A report on the Workshop "Towards Materials and Manufacturing Commons - the enablers Digital Marketplaces, FAIR Principles and Ontologies"

Joana Francisco Morgado[1], Silvia Chiacchiera[2], Yann Le Franc[3], María Poveda-Villalón[4], Hedi Karray[5], Monika Gall[1], Yoav Nahshon[1], Dirk Helm[1], Alexandra Simperler[6] and Gerhard Goldbeck[6]

[1] Fraunhofer Institute for Mechanics of Materials IWM, Germany

[2] UK Research and Innovation, United Kingdom

[3] e-Science Data Factory, France

[4] Universidad Politécnica de Madrid

[5] ENIT Toulouse, France

[6] Goldbeck Consulting Limited, United Kingdom

OntoCommons - Ontology-driven data documentation for Industry Commons, has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement no. 958371.

Table of Contents

1. About the Workshop

The ambition to facilitate data sharing and interoperability within the Materials and Manufacturing domains was the core motivation for this event. Stemming from the OntoCommons[1] H2020 project activities, this focused workshop provided a platform for academic researchers and industrial practitioners to meet and discuss about the Materials and Manufacturing Commons key enablers, i.e., digital Marketplaces, FAIR Principles and Ontologies.

In this workshop, the Digital Marketplaces concept and its current status of implementation was shown in order to continue the discussion about requirements and challenges using ontologies. Tools supporting data documentation and interoperability were showcased, and concrete challenges, success stories, as well as experiences using ontologies were shared.

Initiated by the EU, a session and panel discussion on the future developments of Materials and Manufacturing Commons with focus on materials and manufacturing data spaces followed by interactive input collection from participants rounded up the first part of the workshop.

In the second part, the FAIR principles were introduced and existing tools and guidelines to leverage the FAIR principles in industrial context were identified and discussed together with experts and participants.

The third day of the workshop was dedicated to participants' input, feedback, and questions, (see Appendix A) including an open pitch session for participants. Demos of the tools provided by the ontology commons ecosystem, a virtual tour of digital marketplaces and hands-on working sessions for enhancing the FAIRness score of participants' own ontologies provided a tangible take-away result from the workshop.

This workshop has received high attention from the European Commission (EC) and was very relevant to support the collection of feedback required for further EC incentives/initiatives related to the implementation of Materials and Manufacturing Commons.

The workshop took place on-site at the Fraunhofer-Forum Berlin (Germany) from 4th to 6th of April 2023. The workshop agenda (Appendix B) including links to the presentation slides and the recordings are available online.[2, 3]

[1] https://ontocommons.eu
[2] https://ontocommons.eu/news-events/events/towards-materials-and-manufacturing-commons-enablers-digital-marketplaces-fair
[3] https://www.youtube.com/playlist?list=PL-cwgiwYXckMVegcUtTIp7QbAXNvffx_c

2. Towards Implementations of Materials and Manufacturing Commons

2.1 Key elements of Materials and Manufacturing Commons

The term "Commons" originates from the traditional English legal term "common land". In this original context, resources like land have been made available for joint use in a community. In addition, there is a more general view on commons available in the society with different domain specific interpretations.[4] A major objective of OntoCommons was to create a Commons for data and data spaces that will provide the foundation for applications such as artificial intelligence (AI) and digital twins used by the manufacturing industry. There is a current media hype about AI, how dangerous it is, etc. This may be related to the fact that is not yet based on a Commons and mature technologies. Using OntoCommons' findings, based on the semantics and meanings of data, we aim to provide solutions that are a service to society and industry and lead to economic prosperity in Europe and beyond.

In the workshop we discussed the term "Materials and Manufacturing Commons" with special focus on "Materials Commons", which was introduced during the workshop by the EC and the Advanced Materials Initiative 2030 (AMI2030) presentation. In summary, the Materials Commons knowledge sharing across the materials ecosystem is based on a digital place, in the sense of a trusted system for all stakeholders to make data related to materials accessible. This includes data from modelling, characterisation, and operational use. Materials Commons does not mean that everything is a true common: e.g., data ownership is respected. Major Materials Commons principles are inclusiveness, transparency, accountability, and FAIR. The commons in Materials Commons is making the data FAIR through a common information system based on harmonised documentation of data through glossaries, taxonomies, and ontologies. The current AMI2030 Roadmap[5] also refers to the Materials Commons as " *... a common framework (the so-called 'MaterialsCommons') for all stakeholders (e.g. as materials researchers, developers, manufacturers, up takers as, and end users (B2B, B2C)) supporting their collaboration on advanced materials in a systemic approach across different innovation markets*."

Based on the definitions above and the discussions during the workshop, the key elements of Materials Commons, without any claim to completeness, should be the following:

- **Commons place:** Materials commons is a common place to share knowledge via a trusted ecosystem. It consists of digital places but also has a strong grounding in people, organisations and physical places. Hence initiatives such as European Materials Modelling Council (EMMC)[6], AMI2030, standardisation committees, etc. are of particular importance.

[4] https://en.wikipedia.org/wiki/Commons
[5] https://www.ami2030.eu/roadmap/
[6] https://emmc.eu/

- **Whole life cycle:** Materials Commons supports the interaction and integration of all parts and aspects of the materials life cycle, from raw material, materials design, development, use, re-use, and re-cycling.
- **Digital Materials Commons ecosystem:** The digital materials commons ecosystem does not only include materials data from modelling, characterisation, and operational use as well as general knowledge, but also material related applications itself and their executability. I.e., central elements are data management and data exploration, data/knowledge generators like physical or AI-based models, and digital products like decision making tools. In order to make the digital capabilities available, app stores and code sharing environments will be quite important. The Commons ecosystems will be a holistic environment with marketplaces and data spaces, based on common standards for harmonising the digital technologies.
- **FAIR** principles[7] (see also Section 3), i.e., **Findability**, **Accessibility**, **Interoperability**, and **Reusability** of digital assets. In particular, Materials Commons must achieve the findability of materials related data and information, if permission is granted by the data sovereignty principle, make materials related information throughout the life cycle and from all types of sources like modelling & simulation, characterisation and monitoring accessible under clear conditions, support interoperability across all sectors including manufacturing and Industry 4.0, and of course improve and enhance reuse of data for wider valorisation.

In addition, the following topics are key to a successful materials commons including all stakeholders, including academia, industry, and citizens.

- **Data sovereignty:** For data and information, the data sovereignty is an unbreakable fundamental principle.
- **Trustability:** The trustability of data and information must be granted in the Materials Commons ecosystem by suitable measures.
- **Inclusiveness:** Materials Commons is open to all stakeholders and positively supports stakeholder involvement via education and training
- **Transparency:** Materials Commons activities should be transparent. This includes also the transparency and the traceability defined for data handling and interoperable data usage.
- **Accountability:** The Materials Commons is responsible for the management of the common resources.

The consequences of the above listed key Materials Commons elements are that digital technologies implementing the FAIR principles, as well as data sovereignty, data trustability, etc. must be considered and further developed for materials related applications by the Material Commons community. One also will have to go beyond the traditional models of a commons, where an individual may try and gain as much as possible out of a limited resource. In the Materials Commons, one of the most important things is that in terms of data sovereignty, in terms of decision-making processes and in terms of governance, everything is evenly distributed. So, if an individual brings a novel idea, an additional resource, or knowledge that is needed by other members of the commons, cooperation becomes a necessity. Ideally, the knowledge is brought together from various domains, which is essential for common incentives and a huge motivation for cooperation to happen. Furthermore, inclusivity and participation by all stakeholders must be supported by education, and training. Due to the fast-moving pace of digital technology development, there are already considerable skills gaps and shortages of staff trained in digital technologies.

[7] https://www.go-fair.org/fair-principles/

Materials Commons requires incentives for data sharing and support for the valorisation of knowledge exchange. Following FAIR principles is of course an important foundation. However, as FAIR does not mean free and open access data, a system of agreements regarding intellectual property is extremely important. It should enable combining data from different domains to support current and emerging market possibilities. This will act as an incentive to get large players on board as they can see that releasing their data to a commons rewards them with knowledge beyond the firewall of their organisation. Another incentive is the Green Deal[8] since only a Materials and Manufacturing Commons can offer materials and product data together with environmental boundaries. Data standardisation in this case is extremely important and a top-level metadata standardisation would be highly recommended to make a Materials and Manufacturing Commons FAIR.

During the last decade, different technologies have been developed, which are compatible with the philosophy and the spirit of material commons:

- Data model and data documentation: Fundamental elements in the digital transformation of materials and manufacturing domain are standards for data documentation like modelling data (MODA) (CWA 17284, 2018) and characterisation data (CHADA) (CWA 17815, 2021) as well as vocabularies, taxonomies, and ontologies. From the European projects, the European Multiperspective Material Ontology (EMMO)[9] has been developed and is continuously maintained by the EMMC. In addition, there are also national initiatives like the Platform Material Digital (PMD), which released recently the PMD-Core.
- Dataspaces: The EC is proposing Common European Data Spaces[10] to be this ecosystem that address those barriers by providing solutions to deal with multidimensional data, to release the data and to provide trusted access and trusted tracking of data use. In a common space, we still require data provenance, i.e., where did this data come from, what happened to it, and who owns it.
- Marketplaces: The marketplaces should also provide access to persons who can help with all of this and translate industrial challenges into something that can be modelled. We also can envisage to let computers access the platforms, so a framework for the integration of AI was desirable. All these ideas were implemented in the platform MarketPlace[11] (Goldbeck, et al., 2023) which can be considered as a one stop shop for materials modelling and all services around it. Another important project is DOME 4.0,[12] which stands for digital open marketplace ecosystem. It aims to work with different marketplaces, open simulation platforms, and open translation environments. The players in this ecosystem are data consumers and data owners, and data service providers. The aim of the project is to build a semantically interoperable ecosystem that connects seamlessly with different heterogeneous knowledge bases. Some of the already established connectors tap into Market 4.0,[13] NOMAD,[14] AIIDA,[15] PubChem,[16] and

[8] https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en
[9] formerly European Materials & Modelling Ontology, https://github.com/emmo-repo/EMMO
[10] https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en
[11] https://cordis.europa.eu/project/id/760173, https://www.the-marketplace-project.eu/
[12] https://cordis.europa.eu/project/id/953163; https://dome40.eu
[13] https://cordis.europa.eu/project/id/822064; http://market40.eu/project/concept/
[14] https://nomad-lab.eu/nomad-lab/
[15] https://www.aiida.net/
[16] https://pubchem.ncbi.nlm.nih.gov/

CAMEO.[17] The status quo is that users of DOME 4.0 can access, search, and query these data sources; however, in the future more and more marketplaces and knowledge bases may be added.

