurbine, HydroReservoir, ElectricityGrid, Weather, and MLModel. A distinguishing feature of this template was the prevalence of dynamic attributes carrying both Historic and Live temporal variants, reflecting the dual requirement of historical data for model training and live feeds for real-time scheduling. The ontology was extended to accommodate these components, with particular attention to the temporal data handling: the same physical quantity (e.g., reservoir inflow) needed to be simultaneously represented as a historical training dataset and a real-time operational feed.

LightGBM-based machine learning models were trained to produce 48-hour rolling forecasts of energy demand and water inflow. Demand models incorporated short-term and medium-term lags, calendar features, and weather forecasts. Inflow models used 48 separate estimators predicting incremental change from current values using date/time features, past measurements, and rainfall forecasts. For SES, installed PV capacity was included as an additional predictor to separate the effect of distributed generation from underlying demand. The scheduler optimised hydropower production over a 48-hour horizon, dispatching water during peak demand periods to minimise monthly peak loads.

For the battery sizing study, a simulation environment modelled the 10-household community with systematic combinations of PV capacity (51–205 kWp) and battery storage (0–200 kWh) across three seasonal months (April, August, December), with the MPC controller managing charging and discharging to maximise self-consumption while maintaining self-sufficiency.

2.2.2 Digicities learnings for hydro scheduling use cases

The hydro scheduling use case provided the most comprehensive test of the platform's data structures for ML-driven operational applications, and yielded several findings relevant to future development.

The ontology's temporal data handling proved well suited to the ML workflow. Training required historical datasets, real-time operation required live feeds, and the scheduler produced future projections, all for the same physical quantities (demand, inflow, production). The ability to represent these as temporal variants of a single attribute rather than as separate attributes avoided structural duplication and maintained a clear semantic link between the training data and the operational data it was trained on.

The forecasting results confirmed the value of structured, multi-source data integration. The LightGBM predictor consistently outperformed baseline models (statistical autoregressive, naïve constant-value) across all forecast horizons, with the improvement particularly pronounced when contextual features were included. This validates the platform's premise that organising and standardising diverse data sources through semantic layers enables analytical capabilities that are difficult to achieve when data remains fragmented across ad-hoc formats.

The quantitative results demonstrated financially significant impact. For AEM, the scheduler achieved an estimated average monthly saving of approximately 16 kCHF over four test periods. For SES, savings ranged from 96 to 220 kCHF/month depending on scheduling frequency, with substantially greater improvement when decisions could be updated every 15 minutes rather than daily — highlighting the operational value of automated control over manual scheduling. The battery sizing simulations showed that the Large PV / Large Battery (205 kWp / 100 kWh) configuration offered the best cost-effectiveness with full daily utilisation, while larger batteries exhibited diminishing returns. Forecasting accuracy proved critical: the LGBM model avoided erroneous winter battery charging from the grid that was observed with simpler models.

A practical learning concerned data quality and preparation. Weather forecast integration required careful temporal alignment and missing data handling relied on LightGBM's native capacity to work with gaps rather than imputation. The data processing effort for the first utility was approximately six FTE months, reducing to two months for second utility and two weeks for LIC as workflows and tooling matured. This learning curve suggests that standardised data product templates and reusable processing pipelines, offered by Digicities could significantly reduce the effort required for subsequent deployments at new sites.

Finally, the use case informed the ontology design directly. The five-entity structure developed for the scheduling service (renewable energy source, electricity grid, controllable asset, weather, ML model) acts as a reference pattern for how operational control applications can specify their data requirements through the platform's YAML service registration. The JSON structure used to call the scheduling service references both static parameters and dynamic data sources, supporting a clear separation between configuration and operational data that aligns with the ontology's attribute classification system.

2.3 Regional Energy Planning (Switzerland)

This use case addresses municipal planners' needs for transparent, updateable processes that minimize forecast uncertainty and accommodate evolving data sources (e.g., demands from EV charging stations) required for strategic energy planning. This use case will connect to an API of a software provider offering energy system optimisation. This use case addresses the needs of two stakeholders. Firstly, it provides municipal stakeholders a place to securely manage the data and prepare scenarios for energy studies. Secondly, Digicities aims is to minimise data integrations burden that typically complicates service provider interfaces. A summary of the overarching objectives of this use case is provided below:

- Support mandatory municipal energy reporting
- Track scenario assumptions and quantify uncertainty
- Provide planners with best-available, transparent data
- Enable pathway progress tracking and validation
- Facilitate evolution with new datasets

2.3.1 Methodology

The energy planning need-owners, a utility company and a municipal planner, provided reports and data from past energy modelling studies. In addition to this, they also shared the configuration files for their in-house energy model. The goal was to accommodate these processes using the Digicities platform.

The data and the reports helped the development of the ontology to accommodate their use case. In addition to this, the configuration files helped develop the core internal ontology driving the platform. The existing data was processed and transformed into a graph, timeseries and geospatial data was converted into csvs and geoJSON for storage on NextCloud. This involved storing timeseries into individual files. A knowledge graph of the existing data was created using the available data. Sympheny programmed an API for regional energy planning, and this was directly integrated into the API module of the platform. A service requirements YAML template was generated that referenced the components and attributes in the ontology, shown in Figure 30.

2.3.2 Digicities learnings for energy planning use cases

Energy planning is particularly suitable for the application of semantic technologies due to the inherent heterogeneity of influencing factors, and the network infrastructure is well-suited to representation using linked data. Energy planning studies integrate diverse data types, often spanning temporal horizons (historical performance, real-time operations, future projections), spatial scales (building-level consumption, regional generation portfolios) and disciplinary domains (technical specifications, economic parameters, regulatory constraints), creating semantic complexity that can be handled using ontologies.

The process of defining service requirements through YAML specifications revealed important lessons regarding abstraction levels. Initial prototype implementations employed specific component types limiting data capture scope to precisely matching instances. Future development should prioritise more general component and attribute specifications enabling broader compatibility. This places more responsibility on the role of reasoning through inheritance/inference of attributes. This evolution will proceed incrementally if service developers create APIs accepting data conforming to Digicities ontology and publishing semantic data products that broaden the catalogue's coverage. Service providers specifying requirements at higher taxonomic levels automatically capture a broad range of instances without needing to be overly specific, whilst semantic reasoning over external ontology mappings could further expand compatible data sources.

Scenario management proved conceptually familiar, as alternative system configurations already constitute established practice in energy modelling workflows. Planners routinely evaluate infrastructure investment options, policy interventions, and demand evolution pathways through comparative scenario analysis. The platform's capability to store multiple scenario variants in parallel alongside baseline system representations offers substantial value; however, practical questions emerge regarding scenario lifecycle management as scenario collections accumulate over iterative planning cycles.

Future development priorities include enhanced abstraction handling to accommodate varying detail levels. Often, strategic studies require aggregated regional representations, whilst operational planning demands component-level granularity. Supporting seamless transitions between abstraction levels through ontology hierarchies and automated aggregation mechanisms remains an open challenge.

Ultimately, success depends critically on ecosystem development: expanding semantic data product catalogues providing immediately useable datasets and diversifying service offerings that consume standardised semantic inputs. Ongoing deployment use cases will refine these workflows, validating the proposition that semantic interoperability fundamentally transforms energy planning accessibility and collaborative capacity.

3 NEED OWNER EVALUATION OF SERVICE PROVIDERS

The goal of Digicities was to connect different stakeholders across the value chain. This section looks at the success of Digicities through the perspective of the need-owners that collaborated on providing Digital services in the project.

