Identification of Information Entities

An Industry IoT Consortium White Paper

2022-08-23

Authors

The Industry IoT Consortium Technology Working Group.

CONTENTS

FIGURES

TABLES

Unabated growth in the numbers and types of connected devices and applications in all industrial and business sectors together with technology advances, especially ease of connection at the network level, are driving and are driven by the exploitation of new data sources for new business applications. More and more technology providers and users are contributing to the process and are benefiting from this rapid growth.

Federated data infrastructures (see *Section 8*) are emerging to ensure scalability and sustainable operations. A federated architecture is a realization pattern that allows interoperability and information sharing between semi-autonomous, de-centrally organized IT systems and business applications.

This raises the fundamental question about interoperability of devices and applications. Standardized (instead of proprietary) interfaces are needed to bound the cost of APIs. An essential requirement for interoperability and standards is to have an identifier scheme for identifying all system objects (digital and physical). Identifiers also play a key role in interoperability between federated systems managed by different companies. In a nutshell, all objects need to know who they are and with whom they are communicating so as to share data in a trustworthy way.

This report is structured as follows:

- Section 1 (Introduction) presents briefly what identifiers are and what they are for.
- Section 2 (Properties of Identifiers) describes the key characteristics of identifiers such as their syntax (how they are represented), governance and assignment.
- Section 3 (Basic Identification Standards) mentions several foundational, low-level international standards for object identifiers that are now in common use throughout web-based applications. These standards are used by the higher-level identifier standards presented in the subsequent sections.
- Sections 4 ~ 9 give an overview of identifier solutions as adopted in several major international initiatives from around the globe: OpenID Foundation's OpenID Connect, World Wide Web Consortium's Decentralized Identifiers (DID), Plattform Industrie 4.0, GAIA-X and IDS, Industrial Value Chain Initiative and Alliance of Industrial Internet (AII) in China. These exemplary solutions have been selected based on their openness, commitment to standards and wide-spread recognition and usage in major companies. Such initiatives are expected to be the melting pot for new foundational standards in the future.
- Overall observations and conclusions are discussed in Section 10, Observations and Conclusions.

1 INTRODUCTION

1.1 WHAT ARE IDENTIFIERS?

Identifiers are how humans label things to make sense of the world and exchange information. They act as labels for physical and digital objects and point to meaningful information. Digital identifiers work as pointers to information available online or stored in a variety of systems.

Digital identifiers are part of local, regional and global digital infrastructure. Overall, ID infrastructure consists of data assets, such as labels, technologies that help to manage and use them, policies that govern how they are used and the organizations that curate and maintain them. Identifiers are a glue for integrating data within organizations, between business partners and across industries. Use of common, standard identifiers helps to cut API development costs and increase efficiency. Industry 4.0 and digital supply chains rely on identifiers.

1.2 WHAT IS THE ID FOR?

It is a unique reference or 'name' of an object, in particular for use by machines and programs. In everyday life, a person will have a name and an address. But a person's address (postal or email) is rarely constant over the entire lifetime and the name is rarely unique. Hence for governmental matters, agencies often use additional metadata or identifiers to identify the person uniquely (e.g. social security number or tax identification number). Similarly, there are uniform identifiers for bank accounts that are linked to the person's identifier.

Persistent identifiers, such as tax identification numbers, are characterized by life-long existence and are usually maintained by special governance bodies. Product bar codes may have temporary identifiers.

In general, an industrial internet identification system must provide registration, resolution and discovery services for physical or virtual objects. The identifiers are used to:

- retrieve correct information objects with certainty,
- retrieve further information about the object, find linked information objects (e.g. documents and data to install, service and maintain a device) and
- define access and usage rights of entities to entities.

2 PROPERTIES OF IDENTIFIERS

Creating and provision of identifiers as described in [ODI] requires design decisions, policies and governance that describe how a particular set of identifiers are assigned and used.

2.1 SHARED VOCABULARY

A scheme will describe things like the syntax of the identifiers (e.g. whether it is a human-readable label or bar code), how and when they are assigned, and how they are licensed. An

identifier scheme is an agreed shared vocabulary of identifiers used to describe people, places, things or concepts. It is also a holistic term referring to the data standards (authoritative lists like registers), organizations (such as registration agencies) and processes (like assigning identifiers) required to create, implement and maintain identifiers.

2.2 SYNTAX

Identifiers are the labels or tags assigned to specific people or things through an identifier scheme. The syntax of an identifier can be made up of words, letters, numbers, symbols, or a combination of these. For example: DOI[1] identifiers such as 10.3352/12345, have 10.3352 as prefix, separated by "/" from the suffix. An email address such as "info@example.com" is an email identifier, with name separated from domain name by the "@" sign. Delimiters are often used to lookup for a particular resolver for an identifier.

Unlike DOI or email addresses, "flat" identifiers are just symbols. For example: "00:00:5e:00:53:af" as a MAC address with organizationally unique identifier assigned to a manufacturer (a 24-bit number), and "123e4567-e89b-12d3-a456-426614174000" as a Universally Unique Identifiers (UUID) without any assignments.

The syntax of an identifier depends on how many unique identifiers a scheme needs, whether the scheme needs to be human-readable or machine-readable (or both) and how the scheme is governed and how identifiers are used. For instance, UUIDs are 128-bit machine readable identifiers that are effectively guaranteed to be unique and are easy to create because the only coordination needed is an algorithm providing reasonable global uniqueness at the local domain.

2.3 SEMANTICS

Identifier semantics are generally human-readable, structured or unstructured. A domain name identifier (such as example.com) is human-readable. The purpose of domain name introduction was to give users simple names to remember instead of addressing the resource by hard-to-remember IP addresses. Structured identifiers carry a sequence of characters from which the type or place of the addressed resource can be generally understood. For example, 1 (USA), 7 (Russia) are used in the telephone numbers to indicate "country codes".

Unstructured identifiers have no semantics, for example, UUID identifiers. Semantics can be mixed, for example 77.IOT.1/7280157 is a DOA[2] identifier that has properties in prefix: 77 - Russia, IOT - some sub prefix, 1 - some unknown client, and 7280157 some unknown digital object

[1] Digital Object Identifier (DOI), ISO 26324. The DOI system provides a technical and social infrastructure for the registration and use of persistent interoperable identifiers, called DOIs, for use on digital networks.

[2] *https://www.dona.net/digitalobjectarchitecture*. The Digital Object (DO) Architecture (also known as the DO Architecture or simply the DOA) is a logical extension of the Internet architecture that addresses the need to support information management.

or device. Only after identifier resolution, can a human or system get the property of an information entity.

2.4 GOVERNANCE OF IDENTIFIERS

Any group of people or organizations can work together as a community to set up and maintain an identifier scheme. The community needs a shared understanding of the value of common identifiers (often including a register of the identifiers) and sufficient agreement, funding and resources to manage it. Top-down governance is needed in any state-controlled identification method, such as a tax identification number or a social security number. Bottom-up governance runs on community agreed rules and methods.

For example, the International Corporation for Assigned Names and Numbers (ICANN) declares a multi-stakeholder model for domain name governance. A community can also be in closed or proprietary forms, such as Data Universal Numbering System (DUNS) is a proprietary system developed and managed by Dun & Bradstreet (D&B), or in open form such as the perma-id (PID) group is. PID existed before the internet for citation of paper-based documents; today's examples include ISNI, ISBN, etc.

In many data ecosystems, the shared vocabulary of an identifier scheme has become so important that dedicated organizations have emerged to govern them, often with a focus on openness and independence. For example, the Legal Entity Identifier (LEI) scheme was set up as a result of the global financial crisis originally, but now identifies over 1.7 million entities in over 200 countries and is overseen by the Global LEI Foundation, *GLEIF*. Alternatively, an existing organization can take on the governance of essential identifiers. For instance, the Institute of Electrical and Electronics Engineers (IEEE) oversees identifiers for the electronics industry, with a dedicated registration authority for this purpose.

2.5 ID ASSIGNMENT

Most identifier schemes have 'registration authorities' that manage how identifiers are assigned under governance rules. Identifier assignment schemes can be centralized, decentralized or federated. While the standards and syntax governing a scheme may need to be developed through a central coordination, the assignment of identifiers using those standards can be decentralized. Centralizing the assignment of identifiers can create bottlenecks if you need to assign lots of identifiers quickly. It also creates a dependency on a single central entity, whereas decentralized assignment shares the work among organizations.