However, there never seems to be enough data, and interoperability between data from different domains is missing. Also, data can often be proprietary and not easily shared in a common system. What made data owners more willing to share their data was the Internet of Things Initiative, which could provide viable business insights. The European Effort Alliance for Internet of Things Innovation (AIOTI) is now "*driving on behalf of their members business, policy, standardisation, research and innovation development in the IoT and Edge Computing and other converging technologies across the Digital Value Chain to support European digitisation and competitiveness.*"[18]

A report of the World Economic Forum in 2020 (World Economic Forum, 2020) speaks of $100 billion that could be gained through manufacturing processing optimisation by data sharing. This should be very encouraging to organisations to see their data as an asset and take some actions to share and monetise it. Still, 70% of production generated data are not used and even if they may be useful to a 3rd party, there is a lack of trust and some financial investment is needed to make data shareable.

But just sharing data is not enough, one needs to understand the data. This is, where semantic interoperability comes in. Interoperability is to share something and at the same time, to be able to understand and exploit it. This is where **ontologies** enter the semantic arena. We may describe ontologies as a bridge between the human view of reality and the machine's understanding of this reality; this makes ontologies a model for relating data to reality. They can be seen as a key enabler to interpret and understand the meaning of data, and data documentation must be ontology-based because we need that in order to extract value from data.

However, the acceptance and the attraction of ontologies in industry is not a given, because they are seen as too difficult and time consuming. Over the last two decades, a lot of ontologies have been developed during European projects. What they have in common is that they are not FAIR as there is a lack in terms of methodology, and a lack of a reuse strategy from the start, which leads to the perception that working with ontologies is too time consuming. Hence, OntoCommons is important as we are developing the Ontology Commons EcoSystem (OCES) (d'Aquin, 2021) as a foundation for data documentation.

OCES is a combination of fully harmonised ontology artefacts from top to domain level, and tools and methodologies for building ontologies. So, for future ontologies, OCES provides a complete solution for data documentation in different domains. We started with harmonising existing top-level ontologies and using this to harmonise mid-level and domain level ontologies. This offers two types of interoperability: (i) the internal topology interoperability, which is interoperability by design, which means we have a branch of interoperability from the top level to the domain level and (ii) the cross-ontology interoperability, where we have ontologies from different top-level branches but in the same domain. OntoCommons is a Commons build by the community and it encompasses the OCES and a road map. The latter is important to identify the industrial needs, the gaps and then formulate recommendations for the near future.

Are all materials manufacturers now embracing the enablers for joining a Commons and working with ontologies? There is certainly the use of computers to virtually design processes and to analyse,

[17] https://cameochemicals.noaa.gov/
[18] https://aioti.eu/

evaluate and optimise their processes *in-silico*. We can talk about a single process or a whole process chain, which is a very complex simulation challenge. Also, sustainability goes beyond the processes and we need to think about the lifetime of a manufacturing product. Thus, we need a sort of platform which supports the engineers to reach a more sustainable design during and beyond the processing. Hence, to make sustainability possible one needs a lot of data, many different virtual tools and simulation engines to enable such a sustainable design. This may be only possible for very large enterprises, but certainly not for small and medium enterprises (SMEs), who we also want to contribute to sustainability.

There are also providers of data and software solutions for individual steps of a process. All these providers and consumers should be brought together on a **marketplace**, where they can stitch together modelling solutions for complex processes and SMEs can go there and find short term solutions provided by 3rd parties for an as-needed basis. Ideally, this marketplace is a one-stop shop for software for materials modelling, data validation and benchmark services, catalogues with experts and service providers. It should not only list software but also make it available for testing, running, and combining it with other software and tools to build whole workflows.

A last key element to materials and manufacturing commons are the FAIR principles. (Wilkinson, Dumontier, Aalbersberg, & al., 2016) FAIR means findable, accessible, interoperable, and reusable and these principles should provide a solution to better organise and arrange data we are sharing to enable data-driven science. The latter should not only be done by humans but also by computers. Every community in science and in industry has their own view on what is fair and what is not, and especially concerning metadata. There are social barriers that need to be addressed for the implementation of the FAIR principles. The FAIRsFAIR Project[19] worked hand in hand with many communities, gathering a lot of feedback to then generate common documents. GO FAIR[20] started to work with the communities to define some practical implementations or possibilities to implement the FAIR principles. It became clear that it was vital to understand the diversity of approaches before one could design a common one.

Also, the Materials and Manufacturing Communities are very diverse and there are scattered materials data and knowledge resources, and data of various quality levels in a wide range of repositories and publications. Many of them lack the context and meaning to make them reusable by others. The methods by which we generate data are often not interoperable in themselves and are not really digitalised. There is no standard regarding metadata and each project has its own way of developing certain annotations and schemas and so on.

In OntoCommmons, we plan to bring in Knowledge Management Translators (Goldbeck, et al., 2022) who can bridge the gap between digital expertise and materials sciences. This will require a collective effort of the diverse communities, materials science, data science and semantic technology, expertise, and innovations. In this context, AMI2030 was launched with a manifesto presented to Commissioner Gabriel [21] in 2022. Signed by several organisations, this manifesto pointed out the key role of materials in safety and sustainability, and was addressing many of the issues facing society today.

An important working group of AMI2030 is the "working group on materials digitalisation" which has the objectives of delivering a common digital ecosystem, meaning the interoperability between

---

[19] https://www.fairsfair.eu/
[20] https://www.go-fair.org/
[21] Mariya Gabriel, European Commissioner for Innovation, Research, Culture, Education and Youth (2019 – 2023)

all the relevant technologies that can generate data. Then, we can capitalise on these commonalities and integration synergies and contribute to a federated, highly harmonised interoperable materials data infrastructure. It will be vital in the future to cooperate beyond Europe and work with all citizens to reach a truly global Materials Commons.

2.2 Digital Marketplaces

The **Digital Marketplaces** are supposed to be an **ecosystem that enables** a materials and manufacturing commons, i.e., where all protagonists are coming together. They want to bring and take knowledge with them, but also build new knowledge from different sources. These sources can be databases but also tools that create live data, such as materials modelling software or characterisation techniques or sensors. All sources need to be **interoperable** and semantics will have to play a role in this.

DOME 4.0 is building its platform around nine showcases with different levels of digital maturity and the beneficiaries of this project represent a diverse ecosystem with data providers, data consumers, and digital service providers. The individual showcase decides what type of interoperability between sources is needed, and it can be syntactic if sufficient. The project partners also investigate which connections are wanted; for example, one of their cases requires maritime, environmental and nanoparticle data for air quality in port cities. DOME 4.0 is promoting hackathons as a very good tool for co-creating solutions on how to build these data connections.

Ontologies are part of DOME 4.0's semantic solutions for connectivity and the DOME 4.0 team worked together with OntoCommons in a dedicated work package, so OntoCommons' OCES was adopted wherever feasible. The ontologies are now adding qualitative value to the project. This encompasses linking decentralised data sources/software and enabling interoperability. However, the solutions are very specific; they are not easily adaptable to changes/updates in the ontologies used and adding missing ontology terms requires human intervention. However, FAIR principles are recognised as being crucial for such a development, and have been successfully demonstrated in one of the showcases.

For the materials and manufacturing communities, DOME 4.0 could demonstrate how to search for material data in OPTIMADE[22] and get corresponding machines to process a material from the Market 4.0 product catalogue using semantic search tools.

The **MarketPlace** platform is dedicated to the community of Materials Modellers who intend to optimise materials, materials design processes, process design, component design and product design. It aims to be the one-stop-shop for materials modelling and simulation. The Materials and Manufacturing Commons is invited to explore, interact, create, and execute on this platform. Create and execute is one important pillar of the MarketPlace because the community was interested to perform actual operations on data using tools on the platform. Therefore, the platform hosts an App Store and can make a workflow manager accessible. Existing open simulation platforms such as AIIDA are incorporated.

On the MarketPlace platform, ontologies are used in various aspects such as platform management for software and workflow communication. This enables seamless communication between different types of applications. MarketPlace provides some integrated services, such as its ontology-based

[22] https://www.optimade.org/

knowledge service. Several application ontologies have also been developed, such as an expert ontology and a software ontology. These ontologies are partly based on EMMO to comply with standards driven by the materials and manufacturing community.

Like DOME 4.0, the MarketPlace consortium has discovered that it is sensible to use a little semantics at a time. To generate an environment for a Commons, it is important to on-board tools so there is the possibility for a software tool to be registered on MarketPlace without an ontology but using an application programming interface (API). However, there is a possibility for a semantic integration of the software tools where one can engage ontologies, which was showcased by some of their use cases. Ontologies are also communication vehicles between different experts such as physicists, engineers, materials scientists, etc., who do not speak the same natural language. Making their jargon to an instance in an ontology can aid like a universal translator.

Simultaneously to MarketPlace, also the **VIMMP**[23] consortium has been tasked with developing a marketplace for materials modelling that facilitates exchanges between providers and users. Typical providers on this platform are software owners or translators or modelling experts and typical users are industry researchers or research managers. Again, these protagonists are part of the Materials and Manufacturing Commons.

The interoperability between different components of this marketplace is also handled by ontologies which can be found under GPLv3 licence on the GitHub repository including a series of documentation, papers, etc. A great achievement was the development of the upper ontology for the virtual-marketplace framework, the European Virtual Marketplace Ontology (EVMPO). (Horsch, et al., 2020)

Like both other marketplaces, they use interoperability both on syntactic and semantic levels within the project. Normally there are technical constraints when implementing tools, and it will be vital to understand those as soon as possible. This will provide some guidance whether syntactic or semantic interoperability should be attempted.

2.3 OntoCommons

OntoCommons, as its name suggests, is a major enabler for a future Materials and Manufacturing Commons. It develops ontology standardisation guidelines and alignment tools, maintains several use cases to demonstrate how ontologies and semantics can help organisations, and, finally, is also empowering human resources to actively work with organisation who wish to transform their data into knowledge.

OntoCommons is not creating new ontologies; it is rather taking stock on what is already there, it harmonises and aligns. It also aims to enable cross-ontology interoperability. There are several top-level ontologies which are admitted into this **ecosystem** (pluralistic approach) and this is something unique as most of the ontology related projects normally subscribe to only one top level ontology. The reason is that many existing ontologies are compliant with one of those many top-level ontologies. Thus, by admitting all of them, we can integrate all these different domain specific ontologies into one semantic ecosystem. This is vital for materials and the manufacturing semantic concepts as they are extremely diverse and different. In addition to the top-level ontologies, there are mid-level ontologies which are sort of domain neutral, and are used as a bridge concept. They

[23] https://cordis.europa.eu/project/id/760907, https://www.vimmp.eu/

conform to a top-level ontology but allow the domain or application developer to work with a less complex semantic construct. Several templates are available to help ontologists with alignment, and the demonstrator cases do use these templates and showcase them.

