

DIGIT / Digital Europe Programme

D03.02:

Finnish Roadshow Report

**DEP Action on Promoting semantic interoperability
amongst the European Union Member States (SEMIC)**

Specific Contract 098 under Framework Contract BEACON

Date: 23/05/2024
Doc. Version: v1.00



Document Control Information

Settings	Value
Document Title:	Finnish Roadshow
Project Title:	DEP Action on Promoting semantic interoperability amongst the European Union Member States (SEMIC)
Doc. Version:	0.01
Sensitivity:	Limited
Date:	23/05/2024
Status:	For review

Disclaimer

This report was prepared for the Digital Europe programme by PwC EU Services. The views expressed in this report are purely those of the authors and may not, in any circumstances, be interpreted as stating an official position of the European Commission. The European Commission does not guarantee the accuracy of the information included in this report, nor does it accept any responsibility for any use thereof.

Reference herein to any specific products, specifications, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favouring by the European Commission.

All care has been taken by the author to ensure that s/he has obtained, where necessary, permission to use any parts of manuscripts including illustrations, maps, and graphs, on which intellectual property rights already exist from the titular holder(s) of such rights or from her/his or their legal representative.

TABLE OF CONTENTS

Table of Contents	3
1 Executive summary	4
2 Meeting summaries Tuesday 20 February	4
2.1 Welcome of the roadshow	4
2.2 SEMIC Service Offering - The SEMIC team	5
2.3 Data quality framework – The Finnish team	6
2.4 Finnish Service Catalogue (Suomi.fi) - The Finnish team	7
2.5 Presentation of X-road and ABR - The Finnish team	10
2.6 Core vocabularies and Application profiles - The Finnish team	12
2.7 Semantic registry - The SEMIC team	14
3 Meeting summaries Wednesday 21 February	14
3.1 Interoperable Europe Act - The SEMIC team	14
3.2 Governance interoperability mechanism - The Finnish team	15
3.3 Life event-based digitalisation - The Finnish team	16
3.4 AI for Interoperability / Interoperability for AI – The SEMIC team	18
3.5 Data spaces – SEMIC Team	20
3.6 Closing	20
4 Appendix	21
4.1 Agenda	21
4.2 Group photo	23

1 EXECUTIVE SUMMARY

The SEMIC Roadshow in Finland, held on 20-21 February 2024, continued the series of roadshows done by SEMIC aimed to promote semantic interoperability among the European Union Member States (SEMIC). The event, organised by the European Commission and the Finnish Ministry of Finance, focused on facilitating knowledge sharing, collaboration, and networking opportunities for effective implementation.

The roadshow featured presentations on various topics related to semantic interoperability. The SEMIC team presented their service offering, which included support in areas such as data spaces, catalogue of services, base registries, artificial intelligence and interoperability policy implementation. This was complemented by SEMIC solutions in the categories of specifications, pilots, toolkits and the knowledge hub. Next, the SEMIC team dove into the different topics and welcomed the comments and feedback received from the Finnish counterparts.

The Finnish ministry presented different strands of work on their side by explaining work done, context, any challenges, future outlook. They presented their data quality framework, which aims to improve the usability of public data resources and enable better decision-making. They also discussed the Finnish Service Catalogue, which provides access to various services through an open API. Next, the presentation of X-road and Access to Base Registries highlighted the secure data transfer between public sector bodies in Finland. The Finnish team also discussed their interoperability platform, core vocabularies, and application profiles. Throughout these presentations, interactive discussions with the SEMIC team took place to explore collaboration opportunities and further enhance semantic interoperability.

Finally, the Interoperable Europe Act was presented by the Commission. It will promote the discoverability and usage of interoperable solutions across Europe, making sure that interoperability is considered before developing solutions through ex ante assessments.

Overall, the SEMIC Roadshow provided valuable insights for both the SEMIC team and Finland. It strengthened practical knowledge for promoting semantic interoperability in Finland and built a basis for future collaboration between SEMIC and Finland.

2 MEETING SUMMARIES TUESDAY 20 FEBRUARY

2.1 Welcome of the roadshow

The heads of the Finnish ministry of Finance and SEMIC opened the roadshow. The SEMIC team communicated the objectives of the roadshow and the importance of knowledge sharing and collaboration. The objectives include the sharing of best practices and learn from each other, promoting networking opportunities among stakeholders and providing practical insights to empower member states with knowledge and skills for effective implementation. The SEMIC team then summarised which focus areas SEMIC is active in, with a short introduction of each area:

- Data Spaces
- Catalogue of Services
- Base Registries
- Support in interoperability policy implementation
- AI4interoperability4AI

2.2 SEMIC Service Offering - The SEMIC team

The SEMIC team presented the [SEMIC service offering](#) in more detail, zooming in on each focus area. SEMIC delivers pragmatic support in these areas through specifications, pilots, a toolkit and the knowledge hub.

Following this presentation, several questions were asked by the audience. The following questions, or remarks were brought forward:

- An audience member with over 20 years of experience praised the long-term value of the work on interoperability, but pointed out that the core challenges remain unchanged, exacerbated by increasing legislative demands. There are also concerns about integrating AI with data spaces considering the AI Act requirements and using AI to develop semantics, which in turn would be used by machines. In response, The SEMIC team acknowledged the ongoing complexities, especially related to human language and data abundance, and clarified that AI is currently used by SEMIC as a supportive tool rather than for creating entirely new models, with its role in the application layer of Data Spaces potentially evolving.
- Another point raised by the audience was the varying (financial and human) resources among Member States for developing Access to Base Registries (ABR), with Finland highlighted for its advanced ABR systems. The SEMIC team mentioned the SEMIC Support Centre's (SSC) extensive resources on ABR, and the willingness to recommend advanced solutions. The SEMIC team added that SEMIC's landscape analysis aims to identify silos and varying maturity levels among Base Registries (BRs) to foster synergies and knowledge sharing.
- Another audience member asked about the relation between DCAT-AP 3.0 and data spaces. It was also highlighted that there are many variations of DCAT-AP today and Finland would therefore benefit from guidance on how to manage this complexity.
 - The SEMIC team replied that there are significant efforts being done to keep this coherence. SEMIC is not inventing new specifications around DCAT-AP but extensions which respect the definitions of the main model.
 - The SEMIC team added that there are webinars focused on DCAT-AP, and many others on the separate extensions such as GeoDCAT-AP or, for instance, the health data space that the SEMIC team is also supporting. SEMIC provides guidance and different levels of support based on the actual needs in different domains.
 - The SEMIC team complemented by emphasising that, in some cases, other parties such as DG MOVE started working on an extension themselves based on the SEMIC guidelines.