Assignment of identifiers that are part of the public sector, commercial or community initiatives tend to be more centralized, whereas IEC/ISO-governed schemes tend to be federated (e.g. the digital object identifier for books or articles), often with the international standard being implemented by national agencies with local knowledge and relationships (e.g. LEI, unique global ID for legal entities participating in legal financial transactions). Assignment can be even more

decentralized, for instance to the local or individual level e.g. the Unique Property Reference Number (UPRNs) in UK created by Ordnance Survey and assigned by local governments or automated (e.g. UUIDs, which are arguably centralized by local algorithms). Global uniqueness requires a more complex administrative and technical infrastructure. There may be significant costs involved. Uniqueness within a domain is in general easier to implement, but it will require a way of registering related domain identifiers if interoperability across domains is needed.

2.6 ID RESOLUTION

Resolution is the process of translating an ID such as a URI into a means to locate the entity and retrieve information from it (e.g. through a URL) ID; resolution is an integral part of any identification system. Resolution mechanism follows the architecture of the identification system. It is performed locally, within a given domain, or globally. The information entity addressed by a URI can be of a local or global IT system.

2.7 PHYSICAL OBJECTS WITH IDENTIFIERS

Physical tags with identifiers are used to correlate objects with their properties and belonging. Barcodes, QR code, matrix 2- and 3-d codes, RFID and NFC are some of the well-known tags. These identifiers can be resolved into correlated IT systems by optical scanning (barcodes, QR code, matrix) or radio scanning (RFID and NFC). In the physical world, bar codes dominate the use of the tags due to the common use in logistics and retails. GS1[3] as a registry of barcode identifiers that has been operated since the 70s and now verifies 150 million products.

All devices with a network interface have a Media Access Control (MAC) address. The globally unique MAC address is issued by the IEEE Standards Association Registration Authority[4] to the vendor of the network interface. It is realized in hardware and remains constant for that device. As shown in Table 2-1, these properties enable registration, resolution and discovery of identifiers.

[3] *https://www.gs1.org/*

[4] *https://standards.ieee.org/products-services/regauth/index.html*

Syntax	Letters and numbers (ASCII, Unicode). Structured and unstructured (prefix, symbols separation, character count or flat)
Semantics	Humans can understand or apprehend; or only machine readable
Governance	Managed within organization, association, government or global governance
Assignment	Assigned within a domain or global assignment
Resolution	Local resolution into local IT system or global resolution into local or global IT system

Table 2-1. Properties of Identifiers.

3 BASIC IDENTIFICATION STANDARDS

Several standards for object identifiers that are in common usage.

IETF RFC 3986[5] defines the basic generic resource identifier, Unique Resource Identifier (URI), in the world wide web. A Uniform Resource Locator (URL) is a URI that describes the location of the resource (e.g. a network address). A Uniform Resource Name (URN) is a URI according to the RFC 8141[6] scheme. The URN comprises a Namespace Identifier (NID) and a Namespace Specific String (NSS). The NID denotes the organization responsible for the NSS. The URI syntax is a sequence of US-ASCII characters, thus posing limitations for other alphabets with different characters.

For example, *https://www.iiconsortium.org/pdf/IIRA-v1.9.pdf* is the URL linked to the IIRA (Industrial Internet Reference Architecture) framework document of IIC. It follows the pattern structure scheme, domain name and path. A URL need not be persistent or always refer to the same information resource, leading to the well-known problem of 'broken links.'

IETF RFC 3987[7] defines the Internationalized Resource Identifier (IRI) as a complement to URI. The allowed character set is UCS [ISO10646], covering characters from multiple languages. There is a defined mapping from an IRI to a URI based on a mapping of a character sequence to an octet sequence. This mapping is needed when the IRI is used for resource retrieval.

An International Registration Data Identifier (IRDI) is based on the international standards ISO/IEC 11179-6 (Metadata registries - Part 6: Registration), ISO 6532 and ISO 29002. For example, the IRDI 0173-1#02-AAT096#001 comprises the following elements:

- 0173 is the International Code Designator (ICD) followed by 1 for the Organization Identifier (OI), here of ECLASS. The ICD and OI schemes are specified in ISO 6523,
- 02 is the Code Space Identifier, here 02 stands for a property,
- AAT096 is the 6-digit identifier of the ECLASS Concept Code (here it stands for the property "Number per minute" defined by ECLASS). ISO 29002 defines the syntax and

[5] *https://tools.ietf.org/html/rfc3986*

[6] *https://www.rfc-editor.org/rfc/rfc8141*

[7] *https://tools.ietf.org/html/rfc3987*

requirements for concept identifiers and
- 001 is the version identifier.

The concept identifiers are managed by the organization designated in the prefix. Concept identifiers are important for semantic interoperability.

The Digital Object Identifier (DOI) specified in ISO 26324 is a globally unique, persistent identifier for a document. The DOI system is implemented by a federation of registration agencies following the regulations of the International DOI Foundation. A DOI may be resolved in a web browser as a URL, thus providing access to the document on the internet.

ITU standard X.509 defines public key certificates. An X.509 certificate uses a digital signature to associate an identity to a public key. A certificate contains an identity (a hostname, an organization or an individual) and a public key and is either signed by a certificate authority or is self-signed. Essentially, the certificate authority (CA) provides confidence in the validity of the identity. X.509 is now maintained by multiple bodies and exists within the CA chain of trust.

IETF RFC 4122[8] defines the universally unique identifier (UUID), also known as GUID (Globally Unique Identifier) as a 128-bit label for information objects in computer systems. A UUID requires no central registration process and is, to a high probability, unique. The roots of UUIDs were originally from the Apollo Network Computing System and later in the Open Software Foundation's (OSF's) Distributed Computing Environment (DCE) and then in Microsoft Windows platforms.

There is an ITU-T Recommendation[9] and ISO/IEC Standard[10]. The structure of the UUID is defined in RFC 4122. However, recommendations, methods and use of UUIDs are also specified for various platforms (Microsoft, Apple, etc.). There are several tools available on GitHub for UUID generators and resolvers.

The basic identification standards referred above are applied as a foundation in many industrial IoT systems requiring a higher level of semantics for interoperability. In the following sections we give a short summary of several solutions for identification of system entities:

- OpenID Connect,
- Decentralized Identifiers (DID),
- Plattform Industrie 4.0,
- GAIA-X and IDS,
- Industrial Value Chain Initiative (IVI) and
- Alliance of Industrial Internet (AII).

[8] *https://www.rfc-editor.org/info/rfc4122*
[9] *http://www.itu.int/ITU-T/asn1/uuid.html*
[10] *https://www.iso.org/standard/53416.html*

These solutions have achieved a high degree of acceptance in multi-vendor ecosystems with open, maintained specifications and concept definitions for independent implementations.

4 OPENID CONNECT

OpenID Connect and DID, elaborated in the next section, are the recent common technologies and standards for identification of information entities. OpenID Connect is an identity layer on top of the OAuth 2.0[11] protocol. It allows Clients to verify the identity of the end-user based on the authentication performed by an authorization server, as well as to obtain basic profile information about the end-user in an interoperable and REST-like manner. OpenID Connect allows clients of all types, including web-based, mobile and JavaScript clients, to request and receive information about authenticated sessions and end-users.

The specification suite is extensible, allowing participants to use optional features such as encryption of identity data, discovery of OpenID Providers, and session management. OpenID is managed by OpenID Foundation[12] and exercises open source and open governance principles. OpenID with OAuth is widely used not only in "human" authentication/authorization, but for machine communications as well to create a trusted environment.

The following key entities are specified in OpenID Connect and a high-level workflow is shown in Figure 4-1 below.

- the OpenID Provider that holds the user information and grants access,
- the end user whose user information the application needs to access and
- the relying party needs to access end users' identity information.

