PROJECT DELIVERABLE REPORT

Grant Agreement Number: 101058732

Joint Industrial Data Exchange Platform

Type: Deliverable Report

D2.2 Interoperability specifications and FAIRness compliance

Issuing partner	University of Cambridge (UCAM)
Participating partners	Universita di Trento (UNITN) Technovative Solutions (TVS)
Document name and revision	D2.2 Interoperability specifications and FAIRness compliance
Author	Dr Feroz Farazi
Deliverable due date	31 May 2023
Actual submission date	31 May 2023

Project coordinator	Vorarlberg University of Applied Sciences
Tel	+43 (0) 5572 792 7128
E-mail	Florian.maurer@fhv.at
Project website address	www.jidep.eu

Dissemination Level		
PU	Public	✓
PP	Restricted to other programme participants (including the Commission services)	
CO	Confidential, only for members of the consortium (including the Commission services)	
SEN	Sensitive, limited under the conditions of the Grant Agreement	

Contents

European Commission

Disclaimer

JIDEP has received funding from the European Commission under the Grant Agreement no.101058732. The content of this document does not represent the opinion of the European Commission, and the European Commission is not responsible for any use that might be made of such content.

Acronyms and Abbreviations

AML	Automation Markup Language
EAN	European Article Number
FAIR	Findability, Accessibility, Interoperability, and Reusability
GTIN	Global Trade Item Number
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IOF	Industrial Ontologies Foundry
ISO	International Organisation for Standardisation
ITU	International Telecommunication Union
MIMOSA	Machinery Information Management Open Systems Alliance
OLE	Object Linking and Embedding
OPC UA	OLE for Process Control Unified Architecture
OIIE	Open Industrial Interoperability Ecosystem
OWL	Web Ontology Language
RAMI	Reference Architecture Model for Industry
RDF	Resource Description Framework
SDAIR	Structured Digital Asset Interoperability Registry
SPARQL	SPARQL Protocol and RDF Query Language
STEP	Standard for the Exchange of Product model data
TC	Technical Committee
TVS	Technovative Solutions
UCAM	University of Cambridge
UNITN	University of Trento
URL	Uniform Resource Locator
UUID	Universally Unique Identifier
W3C	World Wide Web Consortium
XML	Extensible Markup Language

1. Executive Summary

One of the two goals of this deliverable is to specify how the JIDEP project achieves interoperability to enable the effective sharing of recyclable product, component and material data to meet the data publishing needs of suppliers and information consumption needs of manufacturers to be able to make an informed decision regarding the type of reusability and circularity. Supplier-published data should fulfil specific criteria concerning openness, licensing, machine processability, etc. Hence, the other goal of this deliverable is to define principles that ensure that the data published by data-providing partners and collected from similar projects conform to the FAIR principles that make the data findable, accessible, interoperable and reusable. To achieve the goals, we have analysed existing industrial product, component and material data publishing, exchange and interoperability standards to specify the required criteria for JIDEP. We have reviewed the FAIR principles and identified the principles the JIDEP project will adopt. We have implemented a workflow to verify the compliance of distributed data contributions via the data storage units with these principles.

2. Introduction

The overall objective of this deliverable is to define standards and principles for data contributors in a distributed computing environment so that the JIDEP project can enable data interoperability for effective data exchange and share and ensure compliance with the FAIR principles for finding, using and reusing product, component and material data. The International Organisation for Standardisation (ISO) specified ISO 10303 as a Standard for the Exchange of Product model data (STEP). ISO 10303 provides the specification for the representation of the complete life-cycle data of a product [33]. It has many parts, each of which is a standard developed for a specific purpose. For example, ISO 10303-41 refers to Part 41, designed to specify the constructs for fundamentals of the product description, management and support, and ISO 10303-11 relates to Part 11 defines a language called EXPRESS for describing different aspects of products, data types and constraints.

Object Linking and Embedding for Process Control (OPC) Unified Architecture (UA) is a standard defined by the OPC Foundation established to develop industrial interoperability standards [34]. OPC UA is a platform-independent standard providing the infrastructure for seamless data exchange between machines, from machine to enterprise and across the enterprise. MTConnect is a standard created for data and information exchange between manufacturing operations linked with machine tools, manufacturing equipment, sensors and software applications [35]. It provides a vocabulary for manufacturing equipment to communicate the semantics, structure and context of data being exchanged. The Reference Architecture Model for Industry (RAMI), a German standard, offers a common structure for uniformity to overcome the complexity barrier, develop a shared understanding, and achieve interoperability for data exchange [21]. The Reference Architecture Model has three dimensions representing product, production, maintenance, asset, integration and communication, etc. [8]. Mapping RAMI with standards developed in other countries and reference architectures has the potential to enable interoperability in industries worldwide.

The Machinery Information Management Open Systems Alliance (MIMOSA) standard is specified for improving data interoperability in managing the lifecycle of assets and critical infrastructure in a heterogeneous cross-domain computing environment [23]. MIMOSA collaborates with other industry associations to create collaborative standards and specifications for achieving interoperability [24]. FIWARE is a framework for standardisation consisting of open-source software components developed to provide instant, interoperable access to data to allow the creation of intelligent solutions for different application areas, including cities and manufacturing industries, to convert them into smart cities and smart manufacturing industries, respectively [28].

Industrial Ontologies Foundry (IOF) is a working group formed to develop reference ontologies for describing the general domain knowledge of manufacturing and engineering industries [25]. IOF sets principles and best practices for defining quality ontologies to enable interoperability and data exchange across industrial domains. As ontologies themselves may not be interoperable, by following the successful strategy of the Open Biomedical Ontologies (OBO) Foundry IOF is formed to offer guidelines in defining high-quality ontologies that will lead to the development of a suite of interoperable software tools allowing consistent access to and reasoning over data at any stage of the product life cycle [3]. The Findable, Accessible, Interoperable and Reusable (FAIR) data principles are technology-agnostic guidelines that encourage data owners to prepare and publish data to be FAIR compliant and allow machines to find and use data automatically [6].

The rest of the document is structured as follows. Section 3 describes the specification details of widely known and used standards for industrial data exchange. Section 4 reports on the interoperability and standardisation approaches implemented in the related projects. Section 5 describes the FAIR principles and a workflow for verifying FAIR compliance. Section 6 defines the standards and principles required to enable interoperability in the project, and Section 7 concludes the report.