Summary of actions

Actions

- SEMIC to continue and share (once ready) the ABR landscape analysis to identify silos and maturity levels of different Base Registries;

2.3 Data quality framework – The Finnish team

The Finnish team introduced the data quality framework that the Finnish ministry has worked on for several years. There is a lot of public sector data available that would benefit from reuse, and the decision-making process depends on the quality of the data. Consequently, Finland has developed a national data quality framework so that data can be used for better decision-making and to improve the usability of public data resources.

The framework is a tool for data quality assessment that includes criteria, indicators, models and tools. Many stakeholders co-created it, and some pilots took place with these stakeholders. The scope of the framework includes open data, public sector exchange and other types of data exchange.

The criteria proposed by the framework are as follows:

- How well does information describe reality?
 - Correctness;
 - Accuracy;
 - Completeness;
 - Consistency;
 - Currentness.
- How has the information been described?
 - Traceability;
 - Understandability;
 - Compliance.
- How can I use information?
 - Portability;
 - User rights;
 - Punctuality.

Its implementation requires to focus on multiple levels: national level, organisational level and dataset level. The framework has considered FAIR and ISO principles. As for data governance, Finland faces a challenge in establishing a data governance that conforms to the Data Governance Act (DGA).

The results achieved by Finland with this data quality framework are relevant in data ecosystems and data spaces, internal market of data, and trustworthy AI. The Finnish team also inquired about the data quality governance within the Commission. The SEMIC team replied that ESTAT is the expert body but that the responsibility is spread across Directorate-Generals depending on their various domains.

The Finnish team emphasised that a significant challenge faced by the ministry is how to increase the awareness of the data quality framework. This necessitates the implementation of well-defined change management practices.

The presenters further discussed the relationship between vocabularies and the framework. They explained that the framework allows for the description of metadata standards, accommodating both existing standards and the reuse of core vocabularies. The framework's primary design goal is to streamline this process of metadata standard description.

Summary of actions

Actions

- Finland may be interested to explore the role of the European Statistical System (ESTAT) as the expert body responsible for data quality, and how it relates to other EU bodies such as DG CONNECT and DIGIT;
- SEMIC to raise awareness about the data quality framework among relevant stakeholders, emphasising its importance and potential benefits for data ecosystems.

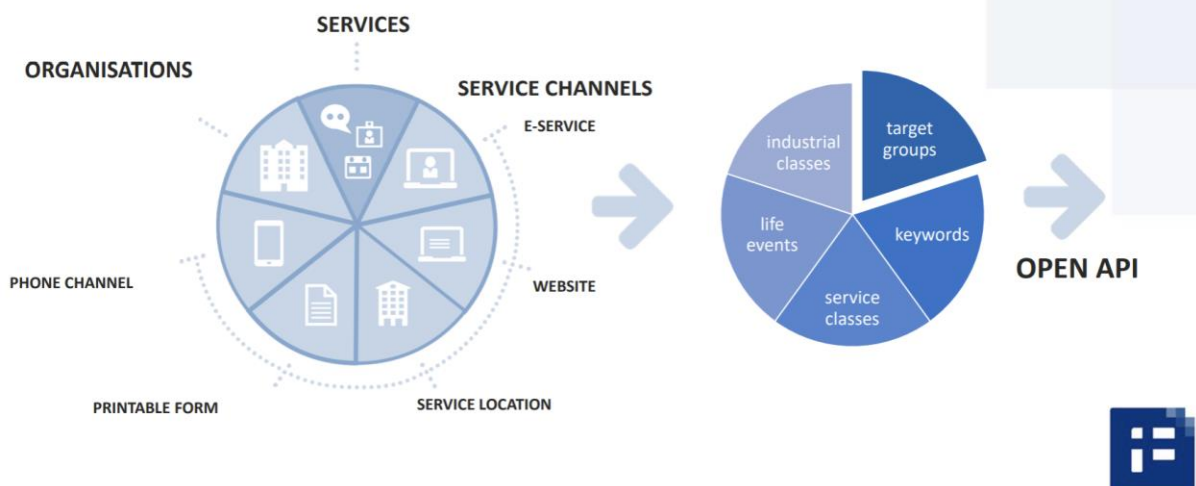
2.4 Finnish Service Catalogue (Suomi.fi) - The Finnish team

Finland presented the Service Catalogue, supported by the national act on '*central government's joint e-services support services*', which obliges municipalities and all public organisations to publish their service descriptions in one central repository (PTV).

Public organisations can update this information directly in the portal, i.e. in one central place from which the information can be used in all relevant web pages.

The Finnish service catalogue defined target groups, industrial classes, life events, service classes and keywords to find and access services. This information can be accessed through an open API.

Finnish Service Catalogue: centralised data repository



Key figures on the platform were shared:

- 25300 services
- 60200 service channels
- 1300 organisations
- 7600 users (civil servants) who update the data
- 325 million API requests
- 120 web services

The Finnish Service Catalogue (FSC) is based on an earlier version of CPSV-AP. There are no big differences between their model and CPSV-AP, but not all parts of CPSV-AP are reused in the FSC. With respect to updates, the FSC representatives expressed a preference for

minimising significant alterations to their data model. The FSC uses the Finnish Interoperability Platform for its code lists, data model and the FSC Glossary.

The FSC's data is used in different web services, for instance by municipalities' websites or public health services organisations.

On top of the FSC, Suomi.fi is a public web service that brings together all the data from the FSC. FSC's data helps, for instance, to answer requirements from the SDG, the Service Directive and Professional Qualifications Directive. It also makes use of a permit wizard.