[11] *https://datatracker.ietf.org/doc/html/rfc6749*

[12] *https://openid.net/foundation/*

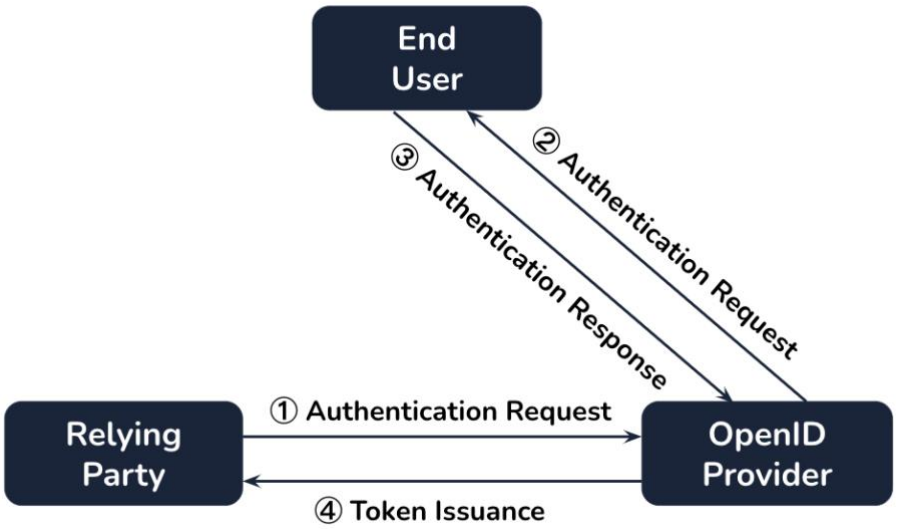

Figure 4-1. A High-Level Workflow of OpenID Connect.

In *Step 1*, an end user attempts to start a session with a client application (i.e., a relying party) and is redirected to the OpenID provider for authentication. In *Step 2*, the OpenID provider prompts for user authentication and shows the end user which user attributes the relying party is requesting. In *Step 3*, the end user authenticates to the OpenID provider and grants or rejects access to the requested user attributes. In *Step 4*, the end user is redirected back to the relying party by the OpenID provider. The requested attributes or a temporary token may return to the relying party, depending on the specific OpenID connect authentication flow.

5 DECENTRALIZED IDENTIFIERS (DID)

Decentralized Identifiers (DIDs) version 1.0[13] is the proposed recommendation of W3C. The core difference from centralized systems is that DIDs can be decoupled from centralized registries, identity providers and certificate authorities. DIDs are a new type of globally unique identifier (URI) that enables verifiable, self-sovereign digital identity. As shown in Figure 5-1, a DID, which is created via a DID method and recorded on a verifiable data registry (e.g. a distributed ledger, a decentralized network, etc.), identifies a DID subject (e.g. a person, organization, thing, data model and abstract entity).

A DID method specifies how a DID is created, read (resolved), updated, and deleted (revoked). As a resolvable identifier, a DID can be resolved to the associated DID document with the aid of a DID resolver. In practice, a DID document, which is controlled by a DID controller, usually contains only the minimal amount of machine-readable metadata (e.g. cryptographic material, verification methods, service endpoints) required to enable trustable interaction with the

[13] *https://www.w3.org/TR/did-core/*

associated DID subject. A DID might provide the means to return the DID subject itself, if the DID subject is an information resource such as a data model.

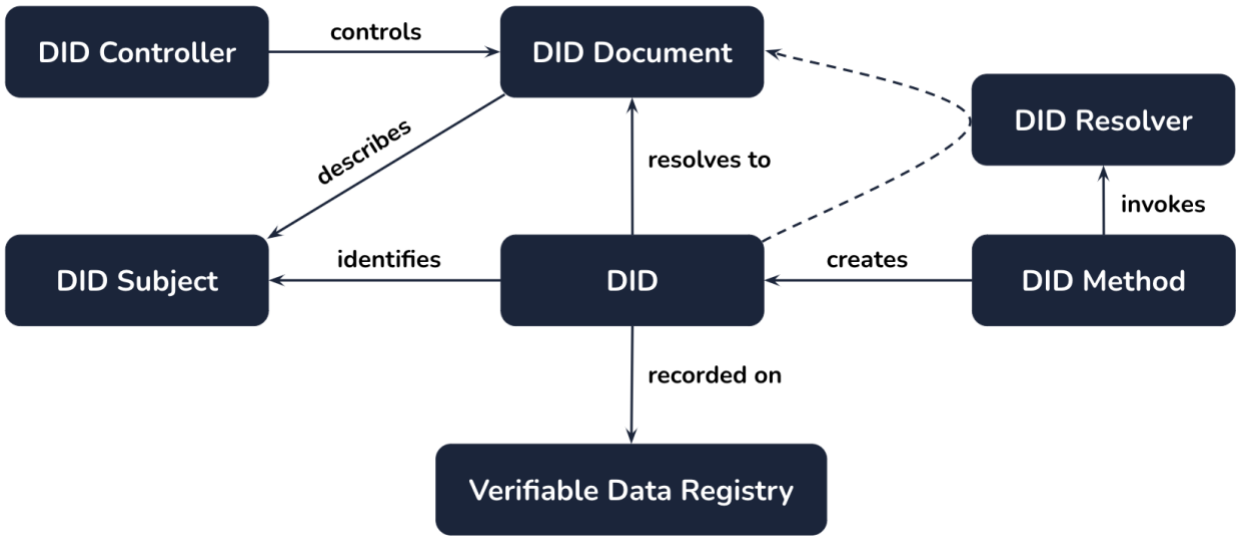

Figure 5-1. An Overview of Decentralized Identifiers (DIDs).

Compared to other identifiers, DIDs have the following four core properties:

- decentralized: DIDs are designed to function without a central registration authority,
- cryptographically verifiable: DIDs are designed to be associated with cryptographic keys that allow the entities controlling DIDs to use those keys for proving ownership,
- permanent: DIDs are persistent and non-re-assignable identifiers and
- resolvable: DIDs can be looked up to discover metadata.

Note that DIDs represent only the base layer of the decentralized identity infrastructure and do not provide meaningful identity attributes. The trust relationship between DID-identified entities can be established through the exchange of verifiable credentials (VCs), which are digitally signed electronic credentials about identity attributes as specified in the Verifiable Credentials Data Model v1.1[14] by W3C.

6 PLATTFORM INDUSTRIE 4.0

Plattform Industrie 4.0 has specified identifiers for all elements in its Asset Administration Shell (AAS). The identification of elements in the AAS Unified Modeling Language (UML) model is

[14] *https://www.w3.org/TR/vc-data-model/*

described in section 4.4 of [AAS-2020]. Table 2 of [AAS-2020][15] defines the attributes and allowed identifiers for each element type.

The attributes considered are *id*, *idShort* and *semanticId*. Depending on the element type, the *id* may be expressed as an IRDI or IRI or as an internal custom identifier such as a manufacturer's GUID. The *idShort* is a string, typically a short name in English. The *semanticid* is a reference to an external information source with further definitions or explanations of the meaning of the element. The following elements have a mandatory ID:

- Assets: physical or logical objects owned by or under the custodial duties of an organization, having either a perceived or actual value to the organization. [RAMI4.0-2016] gives a detailed background explanation of the central concept 'asset' and asset types as defined in the Reference Architecture Model of Plattform Industrie 4.0.
- AAS: The Asset Administration Shell is a digital representation of an asset.
- Sub-model instances and templates. An AAS may have one or more sub-models. A sub-model is a digital model of asset aspects describing technical functionality of the asset. Templates are standardized sub-models.
- ConceptDescription.

The ID of the ConceptDescription may reference through an IRDI an external dictionary such as ECLASS or IEC 61360 CDD (Common Data Dictionary). IRDIs are set by the external repository. URIs and URLs are chosen and managed by the AAS developer. [AAS-2020] describes a best practice for creating URI identifiers.

The white paper [ECLASS-2020] considers how to manage identifiers in a decentralized AAS registry in a distributed network. The Distributed Ledger Technology (DLT) and DIDs (W3C Decentralized Identifier) form the basis of the solution concept. This approach goes beyond the essentially central approaches adopted so far in Plattform Industrie 4.0. It is guided by five use cases involving identity management:

- use case 1: decentralized service registry,
- use case 2: multiple AAS referenced by one identifier,
- use case 3: transfer of ownership,
- use case 4: decentralized identity and access management and
- use case 5: qualification of an asset administration shell.

[15] *https://www.plattform-i40.de/PI40/Redaktion/EN/Downloads/Publikation/Details_of_the_Asset_Administration_Shell_Part1_V3.html*

See also DIN SPEC 91406[16] (Automatic identification of physical objects and information on physical objects in IT systems, particularly IoT systems) for further information on identification.