3. Industrial Data Exchange Standards

3.1 ISO

ISO 10303 [12] is a machine-processable product data representation and exchange standard that can describe products at any phase of the product life cycle. In ISO 10303, a product is defined as a thing, object or substance produced by a natural or artificial process. One example of a product is a FIAT car. In an automotive monocoque, there can be several crossbeams, and crossbeam is a class of products. A component is defined as a product that is not subject to decomposition from the perspective of a specific application. One example of a component is the automotive windscreen.

EXPRESS is a data representation language consisting of language constructs or elements allowing unambiguous data definition with specified constraints [13]. The product, class of products and component are represented using the product defined as an entity or entity data type in the EXPRESS language [14]. A partial representation of the product provided in ISO 10303-41 is shown below.

```
ENTITY product;
  id : identifier;
  name : label;
  description : OPTIONAL text;
END_ENTITY;
```

In this representation, id, name, and description are product attributes. The id attribute refers to the identifier of the data type string that distinguishes the product. The name attribute refers to the label of the data type string and the designation by which the product is known to the world. The description attribute is an optional attribute that refers to the text of the data type string and that informally defines and characterises the product when provided.

A product relationship is defined as an entity forming a relationship between two products providing an identification, name and description of the relationship [14]. This relationship may be parent-child, part-whole or any other relationship meaningfully linking instances of the entity product. The specification of product relationship as described in ISO 10303-41 is shown below.

```
ENTITY product_relationship;
  id : identifier;
  name : label;
  description : OPTIONAL text;
  relating_product : product;
  related_product : product;
END_ENTITY;
```

The relationship between two products is described with the attributes id, name, description, and relating/related product. The id attribute distinguishes one product relationship from another, the name attribute labels the relationship, and the description attribute characterises the relationship. The relating_product attribute refers to one of the two products participating in the product relationship. The related_product attribute refers to the other product that is linked with relating_product. One of the products in this relationship may be dependent on the other and in this case, related_product should be used to represent the dependent one.

Material is defined as one or more substances used to manufacture a product [15]. Stock material is defined as the material typically obtained from a supplier to manufacture a part [16]. A discontinuous fibre assembly is a stock material consisting of a homogeneous material matrix and a collection of discontinuous fibres. Braided assembly, anisotropic material, isotropic material, filament assembly and woven assembly are also specific types of stock materials. All these particular types of stock material are specified as entities in ISO 10303-1771 [16].

The measure is defined as the description of physical quantities, including length, area, mass and volume in ISO 10303-41 [14]. A length measure refers to the numerical representation of a distance, and an area measure refers to the numerical representation of the size of a surface. A mass measure is the numerical representation of the quantity of matter a body or product contains. A volume measure refers to the numerical representation of the solid content within a body or product.

A unit represents a physical quantity that has a constant value of one [14]. The SI units included in ISO 10303-41 with reference to ISO 1000 relevant for representing product and material data are metre, gram, second, kelvin, degree celsius, mole, newton, pascal and joule. An SI prefix refers to the name of a prefix that may be connected with an SI unit. ISO 10303 supports using SI prefixes: exa, peta, tera, giga, kilo, hector, deca, deci, centi, milli, micro, nano, pico, femto, and atto.

An organisation is defined as an entity describing administrative structure using the attributes id, name and description. The organisation specification, as described in ISO 10303-41, is shown below.

```
ENTITY organization;
  id : OPTIONAL identifier;
  name : label;
  description : OPTIONAL text;
END_ENTITY;
```

The id attribute is optional, but it distinguishes the organisation from the rest when provided. The name attribute provides the label by which the organisation is recognised, and the description attribute includes textual information that characterises the organisation.

Organisation_relationship is defined as an entity establishing a relationship between two organisations, with the attributes name, description, relating organisation and related organisation [14]. The organisation_relationship entity can establish a relationship between a department of an enterprise and the enterprise itself, where the department is a part of the whole enterprise. This entity may be used to establish parent-child and part-whole relationships. The specification of organisation_relationship described in ISO 10303-41 is shown below.

```
ENTITY organization_relationship;
  name : label;
  description : OPTIONAL text;
  relating_organization : organization;
  related_organization : organization;
END_ENTITY;
```

The name attribute denotes the label by which the organisation_relationship is recognised. When provided, the optional description attribute refers to the textual description that characterises the organisation_relationship. The relating organisation attribute refers to one of the two organisations participating in the relationship. The related_organisation attribute refers to the other organisation participating in the same relationship. While describing the dependency of one organisation on the other, the related_organisation attribute refers to the dependent one.

A person is an entity describing an individual human being with the attributes id, last name, first name, middle name(s), prefix titles, suffix titles, etc. A partial representation of the person described in ISO 10303-41 is shown below.

```
ENTITY person;
  id : identifier;
  last_name : OPTIONAL label;
  first_name : OPTIONAL label;
  middle_names : OPTIONAL LIST[1:?] OF label;
  prefix_titles : OPTIONAL LIST[1:?] OF label;
  suffix_titles : OPTIONAL LIST[1:?] OF label;
END_ENTITY;
```

The id attribute refers to the identifier that uniquely recognizes a person. Last name and first name refer to the surname and first forename of the person, respectively. The middle names attribute refers to all the forenames but the first. Both the prefix titles and suffix titles attributes represent the text that indicates the social or professional standing of the person. Prefix titles appear before the names, and suffix titles appear after the names of the person. Except for the id attribute, all the attributes are optional, meaning these may or may not be provided.

Person_and_organisation is defined as an entity referring to a person in an organization with the specification of attributes of the person, the organization, name and description. A partial representation of the person_and_organisation entity described in ISO 10303-41 is shown below.

```
ENTITY person_and_organization;
  the_person : person;
  the_organization : organization;
  name : label;
  description : text;
END_ENTITY;
```

In the person_and_organisation entity, the person attribute refers to the person who is related to an organisation; the organisation attribute refers to the organisation the person is related; the name attribute refers to the label by which the entity is recognised, and the description attribute refers to the text characterising the entity.