Data quality was raised as a critical point, as the FSC's data is used in many web services. Finland also implemented some tools that help users to create quality data:

- Text checking tool, correction suggestions, data model validation
- Automatic email notifications for users to update the information.

There was a discussion on the benefits of having the FSC as a single source of truth in place. They have not quantified these benefits, but this could be a useful exercise, for example for communicating the benefits to other Member States.

They did observe that, instead of publishing data on their decentralised portals, municipalities are now used to upload their information to the national portal, after which they mirror the published data to their own portal.

For the FSC to succeed, Finland not only received the support from the top management but actively and persistently engaged with local actors (e.g. municipalities), building the catalogue together with some volunteering municipalities (i.e. champions) receiving dedicated support from the ministry. This approach, combined with sustainable support in the long run to municipalities to connect and use the FSC, enabled the ministry to reach the levels of usages introduced before.

Further, Finland gave a short presentation on OOTS.

In the context of the SDG Regulation, many interactions have taken place within Finland:

- Finland has had 25 events for 8 ministries to increase understanding and implementation support;
- They have had 150 workshops with competent authorities;
- They have held 120 regular meetings with Competent Authorities regarding the OOTS requirements;
- They have organised two SDG events (overarching webinar once a year to report on the current situation and share best practices) gathering more than 350 people.

Several challenges and findings were shared:

- Technical documentation from the European Commission is ever evolving, requiring flexible approaches;
- There are important volume differences in usages by citizens: some services have many thousands of users per year while others have only a few;
- Interoperability and the use of common standards; and
- The importance of collaboration and working together.

For Finland, 'Population' is one of the priority domains in the context of the SDG. Overall, the SDG national coordinator agreed with that the current approach and roadmap for the different

subgroups (including Population, Education and Vehicle as priority domains) as currently presented.

Working towards OOTS 2.0, several needs are identified:

- **Stronger identities:** Identity matching is challenging in Finland. There is a need for permanent identifiers, supported by additional attributes on top of eIDAS, but this was not a viable option for all European actors;
- **European-wide mandates for public solutions for citizens and businesses:** otherwise, it requires Finland to create this mandate with new national laws;
- **Structured data thanks to common data models, vocabularies and classifications** which would be cross-border by default;
- **Legal basis for data processing**, supported by the European mandate.

The SEMIC team initiated a series of inquiries directed towards the Finnish team. They began by asking whether the rate of change associated with the CPSV-AP, alongside the need to ensure backward compatibility, posed a challenge for the FSC model. The Finnish team responded that this was not a concern, as they had reviewed new versions of the CPSV-AP but found no need to implement corresponding changes within their own system.

Next, SEMIC inquired about the impact of Open API usage on the frequency of updates required for the FSC data model. Finland explained that updates are infrequent and that the primary responsibility for users of their web services is to adhere, between others, to Open API requirements.

The SEMIC team then explored the strategies Finland used to navigate change management and engage decentralised stakeholders. Finland acknowledged initial difficulties but noted that the process became smoother once successful implementations provided positive examples. While budgetary support was helpful and necessary, they emphasised that the most significant factors in gaining stakeholder buy-in were a highly supportive work environment that fostered champions and the organisation of numerous stakeholders contact points, such as webinars. Key benefits highlighted to stakeholders included the ability to manage information in a single location and access to additional services like translation and intuitive service description templates.

Shifting focus, SEMIC asked about Finland's expectations. Finland expressed interest in SEMIC providing automated notifications mechanisms for service description maintenance and facilitating the exchange of best practices among Member States. They also suggested that SEMIC could publish a success story or roadmap highlighting Finland's digital public service catalogue development process. Additionally, they proposed the idea of SEMIC clustering countries with similar contexts, acknowledging the risk of isolating clusters but emphasising the value in initiating discussions.

SEMIC then asked whether Finland had needed to update the FSC's data model to meet the SDG requirements from Your Europe or OOTS. Finland reported that minimal updates were necessary, allowing them to reuse the model without issue.

Finally, SEMIC explored the primary challenges hindering the transition to OOTS 2.0 and how SEMIC could offer support. Finland cited funding as a major obstacle and indicated that they will observe adoption trends among other countries before making significant decisions.

Summary of actions

Actions

- SEMIC to investigate opportunities to publish a success story based on the Finnish Service Catalogue, on which Finland could collaborate;
- SEMIC will keep the need to cluster countries with similar contexts for more relevant exchanges between them as a potential next step;
- SEMIC to share concrete best practices (i.e. data quality enhancing measures, automated notifications mechanisms).

2.5 Presentation of X-road and ABR - The Finnish team

The X-road environment in Finland is called the Suomi.fi Data Exchange Layer. This means that public bodies in Finland are mandated by law to use the Data Exchange Layer for data transfers between public sector bodies. X-road is a safe and secure way to transfer data between repositories. For every exchange, there is an agreement between the two parties that are both members of the X-Road (authentication). The data transfer happens via the public net, but X-road proposes constraints to enhance security.

Private and third sector societies, trusts etc. can use the data exchange layer if they consider it beneficial. One third of user organisations today are non-public sector players.

The main use cases are as follows:

- Real time access to base registries and centralised repositories;
- Connection between services, service providers and service consuming organisations together;
- Minimisation of the need to store and duplicate data locally;
- One of the enablers for the OOP and smooth public digital services.