7 GAIA-X AND IDS

Recently raised federated data infrastructures are based on interoperability and information sharing between autonomous and semi-autonomous decentralized IT systems with a main focus on data sovereignty and security.

GAIA-X European Association for Data and Cloud AISBL[17] is a non-profit association founded in January 2021 with the aim of developing technical solutions and a regulatory framework for federated and secure data services in business ecosystems [GAIA-X-2021a]. GAIA-X addresses different sectors including mobility, energy, manufacturing, finance, agriculture, public services and healthcare.

As stated in [GAIA-X-2021b], the core values of GAIA-X include data sovereignty, privacy and confidentiality, security, technology neutrality and interoperability. GAIA-X is not a European cloud solution, but instead provides an architecture standard to connect various cloud service providers while complying with European data protection guidelines.

[16] *https://www.beuth.de/en/technical-rule/din-spec-91406/314564057*

[17] *https://www.gaia-x.eu/*

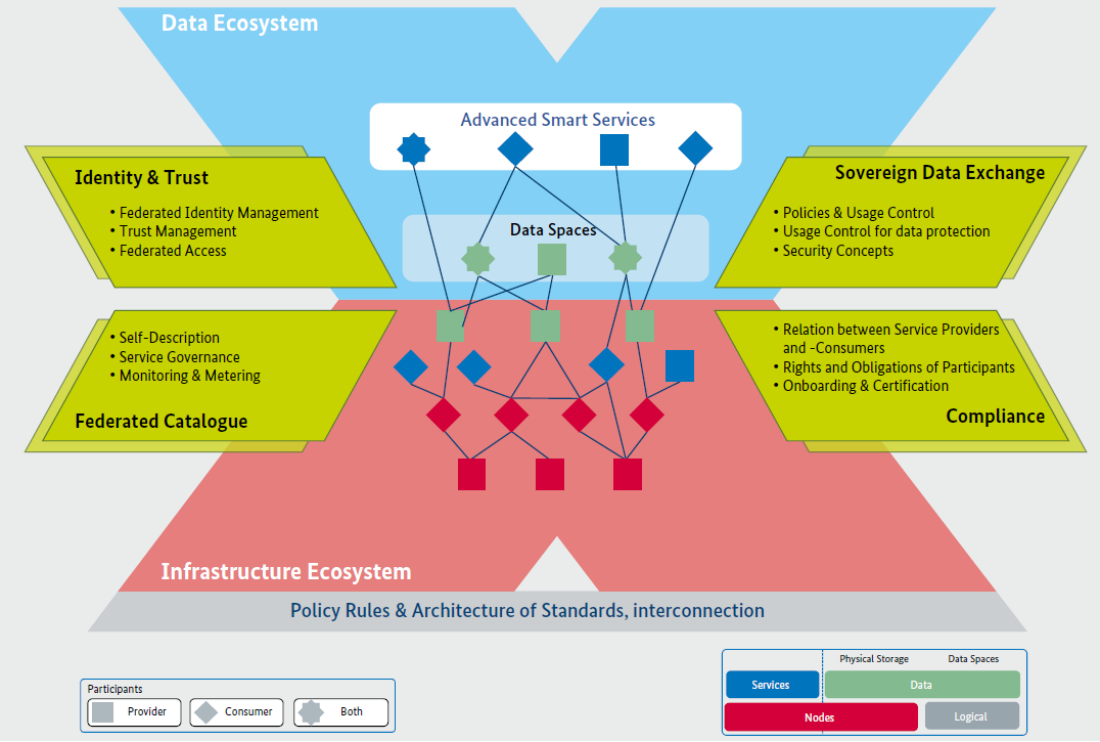

Figure 7-1. GAIA-X Ecosystem.

The International Data Spaces Association (IDSA)[18] is a founding member of GAIA-X. IDSA's mission is to create a secure, sovereign system of sharing in data spaces. The IDSA design concepts on trust and usage control in data spaces are an integral part of GAIA-X, cf. the overview in *https://internationaldataspaces.org/we/gaia-x/*.

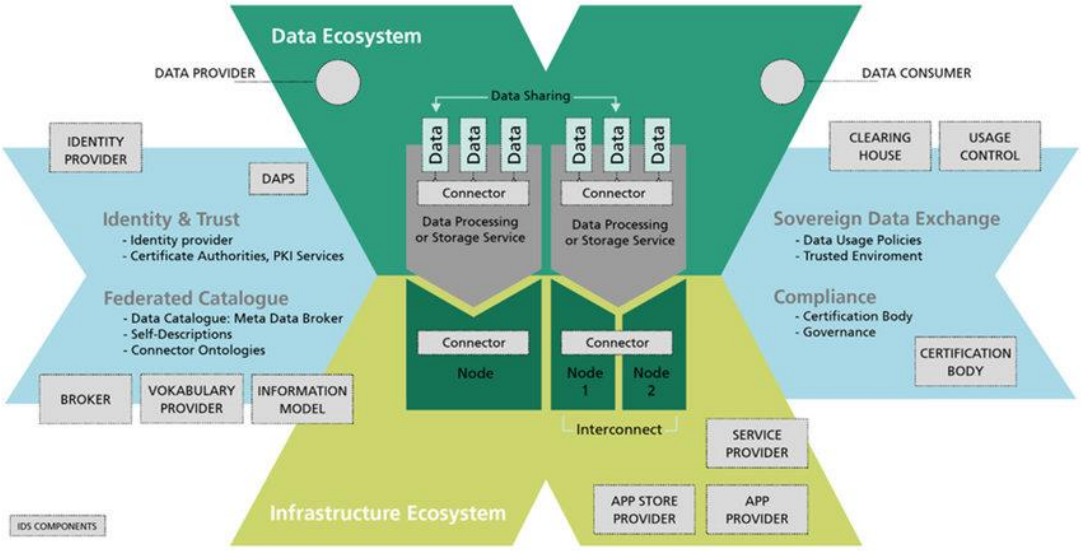

Figure 7-2. Mapping of IDS Components into the GAIA-X Architecture. *Source: IDSA-2021.*

18 *https://internationaldataspaces.org/*

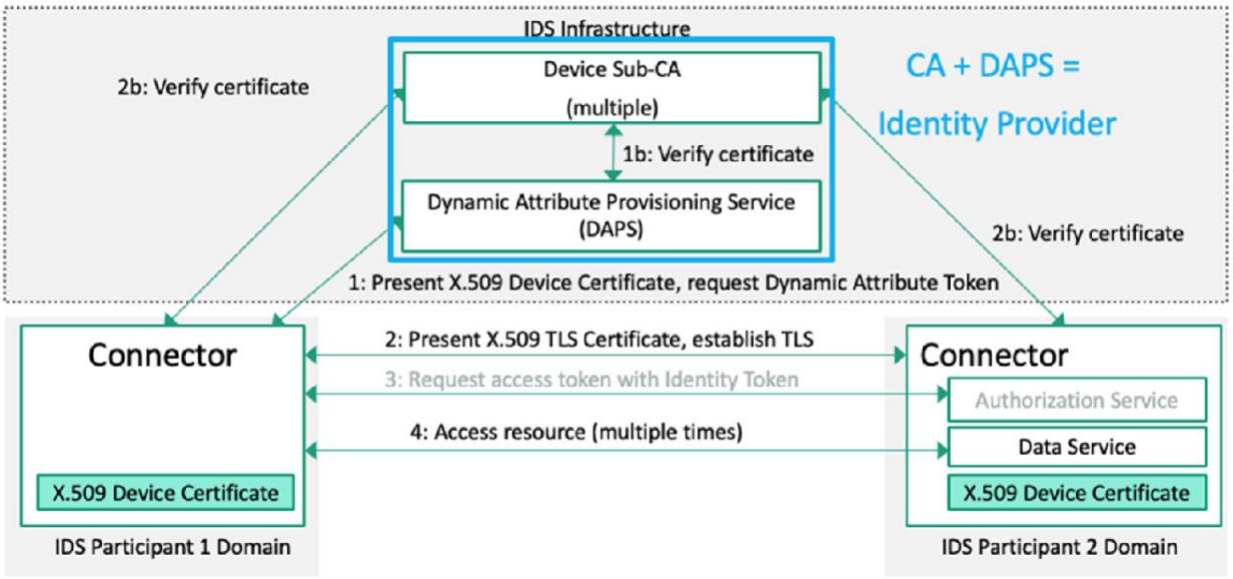

Figure 7-3. IDS Identity Concept. *Source: IDSA-2021.*

7.1 GAIA-X Conceptual Model

The GAIA-X conceptual model describes all concepts in the scope of GAIA-X and relations among them. Minimum versions of important core concepts include mandatory attributes as well as their types and cardinalities. Such as Participant, Provider, Service Offering, Asset, Data Asset, Data Service Offering, Software Asset, Node and Interconnection.