Two of these demonstrator cases are by BOSCH and are about smart manufacturing/welding and smart materials/plastics, respectively. In the case of welding, the organisation wants to combine data from multiple sources to reduce manufacturing errors and in the case of plastics, they want to combine data from multiple labs. They **use ontologies** to unify their welding data and knowledge coming from machine learning routines and AI. For plastics simulation, they use ontologies to optimise the simulation processes to make them faster and cheaper. BOSCH is aiming for high quality standardised semantic models and besides being active in OntoCommons they are partner in Catena-X[24], where the aim to enable the digital flow of information across the entire supply chain in the automotive industry. BOSCH worked with OntoCommons to adapt their processes of ontology development, making the ontologies FAIR and compliant to relevant top-level ontologies. They are embracing the ideas fostered by OntoCommons and are planning to form a working group with other organisations to keep the momentum going. They intend to make most of their in-house data interoperable but also bring in external data sources to drive their innovation. Marketplaces are deemed as being valuable receptacles to access and share data. BOSCH's manufacturing data are described and made available through the standardised vocabulary ontologies which they develop in OntoCommons and they are made discoverable through the ontologies developed through DOME 4.0. If every manufacture does this pre-competitively, everyone can "shop" for the right data, and then take this right data and push it into AI and machine learning (ML).

DOME4.0 and OntoCommons have a use case together where they investigate **materials databases integration** using ontologies. Nowadays, different software programmes are used for calculation on materials and provide reliable materials data worthy to be entered into databases such as the Materials Project[25], the Open Quantum Materials Database (OQMD)[26], or NOMAD[14]. If you search for a compound in each of them, you will get three different ways in how the search result is reported back to you. "Formula" in one report corresponds to "Composition" in another one, and so forth. This can be correctly interpreted by a human operator, but may cause trouble for a computer. The solution, to make these databases interoperable, is a common API. The starting point was to develop a domain ontology for data access and integration. There are still challenges; data can be modelled and shared in very different ways. For instance, data may follow a relational model or a non-relational model. It may be shared and queried by different APIs and in different formats such as JSON or csv.[27] So first, the DOME 4.0 team developed the Materials Design Ontology (MDO)[28] starting from a requirements analysis and identified the use cases and computation questions. Some concepts from existing ontologies, such as the provenance ontology[29] and Quantities, Units, Dimensions, and Type Ontologies (QUDT).[30] The MDO is capable to represent basic domain knowledge for the materials design domain and can be used to generate some mappings among different materials databases.

[24] https://catena-x.net/
[25] https://next-gen.materialsproject.org/
[26] https://oqmd.org/
[27] A comma-separated values (CSV) file is a text file that uses a comma to separate values, giving rise to the name of this file format. https://en.wikipedia.org/wiki/Comma-separated_value
[28] https://www.semantic-web-journal.net/content/materials-design-ontology-0
[29] https://www.w3.org/TR/prov-o/
[30] https://www.qudt.org/

The next step will be to develop a web interface so that users can write some simple queries. DOME 4.0 is also developing the Semantic Data Exchange Ontology to enable the exchange of data between **all data providers and consumers**, and thus contribute to a Materials and Manufacturing Commons.

We will need to bring in **human resources** to aid all organisations who wish to profit from novel semantic technologies. There are two existing roles, we are expanding on. One is the Materials Modelling Translator, (Klein, et al., 2021) a role which was developed as a bottom-up approach by people engaged with the EMMC to close the knowledge gap between industrial stakeholders and material modellers. The second one is the Analytics Translator, whose role is to close the knowledge gap between industrial stakeholders and data experts. Out of these two roles we shaped the Knowledge Management Translator, (Goldbeck, et al., 2022) who will be interfacing between engineering (data, ontology engineering) and the business team. In detail, they will be working closely with the business team to map out innovation cases, and with ontology curators, ontology engineers, devops experts and data engineers to coordinate solutions. They also will have to coordinate with domain experts for terminological curation and engineers for the logical modelling.

OntoCommons created a lot of value by harmonising what exists, adding what was missing and using demonstrators to make ontologies more tangible and a valuable common good. We will have to have the diligence to build communities or bodies that they will govern this further and policy departments within the EC and beyond will have to play an important role.

2.4 Dataspaces

DOME 4.0 will enable data consumers to find data providers; however, data will have to be housed in spaces that either owned by the providers or ideally, common data spaces.

One of these common spaces is **Data Space 4.0**, [31] which oversees preparing the deployment of the manufacturing data space in Europe. The concept of a data space is a decentralised infrastructure enables data transactions and their governance. The Common European Data Spaces [32] are a collection of twelve initiatives that are emerging in different sectors, and Data Space 4.0 is one of them, focussing on the manufacturing sector. It is based on three high level missions where one is to collect all the existing tools, best practices and assets around the manufacturing domain that can be useful for building this data space. Another on is to deliver a set of blueprints with guidelines, recommendations, and the building blocks needed for the implementation of the data space. The final mission is to provide a minimal viable framework for every data space in this domain compliant with Industry 4.0. The manufacturing industry is very diverse and so are their data. Data Space 4.0 is working on smart data models within their project to encompass the diversity and with existing ontologies to make them interoperable.

Platform MaterialDigital (PMD) [33]is funded by the German Ministry of Research and Education, and the idea was to strengthen Germany as a business location. In a first round of funding, 13 projects comprising a very broad spectrum of materials were chosen to create data and fill the platform. The projects develop ontologies and some of them are compliant with the EMMO and some are developed independently. However, it was beneficial to use the PMD core ontology (PMDco) [34]as an

[31] https://manufacturingdataspace-csa.eu/
[32] https://dataspaces.info/common-european-data-spaces/#page-content
[33] https://www.materialdigital.de/
[34] https://github.com/materialdigital/core-ontology

umbrella for all those products on materials science, engineering and to map to different top-level ontologies.

The **European Open Science Cloud** (EOSC) [35] is an environment for hosting and processing research data to support EU science. Its infrastructure is supporting cross-disciplinary science by offering access to tools, procedures, and technologies for researchers all over Europe and later, the world. A strong building block is the notion of Open Science as a vehicle to create new technologies and new tools to induce even more science. EOSC is supposed to expose data to persons interested in them, but also infrastructure to assist in searching this data and extract knowledge from them. There is also a need of supporting services around the users as services like accounting, monitoring access, order management, etc. must be implemented as well to grant a professional service delivery. The EOSC users tap into the EOSC Resource Catalogue where all the services and the research products are tied into. Providers are adding their services to the catalogue which also is a service provision. The perspective on science has changed as well since we are moving away from discipline centric towards cross-disciplinary. A researcher searches not anymore in domain specific databases but cross domain with complex research questions. The EOSC resources demanded a specific ontology that can handle its three important functions; discoverability, order management, and composability. Composability means to assemble something from smaller, independent components, i.e., many resources are needed to answer a complex research question. The EOSC catalogue is expecting to host millions of data records originating from different disciplines, countries, policymakers, funding bodies, etc. Ontologies must help to make search results relevant and information discoverable. But also, the data provider will have to use appropriate data schema and well curated data, to make ontologies work.

It becomes clear that Data Spaces need **management systems** and Fraunhofer IWM did develop such a Dataspace Management System, or DSMS in short. Materials sciences come with a large volume of data with a diverse origin such as experimental, simulation, and literature data, and all stored in different silos. Also, the data format is not unified and covers everything from an excel spreadsheet to pdf files. Data can age so it will be relevant to document their contemporary relevance. Metadata are important, especially experimental conditions, otherwise the actual data can be made redundant.

The first step is always to access data which can be done elegantly with some scripts and often less elegant with a manual copy and paste. Thereafter, the data must be processed with further scripting to finally have a purpose. The Resource Description Framework (RDF) is often used to integrate data from multiple sources. The data, invisible to an end user, can then be stored enriched with semantic meaning. Accessing them may require SPARQL query language, which can be trained to a non-data expert. The DSMS offers entrance points to more advanced users as well who are curating the data space. It also offers control over who can access which data. The latter is very important for common data space as some spaces due to Intellectual Property rights may be more regulated than others.

2.5 Combining Views on Materials Commons

The Bureau of European Design Associations (**BEDA**) [36] comprises members who are publicly funded design organisation professionals and members of trade associations. In the European Community, BEDA assists in framing and solving the challenges regarding climate change and sustainability,

[35] https://eosc-portal.eu/about/eosc
[36] https://beda.org/about-us/this-is-beda/

digitalisation of our common future and the design for growth and prosperity in Europe. BEDA is also a member of AMI2030 as they see they see their design community also profiting from being associated with a Materials and Manufacturing Commons.

For example, the **Basajaun** project [37] is looking into the full supply chain starting from the forest to the construction of wooden buildings. They became a demonstrator in OntoCommons as they are digitally tracking all the materials and components along the supply chain up to constructing wooden buildings. They are also looking into circular economy, as buildings are assembled with components which may need to be disassembled at some point in the future. The deal with data originated from a saw mill and those from the manufacturing or processing industry and thus, cover many different domains. Their next step will be an infographics-based supply chain visualisation.

Organisations with a **design perspective** are very good in knowing what to do, how to produce things and to find viable solutions for it. However, when it comes to sustainability they have to think in new ways about a product. Customers do like to find reused materials in a product as it makes them feel less wasteful wit resources. However, the materials must still enable a particular design or functionality. Designers need to have more knowledge of material properties before they start, so is we can let them in to a Materials and Manufacturing Commons, they may find their data. Customer acceptance is also very important, and if the sustainable materials reliably can replace a current, less sustainable material. Designer would like extract from data how to perceive, produce and combine sustainable materials, trace the materials, and communicate to the end users in a trustful way.

Another important industry working with materials and seeking sustainability is the **fashion industry**. Their supply chains lead from materials to fashion and apparel, but also covers labour and trade secrets. Again, and industry with versatile and many different sources of information. Data specialists in this industry were aware that knowledge graphs are needed, handling of different taxonomies and regulation of how information is shared. Ontologies would help but they need to be explained to the fashion industry so that lay people understand the concepts.

Now, the Materials and Manufacturing Community has the air of "Scientists only" but as there is great interest from other communities one should consider more inclusivity and make ontologies and semantics accessible to everyone who aims for finding knowledge, no matter what their background is.

3. FAIR Principles for Industry

As the Materials and Manufacturing Commons will be based upon sharing data, complying with the FAIR principles will be paramount. This section will discuss what these principles are, how OntoCommons embraces them and finally introduce some of the FAIR related initiatives that would be relevant for building such Commons and supporting industry to implement the FAIR Principles.

3.1 What are the FAIR Principles?

The FAIR principles originated from the need to define a set of common technology-agnostic guidelines for managing research data and for transforming data assets in machine actionable units

[37] https://cordis.europa.eu/project/id/862942; https://basajaun-horizon.eu/

of information. Fifteen individual guiding principles were defined and published in a paper in 2016 (Wilkinson, Dumontier, Aalbersberg, & al., 2016) which has over 10k citations now.