3.1 Digicities and BeyondCivic

3.1.1 Scope of collaboration

TrustRelay is a product from Beyond Civic AG - a comprehensive platform to connect data providers and consumers. The TrustRelay platform provides a fair and trustworthy data ecosystem where users can govern their data, onboard users and perform transactions across data assets. TrustRelay was considered as a key part technologies integrated in Digicities, the role of TrustRelay is shown in Figure 1.

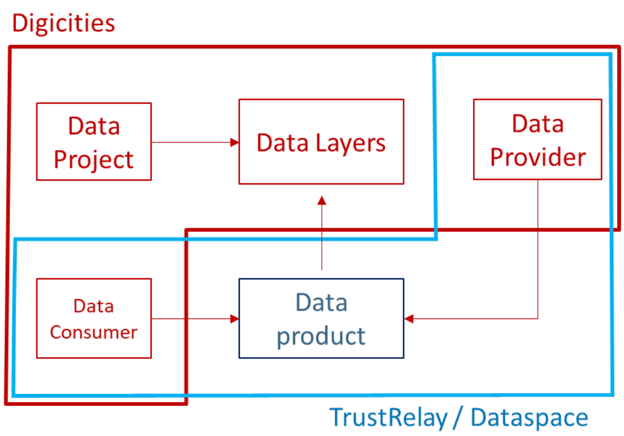

Figure 1: Overview of the intersection between Digicities and the TrustRelay.

3.1.2 Methodology

An environment was configured in TrustRelay and integrated into an early prototype of the Digicities platform. This allowed users to explore and ingest published data products listed on TrustRelay, into the workspace of Digicities. The process of ingesting data products is shown in Figure 2.

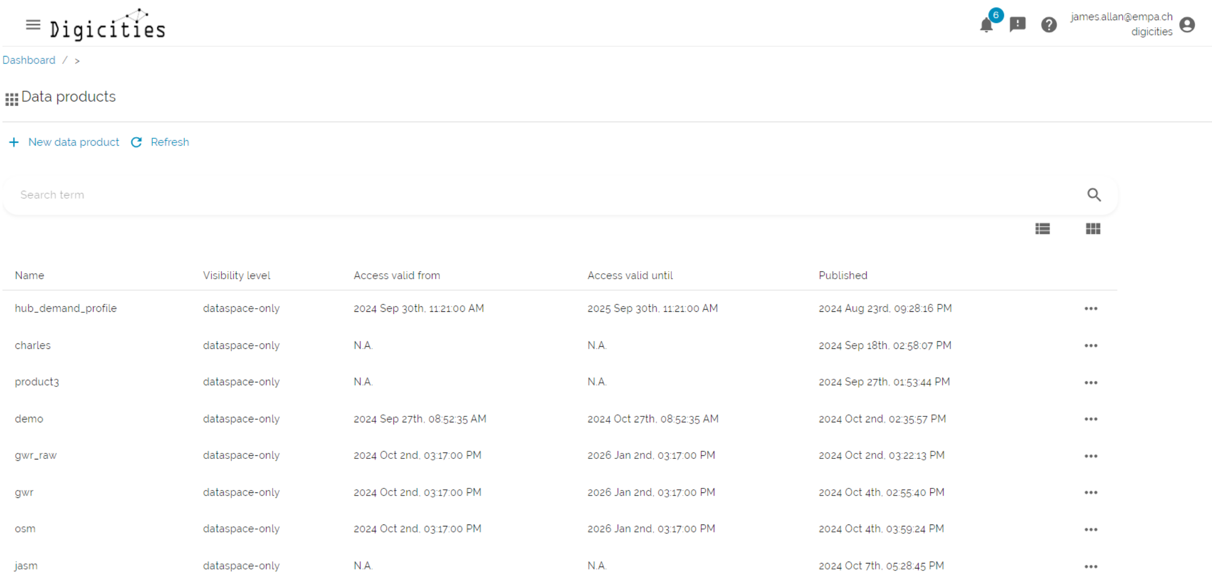

Figure 2: Ingesting data products into Digicities via the TrustRelay middleware

The user could also link to the TrustRelay platform to set-up user agreements, see Figure 3, view additional metadata, see Figure 4, and query the data, Figure 5.

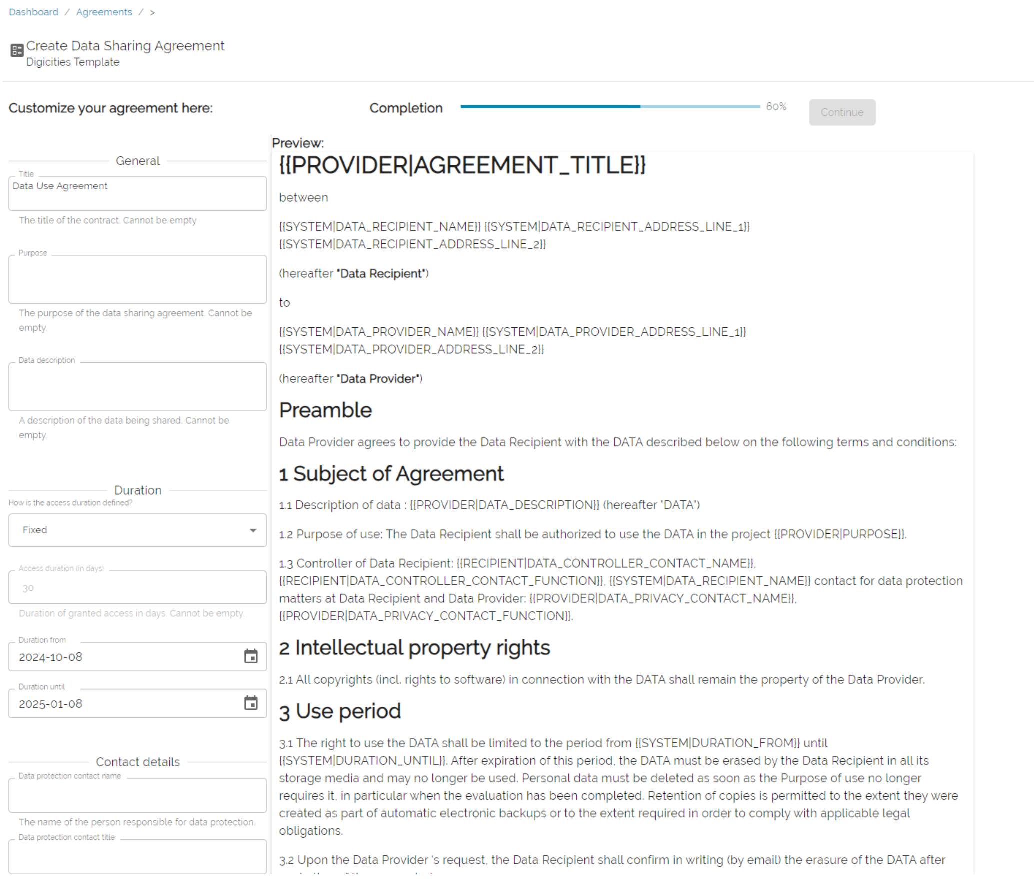

Figure 3: Setting up a data sharing agreement for exchanging data products on Digicities