The GAIA-X Architecture document [GAIA-X-2021a] identifies and describes the concepts of the targeted federated open data infrastructure as well as the relationships among them. It describes how GAIA-X facilitates interconnection, interoperability and integration among all participants in the European digital economy, relative to both data and services.

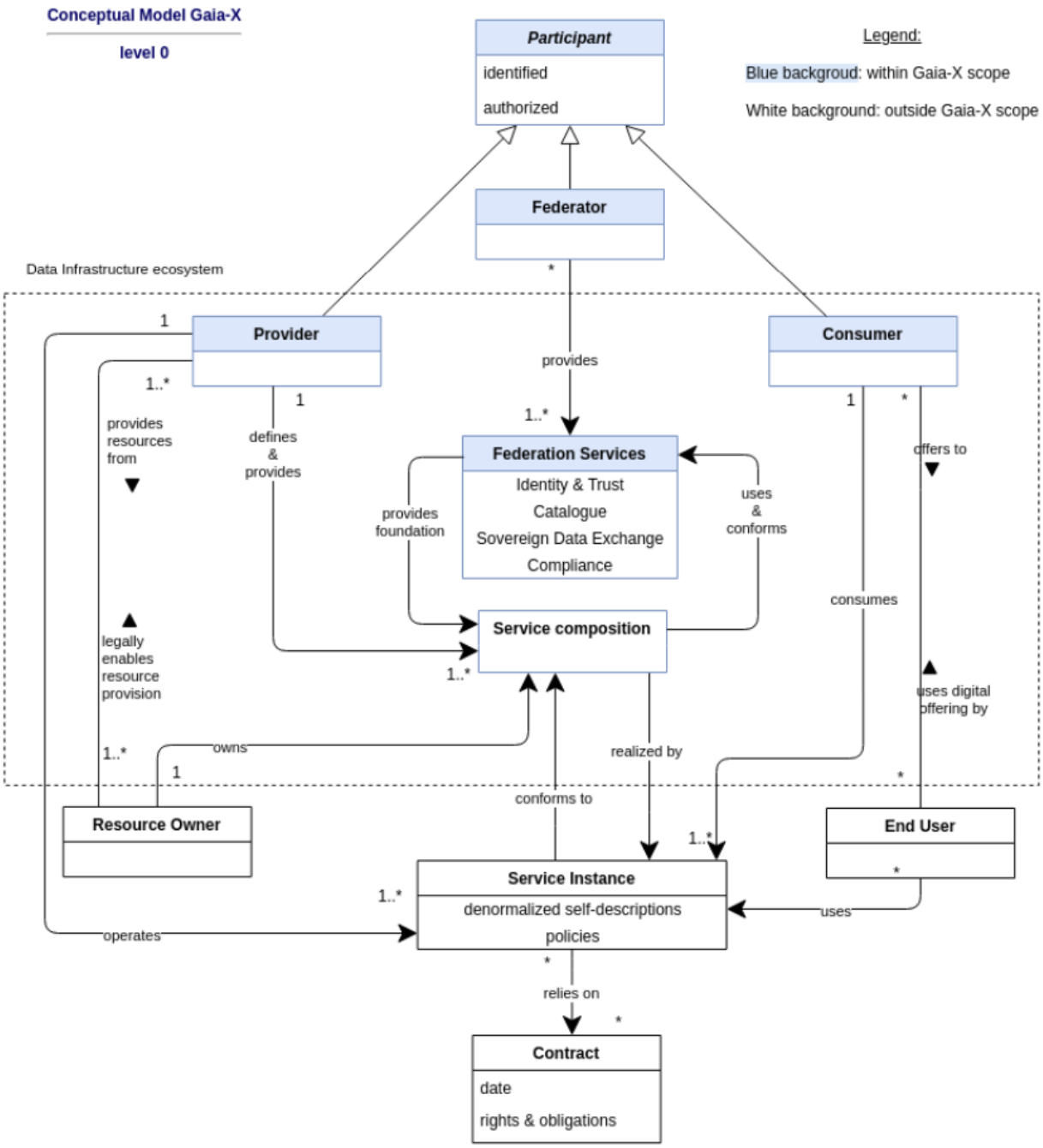

Figure 7-4. GAIA-X Conceptual Model. *Source: GAIA-X-2021a.*

7.2 GAIA-X IDENTIFIERS

GAIA-X Identity System provides additional attributes to the identity of the Gaia-X Principal. (A principal is either a natural person or a digital representation which acts on behalf of a Gaia-X Participant) and forwards this identity to the requestor. A Gaia-X accredited Identity System follows a hybrid approach and consists of both centralized components, like company identity management systems, and decentralized components like Decentralized Identifiers (DIDs).

Requirements for identifiers within the GAIA-X scheme are strictly framed in a set of rules applied across the ecosystem. GAIA-X by itself is not an identifier issuer. It is stated: "There SHALL NOT be GAIA-X issued Identifiers." The concept of reuse of the existing identifiers set the following requirements:

- *RI1.* Identifiers used in GAIA-X must be unique in the context of GAIA-X. This is achieved by combination of a schema and the protocol-specific ID.
- *RI2.* Identifiers are controlled by the identity system of a participant, therefore there shall be no central repository where identifiers must be published.
- *RI3.* Any GAIA-X participant's identity system can generate identifiers. The structure of an identifier has to ensure its uniqueness. It is solely the responsibility of a participant to determine the conditions under which the identifier will be issued.
- *RI4.* Identifiers can be referenced without publishing the identifier beforehand in a separate system.
- *RI5.* Identifiers shall be derived from the native identifiers of an identity system without any separate attribute needed. The identifier shall provide a clear reference to the identity system technology used. OpenID Connect and DID shall be supported. Any scheme for identifiers must permit future extensions to the scheme.
- *RI6.* It is intended that the lifetime of an Identifier is permanent. That is, the Identifier will be globally unique forever,[19] and may be used as a reference to a resource well beyond the lifetime of the resource it identifies or of any naming authority involved in the assignment of its name. Reuse of an identifier for a different entity is forbidden.
- *RI7.* An identifier shall not impede resolution. To be more specific, for identifiers that have corresponding URLs or other resource access protocols, there must be some feasible mechanism to translate an identifier to an address of the resource [RFC1737].[20]
- *RI8.* The identifier shall be comparable in the raw form. A transformation must not be needed to compare two identifiers and tell whether they are the same.
- *RI9.* Identifiers should not contain more information than necessary (including personal identifiable information).

A central matching table is not necessary (as there is no central ID authority within GAIA-X). Participants identity systems can self-issue valid principal identifiers. In line with the above, it is

[19] Persistence of identifiers is defined in [RFC1737, lifetime, Section 2] and [ITU-T (X.1251)].

[20] RFC 1737, Functional Requirements for Uniform Resource Names, *https://datatracker.ietf.org/doc/html/rfc1737; Persistence of identifiers is defined in section 2.*

possible to determine the technology and the unique reference to the identity based on the identifier.

8 INDUSTRIAL VALUE CHAIN INITIATIVE

Industrial Value Chain Initiative (IVI) is a forum and platform for activities to design a new industrial world, where manufacturing and IT are integrated. IVI is also a collaborative area, in which companies can both cooperate and compete by taking advantage of their own technologies. IVI organizes and shares the collaborative area as a reference model to integrate a system in which unique technologies of companies are interconnected with each other.

From the architecture standpoint the Connected Industries Open Framework (CIOF)[21] for manufacturing aims to digitalize production sites in the manufacturing industry. It enables an agreed data exchange as a trading profile between sites to configure and distribute data easily. The specification document defines the basic user and implementation requirements for the system and its subsystems.

IVRA (Industrial Value Chain Reference Architecture)[22] is a three-dimensional reference architecture model of smart manufacturing that inherits CIOF functional modules and terminology. In comparison with similar models (Plattform Industrie 4.0, IDS, IIC) it has:

- an asset view (personnel, plant, product, process or man, machine, material, method),
- an activity view (plan, do, check, act) and
- a management view (quality, cost, delivery, environment).