Four main aspects must be considered for data to be FAIR compliant: (i) data and the associated metadata should have a globally unique, persistent, and resolvable identifier (GUPRI); (ii) data should be richly documented with detailed metadata which is crucial to make data machine actionable, findable, and reusable; (iii) data and metadata should be accessible through free, open, and universal access protocols and (iv) data and metadata should come with a clear licence for both human and machines.

The most important aspect of the FAIR principles is the use of metadata. Metadata should provide as much information as possible describing the content of the data but also its provenance. Metadata must use domain specific metadata standards and should be published using common formats.

GUPRIs must be used to identify both the data and their associated metadata records. In addition, it is crucial that GUPRIs resolve to only one intended defined meaning. Thus, a machine "knows" what is meant and this main goal can be achieved by multiple standards and the extensive use of semantic artefacts i.e., ontologies, controlled vocabularies, etc., which should themselves be FAIR.

Another aspect of the FAIR principles focuses on the accessibility of data through well described open and standardised APIs. However, the FAIR principles do not imply that the data should be open. People working with sensitive data tend to turn away from FAIR quite readily. However, "A" stands for accessible under well-defined conditions, meaning that the metadata of the data includes the consent for reuse or not. Hence, the metadata should be openly accessible, and the data access terms should be made clear with human and machine-readable/machine-actionable licenses.

The FAIR principles are not only technical but also involve a social aspect as one must agree what this persistent identifier is, how long it shall remain unchanged, and what makes metadata "rich" and what is meant by it. These agreements must be defined within communities. The same applies to the rules for accessing the data as well as the associated licencing.

Machine actionability of the data and metadata, emphasised by the FAIR principles, is essential to be able to cope with the ever-increasing volume of data generated in Science. It becomes clearly impossible to be able to harness the content of such large amount of data without considering automation and use of dedicated services to support scientists. For example, some problems in medicine are very complex such as why some people react badly to vaccines or how to model diseases and associated potential treatments in the context of digital twins. Uncovering an answer to such complex questions requires to integrate data coming from different scientific domains, different temporal, and spatial scales and most importantly to a wide range of stakeholders distributed around the world. The adoption of the FAIR principles will enable to ease the integration process and the identification of vast and complex data patterns, using complex algorithms such as ML/AI. Data for such global problems will have to come from beyond Europe; certain nations may not want to share with other. However, if all data are FAIR, data can be visited and analysed, and feed still into a global procedure of creating knowledge. Several nations are recognising the power of FAIR data and do actively take part in respective meetings.

3.2 Examples from Industry

Although the adoption of FAIR is mostly promoted and required in the context of international Open Science initiatives such as the European Open Science Cloud, industry also identified several major benefits to implement them. First, FAIR principles are pushed by several megatrends in digitalisation such as data driven discovery and innovation, hyper personalisation, enhanced decision making, massive data integration and ML/AI. These various trends require to change perspective and think of information architecture not as service centric but rather as data centric. In this data centric viewpoint, companies should harness the connections between the various types of data. To build such data centric architecture, it has become obvious that the FAIR principles should be implemented and should also linked to the aspects dealing with the data quality. Successful and value-generating digitalisation requires true machine actionable data, machine readability alone is not sufficient.

Major companies are realising that the main cost of data management is hidden and encompasses various processes such as search and access, data curation, semantic data integration, data cleansing, extract, transform, and load (ETL) processes, and a flexible IT infrastructure. The pharmaceutical company Roche has identified the business value of prospective FAIRification and high data quality. In this approach, the hidden cost of data management should be delegated to the data producers. For this, they created the Roche Data Commons (Zicari, 2018) which relies on community standards as well as the semantic web standards to implement the FAIR principles. This data commons relies on terminology, metadata, dataset models and ontologies, well defined APIs, and URI schemes as GUPRIs. In addition, existing tools such as the FAIR Toolkit, developed in the context of the Pistoia Alliance[38], are leveraged to assess the FAIR maturity of data assets. One of the main interoperability challenges is due to the large number of standards that exist in the biomedical community. To solve these problems, they leverage the use semantic mappings. On top of this data commons, reference services to manage terminology, metadata and data models, and conceptual models were developed and are used to populate a FAIR unified domain knowledge graph. In this context, the implementation of the FAIR principles leverages semantic web technologies (usage of semantic artefacts, knowledge graphs, etc.)

Similar issues are faced at Bosch, which needs to integrate and connect data from various sources to optimise their production processes and generate wisdom from domain-specific knowledge. Their implementation of the FAIR principles follows similar principles than Roche and heavily rely on the usage of semantic artefacts, such as the Industry 4.0 Core Information Model [39] for Manufacturing to create a harmonised data fabric and data mesh on which they can build digital twins of their factory and their products. This enables them to tackle issues regarding manufacturability analysis and production line management.

Although the added value of implementing the FAIR principles in large industries or within the OntoCommons demonstrators has become obvious, there are still barriers that slow down their integration in a wide range of industries due to the lack of awareness and understanding of the competitive advantage gained by this investment. The importance of semantic technologies is valued across hierarchies in some organisations and even CEOs become aware of them. The FAIR principles are more readily accepted than the use of knowledge graphs and ontologies, respectively. The OntoCommons demonstrators help organisations to work with something specific and allow to start with a set of ontologies capturing their specific use cases. They work like a proof of concept and show value. This is the first step in working towards making ontologies available across the whole organisation. The FAIR principles are implemented when it is becomes necessary to enable reuse and interoperability, i.e., preparing data for ML and AI.

---

[38] https://www.pistoiaalliance.org/

[39] https://openmanufacturingplatform.github.io/sds_blog_Industry_4_0.html

Entrepreneurs will never ask for FAIR; they are more interested in markets, customer needs, costs, revenue, etc. One then must convince them that the answer to all of this lays in their data. There will be soon a realisation that these data must be made accessible, maybe even to third parties. As one can see, in such circumstances the "A" in FAIR is the biggest bottleneck. This is the point of establishing a strong business case to convince an organisation to give access to their data. This business case must evidence how interoperability brings value and that reusability saves effort and money. FAIR enables to become part of a larger pre-competitive ecosystem which gives more visibility and may lead to more costumers. The entrepreneurs must be made aware that FAIR is also a matter of return on investment; ML and AI could be much faster deployed and time and money on the road to the next big innovation can be saved.

3.3 FAIR Principles in OntoCommons

OntoCommons is documenting industrial data with ontologies; hence we must make sure that the ontologies are themselves FAIR. To understand the situation in the Material and Manufacturing domain, OntoCommons initiated an ontology landscape analysis (Le Franc et al., 2022). The study collected 130 ontologies which were classified into 5 domains:

- Physics and Chemistry
- Mechanical and Industrial Engineering
- Thermal and Process Engineering
- Materials Sciences and Engineering
- Computer Sciences, Systems and Electrical Engineering

The analysis revealed that 84 ontologies out of the 130 are machine readable. For example, mechanical engineering has the most of ontologies with a reasonable ratio of machine readable versus non machine readable. An in-depth analysis of the collected ontologies revealed few top-level ontologies which are highly formalised and many specialised domain ontologies with not much logic in them. The landscape analysis revealed that machine readable ontologies are published using different formats, such as RDF XML, Turtle, OWL, etc. Also, different communities tend to have their own best practices to build their own ontologies. Currently, that hampers attempts to find, reuse, and make ontologies interoperable both within the same and across domains, respectively. In summary, semantics in Materials Science and Manufacturing seems to be insufficiently FAIR. For OntoCommons it was necessary to take stock and check if the existing ontologies, identified in the landscape analysis are FAIR (Le Franc, 2022). For this purpose, we considered two approaches described below.

The first approach was to leverage the outputs of the FAIRsFAIR[40] project, which was involved in the regular Knowledge Exchange Space (KExS) meetings organised by OntoCommons beneficiaries. The project introduced the overarching concept of semantic artefacts which are defined as "*machine-actionable and readable formalisation of a conceptualisation enabling sharing and reuse by both humans and machines. These artefacts may have a broad range of formalisation, from a loose set of terms, taxonomy, or thesauri to high order logics. Moreover, semantic artefacts are serialised using a variety of digital representation formats, which could be Turtle, XML, RDF XML, JSON, etc.*" This definition may not be necessarily perfect, but at least it allows to reconcile the different viewpoints and to start discussing how we want to make these different semantic artefacts FAIR. FAIRsFAIR

[40] https://cordis.europa.eu/project/id/831558; https://www.fairsfair.eu/

established 17 generic recommendations (Le Franc, Bonino, Koivula, Parland-von Essen, & Pergl, 2022) to make semantic artefacts FAIR. Each recommendation is aligned with one or more of the individual FAIR principles. The wider community was engaged in several workshops which permitted to align these recommendations with the RFC 2119,[41] which defines conformance keywords. Nine recommendations are a "must", which means they are mandatory if one wants to be FAIR. Seven other recommendations were deemed as "should" be followed, i.e., recommended, and one as "optional".

This idea is domain agnostic; life sciences, materials science manufacturing or any other kind of industry are building ontologies, and we want to see FAIR ontologies across domain and start breaking down silos and agreeing on common general principles. The Research Data Alliance (RDA) proved to be an excellent platform to have an international outreach and a very diverse number of people from Europe, Australia, US, and China became active there. Two task groups were created: the RDA VSSIG Task Group on Minimum Metadata and the RDA VSSIG Task Group on FAIR Semantic repositories.

Based on these recommendations, OntoCommons created a simple evaluation matrix to assess the degree of compliance of ontologies manually with respect to the FAIR principles. For this, we considered a subset of the FAIRsFAIR recommendations covering specifically the ontologies:

- usage of **GUPRIs** for semantic artefacts, their content (i.e., concept/term/class and relation) and their version,
- machine-readable **metadata** to describe the semantic artefacts themselves and their content,
- usage of **repositories** to share, publish and retrieve semantic artefacts and their content
- defining **common APIs** to access and index semantic artefacts and their content,
- **interoperability approaches** to make sure that semantic artefacts of various degrees of complexity and encoding format should work together including publishing mappings and crosswalks between semantic artefacts,
- semantic artefacts and their content should be retrievable through **search engines**.

The second approach to evaluate the FAIRness of the ontologies identified in the landscape analysis, leveraged that FOOPS![42], an ontology pitfall scanner for the FAIR principles, working for both OWL and SKOS vocabularies, may be used to check if ontologies are FAIR. (Garijo, Corcho, & Poveda-Villalón, 2021) This tool uses the recommendations proposed in (Poveda-Villalón, Espinoza-Arias, Garijo, & Corcho, 2020) and the FAIRsFAIR recommendations (Le Franc et al., 2022). FOOPS! enables the automation of the FAIRness evaluation. In addition to FOOPS!, the O'FAIRe tool (Amdouni, Bouazzouni, & Jonquet, 2022) has been developed and integrated with the AgroPortal ontology repository.[43] However, the later tool was not deployed to evaluate the FAIRness of the ontologies identified in the OntoCommons landscape analysis.