Person_and_organisation_role is defined as an entity referring to the role of a person in the context of an organisation. For example, the same person could have a buyer role in their employer's organisation and a customer role in the organisation that sells products to them. A partial view of the person and organisation role entity described in ISO 10303-41 is shown below.

```
ENTITY person_and_organization_role;
  name : label;
  description : text;
END_ENTITY;
```

The person_and_organisation_role entity consists of the name and description attributes, where the name refers to the label by which the entity is recognised, and the description attribute refers to the text characterising the entity.

3.2 OPC UA

Object Linking and Embedding for Process Control (OPC) Unified Architecture (UA) refers to a standard defined by the OPC Foundation to enable interoperability in process automation [17]. OPC UA is a platform-agnostic open standard ensuring seamless data exchange via open communication technologies such as TCP/IP and HTTP among devices procured from multiple manufacturers [18]. OPC UA provides a standard infrastructure model built with the information model, message model, communication model and conformance model for the exchange of information [19]. The information model provides structure and semantics of the data and information being exchanged. The message model supports the interaction between applications or systems to exchange data and information. The communication model enables the transfer of data and information between two endpoints. The conformance model ensures interoperability between applications or systems.

The OPC UA [19] standard supports communications between various devices and systems through request and response messages over different types of networks. It applies the client-server or publisher-subscriber model for robust, secure and attack-resistant communication to ensure the identity of interacting applications. In a client-server model, clients can be informed about the services individual servers can offer using a set of data types defined by themselves and vendors. Servers represent object models that are dynamically accessible to clients. In the OPC UA, clients can access server-provided latest and historical data, alarms and events to be aware of crucial changes. OPC UA supports several communication protocols, including OPC UA TCP, HTTPS and Web Sockets and the trade-off between portability and efficiency as it enables data encoding in various formats. In the context of the publisher-subscriber model, OPC UA provides the PubSub model for publishers to send information to subscribers.

The design goals of OPC UA are manifold [19]. By design, it has an integrated service model and address space that maintains the consistent behaviour of systems and devices. It employs a single server to integrate current, historical data and messages communicating alarms and events into the address space. It provides a combined set of secured services to access the content of this address space. Clients receive server-provided objects with type definitions from the address space contributing to content comprehensibility. OPC UA data can have various transportation formats, such as XML, JSON and binary, that OPC, other standards or vendors can define and that clients can retrieve from the address space by sending a query to the server. This offers the advantage of learning the data format at runtime for effective data consumption.

3.3 MTconnect

MTConnect [20] is a standard for data and information exchange between manufacturing operations linked with manufacturing equipment and software applications. MTConnect relies on a data dictionary to describe the meaning of terms required for describing information related to manufacturing operations. It provides a set of semantic data models to represent the relationship between a manufacturing operation and the associated data to explain the data's origin unambiguously, allowing straightforward interpretation of data from different sources. It is an extensible standard; therefore, any implementation can address the need for more coverage of the term and meaning of data in the data dictionary and semantic models by extending these digital resources.

The design goal of MTConnect is to improve the equipment data acquisition in manufacturing industries to help them implement data-driven decision-making and to create a plug-and-play environment for manufacturing equipment and software applications to minimise the cost of software integration with manufacturing systems.

The communication methods of MTConnect are Request/Response and Publish/Subscribe. HTTP and XML dominate as the transport protocol and representation language for semantic data models in MTConnect, respectively. However, the standard specification does not impose any restrictions in selecting other protocols or languages. MTConnect leverages the most widely known and used standards in the manufacturing and software sectors to ease industry implementation and enable data interoperability across manufacturing operations.

3.4 RAMI

The Reference Architecture Model for Industry (RAMI) [21] is a German standard proposing a typical structure for uniformity in system architectures' conceptual and concrete design to enable experts with different backgrounds to speak a common language, overcome the complexity barrier, and develop a shared understanding. The Reference Architecture Model has three dimensions representing the crucial aspects of Industry 4.0. The dimension that coincides with the x-axis refers to the product, field device, control device, station, work centres and connected world hierarchy levels. The hierarchy levels strongly align with IEC 62264, an international standard for enterprise control-system integration. The dimension that coincides with the y-axis refers to the life cycle and value stream, including type (development, maintenance usage) and instance (production, maintenance usage). The type represents the prototype development of a product, and the instance refers to the product manufactured in the production department. This dimension strongly aligns with IEC 62890, an international standard for industrial-process measurement, control and automation life-cycle management for systems and components. The dimension that coincides with the vertical axis represents the asset, integration, communication, information, functional, and business layers. Mapping RAMI with nationally developed standards and reference architectures has the potential to enable interoperability in industries worldwide.

RAMI includes different user perspectives and allows industry associations and standardisation committees to address the requirements of use cases in various domains, from manufacturing and mechanical engineering to process engineering [22]. It is a 3D map of Industry 4.0, allowing the visualisation of industrial requirements against country-specific and international standards to recognise the overlapping standards, identify, and address the issue of gaps.

3.5 MIMOSA

Machinery Information Management Open Systems Alliance (MIMOSA) [23][24] is a standard developed for enhancing the interoperability of information in managing the lifecycle of assets and critical infrastructure in a cross-sectorial, highly heterogeneous computing environment. MIMOSA collaborated with other industry associations, established the Open Industrial Interoperability Ecosystem (OIIE) and defined collaborative standards and specifications for achieving interoperability at a broader scale. MIMOSA provides collaboratively developed standards for developing and operating supplier-agnostic digital twins to identify the context of various sensor-generated big data and to perform analytics.

A Service Directory [23] is provided for centrally managing the service configuration of individual systems forming the OIIE ecosystem. One example of such a system is a software application. A service directory allows consumers to recognise a service configuration by specifying the relationship between a system and its configuration. The configuration has several properties, i.e., endpoint, endpoint type, channel, topic and security token. Some examples of endpoint types are provider publication and consumer request. Structured Digital Asset Interoperability Registry (SDAIR) [23] within MIMOSA manages object identification, distributed data sources and metadata to enable interoperability across different systems, e.g., engineering and operation & maintenance systems. The Service Directory and SDAIR are jointly responsible for managing supplier and application-neutral ecosystem administration functions to create and maintain an instance of OIIE that allows all applications to interact and operate within this interoperable system of systems at any instant.