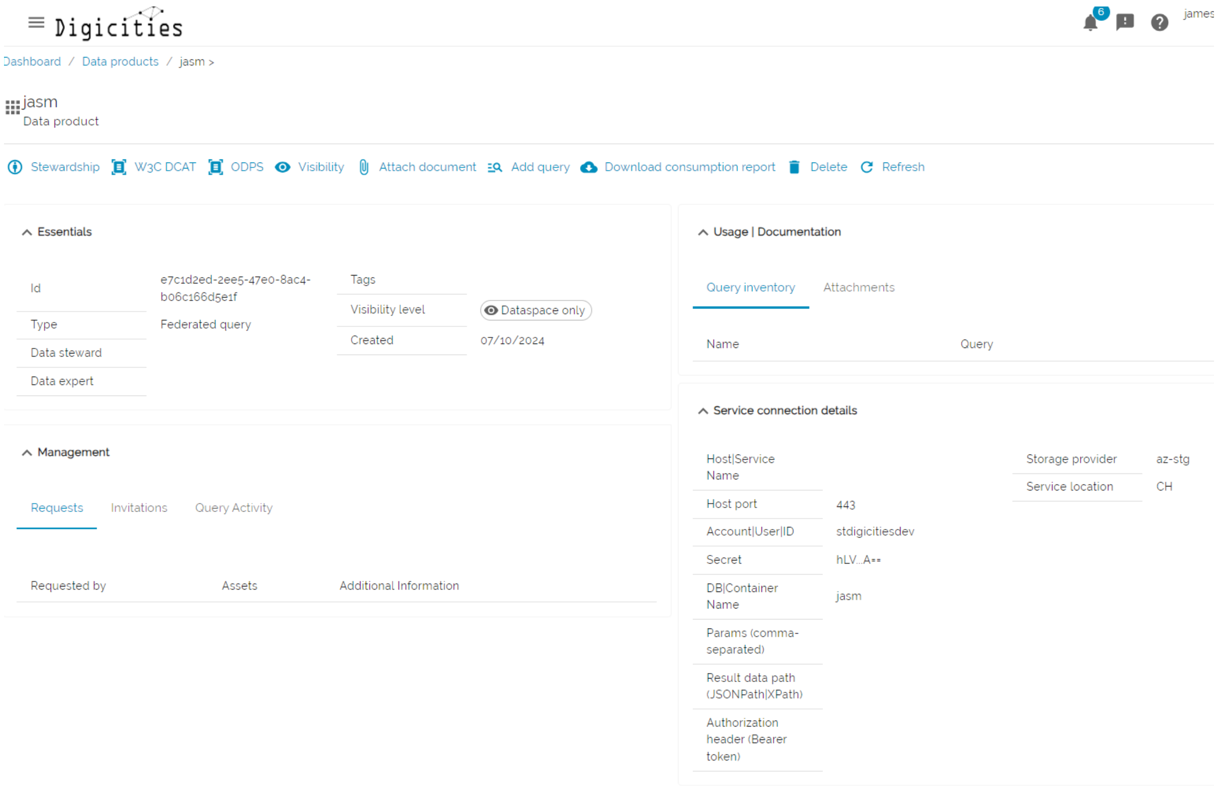

Figure 4: Detailed view of data product metadata within the Digicites environment of TrustRelay

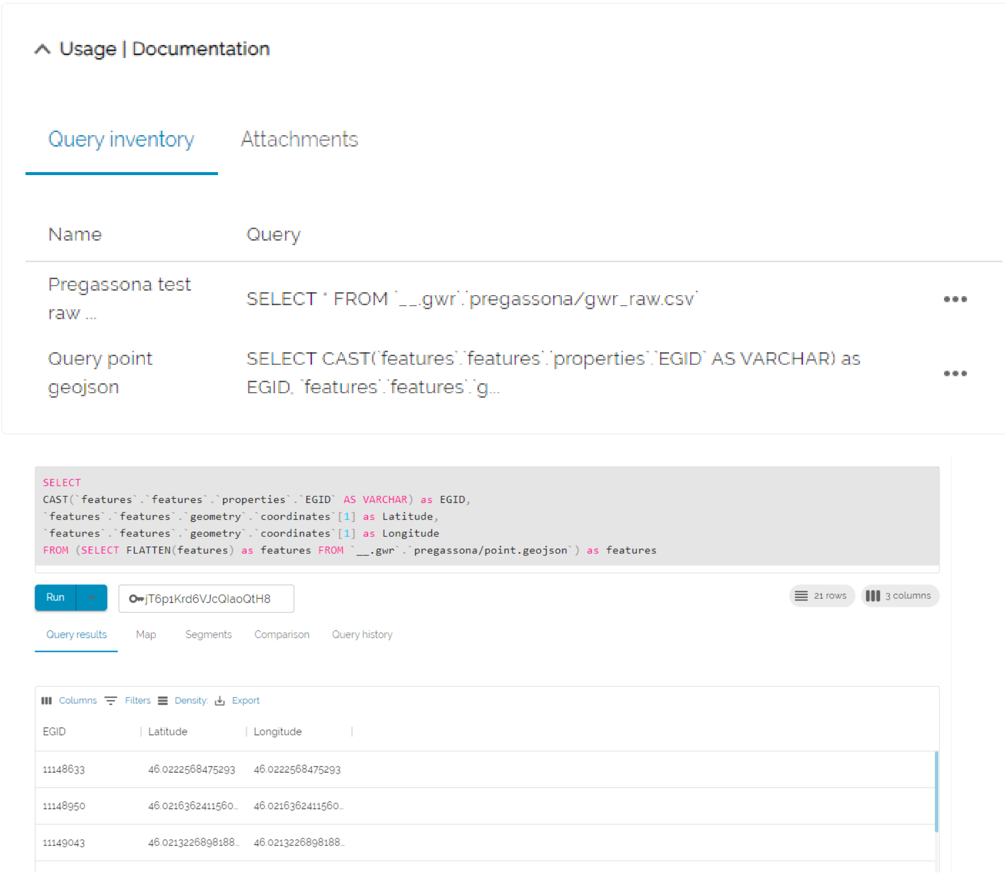

Figure 5: Data querying of data products using TrustRelay

3.1.3 Learnings and future work

In the latter phase of the project, development of the Digicities platform focused on the developing the core functionality. This also involved implementing a new identity management process, replacing the Azure active directory with Keycloak. This meant that it was no longer possible to connect with TrustRelay in the final prototype. Nevertheless, it the role the TrustRelay middleware could play in the Digicities platform was demonstrated. Future work should focus on supporting the new structure of the semantic data products in the data catalogue.

3.2 Digicities and Sympheny

3.2.1 Scope of collaboration

Sympheny is a software provider specialising in energy system optimisation. Their software helps optimise the energy supply and demand based using different technologies based on future scenarios. Their objective is to drive the energy transition by providing energy planners with the tools to make informed decisions on how to develop their assets to meet environmental and sustainability targets.

In addition to their core platform, Sympheny also provide an API for application developers to directly connect to the functions of the platform. The planned interaction between Digicities was planned according to feedback from workshops and exchanges with stakeholders, this interaction is shown in Figure 6.