The first evaluation of 41 ontologies from the landscape analysis dataset resulted that none of the ontologies were FAIR with both the FAIRsFAIR matrix and the FOOPS! tool. Physics and chemistry are the domains with the highest FAIR score on average, but still no ontologies passed the threshold for being minimally FAIR. Despite the difference in the evaluation criteria between the FAIRsFAIR matrix and FOOPS!, the evaluations were quite similar. With the increasing number of tools to measure

[41] https://www.w3.org/wiki/RfcKeywords
[42] https://w3id.org/foops/
[43] https://agroportal.lirmm.fr/

FAIRness for Semantic Artefacts, it has become important to align the various evaluation methods to avoid too much divergence and to standardise the results of the evaluation.

The FAIRsFAIR recommendations have been compared to the ontology engineering community and semantic web practices (Poveda-Villalón, Espinoza-Arias, Garijo, & Corcho, 2020), i.e., "Best practices for implementing FAIR vocabularies and ontologies on the Web", (Garijo & Poveda-Villalón, Best Practices for Implementing FAIR Vocabularies and Ontologies on the Web., 2020) "5-stars for vocabularies" (Vatant, 2012) and "Five stars of Linked Data Vocabulary use". (Janowicz, Hitzler, Adams, Kolas, & Vardeman II, 2014) Several current practices of the semantic web community can be used to make ontologies FAIR in practice, such as using URIs as identifiers, metadata included in the ontology, content negotiation through HTTP/HTTPS protocols, use of Knowledge Representation languages, use of DCAT for ontology collections, etc.

Evaluating ontologies once they have been developed and then updating them to comply with the FAIR principles requires to update the ontology and can involve an extensive use of resources. It would be more efficient to plan for the implementation of the FAIR principles before one begins to develop a new ontology. The LOT methodology[44] has been proposed as a core workflow for building ontologies which can be aligned with various tools to support the development of ontologies from their inception to their publication and use in relevant information systems. FAIRness assessment often happens at the end of the workflow once the ontology is published in a platform. It would be more efficient to develop FAIR-by-design ontologies, i.e., to integrate the technical elements and constraints necessary to comply with the FAIR principles within the ontology implementation step of the workflow. A crucial component of this workflow is the ontology documentation step. Indeed, ontologies need to be properly documented with metadata but also with additional necessary information such as diagrams for the example data, SPARQL query examples, and different formats could be saved such as HTML and pdf for humans and RDF, XML, JSON, Turtle, etc., for machines. It is a wise move to adopt existing practices and technologies and to think about FAIR principles at all stages of the ontology development. But we must be aware that FAIR does not look at the resource quality.

3.4 FAIR Resources for Industry: What is happening in EOSC?

The European Open Science Cloud (EOSC) is an environment for hosting and processing research data to support the EU global strategy on Open Science. EOSC aims at becoming a "web of FAIR Data and Services" for science in Europe. The support the development of common practices within EOSC, the EOSC Association created 13 Task Forces composed of experts working on 4 main topics:

- Metadata and data quality
- Research careers and curricula
- Technical challenges
- Sustaining EOSC

[44] http://lot.linkeddata.es/

As part of these Task Forces, two Task Forces are focusing on key aspects related to the FAIR principles and relevant for OntoCommons and Industry: the FAIR Metrics and Data Quality Task Force [45] and the Semantic Interoperability Task Force.[46]

The FAIR Metrics and Data Quality Task Force aims at investigating ways to harmonise how to evaluate how data and metadata are compliant with the FAIR principles. The FAIR principles propose guidelines to ensure that data are shared in a way that enables an enhance reused by both humans and machines. They do not come with any recommendations on the use of specific technology and standards. There are 23 independent FAIR assessment platforms[47] and most are questionnaire based and several of them are automated and all of them may produce different scores. Hence, a common paper (Wilkinson, et al., 2022) on the governance of FAIR assessment has been written. It advocates a harmonised approach and a definition of a metadata publishing paradigm that will:

- Support all publishers (both large and small, i.e., low complexity)
- Support the agents that are exploring them.
- Work on all types of digital object ("Traditional" data, software, workflows)
- Provide access to the most important metadata: that of the data creator.

One of the key aspects to be considered for this harmonised approach is the disambiguation of resources. For successful traversal of a FAIR Record the unambiguous identification of (i) the globally unique identifiers (GUID)[48] for the records, (ii) of the metadata records, and (iii) of the data record is crucial.

The Semantic Interoperability Task Force has as current focus to propose recommendations for semantic interoperability that build on top of what is called the EOSC Interoperability Framework [49] and assure a deep connection to real-world research problems. Interoperability has many facets and there are legal aspects, technical aspects, semantic aspects, and intra-organisational aspects to be addressed. As the wider community must support this, all knowledge of a wide range of participants to Task Forces and conferences is taken aboard and should ideally converge on some recommendations. The Task Force is organised into three smaller groups which work on topics such as surveying the landscape of semantic interoperability and metadata conventions, defining maturity indicators for semantic artefact catalogues, and collecting some interoperability case studies from various communities. The survey of the interoperability landscape includes some definitions of terms, so one can agree on its basic implementation. The case studies stem from the wider community and cover domain specific situations or an organisation's specific context. Hence, one can assure that EOSC has representations from the wider user community. These use cases can be used to demonstrating value to different kinds of stakeholders and for practitioners to learn how something has been put into practice. To make this collection of case studies and use cases as usable as possible, it is pertinent to find conceptual commonalities that could be used to compare them across the collection and help people identify some general trends that would be useful for them to support implementations and adoption. This is a good recommendation to a future Materials and Manufacturing Commons to have representation of the different special interests. Then "sub-commons" can be formed, and they can extract different components of the ecosystem that meet

[45] https://eosc.eu/advisory-groups/fair-metrics-and-data-quality/
[46] https://eosc.eu/advisory-groups/semantic-interoperability
[47] https://fairassist.org/
[48] globally unique identifier - a 128-bit text string that represents an identification
[49] https://eosc-portal.eu/eosc-interoperability-framework

their needs. Like the EOSC Semantic interoperability task force, one then gets a better handle on what these groups need in terms of tools, data, computing resources, training, interoperability, etc.

Besides the EOSC Association Task Forces, a large number of ongoing EOSC related projects are working on establishing frameworks to implement the FAIR principles in the context of EOSC. The FAIR-IMPACT project[50] is a coordination and support action and has the aim of expanding FAIR solutions across Europe and beyond. This should happen across different scientific communities and scale up to European level. The project partners are seeking existing practices, policies, tools, technical specifications, etc. and translating them to solutions or guidelines for a wide range of domains. The four main domains they work on are social sciences and humanities, photon and neutron science, life sciences and agricultural, food, and environmental sciences grouped together. Each use case partner acts as a bridge between the project and one or more of these scientific domains, and they provide expertise with regards to what are the domain relevant community standards and practises. Semantic artefacts, i.e., ontologies, terminologies, taxonomies, thesauri, vocabularies, metadata schema, and standards, are the key element to achieve FAIR, and these artefacts and catalogues must be FAIR, too. All the recommendations regarding FAIR shall then be extended to other research communities beyond the domains covered by FAIR-IMPACT. This happens via cascading calls to sponsor use cases, the organisation of workshop and collaboration with organisations such as the Research Data Alliance[51]. The outcome of this project is to contribute to the Horizon Europe partnership, meaning basically supporting EOSC readiness through harmonisation, synchronisation, and alignment.

The FAIRCORE4EOSC[52] project focuses on the development and realisation of core components for the EOSC. It is working on the persistent identifiers, metadata and ontologies, interoperability, and research software and develops nine new services:

- Research Discovery Graph (RDGraph) - led by OpenAire[53] who collected metadata on research objects such as publications, funding decisions, datasets, research software, etc. RDGraph connects all those nodes together.
- Persistent Identifier Graph (PIDGraph) – which is focusing on the graph technology, and it will accept only nodes into the graph which have PIDs. It has an API and one can query for example things connected to a specific PID.
- Metadata Schema and Crosswalk Registry – which supports hosting and registering metadata schemas. It has a user interface through which one can visually create crosswalks between metadata schemas.
- Data Type Registry (DTR) – which is a place where one can declare data types, get PIDs for them, and thus, make them referrable from external sources.
- PID Meta Resolver (PIDMR) – which is a system that allows one to put in any type of a resolvable PID and it will handle the resolution irrespective of whether that is a file handle or DOI, URL, etc.
- Compliance Assessment Toolkit (CAT)- which can assess what a tool does and query its compliance with the established policy.

[50] https://cordis.europa.eu/project/id/101057344; https://fair-impact.eu/
[51] https://www.rd-alliance.org/
[52] https://cordis.europa.eu/project/id/101057264; https://faircore4eosc.eu/
[53] https://www.openaire.eu/

- Research Activity Identifier Service (RAiD) – which captures information around research activities of larger research projects which are not constrained by a specific funding decision but can be activities that are spanning 10 - 20 years, and funded by multiple funding bodies.
- Research Software APIs and Connectors (RSAC) – which creates new APIs and connectors to software.
- Software Heritage Mirror (SWHM) – which is collecting source code on research software around the globe and archiving it. It issues software heritage identifiers, so a source code becomes referenceable.

These components are co-developed and tested by various community use-cases such as CLARIN[54], a digital infrastructure which provides easy and sustainable access to a broad range of language data and tools to support research in the humanities and social sciences. The Deutsche Klimarechenzentrum will adopt the listed components and show how they produce added value for their community regarding the climate change. FIZ Karlsruhe tests the components for the mathematics community and to reach European Integration of National-level Services the FAIRCORE4EOSC services are used to firstly pull together information nationally on publications, datasets, researchers, research organisations, etc. and secondly make it accessible Europe-wide.

Within the EOSC-Pillar project,[55] tooling and approaches are developed to support the implementation of the FAIR principles. One of these tools is the Federated FAIR Data Space which leverage existing tools to support FAIR such as the FAIR Data Point (FDP)[56,57]. The rationale for this tool is grounded on the needs from the various scientific use-cases of the project. In these use-cases, data scientists or a computer scientists may want to build data-driven models and will be confronted with the need to use data from multiple databases which all have a different API, different metadata schema and even the internal model of the data is different. To build such data-driven models, it is necessary to leverage the FAIR principles which provide a framework to transform data into machine actionable units of information so that one can automate the process of working with data.