3.6 IOF

The Industrial Ontologies Foundry (IOF) [25][26] working group has been formed to define reference ontologies for representing the more common or general notions of manufacturing and engineering industries. The goals of IOF are to provide a suite of principles-based ontologies that is open to use and that can form the basis for creating domain and application-specific ontologies and to develop principles and best practices for defining quality ontologies that will enable interoperability in industrial domains. IOF-compatible ontologies should be humanly comprehensible via the natural language description and machine-interpretable via a formal language representation, have documentation to allow users to learn the ontology development and provide ontological definitions of classes and properties. For coherency and consistency, IOF adopts the layered approach of ontology organisation. The top layer is the more generic foundational or upper ontology layer. The layer below the top layer consists of the domain-independent mid-level ontology and domain-specific reference ontology. The layer below this is called the subdomain ontology layer, and the bottom layer is called the application ontologies layer.

Ontologies belonging to the top two layers are IOF ontologies. IOF emphasises that an RDF/XML representation of any IFO ontology must be available. An IFO ontology can be available in a formal language, such as Common Logic or OWL, using one of its variants, e.g., OWL2-XML, OWL2-Manchenster Syntax and Turtle. Ontologies within IOF should be freely and openly available to the industrial community. IOF allows the reuse of classes or properties from external ontologies, but using their original IRI is mandatory. Use of an entire ontology through the import statement is permitted if the ontology is published with an open, flexible license that does not impose any restrictions.

The IOF approach recommends only one reference ontology per domain. The scope of coverage and context in which an ontology is developed should be described clearly because such clarification can minimise confusion, enable users to find the intended ontology with reduced effort and enhance possible use and reuse. IOF recommends using the *abstract* property from the Dublin Core (DC) terms [27] to describe the scope and context. Dependency on external ontologies should be explicitly included in the ontology as it contributes to the identification of scope and context.

3.7 FIWARE

FIWARE [28][29] is a framework of open-source software components developed to create smart solutions for different application areas. The goal is to transform, for example, cities, manufacturing industries and utilities respectively into smart cities, smart manufacturing industries and smart utilities. In FIWARE, data representation is digital twin centric and data from different sources, including sensors, information systems and social media, are represented and stored in a way that allows instant access to the data. A digital twin is a digital representation of a real-world object. FIWARE standardises both the data access API and data models. The FIWARE community used the NGSI API for data access, improved it for production-level use and published the new version as NGSIv2. The community and the FIWARE Foundation evolved the API further and released it as NGSI-LD. FIWARE-compatible software architectures use the NGSI-LD API for data integration. A standardised library of data models in FIWARE is described in JSON and JOSN-LD, of which the former is compatible with the NGSIv2 API, and the latter is compatible with the NGSI-LD API.

In FIWARE, data sharing among different parties is managed via a data space that decouples data providers and consumers. FIWARE adopts a common language and data APIs and defines common data models to enable interoperability and efficient data exchange. Data sharing via the data space requires governance consisting of operational, organisational and business agreements clarifying the data protection regulation, terms and conditions, and governance bodies, respectively. The data space supports cross-domain data sharing to allow the creation of innovative services.

4. Related Projects

4.1 QU4LITY

The QU4LITY Horizon 2020 project aims to provide a data-sharing-enabled standardised product and service model to assist industries, primarily SMEs, in manufacturing. QU4LITY [30] produced guidelines for standardisation in several areas including security, interoperability, artificial intelligence, quality and reference architecture. It created a workflow for developing specifications for all these areas. The workflow includes analysing and classifying requirements from the pilots, mapping the list of requirements to reference architecture components, analysis and classification of standards, mapping the identified standards to the related reference architecture components, and recommendations for using standards in the components based on the mapping results. Some example components are distributed trustworthiness middleware, digital models and vocabularies, and control service.

Interoperability can be defined as the technical capacity of a system or application to exchange or exploit data. In QU4LITY, interoperability among different systems, applications, and processes was necessary to fulfil the requirements of different pilot-specific scenarios. While specifying interoperability standards in the context of this project, it was understood that complex systems and applications require several standards to enable interoperability. Occasionally, context-dependent unforeseen standards might be necessary to overcome the interoperability issue. The QU4LITY interoperability specifications were based on the technical and operational standards required for software components and interfaces. The project collected pilot-specific requirements and performed an overall analysis across pilot use cases. It discussed four degrees of interoperability – Exchange of Unstructured Data, Exchange of Structured Data, Seamless Sharing of Data and Seamless Sharing of Information – and proposed the refinement of these degrees to make them technically meaningful and implementable.

4.2 BOOST 4.0

BOOST [31] was a Horizon 2020 project aimed at developing standardised data exchange and sharing supported replicable Factory 4.0 model via several use cases to demonstrate large-scale experimentation with big data to enhance the competitiveness of European manufacturing industries. This project analysed the role of standardisation in manufacturing industries and reported the results of a review of standardisation initiatives in the context of sharing large amounts of data across the supply chain. It recognised the distinction between formal and informal standardisation bodies developed by international, European and national regulatory organisations. Some examples of formal standards are ISO, IEC and ITU, and some examples of informal but widely used standards are W3C, AML and IEEE. BOOST identified a set of ISO standards for smart manufacturing, such as ISO/TC 184/SC 4, specified for the representation, management and exchange of industrial data, ISO/TC 184/SC 5 for interoperability, integration, and architectures for enterprise systems and automation applications, and IEC/SC 3D for product classes and properties and their identification. Standards developed by W3C for web technologies include HTML, CSS, JavaScript APIs, RDF, OWL and SPARQL. The Automation ML initiated the specification of an open vendor-neutral data store and exchange format called Automation Markup Language (AML) for plant engineering information in production systems.

Standards identified in the BOOST pilots dealt with data management, data governance, testing, certification, security, privacy, trust and resilience. Defining semantic data models to create an interoperability-enabled data environment is crucial. The Standard for the Exchange of Product model data (STEP) and Standard Triangle/Tessellation Language (STL) are the most relevant standards for representing data models as they support effective data processing, management and visualisation. The Initial Graphics Exchange (IGES) and STEP are widely accepted standards for defining, managing and exchanging product data, e.g., 3D product design information, in the product definition process.