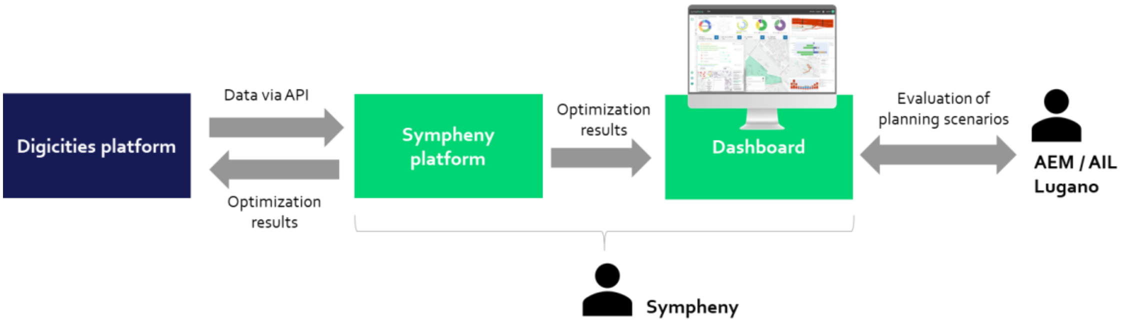

Figure 6: Interaction between Sympheny and Digicities

3.2.2 Methodology

The first step involved extending the ontology to accommodate the terms of a regional planning service hosted on Digicities. The YAML service registration is shown in Figure 7.


```yaml
service_name: Sympheny-Regional-Analysis
scenario_data:
  uri: Scenario.URI
  name: Scenario.label
  Region:
    name: Region.label
    uri: Region.URI
    EnergyConsumer:
      link: CL.Region.EnergyConsumer
      template:
        uri: EnergyConsumer.URI
        ElectricityDemandProfile: EnergyConsumer.ElectricityDemandProfile.hasHistoricTimeSeriesReference
    EnergyConverter:
      link: CL.Region.EnergyConverter
      template:
        uri: EnergyConverter.URI
        CAPEX: EnergyConverter.CAPEXPerRatedPower
        Efficiency: EnergyConverter.Efficiency
    EnergyGenerator:
      link: CL.Region.EnergyGenerator
      template:
        uri: EnergyGenerator.URI
        SolarPotentialProfile: EnergyGenerator.Power.hasHistoricTimeSeriesReference
    EnergyCarrier:
      link: CL.Region.EnergyCarrier
      template:
        uri: EnergyCarrier.URI
        EnergyCost: EnergyCarrier.EnergyCost
```


Figure 7: Service YAML for the Sympheny Regional Energy Planning

The API to receive the data was then configured on the platform. Scenarios were converted to JSON payloads and sent to the Sympheny API. Once the optimisation was complete, the unique dashboard URL was sent back to the Digicities platform and users could view the results. An example dashboard for the Sympheny Regional Analysis created through integration with Digicities is shown in Figure 8.

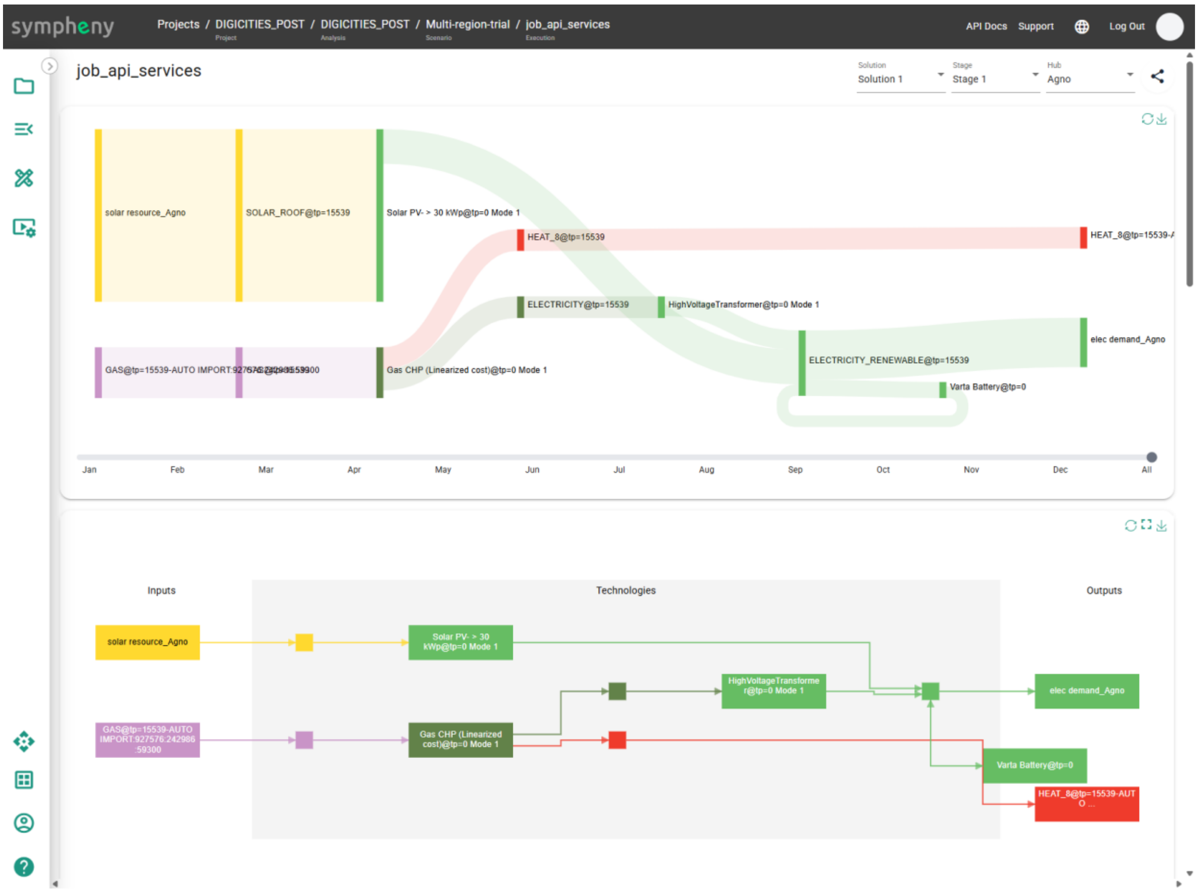

Figure 8: Dashboard generated for a scenario using Sympheny's Regional Energy Planning

In addition to this, the response from the API is stored on storage of the workspace, allowing users to return to previously simulated scenarios. The UI of past scenarios simulated on Digicities is shown in Figure 9.