In an ideal world, if FAIR principles were implemented everywhere, the current heterogeneous and distributed data landscape would be transformed into a FAIR data landscape where APIs would be harmonised, metadata schemas would be made interoperable, and all these data resources would be themselves findable. However, to achieve this ideal world requires both time and financial resources to make all these data repository FAIR, which are currently lacking. Hence, the Federated FAIR Data Space (FFDS) solution has been developed as a mean to create a FAIR data space on top of the existing heterogeneous and distributed data landscape without requiring changes of the existing repositories. The FFDS solution acts mostly at the metadata and the access level of the FAIR principles by allowing the harmonisation of heterogeneous metadata schema into a common schema (DCAT) through the creation of a harmonised API federation based on the smartAPI description and approach[58].

Different pieces of technology have been put together into a simple service architecture with at the reference implementation of the FAIR Data Point, developed in FAIRsFAIR (Behnke, et al., 2022) at the core of the architecture. A sort of API federation is formed by generating a smart API description

---

54 https://www.clarin.eu/content/about-clarin
55 https://www.eosc-pillar.eu/
56 https://specs.fairdatapoint.org/fdp-specs-v1.2.html
57 https://www.fairdatapoint.org/
58 https://smart-api.info/

of the APIs of each of the relevant data repositories that comprises how one extracts the metadata content. Then these machine actionable descriptions of APIs are used by a so-called "metadata Smart Harvester", which is going to take this API description, automatically generates the query and fetches all the committed metadata content from the different repository automatically. Those harvested metadata need to be mapped with a DCAT schema and published it into the FDP. Finally, data analytics tools can thus be linked to the FDP API which provide a centralised access point to the harmonised metadata from the repositories. This procedure can make every data repository FAIR, at least at the level of metadata.

To test the Federated Fair data space solution, the FAIRsFAIR team extended the FFDS solution and created a Proof of Concept (PoC) "search engine" for semantic artefacts across multiple community driven semantic artefact repository which leverage the minimum metadata schema for FAIR semantic artefact developed in FAIRsFAIR (Le Franc, Bonino, Koivula, Parland-von Essen, & Pergl, 2022). Another PoC for a cross-disciplinary semantic index for fast access to existing concepts and relations to enrich data and metadata has been developed in EOSC Pillar (Domínguez, et al., 2023).

Another important resource which would be useful for implementing FAIR principles in industry has been produced by the FAIRplus project[59], an Innovative Medicine Initiative (IMI)[60] funded project: the FAIR Cookbook[61]. Recipes of how to make data FAIR were collected in this FAIR Cookbook from academics and data professionals from the pharmaceutical industry. This effort includes research infrastructures like Elixir[62] which operates in the life sciences. Pharmaceutical industry has adopted FAIR (Wise, et al., 2019) from the very beginning as they want to use it to drive ML and AI. To start with, guidance for implementing FAIR was very generic. However, something more specific was needed to support life sciences with its different data types and different scenarios; but still it had to be endorsed by all players involved.

There is a collaboration with the Pistoia Alliance both communities are referring to each other's content because it has a different level of depth and granularity. The recipes of the Cookbook are very hands-on and are focused on both the technological aspect as well as giving a practical example of implementing data covering all four aspects of FAIR. Each data journey to FAIR is different and will come with its own challenges and the FAIR Cookbook offers a variety of recipes to combat them. People can not only read it but also contribute and is resides on GitHub.[63] All is curated by an editorial group who review the submissions and make sure there is a consistent building and growth of the content. Organisations must be aware that it will take more than one recipe to start one's first journey to FAIR and it is pertinent to find a trade-off between the level of technology and the work involved. This resource can be leveraged by other industry to start their FAIR journey.

3.5 FAIR Resources from GO FAIR

The mandate of GO FAIR[64] are statements about behaviours that would be needed or expected to automate the FAIR principles and the translation of those principles into actual concrete

[59] https://fairplus-project.eu/
[60] https://www.imi.europa.eu/
[61] https://faircookbook.elixir-europe.org/content/home.html
[62] https://elixir-europe.org/
[63] https://github.com/FAIRplus/the-fair-cookbook
[64] https://www.go-fair.org/

implementations. At some point, FAIR could be more and more automated, and humans would not have to be in the loop anymore.

Metadata is everywhere in the FAIR principles and has become a first-class citizen along with data itself. It is really the machine actionable metadata that would help to instruct agents to perform these FAIR operations that the community envision.

There are FAIR principles that would refer to technical implementations and these would be generic services that could be operating on data. This requires working with stakeholders that have a much more technical background and who are thinking about general infrastructure issues as well. Domain experts then have very strong opinions on vocabularies and the terminologies in data practices, etc. Both these groups must work together to realise the FAIR principles into concrete implementations. The domain experts who are data aware have the current vocabulary, a "lingua franca" that is needed to embed vocabularies, metadata, data formats, etc., that will constitute the FAIR data.

GO FAIR focusses a kind of openness to decentralised approaches, and these decentralised architectures are taken as the default and some more centralised approaches are taken when they fit for purpose. FAIR intrinsically relates to openness; however, not all data can be made open and restricted data protocols are a necessity.

GO FAIR's three-point framework maximises the reuse of existing resources, interoperability, and accelerates convergence on standards and technologies supporting FAIR data and services. The first point is metadata for machines (M4M), where domain experts work with metadata experts who know something about FAIR. Their goal is to create FAIR that are machine actionable schema and controlled vocabularies. This happens during M4M workshops which function as a kind of vehicle for the delivery of good metadata solutions. The second point is to create a FAIR implementation profile, i.e., a list of what is called FAIR enabling resources that are delivering on those FAIR principles. Each community and each implementation would have a unique list. GO FAIR formulated a set of precise questions to try and elicit the answers that are going to be most meaningful for a FAIR implementation profile. The latter are the connection between M4M and FAIR data points (FDPs), which are the last point of the verification framework. FDPs can be decentralised and embedded in a growing data network, and they can be shared (physically copied) or visited (never copied). When data are visited, queries have some sort of authentication behind them and there are consent agreements in place that allow those algorithms to go as deep into the data as they are permitted.

4. Conclusions

Our ambition is to move towards a Materials and Manufacturing Commons using the enablers Digital Marketplaces, FAIR Principles and Ontologies. The workshop focussed especially on a Materials Commons, which is envisaged as a trusted place for knowledge sharing for all stakeholders, while respecting data ownership. During the workshop, we concurred that inclusiveness, transparency and accountability are key elements of a Materials Commons. It needs to include the whole materials ecosystem, i.e. support the interaction and integration of all parts and aspects of the materials life cycle, and a digital materials commons ecosystem comprising materials data from modelling, characterisation, also materials related applications itself and their executability. All must be built on the FAIR principles, i.e., Findability, Accessibility, Interoperability, and Reusability of digital assets and of all semantic artefacts.

Knowledge Management Translators will have the mission to bridge the gap between digital expertise and materials sciences. This will require a collective effort of the diverse communities, to capitalise on these commonalities and integration synergies and contribute to a federated, highly harmonised interoperable materials data and knowledge infrastructure and lay the basis for a future global Materials Commons.

5. Acknowledgements

We kindly acknowledge the contributions to this workshop both in planning and execution from Florina Piroi, Katharina Flicker and Marie Czuray (TU Wien, Austria), Arkopaul Sarkar (Ecole Nationale D'Ingénieurs De Tarbes, France), Lan Yang (University of Galway, Ireland), Michela Magas (Industry Commons Foundation, Sweden), Lars Erikson (Jonkoping University, Sweden), Mark Illi (BEDA, Switzerland) and Cristina Mancarella and Rita Giuffrida (Trust-IT Services, Italy).

6. Disclaimer

Any views or opinions represented in this report are personal and belong solely to the authors and do not represent those of people, institutions, or organisations that the authors may or may not be associated with in professional or personal capacity. The information used and statements of fact made are not guarantees, warranties or representations as to their completeness or accuracy and may not reflect the opinion of the participants of the "Towards Materials and Manufacturing Commons - the enablers Digital Marketplaces, FAIR Principles and Ontologies" Workshop 2023. The authors assume no liability for any short term or long-term professional decisions made by any reader based on this report.

7. Acronyms and Abbreviations

AI – Artificial Intelligence

AMI2030 - Advanced Materials Initiative 2030

API – Application Programming Interface

B2B – Business-to-business

B2C - Business-to-consumer

BEDA - Bureau of European Design Associations

CEO - Chief Executive Officer

CHADA – Characterisation Data

csv - A comma-separated values file is a text file that uses a comma to separate values, giving rise to the name of this file format. https://en.wikipedia.org/wiki/Comma-separated_value

DCAT - Data Catalog Vocabulary, an RDF vocabulary designed to facilitate interoperability between data catalogues published on the web.

DSMS - Dataspace Management System

EC – European Commission

EMMC – European Materials Modelling Council

EMMO - Elementary Multiperspective Material Ontology

EOSC - European Open Science Cloud

EVMPO - European Virtual Marketplace Ontology

FAIR - Findable, Accessible, Interoperable and Reusable

FDPs - FAIR data points

GUID - globally unique identifier - a 128-bit text string that represents an identification

GUPRI - globally unique, persistent, and resolvable identifier

HTTP - Hypertext Transfer Protocol

HTTPS - Hypertext Transfer Protocol Secure

IMI - Innovative Medicine Initiative

JSON - JavaScript Object Notation is a lightweight data-interchange format. It is easy for humans to read and write. It is easy for machines to parse and generate. It is based on a subset of the JavaScript Programming Language Standard ECMA-262 3rd Edition - December 1999. https://www.json.org/json-en.html

M4M - metadata for machines

MDO - Materials Design Ontology

ML – machine learning

MODA – Modelling Data

OCES - Ontology Commons EcoSystem

OQMD - Open Quantum Materials Database

OWL - W3C Web Ontology Language

PMD - Platform Material Digital

PoC - proof of concept

QUDT - Quantities, Units, Dimensions, and Type Ontologies

RDF - Resource Description Framework

RDA – Research Data Alliance

SKOS - a common data model for knowledge organization systems such as thesauri, classification schemes, subject heading systems, and taxonomies.

SME – Small and Medium Enterprise

SPARQL - SPARQL Protocol and RDF Query Language, enables users to query information from databases or any data source that can be mapped to RDF.

Turtle - Terse RDF Triple Language

XML - Extensible Markup Language

VSSIG - Vocabulary and Semantic Services Interest Group

8. References

Amdouni, E., Bouazzouni, S., & Jonquet, C. (2022). O'FAIRe: Ontology FAIRness Evaluator in the AgroPortal semantic resource repository. . *ESWC 2022 - 19th Extended Semantic Web Conference, Poster and demonstration* (pp. 89-94). Cham: Springer Int. Publishing.