4.3 Z-BRE4K

The Z-BRE4K [32] Horizon 2020 project proposed to improve the unexpected breakdown by defining strategies for predictive maintenance of systems, machines and components to increase the working life of factories. Z-BRE4K identified, interpreted, analysed and assessed different international and national standards to determine the relevant ones for their technologies. The outcome was discussed with project partners to seek their expert opinion on the impact of applying specific standards. Selected standards include IEC 60812 for analysing system reliability, failure modes and effects, OASIS for production planning and scheduling, EN 15341 for maintenance key performance indicators, ISO/TC39 for product interoperability and integration, IEC 62264 for enterprise control system integration and operations management in manufacturing industries, and OPC and OPC-UA for communication protocols and standards.

In Z-BRE4K [32], a classification of relevant standards was developed under the following top-level categories: interoperability, security, communication, quality, scalability, and ontology. Interoperability enables the ability to share data across applications, and security defines data protection measures and prevents cybercrime and personal and operational data protection. Communication deals with protocols for the syntax, semantics and synchronization of messages transferred between systems and applications. Quality ensures the manufacturing of products and development of services that are fit for purpose, and scalability focuses on efficiently completing large volumes of tasks. Ontologies allow improved knowledge and data sharing through a shared ontological model of concepts and properties. They also eliminate ambiguity via a vocabulary of terms and the definition of contexts via their relationships with other terms. Ontology-related standards data representation are Resource Description Framework (RDF), Structured Query Language (SQL), not only SQL (NoSQL), Web Ontology Language (OWL), JavaScript Object Notation (JSON), Standard for the Exchange of Product model data (STEP) and Extensible Markup Language XML).

5. FAIR Principles

The FAIR principles [6] were designed to guide digital asset owners in publishing assets in a way that makes them Findable, Accessible, Interoperable, and Reusable. To meet the overarching need of data consumers for data reusability, stakeholders from industry and academia created the FAIR Data Principles that can guide data proprietors to publish automatically findable, processable and usable data. These principles can assess the FAIR compliance of published data as they are measurable.

To reuse or directly use data and metadata, humans and machines should be able to find them easily. Metadata published in a machine-readable format is critical to enable finding datasets and services automatically. Hence, for discoverability, data, metadata or both are assigned a universally unique identifier (UUID) that is also persistent; data are described using a rich set of metadata; metadata should always describe the identifier of the corresponding data; and data, metadata or both are stored in a searchable indexed storage or registry.

Following the discovery of intended data, users must have the information about how data can be accessed, possibly requiring access permission and right allocation via authentication and authorisation. Users should be able to access and retrieve data, metadata or both using their standardised protocol-based identifiers. This principle has been further extended to include the following two essential aspects of the standardised protocol: one emphasises the open, free and widely applicable protocol such as HTTP, and the other focuses on the protocol-level support for allowing implementation of authentication and authorisation functions if necessary. Once created, metadata associated with data should be accessible even if the data is no longer accessible due to deletion or other reasons.

Data found in one source often requires meaningful integration with another source for its improved use. Data must be interoperable with relevant applications, frameworks or workflows to allow storage, processing and analysis. Data, metadata, or both should be formalised using a logic-based knowledge representation language to enable interoperability. Publishing of vocabulary used in formalisation should be compliant with FAIR principles. Contextual information included in named references in data or metadata should be as precise as possible when linking to other data or metadata.

The overarching objective of publishing data using FAIR principles is to enhance its reuse. Achieving this requires a detailed description of data and metadata for replication or combined use with other data in different applications. Such characterisation includes as many relevant attributes as possible filled in with correct values in data and metadata. Data and metadata should accompany a license with clearly defined terms for use and reuse and detailed provenance information. In addition, data and metadata must meet the community-expected domain-specific standards.

A summary of the FAIR principles for publishing data and metadata is shown in Table 1. The table describes different aspects of FAIR, i.e., findability, accessibility, interoperability and reusability, and the specific principle within each category and its associated guideline as the recommendation.

***Table 1** A summary of the FAIR principles for publishing data and metadata in a way that ensures their findability, accessibility, interoperability and reusability [10].*

FAIR Aspect	FAIR Principle	Recommendation
Findability	F1	Data, metadata or both are allocated a universally unique identifier that is also persistent
	F2	Data are narrated with a rich set of metadata
	F3	Metadata must describe the identifier of the associated data
	F4	Data, metadata or both are uploaded to searchable indexed storage or registry
Accessibility	A1	Data, metadata or both are retrievable via their identifier created with a standardised internet protocol. This principle covers two crucial aspects of the protocol
	A1.1	The protocol is open, free, and universally implementable

	A1.2	The protocol allows for an authentication and authorisation procedure, where necessary
	A2	Metadata are accessible, even when the data are no longer available
Interoperable	I1	Data, metadata or both use a formal, accessible, shared, and broadly applicable language for knowledge representation
	I2	Data, metadata or both use vocabularies that follow FAIR principles
	I3	Data, metadata or both include qualified references to other data or metadata
Reusable	R1	Data, metadata or both are richly described with a plurality of accurate and relevant attributes
	R1.1	Data, metadata or both are released with a clear and accessible data usage license
	R1.2	Data, metadata or both are associated with detailed provenance
	R1.3	Data, metadata or both meet domain-relevant community standards

Several groups took the initiative to implement the FAIR principles in publishing data for all domains or specific domains. Harvard Dataverse is a FAIR data repository, storing over 60,000 datasets covering many research topics, demonstrating an exemplary implementation of FAIR principles [6]. It is based on the open-source software Dataverse [2] developed for creating data repositories. Harvard Dataverse is the largest across all implementations based on Dataverse. FAIRDOM also implements the FAIR principles relying on SEEK [11] and OpenBIS [1] platforms. It was developed for the Systems Biology domain to support FAIR data and model management [6].

The interoperability aspect of the FAIR principles depends on the adoption of community-specific standards and converging solution approaches in publishing data [4]. Since 2019 an alliance has been working to develop and maintain a matrix for measuring the convergence of communities of practice and FAIR-enabled digital resources published by owners [5]. Achieving interoperability by implementing FAIR is non-trivial and relies on converging solutions and standards used in research communities [4]. A coalition was formed in 2019 to create the Convergence Matrix consisting of communities of Practice and FAIR-supported Digital Resources [5].