This reference architecture is based on four pillars:

- Smart Manufacturing Unit (SMU) with activity view based on the Plan-Do-Check-Act (PDCA) cycle, asset view (4P: Personal, Plant, Product, and Process, each of which corresponds to 4M: Man, Machine, Material, and Method respectively), management view addressing quality, cost, delivery and environment.
- Three-axis value chain: product axis, service axis, knowledge axis.
- Portable Loading Unit (PLU): value container, cyber container and physical container.
- IVI ontology: actor, information, thing and data.

[21] *https://iv-i.org/wp-content/uploads/2019/01/CIOF_SystemRequirements_pr.pdf*

[22] *https://ivi-iot.sakura.ne.jp/iviwp_renewal/wp-content/uploads/2018/04/IVRA-Next_en.pdf*

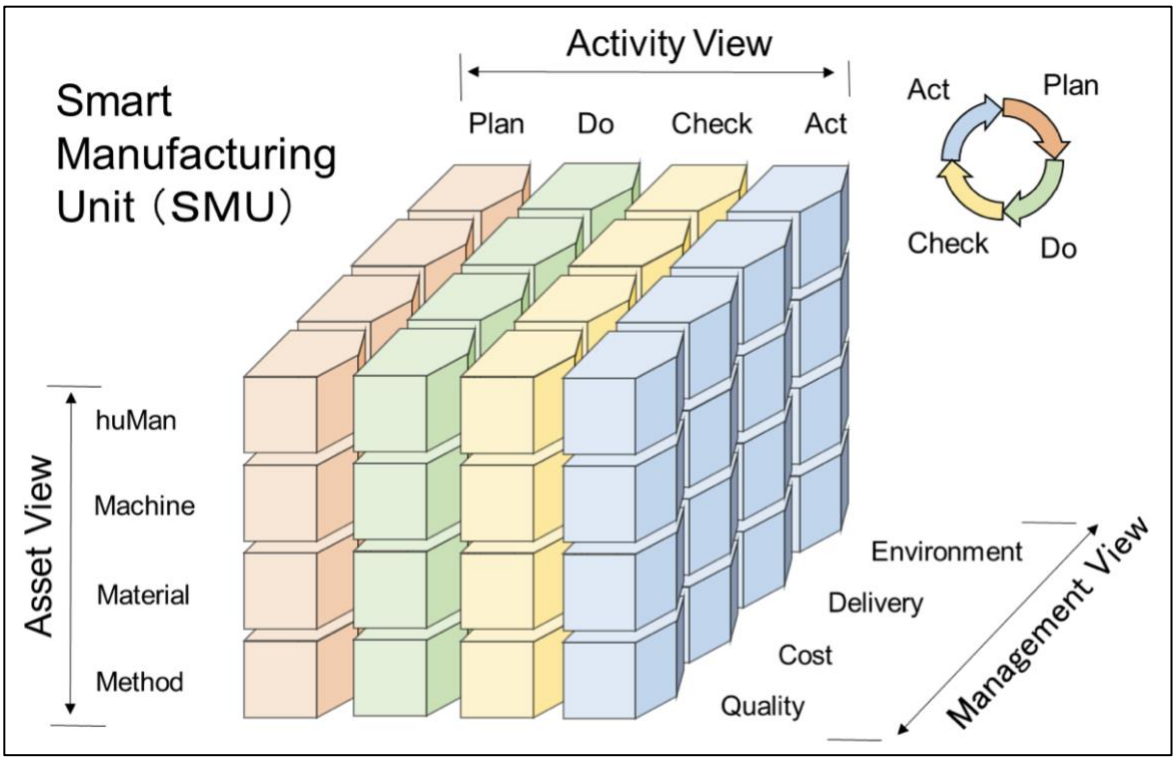

Figure 8-1. IVRA Reference Architecture.

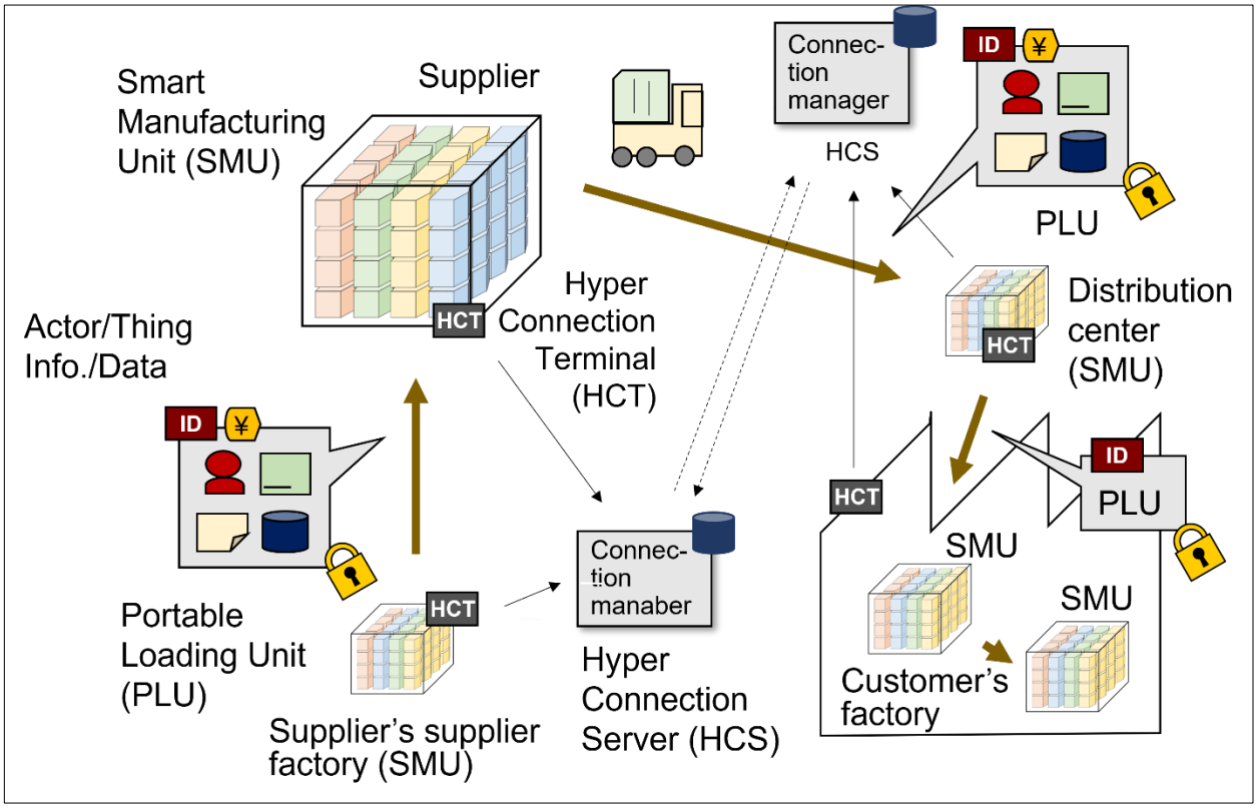

Figure 8-2. IVRA Data Trade as Defined in CIOF.

In the IVRA, both data and objects are traded between SMUs. Procedures for data trading are defined by the Connected Industries Open Framework (CIOF) specification.[23] CIOF enables data trading, as a process of establishing and fulfilling data rights and obligations in advance for both data providers and data users when moving data between different terminals. Trading involves a series of actions, such as making contact, transmission of data, fulfillment of rights and obligations and expiration of the contract. The trading data received from a Hyper Connection Server (HCS) is assigned a unique ID. The HCS ID and trading data ID are globally unique. A hash is generated for trading data and by saving the hash instead of the trading data, it is possible to match the trading data. In addition, this hash is used for traceability of the trading data.