Behnke, C., Burger, K., Järveläinen, P., Le Franc, Y., Tykhonov, V., & Parland-von Essen, J. (2022, Feb 28). *D2.9 Second reference implementation of the data repositories features and client application (V1.0).* Retrieved from Zenodo: https://doi.org/10.5281/zenodo.6699160

CWA 17284. (2018). *CEN/WS MODA - Materials modelling - terminology, classification and metadata .* Retrieved from European Committee for Standardization: https://www.cen.eu/news/workshops/Pages/WS-2017-012.aspx

CWA 17815. (2021, October). *Materials characterisation - Terminology, metadata and classification.* Retrieved from CEN-CENELEC: https://www.cencenelec.eu/media/CEN-CENELEC/CWAs/ICT/cwa17815.pdf

d'Aquin, M. (2021). *Ontology Ecosystem Specification.* Retrieved from Zenodo: https://doi.org/10.5281/zenodo.6504633

Domínguez, J. M., Huschle, T., Toulemonde, B., Claude, A., Sarramia, D., Todor, A., & Lombardo, F. (2023, Jan 23). *EOSC-Pillar D5.5 - Service architecture providing fast access to the concepts held within semantic artefacts (1.2).* Retrieved from Zenodo: https://doi.org/10.5281/zenodo.7622080

Garijo, D., & Poveda-Villalón, M. (2020, Mar 29). *Best Practices for Implementing FAIR Vocabularies and Ontologies on the Web.* Retrieved from arXiv: https://doi.org/10.48550/arXiv.2003.13084

Garijo, D., Corcho, O., & Poveda-Villalón, M. (2021). *FOOPS!: An Ontology Pitfall Scanner for the FAIR principles.* Retrieved from International Semantic Web Conference (ISWC) 2021: Posters, Demos, and Industry Tracks: https://ceur-ws.org/Vol-2980/paper321.pdf

Goldbeck, G., Simperler, A., Bull, L. G., Karray, M., Kharlamov, E., Kiritsis, D., . . . Noeske, M. P. (2022). The Translator in Knowledge Management for Innovation – towards Industry Commons. *Zenodo*.

Goldbeck, G., Simperler, A., de Andres, P., Bierwisch, C., Cloete, H., Di Stefano, D., . . . al. (2023, Sept 9). *MarketPlace - a Digital Materials Modelling Marketplace.* Retrieved from Zenodo: https://doi.org/10.5281/zenodo.8330333

Horsch, M., Chiacchiera, S., Seaton, M., Todorov, I., Šindelka, K., Lísal, M., . . . Leite Cavalcanti, W. (2020). Ontologies for the Virtual Materials Marketplace. *Künstl Intell 34*, 423–428.

Janowicz, K., Hitzler, P., Adams, B., Kolas, D., & Vardeman II, C. (2014, Mar 31). *Five Stars of Linked Data Vocabulary Use.* Retrieved from Semantic Web Journal: https://www.semantic-web-journal.net/content/five-stars-linked-data-vocabulary-use

Klein, P., Konchakova, N., Hristova-Bogaerds, D. G., Noeske, M., Simperler, A., Goldbeck, G., & Höche, D. (2021, Apr 30). *Translation in Materials Modelling – Process and Progress*. Retrieved from Zenodo: https://doi.org/10.5281/zenodo.4729918

Le Franc, Y. (2022). *Report on existing domain ontologies.* Retrieved from Zenodo: https://doi.org/10.5281/zenodo.6504553

Le Franc, Y., Bonino, L., Koivula, H., Parland-von Essen, J., & Pergl, R. (2022). *FAIR Semantics Recommendations Third Iteration (V1.0).* Retrieved from Zenodo: https://doi.org/10.5281/zenodo.6675295

Poveda-Villalón, M., Espinoza-Arias, P., Garijo, D., & Corcho, O. (2020). Coming to Terms with FAIR Ontologies. In C. Keet, & M. Dumontier, *Knowledge Engineering and Knowledge Management. EKAW 2020. Lecture Notes in Computer Science, vol 12387* (pp. 255–270). Cham: Springer.

Vatant, B. (2012, Feb 10). *Is your linked data vocabulary 5-star?* Retrieved from blogspot: https://bvatant.blogspot.com/2012/02/is-your-linked-data-vocabulary-5-star_9588.html

Wilkinson, M., Dumontier, M., Aalbersberg, I., & al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data 3*, 160018.

Wilkinson, M., Sansone, S.-A., Grootveld, M., Nordling, J., Richard, D., & Hecker, D. (2022). *FAIR Assessment Tools: Towards an "Apples to Apples" Comparisons.* Retrieved from Zenodo: https://doi.org/10.5281/zenodo.7463421

Wise, J., Grebe de Barron, A., Splendiani, A., Balali-Mood, B., Vasant, D., Little, E., . . . Hedley, V. (2019). Implementation and relevance of FAIR data principles in biopharmaceutical R&D. *Drug Discovery Today 24 (4)*, 933-938.

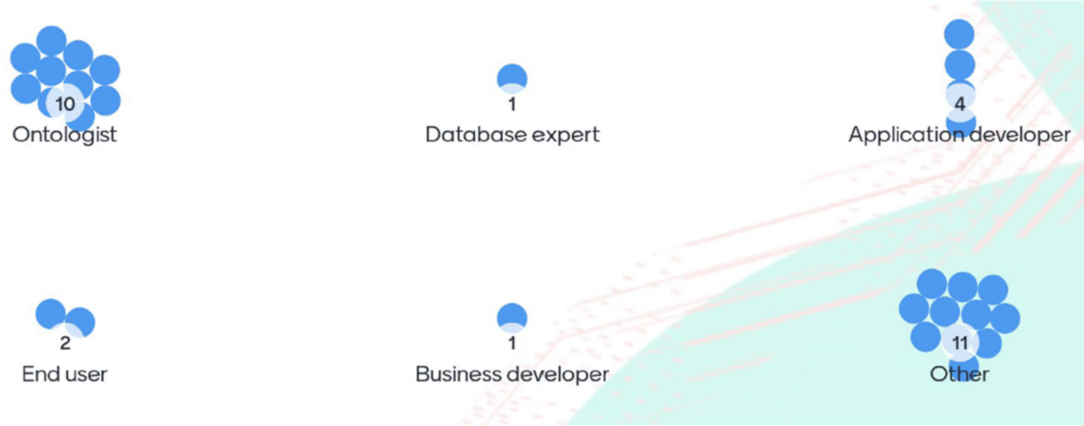
World Economic Forum. (2020, Jan 13). *Share to Gain: Unlocking Data Value in Manufacturing*. Retrieved from https://www.weforum.org/publications/share-to-gain-unlocking-data-value-in-manufacturing

Zicari, R. (2018, Sep 10). *On using AI and Data Analytics in Pharmaceutical Research. Interview with Bryn Roberts .* Retrieved from ODBMS Industry Watch blog: https://www.odbms.org/blog/2018/09/on-using-ai-and-data-analytics-in-pharmaceutical-research-interview-with-bryn-roberts/

9. Appendix A: Who our Participants were

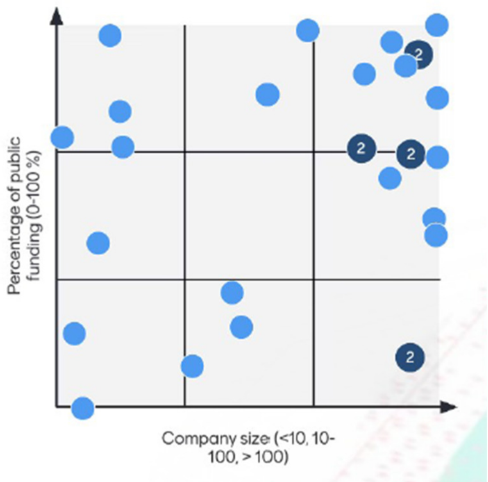
The first day of the workshop ended with an interactive session to learn more about the participants and their expectations form the workshop.

In which role are you here today?



We did have a fair number of ontologists in the meeting and persons of many backgrounds.

Where does your Institute or company sit on this map? We offered on the Y axis the % of public funding from zero to 100% and on the X axis the company size.



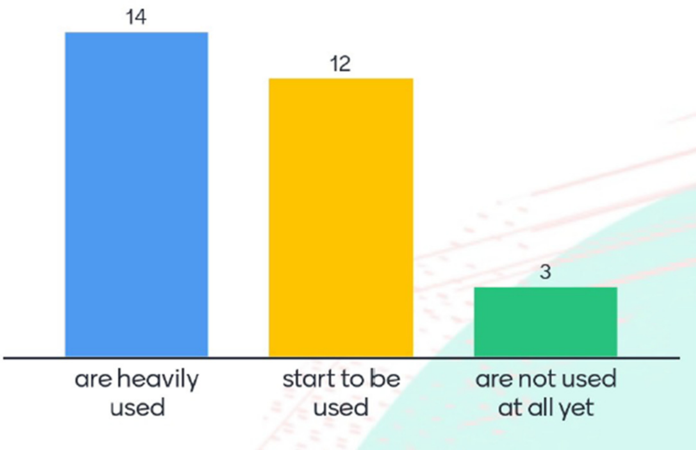
Most persons we hosted are working in organisations with large public funding %.

What are the main application domains?

This was an open question and we generated a word cloud.

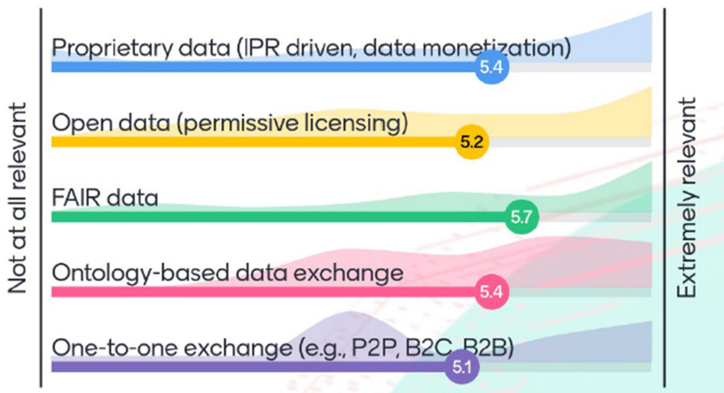
As expected, due to the nature of our workshop, "manufacturing" was a main application domain and material, simulation, and materials modelling featured strongly.

In your institute/company, semantic technologies …



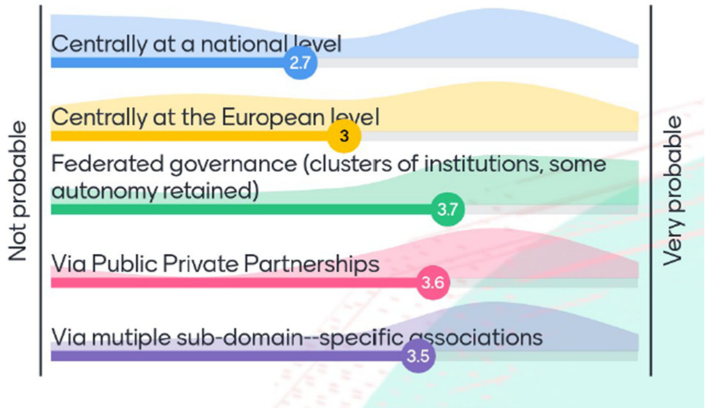
… are either heavily used or start to be used and only 10% of the respondents are not using them at all.