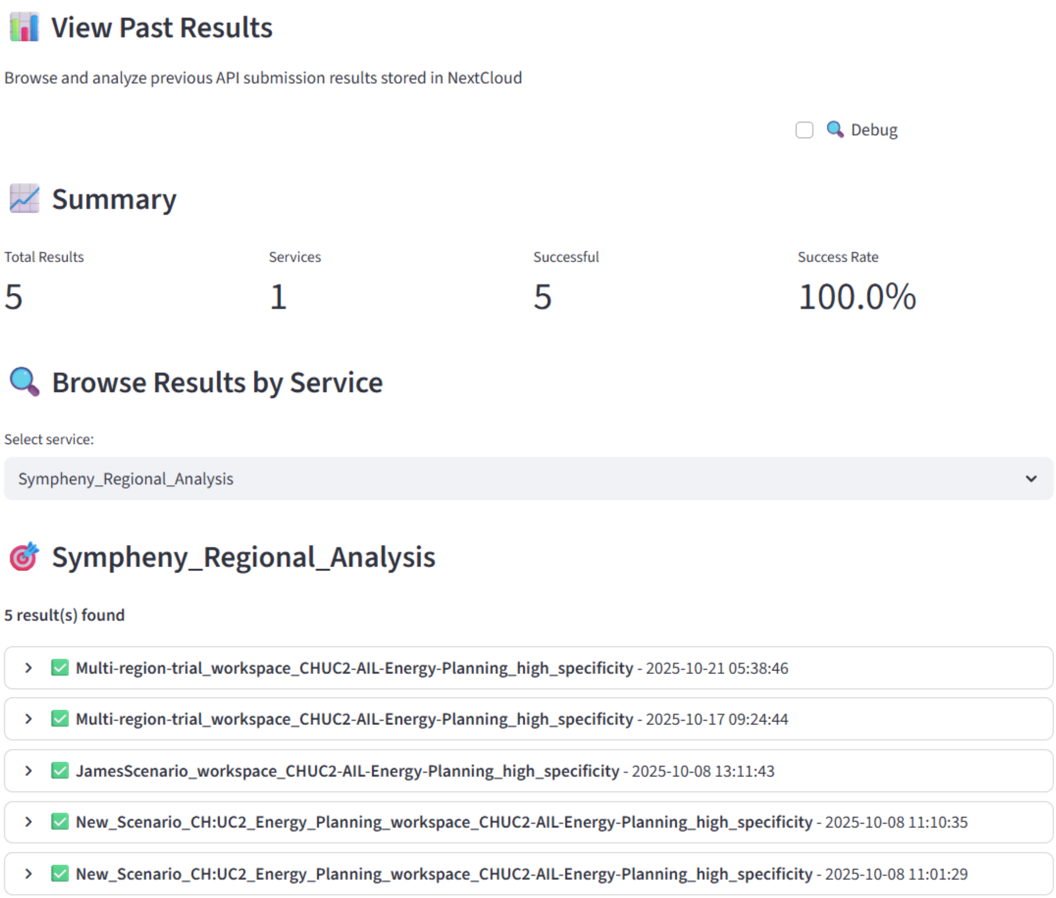

Figure 9: The result of simulations stored on the workspace

A need-owner expressed a need to extract demand profiles for electric vehicles and solar potential for specified regions. Sympheny had an API for such calculations, and we used this as an opportunity to demonstrate how data products could function on the platform and be ingested into different scenarios. To accommodate this, Sympheny configured two APIs for EV and solar façade profile calculation that took a geospatial boundary as an input. These demands would be directly useable in the Regional Planning Service as an instance of solar potential profile and energy consumer. A data product for each municipality in Lugano containing solar building façade potential and EV demand.

3.2.3 Learnings and future work

Sympheny's integration with the Digicities prototype platform demonstrated benefits for specialised energy system optimisation services. By connecting their regional energy planning API to the Digicities platform's standardised scenario submission framework, Sympheny reduces the burden of custom data integration typically required on their own platform. The semantic data product structure enabled clients to prepare input data using ontology-guided templates rather than mastering Sympheny's proprietary input schemas, lowering technical barriers for municipal authorities and regional planners. Future work will focus on extending new services and exploring new products to ingest and publish through the platform.

4 TECHNOLOGY POTENTIAL

This section gives an overview of the technical features of the platform and the value they could add to the current digital value chain.

4.1 Flexible Ontology

The Digicities ontology framework offers substantial advantages for application developers compared to traditional domain ontology adoption. Conventional energy ontologies such as CIM, CGMES, or SAREF present significant barriers: developers must invest considerable time understanding complex taxonomies, navigate overlapping terminology across multiple standards, and encounter situations where specific use case requirements do not align with predefined structures. Maintenance burdens compound these challenges as external ontology updates can break existing implementations, whilst importing multiple subontologies creates unmanageable conflicts and versioning issues.

The Digicities approach fundamentally inverts this paradigm by prioritising service-specific data requirements. Developers specify only the components and attributes their services require, creating lightweight, purpose-built extensions rather than adopting comprehensive domain models containing irrelevant concepts. The attribute-centric design provides consistent structural patterns for diverse data types whilst maintaining semantic precision through QUDT and official currency abbreviations.

Critically, internal flexibility does not preclude external interoperability. The systematic mapping framework enables developers to establish equivalence, hierarchical, or similarity relationships with established standards when integration becomes necessary, bridging internal models with external ecosystems on-demand rather than imposing universal conformance. This proves particularly valuable in multi-standard environments requiring simultaneous SAREF and CIM compliance.

The accompanying Digicities platform amplifies these benefits by abstracting RDF/OWL complexities behind intuitive interfaces. The Ontology Manager enables extensions without mastering semantic web technologies, the Replica Builder populates knowledge graphs through Excel imports, and the Service Registration module defines YAML-based requirements with automatic validation. The Scenario Builder enables rapid analytical variant configuration, whilst the API Submission module handles semantic-to-JSON/YAML conversion, eliminating custom transformation logic. This integrated toolchain transforms semantic modelling from specialised expert activity into accessible workflow, enabling developers to leverage knowledge graph capabilities whilst focusing on domain logic, ultimately reducing time-to-deployment and maintenance costs.

4.2 Decentralised workspace

The integration of Keycloak federated identity management within the Digicities platform architecture presents significant opportunities for future development toward truly decentralised data exchange paradigms. Whilst the current implementation employs centralised NextCloud storage infrastructure for all workspaces, a federated storage model could enable users to connect their own cloud storage services as direct replacements for NextCloud. This architectural evolution would allow users to authenticate through their organisational identity providers via Keycloak, which would then manage credentials for their connected storage backends. The entire workspace, including knowledge graph replicas, ontology extensions, scenario configurations, data products, and historical service results, would reside exclusively within the user's connected storage environment rather than platform-controlled infrastructure.

This approach would reconfigure the platform's data governance model, transitioning from a centralised to a distributed architecture where users maintain full sovereignty over the data of the digital replica. Platform services would operate as computational orchestration layers, executing TTL parsing, scenario generation, and API submissions against user-controlled storage endpoints through Keycloak-managed authentication credentials. Platform administrators would possess no direct access to user data, facilitating semantic interoperability and analytical workflows without data custody responsibilities. All read and write operations performed by platform modules would target user-specified storage locations rather than shared platform repositories.

This architectural pattern aligns closely with dataspace principles articulated by the International Data Spaces Association and European Commission frameworks, which emphasise data sovereignty and secure cross-organisational exchange through standardised protocols whilst maintaining owner control. Users could selectively share specific data products or scenario configurations with collaborators through native cloud storage sharing mechanisms, implementing granular access policies at the storage layer. GraphDB integration could similarly evolve toward federated architectures, where triple stores are deployed within user-controlled infrastructure and accessed through authenticated connections rather than centralised repositories.

4.3 Semantic data products

Semantic data products represent a structured approach to data publishing that addresses fundamental challenges in energy data sharing: inconsistent formatting, ambiguous semantics, and disconnected metadata. Each Digicities data product comprises two integrated components: a semantic graph containing instance data conforming to the ontology structure alongside any necessary ontology extensions, and mandatory resource files stored in cloud infrastructure including geospatial data (GeoJSON), time series datasets (CSV), or domain-specific standardised formats such as EPW weather files. The semantic graph provides formal definitions and relationships between components whilst resource files contain the operational data itself. Critically, resource files are referenced within the graph through explicit links, establishing provable connections between semantic descriptions and actual data assets. The components of a Digicities data product are shown in Figure 10.