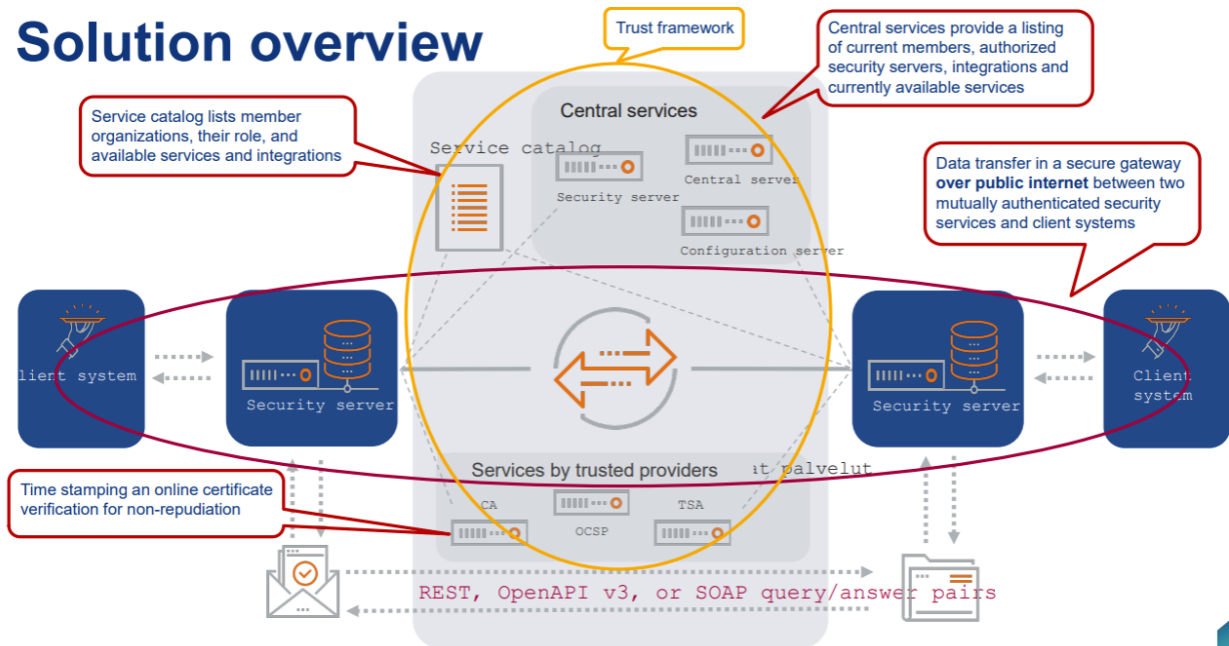
And key numbers were presented:

- Live since 2015;
- 730+ connections to exchange data;
- 330+ orgs connected directly;
- 30+ service providers offer access to services;
- 360+ distinct services;
- Close to 120 million queries in 2023.

The Finnish team shared upcoming developments:

- X-road has a 'cloud first' strategy. For instance, Microsoft Azure will need to be enabled. Finland does not expect major difficulties on that front;
- Considering the volumes, automated certificate management is required. The solution overview is presented below:

Solution overview



23/02/2024

Anssi Ahlberg, Finnish Data Exchange Layer 20 February, 2024

8



The Finnish team highlighted Finland's extensive progress and accumulated experience in the domain of ABR, inquiring about the long-term vision for the field. In response, the SEMIC team acknowledged the variations in legislation across different countries, emphasising the right of each nation to maintain its own distinct legal frameworks and political considerations. They suggested that SEMIC and Finland could collaborate in disseminating success stories and best practices with other countries. However, they clarified that SEMIC does not have the mandate to establish a universal, fixed plan for ABR implementation.

The Finnish team also highlighted some general pitfalls to show that Finland's approach, where public sector bodies enjoy wide access to each other's data by default, has some issues the public sector bodies face when exchanging data with each other when carrying out their duties:

There are many legal obstacles, related to privacy and security in data sharing (both national and cross-border data exchange)

- There is a risk for misuse of data. When data flows easily, we easily get in a grey zone where data can be used for other purposes than what it was intended for.

Despite the challenges described, the SEMIC team highlighted that the public and private connections and queries enabled by the Finnish Data Exchange Layer demonstrates key characteristics of a data space, such as identified in the DSSC blueprint.

Furthermore, on the data spaces, the SEMIC team posed questions regarding Finland's engagement with existing data space protocols. Finland confirmed that they have examined these protocols, indicating that some components of their system will be gradually replaced with those defined at the European level. Others will be modified to ensure compliance or integrated with existing back-end elements, such as identity authentication.

SEMIC then inquired whether Finland maintains a data catalogue analogous to their service catalogue, designed in accordance with the Finnish data quality framework. Finland acknowledged that this exists only in a partial form. They explained that while administrative and service descriptions should theoretically address data content queries, these descriptions are often incomplete. Furthermore, they view their role primarily as a facilitator of data transfer, without in-depth examination of content.

Summary of actions

Actions

- SEMIC to share Finland's experience and best practices in ABR and data exchange with other countries to foster international collaboration and learning. This could also take place in the context of the data spaces.

2.6 Core vocabularies and Application profiles - The Finnish team

Finland presented their interoperability platform, guiding the audience through the main functionalities of the platform. This platform was created with the funding from Ministry of Finance to support semantic interoperability by design across all ministries.

When it comes to semantic interoperability, Finland has a lot of challenges in convincing the political level about its importance. The SEMIC team referred to the legal background and enforcement needed to enforce interoperability. For example, the Interoperability Act might act as a new window to highlight these discussions.

Finland also expressed the need for having Core Vocabularies in a machine-readable and **resolvable** manner. Concrete feedback was shared before the roadshow and the SEMIC team will meet this demand.

The Finnish team presented the Semantic Interoperability Framework, consisting of four main pillars:

- Controlled vocabularies (Terminologies)
- Core Vocabularies
- Application Profiles
- Reference Data (code sets)

Finland shared that their tool could become a part of the reusable solutions catalogued by the Commission.

The Finnish team then presented the beta version of the Data Vocabularies tool, which is one of the tools within the FI-Platform:

- **Core Vocabulary Integration:** Core Vocabularies are embedded within the platform, providing a foundation for Finnish model development.
- **Data Model Creation:** Users can initiate the creation of new data models through a 'Add a new data model' function. This includes options to select languages, define descriptive metadata (name, description, prefix, unique URI/IRI), and specify the information domain. Existing data models can also be imported if they are presented in Turtle, JSON-LD or RDF format..