5.1 FAIR Compliance

Since the introduction of the FAIR principles, assessing the compliance of published digital resources with these principles in a transparent approach has been the goal of many stakeholders. The FAIR Metrics group has been working to create FAIR compliance metrics to allow machines automatically complete the verification [9]. For this purpose, one standard metric has been created for each principle to determine if a published digital resource is findable, accessible, interoperable and reusable. An assessment software has been designed and developed to determine if a digital resource qualifies as FAIR compliant based on the selected subset of the standard metric meeting the FAIR requirements for a community [7]. FAIR maturity indicators have been developed to overcome several issues, such as less awareness of best practices in data publishing, non-standard, simple responses for validating FAIR compliance, etc. [6].

5.2 FAIR Workflow

Figure 1 shows a fully automated workflow developed to verify FAIR principle compliance of use case partner-contributed distributed data published in data-sharing units. Two layers for this workflow development have been defined: the abstraction and implementation layers. The abstraction layer includes the following top-level components:

1. Findability components: metadata-based search for finding data
2. Accessibility components: data access for retrieval or download for further processing
3. Interoperability components: identification of data representation in interoperable ontologies
4. Reusability components: detection of data reusability features

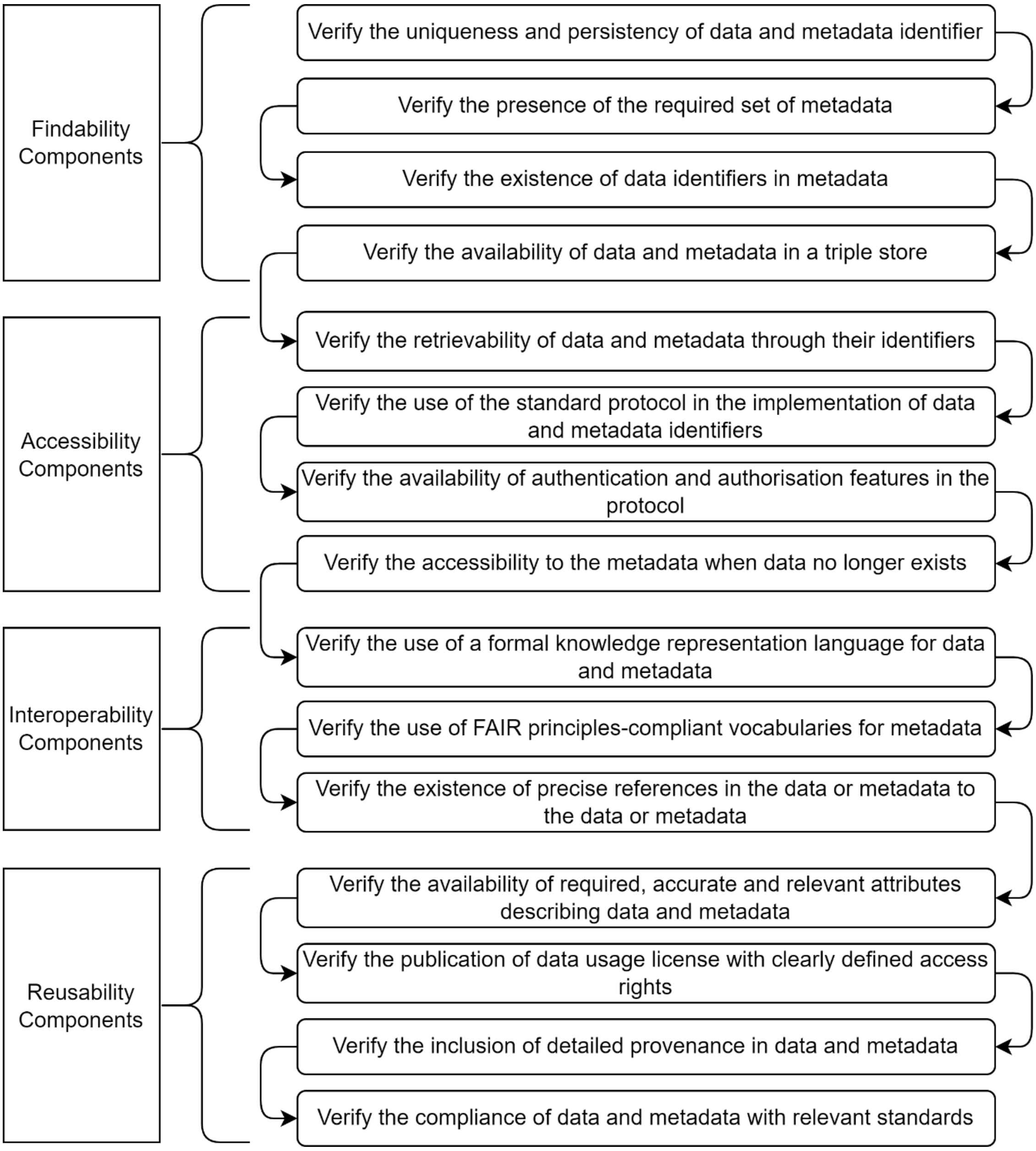

Figure 1. *The workflow for verifying FAIR compliance of data and metadata contributed by use case partners in distributed storage units.*

As shown in Figure 1, implementing findability features requires verifying data and metadata identifiers that are universally unique and persistent. Identifiers that comply with valid Universally Unique Identifier (UUID) fulfil this criterion and remain unchanged even if the data or metadata they refer to is changed or updated. The presence of metadata required for describing product, component and material data for the JIDEP use cases is checked in the data accompanying metadata representation. The existence of identifiers, referring to data, in the metadata is tested, and the availability of data and metadata in a triple store is verified. Similarly, the accessibility, interoperability and reusability components are executed to verify and conclude the compliance of distributed data contributions with FAIR principles.

6. JIDEP Specifications of Standards and Principles

This section specifies data representation requirements of JIDEP and mapping with the most relevant community-specific standard that describes the modelling of such data as an entity, schema, or attribute and that will be used to improve and enable data interoperability and shareability in this project. It also specifies data publishing requirements of JIDEP and mapping with the state-of-the-art principle to ensure findability, accessibility, interoperability and retrievability of published data and metadata.