The overall architecture and process flow is shown in Figure 8-3:

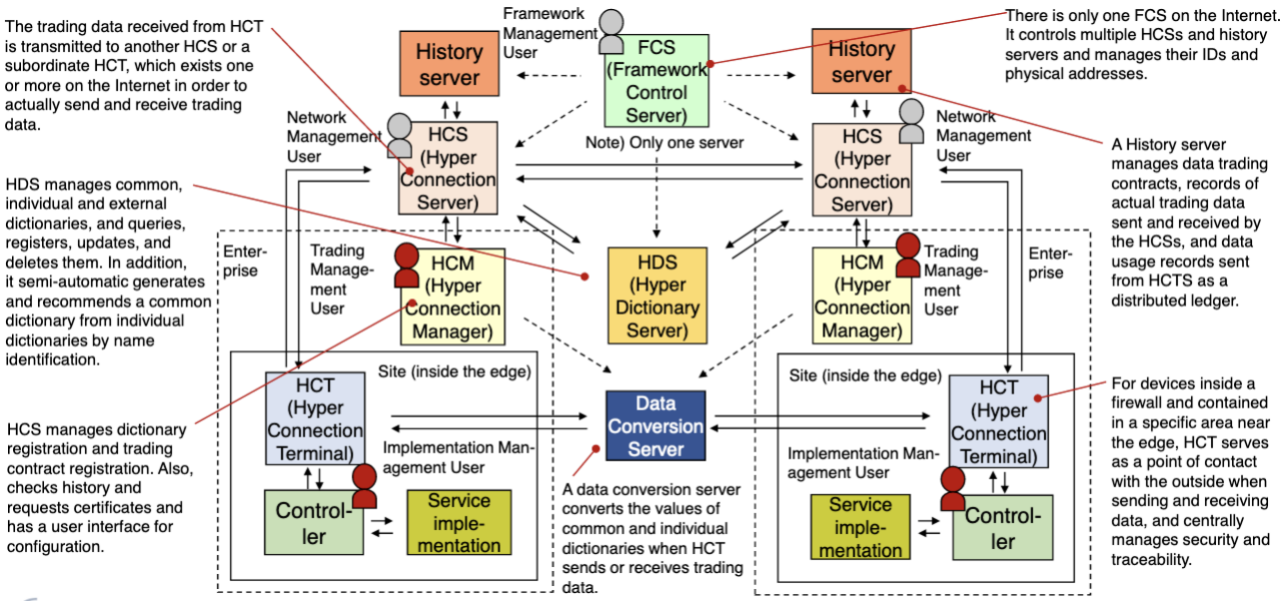

Figure 8-3. Process and Architecture from CIOF.

In CIOF, uniqueness of ID is ensured in a bottom-up style. They are checked and kept unique when SMUs trade data.

[23] *https://iv-i.org/wp-content/uploads/2019/01/CIOF_SystemRequirements_pr.pdf*

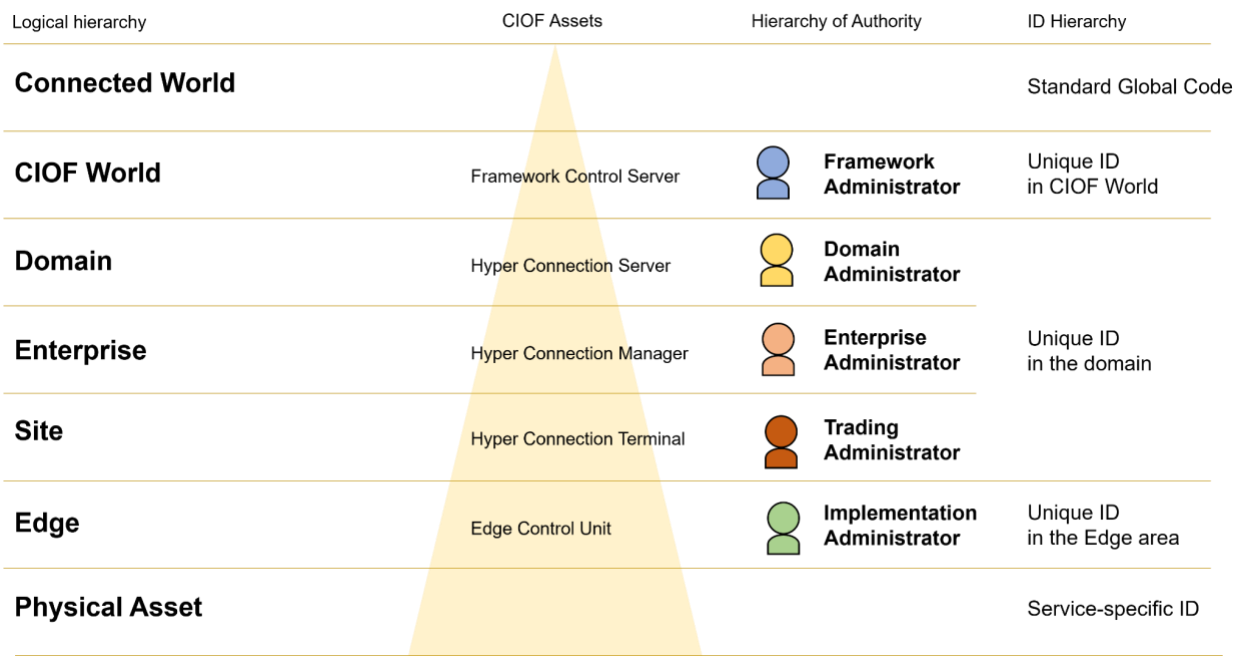

Figure 8-4. Bottom-Up ID Structure of CIOF.

The ID hierarchy has three control levels: global ID, ID within CIOF world, and ID within the local manufacturing environment:

- *Standard Global Code:* CIOF enables connection to any external system by associating a unique ID in the CIOF world with an external global ID.
- *Unique ID in CIOF world:* Since CIOF is a distributed system that connects platforms managed by different enterprises, common ID management is minimized because each platform functions autonomously and manages internal IDs by its own rules
- *Unique ID in a Domain:* Most of the data used by CIOF is managed by IDs issued within a domain. That is, a platform can identify the data for each HCS. If an enterprise or site is transferred to another domain, different IDs will need to be issued. Enterprise IDs and user IDs are unique in a domain.
- *Management ID in a Controller:* IDs that are local inside the edge can be independently issued and managed separately from the IDs managed by CIOF. The edge controller is responsible for the association between the CIOF-managed IDs and the unique internal IDs.
- *Service Specific ID:* Various assets deployed in the field are assigned unique IDs set by each manufacturer and service provider. These IDs can be used as they are management IDs inside the controller.

9 ALLIANCE OF INDUSTRIAL INTERNET (AII) IN CHINA

Alliance of Industrial Internet (AII)[24] was established to accelerate the development of the industrial internet, to gather industrial power in China and abroad, to establish a public platform for collaborative improvement in administration, industry, academia, research and application, to conduct joint research and to facilitate the application of the industrial internet. AII released the *Industrial Internet Architecture 2.0* on October 27, 2019.

To satisfy the requirements of digital transformation, Industrial Internet Architecture 2.0 integrates the latest concepts, values, technologies, functions, paradigms and processes of the industrial internet, forming a comprehensive system framework to guide the multi-level work of the state, society, industry and enterprises.