- **User-Friendly Editing:** The platform features an intuitive 'editing environment' where users can reference external, resolvable data models. This functionality aligns with the desire for resolvable Core Vocabularies.
- **Core Vocabulary Utilisation:** Users have the ability to add classes (e.g., from a Core Vocabulary) directly into the editing canvas. Language selection is streamlined through a dropdown menu, and additional attributes and elements can be incorporated as needed.
- **Model Management:** The platform supports comprehensive model management, including the ability to search for existing models and export them in serialised form. Currently, the serialised forms supported are Turtle, JSON-LD, OpenAPI and RDF formats.

SEMIC would like to be in touch with Finland in the context of tooling for data modelling that does not start from UML. Finland welcomes the style guide and the discussions on data modelling using other methods than UML. Their source code is mostly available on Github.

Summary of decisions and actions

Actions

- The SEMIC team to explore collaborative opportunities in tooling for data modelling that does not rely on UML, potentially incorporating Finland's approach and methodologies;
- SEMIC to share Finland's experience, challenges, and successes in creating the tool with other Member States. This would be published on the SEMIC Support Center;
- SEMIC to look into having Core Vocabularies in machine-readable and resolvable manner.

Nordic Smart Government & Business - The Finnish team

The context of the program was presented. The focus is on sharing business information with a common API between all countries involved (Norway, Sweden, Finland, Iceland, Denmark). They have a common terminology published on the FI-Platform that describes the information in English (111 concepts defined). This was the first step taken. Next, they started modelling the information exchange. Eventually, they convinced their neighbours to jump on their tool and collaborate.

To facilitate this, they reused the CBV as much as possible. Any issues regarding the reuse of Core Vocabularies were solved efficiently with the SEMIC team.

After the presentation, a short discussion took place on the current way of using only one Open API to query data from several registries. The SEMIC team proposed possibilities for using LDES to simplify the way publishers can publish and make their data queryable. The Finnish team is interested to learn more about LDES.

Summary of actions

Actions

- Finland to share a screencast of the demo they presented regarding the Nordic Smart Government & Business.
- SEMIC to organise a meeting around LDES with the Finnish counterparts The Finnish team
- SEMIC to capture the reuse of SEMIC assets by Finland.

2.7 Semantic registry - The SEMIC team

The SEMIC team presented the work that is currently being done within SEMIC regarding a future EU-wide registry of semantic models. The SEMIC team presented the current scope of the registry, and the further approach and workshops planned. The SEMIC team invited Finland to contribute to the shaping of the registry, as they did during the kick-off of the activity.

Upon security concerns shared by the Finnish team, the distinction between the registry and the repository was clarified. The registry will not store the data from the national or external repositories but instead be a registry for all models harvested. The Finnish counterparts agreed with this approach.

The Finnish team inquired about the intended target groups and domains for the registry. The SEMIC team clarified that, in general, the registry will maintain an open approach, accessible across all domains. In response, Finland indicated that several interested parties exist within their country, including the National Health and Wellbeing Institute and the National Library.

Summary of actions

Actions

- SEMIC to stay in close contact with Finland throughout the development of the registry. Bilateral meetings can be planned when necessary.

3 MEETING SUMMARIES WEDNESDAY 21 FEBRUARY

3.1 Interoperable Europe Act - The SEMIC team

The SEMIC team presented the Interoperable Europe Act. It was mentioned that the Act is not yet fully adopted. It aims to improve cross-border interoperability with technology that serves all of Europe.

It wants to help governments, citizens, businesses and the innovation ecosystem through:

- consistent, human-centric EU approach to interoperability from policymaking to policy implementation governance structure to enable public administrations to work together and agree on shared interoperability solutions
- ecosystem of interoperability solutions for the EU's public sector so that public administrations in the EU and other stakeholders can contribute to and re-use such solutions, innovate together and create public value

The SEMIC team then mentioned that there are many synergies with the current EU digital legislation landscape. The (1) Digital single market, (2) Data governance and sharing, (3) Cybersecurity and (4) Innovation and technology were emphasised.

The Team also introduced more detailed elements of the IEA:

- Recognised reusable interoperability solutions
 - Interoperable Europe solutions
 - Mandatory share and reuse
 - Interoperable Europe portal
 - European interoperability framework
- Structured and co-owned EU cooperation
 - Governance (board, committee, competent authorities...)
 - Interoperable Europe agenda
 - Monitoring
- Digital-ready and interoperable-by-design
 - Interoperability assessment
- Strengthened interoperability support
 - GovTech and interoperability regulatory sandboxes
 - Policy implementation support projects, trainings, peer reviews

Several topics were subsequently addressed. An audience member inquired about the cascading of interoperability assessments throughout various official documents stipulating requirements on digital public services. They further questioned whether a solution derived from a previously assessed solution would necessitate a new assessment. The SEMIC team clarified that interoperability assessments are conducted ex-ante, meaning prior to implementing changes to an existing process or developing a new one. However, different stages within the legislative or development lifecycles could trigger the need for new assessments based on the established criteria. However, in such situations the newly triggered assessments are strictly complementary to “parent assessments”, i.e., an assessed requirements are not being re-assessed.

Another question focused on the Commission's perspective regarding the relationship between the IEA (Interoperable Europe Act) and legislation like the Data Spaces Regulation, which also incorporates interoperability provisions. The SEMIC team responded that competent authorities would likely benefit from consolidating the assessments and requirements of the IEA with those outlined in other applicable acts. Regarding the link to Data Spaces, DG Connect would be the primary regulatory body. The SEMIC team emphasised that "interoperability by design" is a guiding principle intended to unify all relevant legislation. Additionally, the acts differ in scope, with the IEA concentrating on public services and others adopting a broader transversal focus.

Summary of decisions and actions

Actions

Policy team to maintain the contact with Finland with regard to various aspects of the Act's implementation, notably the Interoperability Assessments.

3.2 Governance interoperability mechanism - The Finnish team

The Finnish team outlined Finland's national interoperability framework's organisational and governance structure, centred around public sector information management led by the Information Management Board, Ministry of Finance, and relevant administrative branches. The Ministry of Justice plays a key role in drafting and consulting on regulations, while a third pillar focuses on steering common service interoperability. The framework encompasses all shared information resources and data exchanges.