Interoperability Standards for Product, Component and Material Data

JIDEP data representation requirement	Standard meets the requirement	Description
Product	ISO 10303-41	A product is defined as an entity in ISO 10303-41.
Product.Identifier	ISO 10303-41	The identifier of a product. Product.id is defined as an attribute in ISO 10303-41.
Product.Name	ISO 10303-41	The name of a product. Product.name is defined as an attribute in ISO 10303-41.
Product.Description	ISO 10303-41	The description of a product. Product.description is defined as an attribute in ISO 10303-41.
Product.Trade_name	-	The trade name of a product
Product.Brand_name	-	The brand name of a product
Product.GTIN	-	Global Trade Identifier Number
Product.EAN	-	European Article Number
Product.Functionality	-	The functionality of a product
Product.Automatic_tracking_or_scanning	-	If the automatic tracking or scanning is supported by the product
Product.Level	-	The position of a product or component within a product-component part of hierarchy
Product.Part_of	-	Refers to the parent product
Product.Image	-	The image of a product

JIDEP data representation requirement	Standard meets the requirement	Description
Image.URL	ISO 10303-41	The URL of an image. URL is described as a derived attribute of the address entity in ISO 10303-41.
Product.Manufacturer	-	The manufacturer of a product
Manufacturer.Name	ISO 10303-41	The name of a manufacturer. We will apply the name attribute described in the organisation entity in ISO 10303-41.
Manufacturer.Registration_number	-	The registration number of a manufacturer
Manufacturer.Registration_country	ISO 10303-41	The country where the manufacturer is registered. Country is defined as an attribute in the address entity in ISO 10303-41.
Product.Physical_property	ISO 10303-41	The physical property of a product. The measure schema describes the physical quantities in ISO 10303-41.
Physical_property.Density	ISO 10303-41	The density of a product. Density is described as a kind of physical quantity in ISO 10303-41.
Physical_property.Dimension	-	The dimension of a product
Dimension.Height	ISO 10303-42 [36]	The height of a product. Height is an attribute of the data type positive_length_measure in ISO 10303-42.
Dimension.Width	-	The width of a product
Dimension.Length	ISO 10303-41	The length of a product. Length is described as a kind of physical quantity in ISO 10303-41.
Physical_property.Resistance	-	The resistance of a product
Resistance.Compressive_strength	-	The compressive strength of a product
Resistance.Shear_strength	-	The shear strength of a product
Resistance.Tensile_strength	-	The tensile length of a product
Physical_property.Rigidity	-	The rigidity of a product
Rigidity.Shear_modulus	-	The shear modulus of a product
Rigidity.Young’s_modulus	-	The Young’s modulus of a product
Physical_property.Mass	ISO 10303-41	The mass of a product. The mass measure is described in ISO

JIDEP data representation requirement	Standard meets the requirement	Description
		10303-41 as the quantity of matter or substance that a body contains.
Component	ISO 10303-41	Product, class of products and component are represented using the product entity in ISO 10303-41.
Material	ISO 10303-1771	Material is defined in ISO 10303-1771 as one or more substances used to manufacture a product.

Data Publishing Principles

JIDEP data publishing requirement	Principle meets the requirement
Universally unique and persistent data and metadata identifier	FAIR [10] Findability Principle – F1
Required set of metadata associated with data	FAIR Findability Principle – F2
Availability of data identifiers in metadata	FAIR Findability Principle – F3
Data or metadata uploaded to a triple store	FAIR Findability Principle – F4
Retrievability of data and metadata using their identifiers	FAIR Accessibility Principle – A1
Data and metadata identifiers created with the standard protocol	FAIR Accessibility Principle – A1.1
The protocol supports authentication and authorisation	FAIR Accessibility Principle – A1.2
Metadata can be accessed even if the data are no longer accessible	FAIR Accessibility Principle – A2
Data and metadata are represented using the formal knowledge representation language RDF or OWL	FAIR Interoperability Principle – I1
Vocabularies used for representing metadata should be findable, accessible and reusable	FAIR Interoperability Principle – I2
While linking data or metadata more specific references should be used in them	FAIR Interoperability Principle – I2
Required, accurate and relevant of attributes should describe data or metadata	FAIR Reusability Principle – R1
Data usage license should be published with clearly defined access rights	FAIR Reusability Principle – R1.1
Identifiable provenance should be included in data and metadata	FAIR Reusability Principle – R1.2
Data or metadata should meet domain-relevant community-specific standards	FAIR Reusability Principle – R1.3

7. Conclusions

We have reviewed and analysed: i) many industrial data modelling and exchange standards, including ISO 10303 (STEP), OPC-UA, MTConnect, RAMI 4.0, MIMOSA, IOF and FIWARE; ii) outputs of several related projects such as QU4LITY, BOOST 4.0 and Z-BRE4K; and iii) FAIR principles. We have mapped the JIDEP data representation and publishing requirements with the most relevant standards and state-of-the-art data publishing principles to specify the standards and principles for data-providing partners and other data contributors. We have designed and developed an automated workflow to verify the compliance of product, component and material data contributed in a distributed environment by the use case partners and interested parties with FAIR principles. The output reported in this deliverable will enable the implementation and use of standards and principles required for this project.

References

[1] Bauch, A., Adamczyk, I., Buczek, P., Elmer, F.-J., Enimanev, K., Glyzewski, P., Kohler, M., Pylak, T., Quandt, A., Ramakrishnan, C., Beisel, C., Malmström, L., Aebersold, R., & Rinn, B., 2011. OpenBIS: A flexible framework for managing and analyzing complex data in Biology Research - BMC Bioinformatics. BioMed Central.

[2] Crosas, M., 2011. The Dataverse network: An open-source application for sharing, discovering and Preserving Data. D-Lib Magazine.

[3] Karray, M., Otte, N., Rai, R., Ameri, F., Kulvatunyou, B., Smith, B., Kiritsis, D., Will, C. and Arista, R., 2021. The industrial ontologies foundry (IOF) perspectives, in Proceedings of the Industrial Ontology Foundry (IOF)-Achieving Data Interoperability Workshop.

[4] Jacobsen, A., de Miranda Azevedo, R., Juty, N., Batista, D., Coles, S., Cornet, R., Courtot, M., Crosas, M., Dumontier, M., Evelo, C. T., Goble, C., Guizzardi, G., Hansen, K. K., Hasnain, A., Hettne, K., Heringa, J., Hooft, R. W. W., Imming, M., Jeffery, K. G., … Schultes, E., 2020. Fair principles: Interpretations and implementation considerations. MIT Press.