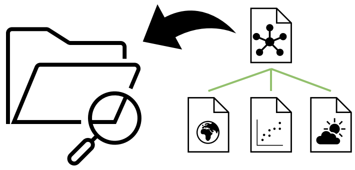

Figure 10: The components of a Digicities semantic data product

This architecture offers advantages over conventional data distribution approaches. Publishers package data with machine-readable semantics, eliminating ambiguity about units, temporal characteristics, spatial references, and attribute meanings that cause problems in ad-hoc CSV or spreadsheet exchanges. Consumers can programmatically validate data product completeness and semantic compatibility before integration. The Scenario Builder leverages this structure to enable direct ingestion of data product instances into scenario configurations, with automatic validation against service requirements ensuring only compatible data products, comprised of the correct components and attributes, can be selected. This reduces data preparation overhead compared to manual schema mapping and transformation workflows.

The separation between semantic descriptions and resource files supports flexible deployment patterns. Graphs can be queried independently for discovery and validation without downloading large time series or geospatial datasets, whilst resource files can be cached, versioned, or stored in optimised formats. A download client component enables third-party services to programmatically access data products ingested into workspaces, retrieving both semantic metadata and associated resource files through authenticated requests. This facilitates automated data pipelines where external analytical tools consume Digicities data products without manual file transfers.

For data publishers, this structure provides a standardised publication framework ensuring datasets are discoverable, interpretable, and useable. For consumers, semantic data products eliminate data wrangling bottlenecks, enabling focus on analysis rather than data preparation.

4.4 Benefits for developers of services

API service providers connecting to the Digicities platform gain advantages that reduce integration overhead whilst expanding market reach. Traditional service deployment requires extensive custom integration work for each client: developing bespoke data parsers, validating inconsistent input formats, mapping client-specific schemas to service requirements, and troubleshooting data quality issues that consume significant support resources. The semantic data product structure and Service Registration framework eliminate these bottlenecks by standardising how clients prepare and submit data.

Service providers define their data requirements using a YAML specification, which references the required ontology components and attributes, which the platform automatically validates when building a scenario. This ensures only semantically correct, complete records reach service APIs. The API Submission module handles format conversion from semantic TTL to service-specific JSON/YAML structures, eliminating the need for clients to understand proprietary input schemas or maintain custom transformation scripts. Services receive clean, validated, properly formatted data without investing in client-side integration tooling.

This standardisation aims to improve user experience. Rather than lengthy integration projects requiring data format negotiations and custom connector development, clients leverage platform tools to map their data to service requirements, with validation feedback identifying gaps before submission. Service providers reach customers who previously lacked technical resources for complex API integration, expanding addressable markets beyond organisations with dedicated integration teams.

Furthermore, the platform's scenario management capabilities enable clients to submit multiple analytical variants efficiently, increasing service utilisation without proportional support overhead. Result persistence in structured formats facilitates downstream analysis and reporting, adding value beyond raw API responses. For service providers, Digicities aims to transform API consumption from high-touch, bespoke implementations into scalable, self-service workflows that reduce acquisition costs whilst improving data quality and client satisfaction.

4.5 Enabling semantic reasoning of requirements

A critical area for future development concerns the implementation of comprehensive semantic reasoning capabilities to exploit the taxonomic hierarchies and external ontology mappings inherent within the Digicities framework. Whilst the current platform supports explicit specification of service data requirements through YAML configurations referencing individual component types and attributes, it does not yet leverage automated reasoning to infer applicable components through class hierarchies or inherit properties through external ontology alignments. Two principal reasoning patterns present substantial opportunities for enhancing platform capabilities and reducing configuration complexity.

4.5.1 Hierarchical Component Reasoning

The first pattern involves enabling services to specify data requirements at higher levels of abstraction within the component taxonomy, with automated reasoning mechanisms identifying all applicable subclass instances. For example, a regional energy analysis service requiring data on energy generation assets could specify requirements for the abstract EnergyGenerator component class rather than enumerating specific technology types. Semantic reasoners would automatically infer that instances of WindTurbine, SolarPanel, BiomassPlant, and other EnergyGenerator subclasses satisfy this requirement, enabling the Scenario Builder to present all relevant generation assets for selection regardless of specific technology implementations. This approach substantially reduces configuration complexity whilst maximising data capture, particularly valuable for services requiring technology-agnostic system representations or comparative analyses across diverse generation portfolios.

Hierarchical reasoning would similarly apply to attribute inheritance, where attributes defined at parent component classes automatically propagate to subclasses. Services specifying requirements for attributes at abstract levels would automatically recognise these attributes across all generator subtypes. This inheritance mechanism reduces YAML specification verbosity whilst ensuring consistency across component hierarchies, eliminating situations where service requirements must manually list identical attributes across multiple related component types.

4.5.2 External Ontology Property Inheritance

The second reasoning pattern concerns leveraging equivalence mappings to established domain ontologies for property inheritance and interoperability enhancement. The current mapping framework enables semantic alignment between Digicities concepts and external vocabularies through owl:equivalentClass assertions; however, these mappings are not currently utilised to infer additional applicable properties from external ontologies. Implementing reasoning over these equivalence relationships would enable substantial interoperability advantages.

Consider the mapping between dici_onto:WindTurbine and cim:WindGeneratingUnit established through equivalence assertions. With advanced reasoning, service developers could specify requirements for attributes defined within the CIM ontology directly, with the reasoning engine automatically inferring that WindTurbine instances possess all properties defined for WindGeneratingUnit within CGMES specifications, see Figure 11.

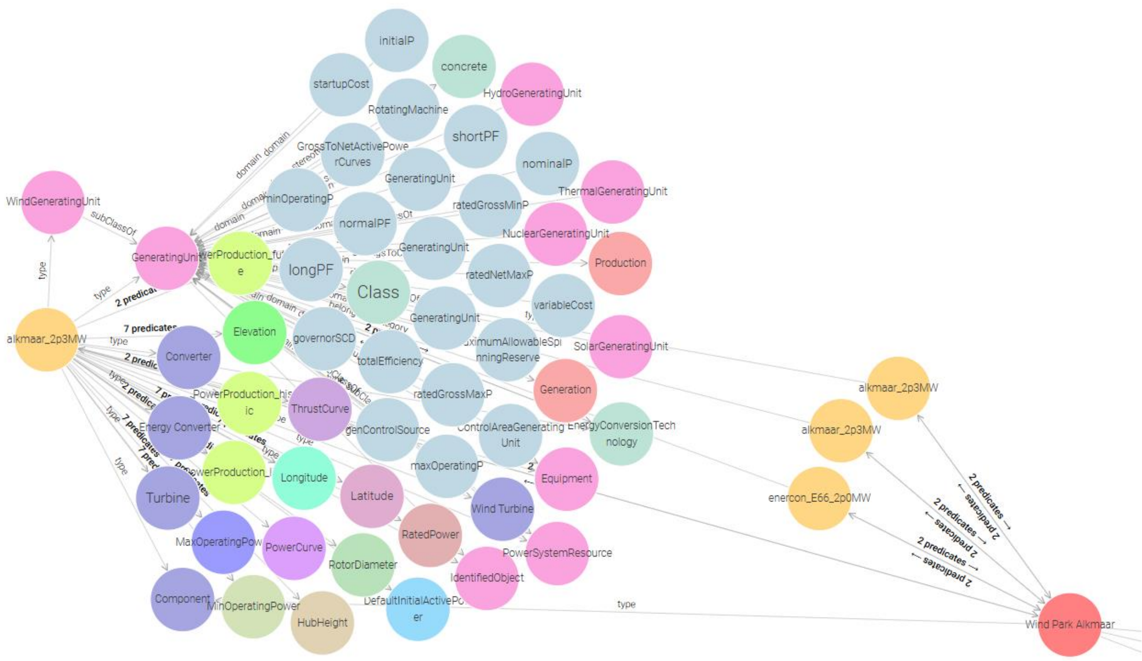

Figure 11: Mapping the instance of dici_onto:WindTurbine to the instance of WindGenerating unit in ENTSOE

This would enable services requiring CIM compliance to specify data requirements using CIM terminology whilst accepting data from Digicities-modelled instances, thus bridging semantic gaps between internal platform representations and external standardisation frameworks. Services could simultaneously specify requirements using multiple ontology vocabularies with reasoning mechanisms determining which Digicities component instances satisfy these heterogeneous requirements through their external ontology mappings.