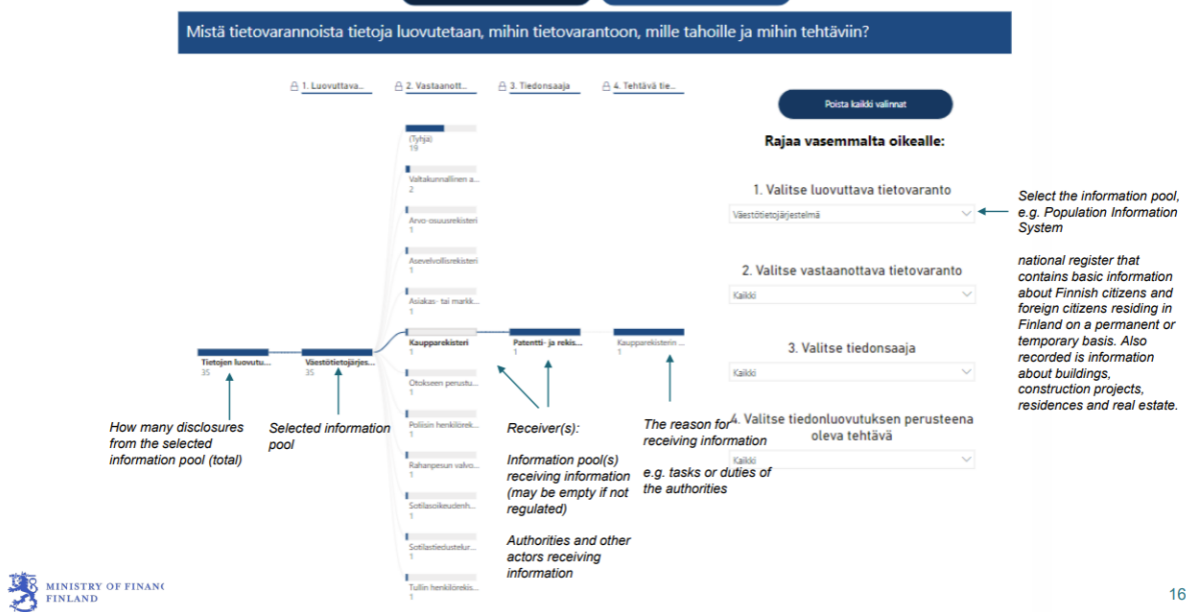
Finland already incorporates assessments into its governance, causing some uncertainty about implementing the Interoperable Europe Act (IEA)'s assessment requirements. The IEA necessitates a national authority to liaise with the European Interoperability Board, serving as the primary link between European and national levels.

Further emphasis was placed on the Advisory Board and the [Information Management Map](#), which offers a comprehensive overview of public administration information resources and their exchanges, managed by the Ministry of Finance in collaboration with other ministries.

- Updates to the map are provided biannually;
- Data is currently collected via Excel and visualised through a dashboard, highlighting the varied approaches to information resource management across ministries.

An example of how the data is visualised is displayed below.

Information exchange (visualisation option 1)



Finland advocates for extending this mapping to the EU level to harmonise the understanding of key information resources, base registries, and public services. Future steps involve enhancing data collection in machine-readable formats and employing AI and analytics tools to aid decision-making in information resource and business process management.

Summary of actions

Actions

- SEMIC and Finland to explore opportunities for support on the extension of the current Information Management Map. There are possibilities for collecting more information (e.g. about existing data flows or flows that are not captured such as the ones not falling under a law), in a machine-readable format, applying AI or other analytics

3.3 Life event-based digitalisation - The Finnish team

The Finnish team presented the work on life event-based digitalisation, which is based on the National Digital Strategy, a government programme of the prime minister and the underlying global trend of life event-based digitalisation.

The 'objective 11' (part of the National Digital Strategy) is to digitalise or automate public services with a human centric approach. So far, 40 of the most significant life event packages have been entirely digitised or automated.

The advantages of life event-based digitalisation of public services were brought forward:

- Dealing with several civil servants, companies or communities becomes easier;
- Information flow or joint use between different actors is optimised;
- Meeting the need for guidance to handle complex situations (e.g. environmental permit).

Several use cases have been identified as initial priorities within this new digitalisation effort. Primary emphasis will be placed on the management of affairs and estates related to deceased family members.

Key challenges to embark on this digitalisation include:

- **Interoperability**
 - There are different ways for actors to describe information and maintain it;
 - Operators have built their own solutions and information systems, e.g. base registries are built separately one to another, requiring considerable efforts a posteriori to reach interoperability;
- **Legislation** that slows down digitalisation because it must be updated, for instance to adapt the access rights to information, or to give the possibility to use electronic signature.

Several broad initiatives are being followed by Finland in relation to this work: data spaces, digital wallet, cross-border activities.

The SEMIC team inquired whether Finland has adopted a common reference architecture for all life events (extending the FSC data model) or if they have concentrated on specific architectures tailored to individual life events or groups thereof. Finland expressed a preference for a comprehensive architecture but acknowledged that, for practical reasons, initial efforts would likely centre on a more specialised approach.

An audience member from Finland questioned whether the public sector necessitates involvement in every data space or if a dedicated public sector data space would be a more effective strategy. SEMIC suggested that exploring the concept of making data more accessible to the public sector could prove valuable and proposed a webinar with a broader group of Member States to facilitate discussion. SEMIC acknowledged that they are also investigating the potential of a public administration data space, encompassing domains where public administrations play a leading role (such as public procurement and policymaking).

Finally, SEMIC referenced the recently published code lists for life events now published by the Publications Office of the EU ([Business events](#) and [Life events](#)). They invited Finland to propose any additional codes for inclusion and provided information on how to contribute to the list. SEMIC will share the [link](#) on how to contribute.

Summary of decisions and actions

Actions

- The SEMIC team, as part of its data space initiative, may organise a webinar for public administrations to exchange best practices.

- Finland may have a look at the life and business events code lists shared by the SEMIC team to see whether their experience and inputs could enrich it for all Member States.