[5] Sustkova, H. P., Hettne, K. M., Wittenburg, P., Jacobsen, A., Kuhn, T., Pergl, R., Slifka, J., McQuilton, P., Magagna, B., Sansone, S.-A., Stocker, M., Imming, M., Lannom, L., Musen, M., & Schultes, E., 2020. Fair convergence matrix: Optimizing the reuse of existing fair-related resources. MIT Press.

[6] Wilkinson, M.D., Dumontier, M., Sansone, S.A., Bonino da Silva Santos, L.O., Prieto, M., Batista, D., McQuilton, P., Kuhn, T., Rocca-Serra, P., Crosas, M. and Schultes, E., 2019. Evaluating FAIR maturity through a scalable, automated, community-governed framework. Scientific data, 6(1), p.174.

[7] Wilkinson, M.D., Dumontier, M., Sansone, S.A., da Silva Santos, L.O.B., Prieto, M., McQuilton, P., Gautier, J., Murphy, D., Crosas, M. and Schultes, E., 2018. Evaluating FAIR-compliance through an objective, automated, community-governed framework. bioRxiv, p.418376. doi: 10.1101/418376

[8] Pisching, M.A., Pessoa, M.A., Junqueira, F., dos Santos Filho, D.J. and Miyagi, P.E., 2018. An architecture based on RAMI 4.0 to discover equipment to process operations required by products. *Computers & Industrial Engineering*, *125*, pp.574-591.

[9] Wilkinson, M.D., Sansone, S.A., Schultes, E., Doorn, P., Bonino da Silva Santos, L.O. and Dumontier, M., 2018. A design framework and exemplar metrics for FAIRness. *Scientific data*, 5(1), pp.1-4.

[10] GO FAIR - *FAIR Principles*. URL: https://www.go-fair.org/fair-principles/. Accessed on 23 May 2023.

[11] Wolstencroft, K., Owen, S., Krebs, O., Nguyen, Q., Stanford, N.J., Golebiewski, M., Weidemann, A., Bittkowski, M., An, L., Shockley, D. and Snoep, J.L., 2015. SEEK: a systems biology data and model management platform. BMC systems biology, 9(1), pp.1-12.

[12] ISO 10303-1, Industrial automation systems and integration — Product data representation and exchange — Part 1: Overview and fundamental principles.

[13] ISO 10303-11, Industrial automation systems and integration — product data representation and exchange — Part 11: description methods: the EXPRESS language reference manual.

[14] ISO 10303-41, Industrial automation systems and integration — product data representation and exchange — Part 41: Integrated generic resource: Fundamentals of product description and support.
[15] ISO 10303-45, Industrial automation systems and integration — product data representation and exchange — Part 45: Integrated generic resource: Material and other engineering properties.
[16] ISO 10303-1771, Industrial automation systems and integration — product data representation and exchange — Part 1771: Application module: Stock material.
[17] GHERASIM, S.A., GĂITAN, A.M., GĂITAN, V.G. and POPA, V., 2010. New trend on OPC Middleware. Development and Application Systems, p.101.
[18] Valenzuela, E., Cano-Delgado, A., Cruz-Miranda, J., Rouret, M., Micciché, G., Ros, E., Arranz, F. and Diaz, J., 2023. The IFMIF-DONES remote handling control system: Experimental setup for OPC UA integration. *Fusion Engineering and Design*, *192*, p.113776.
[19] OPC 10000-1: UA Part 1: Overview and Concepts. URL: https://reference.opcfoundation.org/v104/Core/docs/Part1/. Accessed on 12 May 2023.
[20] MTConnect Standard Part 1.0 – Overview and Fundamentals Version 1.8.0.
[21] RAMI 4.0 (Reference Architecture Model Industrie 4.0). URL: https://www.sci40.com/english/thematic-fields/rami4-0/. Accessed on 14 May 2023.
[22] Hankel, M. and Rexroth, B., 2015. The reference architectural model industrie 4.0 (rami 4.0). *Zvei*, *2*(2), pp.4-9.
[23] MIMOSA. URL: https://www.mimosa.org/. Accessed on 15 May 2023.
[24] Koronios, A., Nastasie, D., Chanana, V. and Haider, A., 2007, June. Integration through standards–an overview of international standards for engineering asset management. In Fourth International Conference on Condition Monitoring, Harrogate, United Kingdom.
[25] The Industrial Ontologies Foundry.URL: https://www.industrialontologies.org/. Accessed on 16 May 2023.
[26] Technical Principles of the Industrial Ontologies Foundry. URL: https://industrialontologies.org/technical-principles. Accessed on 16 May 2023.
[27] Dublin Core – DCMI Metadata Terms. URL: https://www.dublincore.org/specifications/dublin-core/dcmi-terms/. Accessed on 17 May 2023.
[28] FIWARE for Data Spaces. URL: https://www.fiware.org/wp-content/uploads/FF_PositionPaper_FIWARE4DataSpaces.pdf. Accessed on 17 May 2023.
[29] Ahle, U. and Hierro, J.J., 2022. FIWARE for data spaces. *Designing Data Spaces*, p.395.
[30] Meyer, O., Knaak, C. and Ewald, T., 2020. Standards Compliance and Interoperability Specification (Final Version), QU4LITY Horizon 2020 Project, Deliverable 2.8.
[31] Raggett, D. and Ubis, F, 2019. Boost 4.0 standardization & certification v1, Boost 4.0 Horizon 2020 Project, Deliverable 2.7.
[32] Faria, H., Durate, M., Leal, A. S., Moreira, T., Neves, E., Milenkovi, J., Avendano, D. and Gesto, D., 2021. Report on standards used and triggered in Z-Bre4k, Z-BRE4K Horizon 2020 Project, Deliverable 7.1.
[33] Peng, T.K. and Trappey, A.J., 1998. A step toward STEP-compatible engineering data management: the data models of product structure and engineering changes. *Robotics and Computer-Integrated Manufacturing*, *14*(2), pp.89-109.
[34] OPC Foundation OPC unified architecture, 2006. URL: https://opcfoundation.org/about/opc-technologies/opc-ua/. Accessed on 23 May 2023.
[35] MTConnect institution MTConnect, 2008. URL: http://www.mtconnect.org/. Accessed on 23 May 2023.
[36] ISO 10303-42, Industrial automation systems and integration — product data representation and exchange — Part 42: Integrated generic resource: Geometric and topological representation.