4.6 Role in a dataspaces

The integrated Digicities technology stack, comprised of the flexible ontology framework, platform tooling, and decentralised federated storage architecture, which aims to conform to the fundamental pillars of the European Data Strategy[1] and the Rulebook of the International Data Spaces Association[2]. The IDSA have developed a reference architecture with a functional layer describing the principles of trust, security and data sovereignty; as well as an information layer focused on the semantics, covering structure, representation and exchange[3]. The semantics described in the Digicities ontology aims to conform to these requirements.

4.6.1 Semantic Interoperability

Energy Data Spaces must accommodate heterogeneity in data and data sources: diverse stakeholder types (utilities, aggregators, municipalities, industrial consumers) operating disparate systems with incompatible data models spanning multiple energy domains. The Digicities ontology framework directly addresses this challenge. By placing focus on the data requirements of services of the platform, the extendable core ontology provides consistent structural patterns without imposing comprehensive domain coverage, enabling rapid adoption by reducing complexity barriers that impede conventional ontology deployment. Simultaneously, the systematic mapping framework enables alignment with established standards, such as CIM for grid operations, SAREF for smart appliances, allowing stakeholders to maintain internal data representations whilst establishing semantic bridges when cross-organisational exchange becomes necessary.

This selective interoperability proves critical for Data Space viability. Rather than mandating universal adoption of monolithic standards, an approach that has often failed due to complexity and maintenance burdens, the proposed Digicities model enables incremental semantic alignment. Participants define internal structures meeting immediate requirements, then establish mappings to domain models when necessary. Advanced semantic reasoning capabilities would further reduce integration challenges by automatically inferring compatibility through ontology hierarchies and external alignments. This could also enable data discovery across

---

[1] https://eucrim.eu/news/commission-presents-european-data-strategy/ [21/10/2025]

[2] https://internationaldataspaces.org/idsa-rulebook/ [Accessed 21/10/2025]

[3] https://internationaldataspaces.org/offers/reference-architecture/ [Accessed 21/10/2025]

organisational boundaries without exhaustive manual mapping exercises across schemas, which has proven to be challenging [2].

4.6.2 Semantic Data Products as Standardised Exchange Units

The semantic data product structure provides a mechanism for standardised data publication within Data Spaces, addressing known challenges around data quality and usability [3]. It has also been reported that the key to successful data exchange lies in the mutual understanding of the data being shared, which is referred to as semantic interoperability [4]. Achieving this level of interoperability requires consistent semantic descriptions of the shared data.

To address this, the proposed Digicities Semantic data products, package operational data with machine-readable descriptions conforming to ontology structures, eliminating interpretative ambiguity whilst enabling programmatic validation of data completeness and compatibility. Data publishers describe offerings once using ontology concepts, with semantic graphs enabling automated discovery through federated query mechanisms across distributed catalogues.

This architecture naturally extends to federated data product catalogues where multiple organisations publish products to shared semantic registries without centralising actual data. GraphDB federation capabilities enable cross-organisational queries spanning distributed catalogues, with each participant maintaining sovereignty over their published products whilst contributing to collective discoverability. Usage policies and access controls remain enforced at publisher infrastructure, ensuring compliance with regulatory requirements and commercial agreements.

5 VALUE PROPOSITION

5.1 Future challenges

Whilst the Digicities technology stack presents a toolkit of technical solutions for semantic interoperability and decentralised data exchange, substantial barriers threaten adoption and long-term viability.

5.1.1 Ontology Governance and Evolution

The flexible ontology model, whilst reducing initial adoption complexity, introduces significant governance challenges as the user community expands. Multiple organisations independently extending the core ontology risk creating multiple overlapping and incompatible extensions which resembles the heterogeneity in data formats that the framework aims to resolve. Without coordinated governance mechanisms the ontology maintenance may devolve into proliferating incompatible variants requiring extensive mapping efforts. Establishing legitimate governance authority presents further complications: which stakeholder groups possess decision-making power over core ontology evolution, and how are conflicts resolved? Without dedicated resources for ontology maintenance, documentation, and community coordination, the framework risks becoming inconsistently applied, a burden to manage and ultimately abandoned.

5.1.2 Organisational Resistance and Value Chain Coordination

Energy sector organisations possess substantial investments in existing data infrastructure and established workflows. Adopting semantic technologies requires fundamental changes to existing data management practices and staff retraining. Without compelling business cases or regulatory mandates, decision-makers are likely to prioritise incremental improvements to functioning legacy systems over speculative adoption of emerging frameworks.

Realising Data Space benefits requires coordinated transformation across entire value chains, all stakeholders must simultaneously adopt compatible semantic frameworks and establish data sharing agreements. Individual organisations cannot capture benefits without widespread ecosystem participation, yet no individual actor possesses sufficient authority to drive universal adoption. This point was highlighted during the pre-piloting sessions with the energy planning stakeholders.

Current value chains feature misaligned incentives regarding data sharing, with operators viewing certain operational data as commercially sensitive whilst service providers may wish to maximise financial compensation for providing data. To achieve full its full potential regulatory mandates for standardised formats, financial incentives for semantic data creation and publication, liability frameworks that addressing data sharing risks are all required for the platform to realise its full potential.

5.2 Path Forward

The challenges outlined in the previous section reflect systemic obstacles confronting the broader Energy Data Space movement. Success requires recognising that semantic interoperability constitutes not merely a technical problem but an institutional challenge demanding policy interventions, business model innovation, and cultural change. This section outlines several key aspects necessary to keep further develop the platform to reach its full potential.

5.2.1 Alignment with Dataspace initiatives

Achieving the transformative potential necessitates collective action, including industry establishing shared governance frameworks, regulatory authorities mandating standards, public sector institutions providing foundational infrastructure, and educational systems developing semantic technology competencies. This is already being done in initiatives such as GAIA-X [5]. The Digicities proposal outlined an intention to align with such initiatives in this project; however, the platform was too immature to connect with these initiatives in the early phase of development. Now there is a prototype Digicities platform, future work should try to align with dataspace activities, for example, by connecting to dataspace connectors [6].

5.2.2 Integration of data-agreement middleware

Future development should prioritise integration of data-agreement middleware such as BeyondCivic's TrustRelay platform to enhance data product cataloguing and transaction management capabilities. TrustRelay provides comprehensive infrastructure connecting data providers and consumers through fair, trustworthy ecosystems where users govern data assets, manage onboarding, and execute transactions. Earlier prototype integration demonstrated feasibility, enabling users to explore published data products listed on TrustRelay and ingest them directly into Digicities workspaces. Enhanced integration would position TrustRelay as the federated catalogue layer managing data product discovery, licensing agreements, usage policy enforcement, and provenance tracking, whilst Digicities handles semantic validation, scenario composition, and analytical workflows. This separation of responsibilities, with TrustRelay managing data governance and transactions, and Digicities providing semantic tooling, creates a complementary architecture supporting Energy Data Space requirements for policy-controlled data exchange whilst maintaining the platform's focus on semantic interoperability and digital twin functionality.