3.4 AI for Interoperability / Interoperability for AI – The SEMIC team

The SEMIC team shared the SEMIC AI offering, vision and ongoing work regarding AI activities in SEMIC. Three areas were presented:

1. **Proof-of-concepts and Minimum Viable Products:** Concrete pilots or projects demonstrating how AI can complement and automate existing ways of working: extracting data insights in specific domains or facilitating semantic interoperability activities across domains.

Example: Pledge analysis tool that extracts and visualises relevant insights (topics, dates, actions) from a large set of pledges on the Transition Pathway for Tourism.

2. **Studies and research:** Comprehensive research endeavours aimed at investigating various aspects of Artificial Intelligence, interoperability, and advanced technologies. Learning material on how AI can support interoperability.

Example: Conference paper on how the retraining of BERT on tourism data impacts the quality of pledge clustering (latest work - in progress).

3. **Semantic Interoperability:** MLDCAT-AP is a semantic model in the Machine Learning field aimed to extend the use of DCAT-AP. It facilitates standardised descriptions of a machine learning process. The model was developed in collaboration with OpenML.

Regarding the structure of pledges and their associated information, an audience member inquired about lessons learned concerning data input. It was suggested that standardising the input format would streamline the process; however, this remains an ongoing development from the business perspective. To optimise algorithm training and utilisation, the recommendation was to leverage existing or easily accessible data as a starting point.

Presentation from Finland - The Finnish team

Finland presented their work on AI. Finland is actively experimenting with AI. They launched an inter-ministry discussion portal where thoughts and use cases on AI are discussed. They are thinking about use cases such as drafting communications, summarising consultations and machine translation.

Finland shared their findings and a common thread: everyone is interested in AI, but with a cautious attitude. There are questions regarding risks, liability and responsibility.

Finland has four considerations when using generative AI:

- Rapidly developing technology;
- No personal or confidential information as inputs;
- No output used before thorough review;
- No vendor lock-in.

Finland shared a first practical example: a tool named 'Aura', in use since Sept 2021. Aura provides machine translation from and to Finnish, English and Swedish for all civil servants. It is trained using specialised datasets and specific vocabularies.

Further, a second practical example was shared: a tool to summarise the results of a public consultation which will be launched in June 2024. This tool is important as conducting a public consultation is a mandatory phase in most legislative processes. The tool will analyse opinions and sentiments, it will summarise and provide source tracking. Finland is going to contract an external provider, which will probably use GPT4.0.

Next, a third tool, named Semantic Finlex, was discussed. It contains a lot of training data (publication system for legislation) and legislation as open data in a machine-readable format (2800 regulations, 60000 court decisions). Part of Finlex is available in English and as linked open data. The SEMIC team asked about how fine-grained the semantic annotation is. Finland replied that it is quite detailed, based on the national library ontology.

Subsequently, the SEMIC team mentioned that if they are interested, DIGIT is launching an initiative to support public agencies in the procurement of AI services.

Finland then presented a fourth use case, conducted by Vero, the tax administration. It interprets legal texts to answer questions received by the tax administration based on specific instructions. The Finnish development team was able to solve hallucinations or put it at a level where it was acceptable. The team used ChatGPT 3.5.

To the question 'How did Finland deal with accountability of public administrations?' the Finnish team answered humans must be involved in the different stages of AI: development, training, usage and maintenance.

Next, The Finnish team presented the study made by JRC on AI for IOP in the Public Sector. The idea of increasing data quality, and then semantic interoperability using AI raised their interest.

Finland identified two ways of AI for interoperability (IOP) and interoperability for AI:

- A top-down approach where you start from ontologies;
- A bottom-up approach where AI-based structuring can happen through an LLM.

The audience agreed that both approaches could be complementary.

Some challenges were brought forward:

- **Hallucination:** despite the positive results achieved, hallucination remains a general challenge;
- **Degree of certainty for successful execution** is always an open question with no scientific consensus at this stage;
- **Testing methods:** more work is needed but promising results such as by Vero in the tax domain.

The Finnish team then presented the algorithmic similarity analysis which tracks how items (i.e. sentences, set of words or specific concepts) flow across datasets and usages to identify the specific changes. This a concrete application of data lineage which could be an interesting area for data mapping.

They introduced the example of BLAST (Vesanto et al. 2017, 2019), where researchers managed to get some results on how the concepts travel across information sources.

Summary of actions

Actions

- SEMIC and Finland to explore possibilities for sharing success stories on the finalised AI projects done by Finland (e.g. Aura, leveraging the Semantic Finlex);
- SEMIC to follow up with Finland on the ongoing initiative by DIGIT to support public agencies in AI service procurement;
- SEMIC to investigate algorithmic similarity analysis as a method for data mapping, identifying how concepts "travel" across datasets and usages to support semantic interoperability.

3.5 Data spaces – SEMIC Team

The SEMIC team presented the SEMIC vision and service offering regarding data spaces. SEMIC focuses on three areas: (1) DCAT-AP, (2) LDES and (3) Pilots.

After the overview, the SEMIC team went into further detail on each of the three areas, starting with the service offering of DCAT-AP for the data space.

An audience member inquired about the relationship between the DSSC and SEMIC, and how to differentiate their solutions. The SEMIC team clarified that while both are funded or managed by the Commission, they operate as distinct entities. Additionally, DIGIT (responsible for SEMIC) is part of the ISSG and maintains regular communication with DG CONNECT and the DSSC. The SEMIC team emphasised that SEMIC, through its tools and semantic interoperability recommendations, has the potential to make substantial contributions to the development of data spaces. In fact, SEMIC holds the official designation of a strategic DSSC partner, and its tools are featured prominently in the official data space tool catalogue and referenced within the blueprint.

An audience member whose team is actively involved in developing the health data space, expressed interest in future collaboration with SEMIC. It was noted that DCAT-AP for health is a crucial area of focus for SEMIC within the health data space, presenting an excellent opportunity for Finnish contributions.