Materials research data: how relevant will these dimensions be for data exchange in 2030?



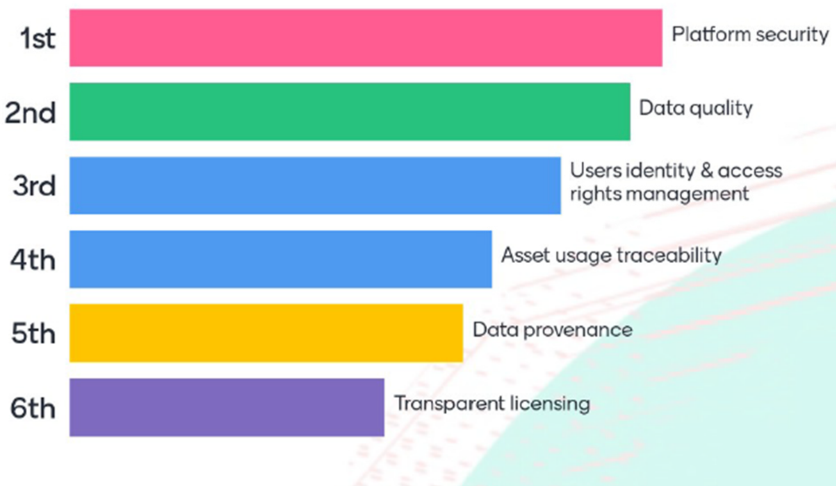
We could say, the first four options put together can be realised with marketplaces and all of them deemed important. The underlying graphs (shaded areas) depict the distribution of the answers. A simple one-to-one exchange is deemed the least important.

Materials Research Data: how will it be governed in 2030?



Federated governance and public private partnerships were in the lead, which means data spaces is the way to go.

As a user/provider of data/software, what features of a common data space are necessary to trust it?

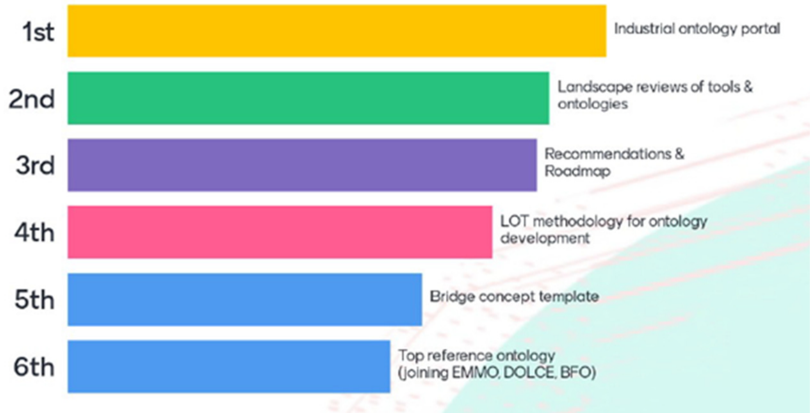


Platform security and data quality where the two most important.

As a user/provider of data/software, which factors prevent you from using a common data space?

The outcome was that extra development is needed, which means higher technology readiness levels. Some other factors, extracted from an open question were how people feel about other people's data. Did the data owners have the right skill set to produce them? Did they use the correct standards? Can they trust this 3rd party data? There was also a notion of dis-jointness across groups, and the complexity of federation. This is also an important point, as this complexity arises both from a technical and organisational point of view.

What outcomes of the OntoCommons project are you most likely to use in the next 5 years?



The Industrial Ontology Portal of appears to be the most like concrete and usable item that we see in the next five years to be used. This reflects the need to find ontologies and assess them. This is followed by landscape reviews of tools and ontologies and the third item to be most likely used will be the roadmap.

10. Appendix B: Workshop Programme

Day 1, Tuesday, April 4th 2023

Welcome and Introduction

Dirk Helm (Fraunhofer IWM), Hedi Karray (ENIT), Laszlo Hetey (European Commission)

Session 1: Key elements of Materials and Manufacturing Commons

Chair: Dimitris Kiritsis (EPFL-UiO)

- **Industry Commons: Towards a Common European Data Space** - Michela Magas (ICF)
- **OntoCommons: standardising materials and manufacturing data documentation to support data sharing in a common data space** - Hedi Karray (ENIT)
- **Digital Marketplaces based on shared data documentation principles** - Amit Bhave (CMCL), Dirk Helm (Fraunhofer IWM)
- **FAIR-Principles** - Yann Le Franc (e-science Data Factory)
- **Towards a Materials Commons: Materials Digitalisation in AMI2030 Roadmap** - Gerhard Goldbeck (Goldbeck Consulting Ltd)
- **Materials Commons: knowledge sharing across the materials ecosystem** - Andrea Ceglia (European Commission)

Session 2: Towards implementations of Materials and Manufacturing Commons - Digital Marketplaces

Chair: Yoav Nahshon (Fraunhofer IWM)

- **Dome4.0: from dataset to data documentation using ontologies** - Amit Bhave (CMCL)
- **MarketPlace** - Dirk Helm (Fraunhofer IWM)
- **VIMMP** - Silvia Chiacchiera (UKRI)

Session 3: Towards implementations of Materials and Manufacturing Commons – OntoCommons

Chair: Gerhard Goldbeck (Goldbeck Consulting)

- **Ontology Commons EcoSystem - ontology standardisation guidelines and alignment tools** - Arkopaul Sarkar (ENIT)
- **Use case DOME4.0 and OntoCommons – lessons learnt for ontology-based commons: Semantically Empowered Industry 4.0 @Bosch** - Evgeny Kharlamov (BOSCH)
- **Use case DOME4.0 and OntoCommons - lessons learnt for ontology-based commons: Materials Databases Integration using the Materials Design Ontology** - Huanyu Li (Linköping University)
- **Alignment of DOME4.0 with OntoCommons** - Emanuele Ghedini (Università di Bologna)
- **Knowledge Management Translator: skills and human resources development** - Alexandra Simperler (Goldbeck Consulting Ltd)

Session 4: Towards implementations of Materials and Manufacturing Commons - Data Spaces

Chair: Michela Magas (ICF)

- **EU Data Spaces - Data Space 4.0 CSA** - Alberto Abella (FIWARE) and Clara Pezuela (FIWARE)
- **The contribution of the Platform Material Digital (PMD) in building up a Materials Data Space – application to glass design and manufacturing** - Pedro Portella (Fraunhofer IWM)
- **EOSC presentation** - Roksana Wilk (EOSC Future)
- **A Data Space Management System** - Tobias Huschle (Fraunhofer IWM)

Session 5: Towards implementations of Materials and Manufacturing Commons:

Chair: Mark Illi (BEDA)

- **Introduction of BEDA** (slides) Mark Illi (BEDA)
- **Basajaun project - tracking the lifecycle of wood** - Andreas Rudenå (Paramountric)
- **The design perspective** - Lars Eriksson (Jonkoping University / BEDA)
- **From Mirabel Slabbnick to Quantifactum** - Christof Ameye (Quantifactum)

Day 2, Wednesday, April 5th 2023

Session 1: What are the FAIR Principles?

Chair: Yann Le Franc (e-science Data Factory)

- **Keynote: Introduction to FAIR Principles** - Barend Mons (GO FAIR Foundation)
- **FAIR Principles implementation at Roche** - Dr. Martin Romacker & Nick Perry (Roche)

Session 2: Examples from OntoCommons

Chair: Yann Le Franc (e-science Data Factory)

- **FAIR Principles implementation at BOSCH** - Irlan Grangel Gonzalez (BOSCH)
- **An entrepreneur perspective on FAIR** - Dermot Doyle (Dynaccurate)

Session 3: FAIR principles in OntoCommons

Chair: Yann Le Franc (e-science Data Factory)

- **FAIR Ontologies: a requirement for interoperability** - Yann Le Franc (e-science Data Factory)
- **How can we evaluate the FAIR Compliance of Ontologies?** - Maria Poveda (UPM)

Session 4: FAIR Resources for Industry: what is happening in EOSC?

Chair: Yann Le Franc (e-science Data Factory)

- **EOSC FAIR Metrics and Data Quality** - Romain David (ERINHA)
- **EOSC Semantic Interoperability Task Force** - Wolmar Nyberg Åkerström (Uppsala University)
- **FAIR Cookbook & FAIRSharing** - Susanna Assunta Sansone (Oxford University)
- **FAIR Impact** - Pascal Flohr (DANS-KNAW)
- **FAIRCORE4EOSC** - Tommi Suominnen (CSC)

Session 5: FAIR Resources from GO FAIR

Chair: Yann Le Franc (e-science Data Factory)

- **M4M Workshops & FAIR Implementation Profiles** - Erik Shultes (GO FAIR Foundation)
- **Creating federated FAIR Data Space with the FAIR Data Point** - Yann Le Franc (e-Science Data Factory)

Day 3, Thursday, April 6th 2023

Session 1: OntoCommons Roadmap collecting Industry needs

Chair: Silvia Chiacchiera (UKRI)

Introduction to the OntoCommons Roadmap - Hedi Karray (ENIT)

Pitches:

- Dr. Sebastian Brückner: **Data Structures and Tools for FAIR Synthesis data** (slides)
- Achim Kohler et al.: **Green Data Lab**
- Nick Garabedian, Ilia Bagov, Christian Greiner: **Collaborative Metadata Definition using FAIR Controlled Vocabularies, and Ontologies**
- Martin Thomas Horsch, Björn Schembera, and Simon Stephan: **XAIR research data and epistemic metadata for molecular methods**
- Dr. Natalia Konchakova: **Virtual Open Innovation Platform for Active Protective Coatings Guided by Modelling and Optimization (VIPCOAT)**

Collect input for updated version of Roadmap from participants - Silvia Chiacchiera

Session 2: Demos

Chair: Joana Morgado (Fraunhofer IWM)

- **Demo Ontology Commons EcoSystem (OCES)** - Arkopaul Sarkar (ENIT)
- **Demo FAIRness score/FAIRification** - María Poveda (UPM)
- **Demo MarketPlace** - Yoav Nahshon (Fraunhofer IWM), Pablo de Andres (Fraunhofer IWM)

Session 3: Hands-on Q&A Session

Chair: Pablo de Andres (Fraunhofer IWM)

- **Q&A: OCES** - Arkopaul Sarkar (ENIT)
- **Q&A: FAIRification** - María Poveda (UPM), Yann LeFranc (e-science Data Factory)
- **Q&A: Digital Marketplaces** - Yoav Nahshon (Fraunhofer IWM), Dirk Helm (Fraunhofer IWM), Amit Bhave (CMCL)