5.2.3 Community data model development

The challenges related to ontology extension are detailed in the accompanying D2 Data Structures and their Application. This explores how such challenges could be addressed. Sustainable ontology evolution requires establishing community governance structures for Digicities core maintenance and extension development. Drawing on learnings from the WeDoWind[4] initiative [7], WeDoEnergy[5] community model could coordinate ontology development across multiple deployment contexts, including the Sweet CoSi[6] research project in Switzerland where Digicities core is being deployed. Such communities would facilitate collaborative discussion of proposed extensions, evaluation of integration into the core ontology versus domain-specific branches, and establishment of quality criteria for extension acceptance. Community governance mechanisms could include working groups for specific energy domains, regular review cycles for proposed modifications, transparent decision-making processes. This would complement the flexible, pragmatic approach enabled by the Digicities approach. This approach distributes ontology maintenance effort across stakeholder organisations spanning the digital value chain whilst ensuring extensions reflect genuine requirements validated through practical deployment, creating pathways for ontology evolution responsive to emerging energy system digitalisation needs.

5.2.4 AI-based data model development

Artificial intelligence tools present promising opportunities for automated ontology quality assurance [1,8,9] and consistency management as community-driven development scales. AI-based agents could continuously validate ontology extensions against core structural patterns, identify inconsistencies across independently developed branches, suggest harmonisation strategies when

[4] https://www.wedowind.ch/ [Accessed 20/10/2025]

[5] https://www.wedoenergy.org/ [Accessed 20/10/2025]

[6] https://www.sweet-cosi.ch/ [Accessed 20/10/2025]

overlapping concepts emerge, and flag violations of naming conventions or semantic patterns. Machine learning models trained on validated ontology structures could assist extension developers by recommending appropriate parent classes, suggesting relevant attributes based on component types, and identifying potential mapping candidates to external vocabularies. Natural language processing capabilities could analyse textual definitions and labels to detect semantic drift or ambiguous terminology requiring clarification. Where inconsistencies cannot be automatically resolved, AI systems could generate prioritised issue reports for community review, substantially reducing manual quality assurance burden. This human-AI collaborative approach combines automated consistency checking with community governance for complex decisions, accelerating ontology evolution whilst maintaining semantic rigour essential for reliable interoperability.

5.2.5 Open-source potential

The possibility of releasing the full software stack as open source was not part of the original project scope, but has emerged as a consideration given the project's outcomes. No formal business model for the platform software was developed during the project, and the stakeholder assessment identified limited willingness among end users to commit to adoption without a broader base of case studies and community validation. A proprietary licensing model would face the challenge of building this evidence base from a narrow user community, while simultaneously requiring sustained investment in platform maintenance and development.

An open-source release would address several of these challenges. It would lower the barrier to adoption for research groups, public-sector organisations, and smaller utilities who are unlikely to commit to proprietary platform licensing but could deploy and evaluate the tooling independently. Community-driven development would accelerate refinement of the ingestion methods, ontology extension workflows, and module-level improvements that the project identified as priorities. Independent validation and extension across a wider range of use cases and geographies would build the evidence base that stakeholders identified as a prerequisite for commitment, more rapidly than a single consortium could achieve alone.

This approach would preclude commercial licensing of the platform software itself, shifting the value proposition toward services built on top of the platform, such as, data product curation, data analytics, bespoke service integration, ontology extension & integration support, and training. This model aligns with the open-source approaches that have proven effective for comparable semantic infrastructure and with the European Data Space philosophy of open, interoperable tooling supporting a commercial data economy. Critically, open-sourcing the platform would not prevent commercial exchange of data products themselves, as the value of curated, validated energy datasets is distinct from the value of the software that manages them. The ontology, its extensions, and the community governance structures described in Section 5.2.3 would also benefit from the broader scrutiny and contribution that open-source development enables, accelerating the identification and resolution of the naming, interoperability, and governance challenges documented in the companion D2 deliverable.

6 CONCLUSION

The Digicities project has demonstrated through three use cases that a semantic data platform can provide practical value across the urban energy digital value chain, from structuring simple dashboard datasets to supporting automated operational control with financially significant outcomes (16–220 kCHF/month in peak reduction) and enabling end-to-end scenario submission to third-party analytical services. The technology potential of the platform lies in four reinforcing capabilities: a flexible ontology that reduces standards adoption barriers while preserving interoperability through mapping, semantic data products that eliminate ambiguity in data exchange, a workspace architecture aligned with data sovereignty principles, and service developer tooling that standardises API integration. Realising this potential beyond the project, however, depends on challenges that are institutional as much as technical: establishing community governance for ontology evolution, aligning with emerging Data Space initiatives, building the catalog of ready-to-use data products that makes the platform immediately valuable to new users, and creating the regulatory and commercial conditions that incentivise participation. The consortium's consideration of an open-source release reflects the recognition that these ecosystem-level challenges are more effectively addressed through community-driven development than proprietary licensing, and that the platform's long-term impact will be determined not by the sophistication of its software but by the breadth and quality of the data exchange it enables.

7 ACKNOWLEDGEMENTS

We would like to thank the need-owners and collaborators that have supported this project from the beginning. The generation of text in the report was assisted by AI tools Grammarly and Claude.ai for readability and structure.

8 REFERENCES

[1] Du R, An H, Wang K and Liu W 2024 A short review for ontology learning: Stride to large language models trend *ArXiv Prepr. ArXiv240414991*

[2] Rahm E and Bernstein P A 2001 A survey of approaches to automatic schema matching *VLDB J.* **10** 334–50

[3] Conde J, Pozo A, Munoz-Arcentales A, Choque J and Alonso Á 2024 Fostering the integration of European Open Data into Data Spaces through High-Quality Metadata *ArXiv Prepr. ArXiv240206693*

[4] Boukhers Z, Lange C and Beyan O 2023 Enhancing data space semantic interoperability through machine learning: a visionary perspective Companion Proceedings of the ACM Web Conference 2023 pp 1462–7

[5] Braud A, Fromentoux G, Radier B and Le Grand O 2021 The road to European digital sovereignty with Gaia-X and IDSA *IEEE Netw.* **35** 4–5

[6] Dam T, Klausner L D, Neumaier S and Priebe T 2023 A Survey of Dataspace Connector Implementations *ArXiv Prepr. ArXiv230911282*

[7] Barber S, Lima L A M, Sakagami Y, Quick J, Latiffianti E, Liu Y, Ferrari R, Letzgus S, Zhang X and Hammer F 2022 Enabling Co-Innovation for a Successful Digital Transformation in Wind Energy Using a New Digital Ecosystem and a Fault Detection Case Study *Energies* **15** 5638

[8] Mücke J and Scherp A 2025 GLaMoR: Consistency Checking of OWL Ontologies using Graph Language Models *ArXiv Prepr. ArXiv250419023*

[9] Zhang B, Carriero V A, Schreiberhuber K, Tsaneva S, González L S, Kim J and de Berardinis J 2024 Ontochat: a framework for conversational ontology engineering using language models European Semantic Web Conference (Springer) pp 102–21

FUNDING

This document was created as part of the ERA-Net Smart Energy Digicities project, which was funded through the through the framework of the joint programming initiative ERA-Net Smart Energy Systems' focus initiative Digital Transformation for the Energy Transition, with support from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883973.