Next, the SEMIC team presented LDES for data spaces. The audience was aware of LDES and recognised its potential. A dedicated meeting could be organised between SEMIC and the interested Finnish parties to discuss LDES and its use cases in more detail. The Finnish team mentioned that the Finnish Ministry of environment could be particularly interested.

Finally, the SEMIC team presented the pilots related to data spaces. The link was made with an audience member who is part of the DeployEMDS action. Collaborations could be explored as part of the Mobility Solid Proof-of-Concept.

Summary of actions

Actions

- SEMIC to follow up with The Finnish team to explore collaborations on the Health Data Space;
- Finland to get in touch with the SEMIC team if concrete LDES use cases would arise. SEMIC to organise a session with Finnish counterparts interested in LDES and particularly the Finnish Ministry of environment who could be interested in collaborating.

3.6 Closing

The SEMIC team closed the session by thanking the Finnish colleagues for a fruitful and collaborative roadshow. The SEMIC team also informed the audience on the GovTech incubator. The European Commission invites experts from the field of Digital Government and GovTech to register in the [European Commission's experts database](#). From this database, the Commission will select candidates with the most suitable profile for the following activities:

- Evaluating [GovTech Incubator call](#) under Digital Europe Programme;
- Monitoring the implementation of funded projects;
- Participate in expert groups activities.

4 APPENDIX

4.1 Agenda



MINISTRY OF FINANCE
FINLAND

AGENDA

1(3)

20-21 February 2024

SEMIC Roadshow, 20-21 February 2024

Participants

European Commission

- Alexandra Balahur, Team Leader SEMIC
- Tomas Vagner, Interoperability Policy Team
- Anastasia Sofou, Senior Expert Semantic Interoperability SEMIC
- Florian Barthelemy, Project Manager for SEMIC from PwC
- Nathan Ghesquiere, Senior Technology Consultant from SEMIC from PwC

Ministry of Finance - all core participants come from Ministry of Finance, Public Sector ICT department

- Antti Helin, Ministerial Adviser
- Suvi Remes, Ministerial Adviser
- Viveca Still, Ministerial Adviser
- Anna-Mari Wallenberg, Senior Specialist
- Riitta Autere, Ministerial Adviser
- Arla Peltola, Coordinator

Tuesday, 20 February 2024

Location: Ministry of Finance, meeting room Paja 9:30-13:00 and meeting room Jakovara 13:00-17.15

Address: Mariankatu 9, 00170 Helsinki

09:30-9:45 Opening of the event

- Maarit Huotari, Ministry of Finance, Head of Unit, Senior Adviser for Legislative Affairs

09:45-10:15 Intro of the day, agenda and objectives

10:15-11:45 SEMIC Service Offering – Large audience plenary

- Alexandra Balahur and Florian Barthelemy, SEMIC-team

11:45-12:30 Data quality framework

- Kirsti Pohjanpää, Group Manager, Statistics Finland
- Olli-Pekka Rissanen, Chief Specialist, Ministry of Finance

12:30-13:15 Lunch at the Ministry of Finance

13:15-14:00 Presentation of Suomi.fi Finnish Service Catalogue

- Anastasia Sofou and Florian Barthelemy, SEMIC-Team

- Riitta Alkula, Chief Specialist, Terhi Tuokkola, Chief Specialist and Anssi Ahlberg, Chief Specialist, Digital and Population Data Services Agency
- Jarmo Kovero, Director, Development and Administrative Services Centre, KeHa Centre

14:00-14:45 Presentation of X-road + Focus on ABR

- Anastasia Sofou and Florian Barthelemy, SEMIC-Team
- Riitta Alkula, Chief Specialist and Anssi Ahlberg, Chief Specialist, Digital and Population Data Services Agency

14:45-15:00 Break

15:00-16:00 Core Vocabularies and Application Profiles + Style Guide

- Anastasia Sofou, SEMIC-Team
- Suvi Remes, Ministry of Finance
- Riitta Alkula, Chief Specialist, Digital and Population Data Services Agency
- Mikael af Hällström, Research And Development Specialist, Finnish Tax Administration

16:00-17:00 Tooling

- Nathan Ghesquiere and Anastasia Sofou, SEMIC-Team
- Suvi Remes, Ministry of Finance
- Riitta Alkula, Chief Specialist, Digital and Population Data Services Agency
- Mikael af Hällström, Research And Development Specialist, Finnish Tax Administration

17:00-17:15 Closing of the day

- Nathan Ghesquiere and Anastasia Sofou, SEMIC-Team

18:30-20:30 Dinner at restaurant Salutorget, address: Pohjoisesplanadi 15, 00170 Helsinki

Wednesday, 21 February 2024

Location: Ministry of Finance, meeting room Paja 9:00-13:00 and meeting room Jakovara 13:00-17.00

Address: Mariankatu 9, 00170 Helsinki

9:00-9:15 Welcome / Introduction

- Antti Helin and Viveca Still, Ministry of Finance

9:15-9:30 Intro of the day, agenda and objectives

- Antti Helin and Viveca Still, Ministry of Finance

09:30-10:45 Large audience plenary - Policy context and interoperable Europe Act

- Tomas Vagner, Alexandra Balahur and Florian Barthelemy, SEMIC-Team

10:45-11:00 Break

11:00-11:30 Governance mechanisms for interoperability in the Information Management Act

- Suvi Remes, Viveca Still and Antti Helin, Ministry of Finance

11:30-12:30 Life event based digitalisation

- Nico Käräjäoja, Project Manager, Ministry of Finance
- Anastasia Sofou, Alexandra Balahur and Nathan Ghesquiere, SEMIC-Team

12:30-13:30 Lunch at the Ministry of Finance**13:30-15:00 AI for Interoperability**

- Antti Helin and Anna-Mari Wallenberg, Ministry of Finance
- Alexandra Balahur and Nathan Ghesquiere, SEMIC-Team

15:00-15:15 Break**15:15-16:30 Large audience plenary - Data Spaces**

- Alexandra Balahur and Florian Barthelemy, SEMIC-team

16:30-17:15 Wrap up and end of the day – Antti Helin, Suvi Remes and Viveca Still, Ministry of Finance

4.2 Group photo