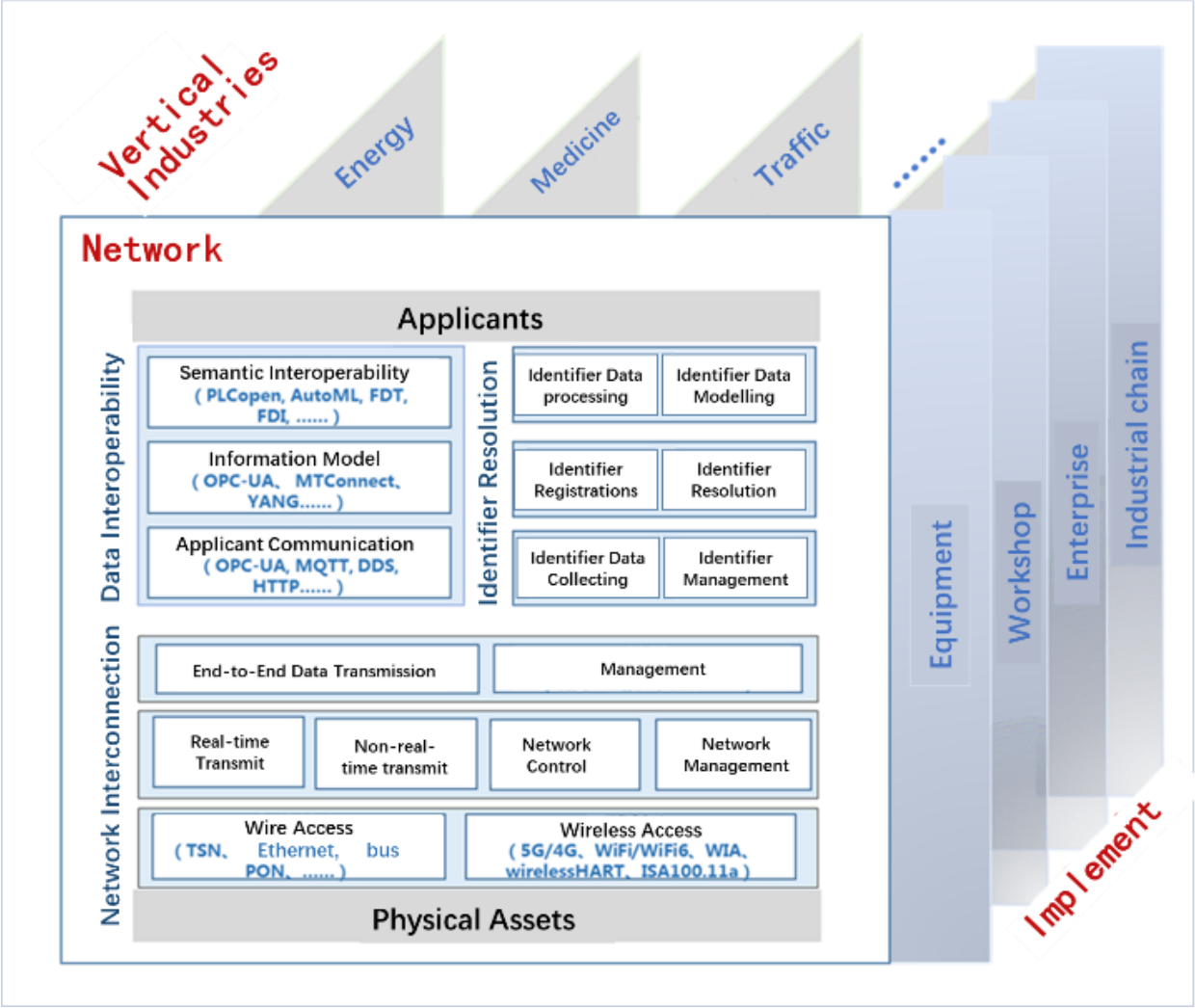

Figure 9-1. The Chinese Industrial Internet Architecture 2.0.

[24] *http://en.aii-alliance.org/index.php*

9.1 THE IDENTIFICATION RESOLUTION SYSTEM

The identification resolution system is the nerve center of the industrial internet. The function is similar to the domain name system (DNS) for the internet. The enterprise information system will realize the data flow through the identification resolution system. The core of the identification resolution system includes the following parts: identification code, resolution system and data service.

- *Identification code:* The identification code can uniquely identify physical resources such as machines and products as well as virtual resources, similar to an *identity card*.
- *Resolution system*: The system device can query the network location of the object or related information according to the identification code. It is the premise and basis for connecting the different enterprise production systems. The detailed operation of resolution systems is shown in Figure 9-2 and Figure 9-3.
- *Data service:* The data service carries out identification data management and data sharing across enterprises, industries, regions and countries using the identification code and the resolution system.

The Chinese Industrial Internet identification resolution system has adopted several standard schemes, such as VAA[25], GS1(ISO/IEC 15459), OID (ISO 9834-1) and Handle (RFC 3652). At present, VAA is mainly used to provide a global unique identification for people, machines and other objects in identification resolution systems.

The VAA is compatible with a variety of identification carriers, including passive identification carriers (such as one-dimensional bar code, two-dimensional code, RFID and NFC) and active identification carriers (such as general integrated circuit cards, chips and modules). Now, VAA has been applied in more than 30 industries and it is fully compatible with all kinds of identification in identification resolution systems.

The Chinese identification resolution system is designed as a hierarchical architecture to establish distributed identification resolution nodes to deal with global and large-scale identification query applications. The identification resolution systems architecture is given in Figure 9-2.

[25] CAICT is the official issuing agency with Issuing Agency Code (IAC) “VAA,” given by ISO/IEC 15459. The IAC indicates an authorized qualification of distributing identifiers with its own allocation rules. *http://www.caict.ac.cn/english/news/202006/t20200629_285130.html*

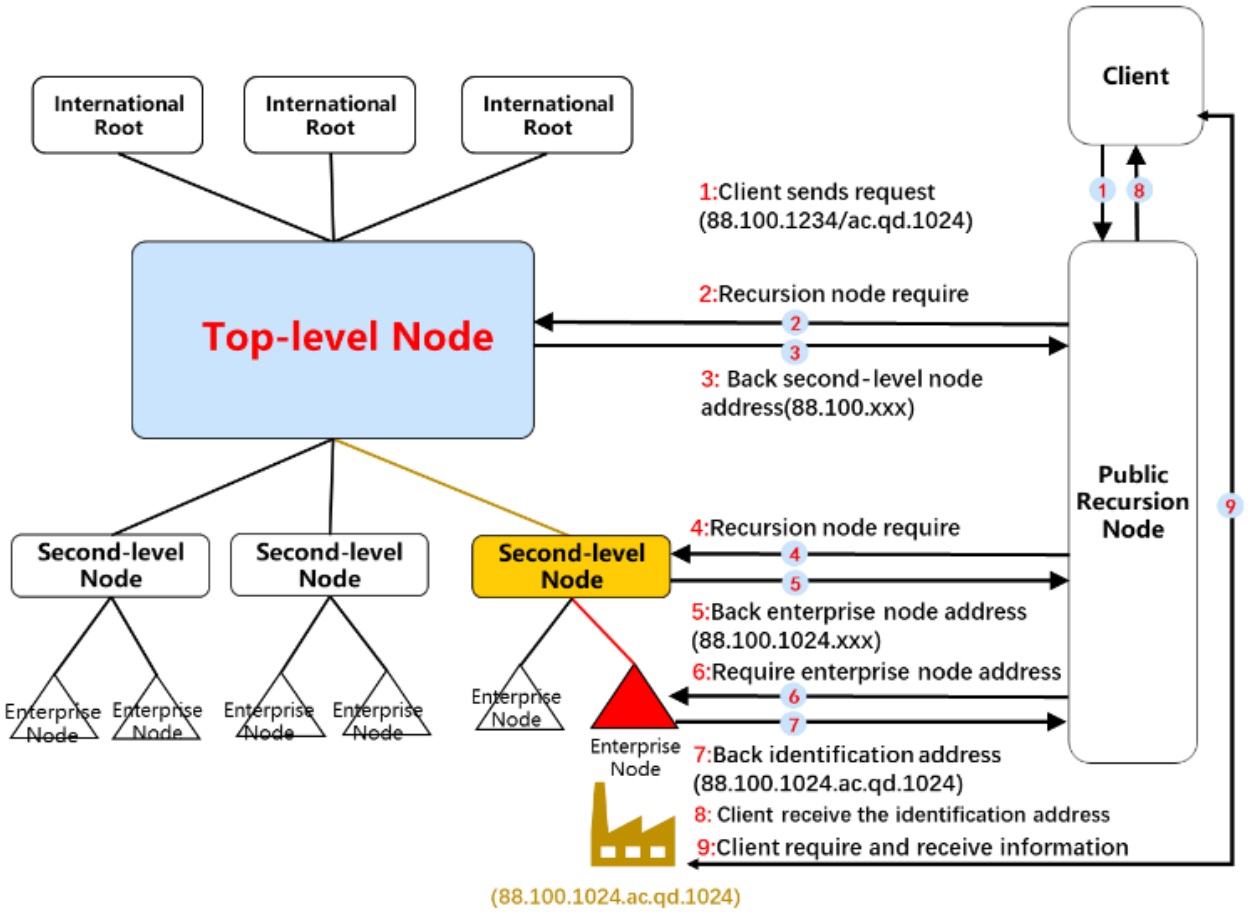

Figure 9-2. The Architecture of the Identification Resolution System.

The industrial internet identification resolution system provides registration, resolution and discovery services for physical or virtual objects. One or more location identifications are retrieved through resolution, and then retrieved according to the location information, which can return location information or content service information. In Figure 9-3, each enterprise node registers information through the secondary node, and the national top-level node assigns a globally unique identification. Through the identification resolution system, the information interaction between enterprise nodes is realized.

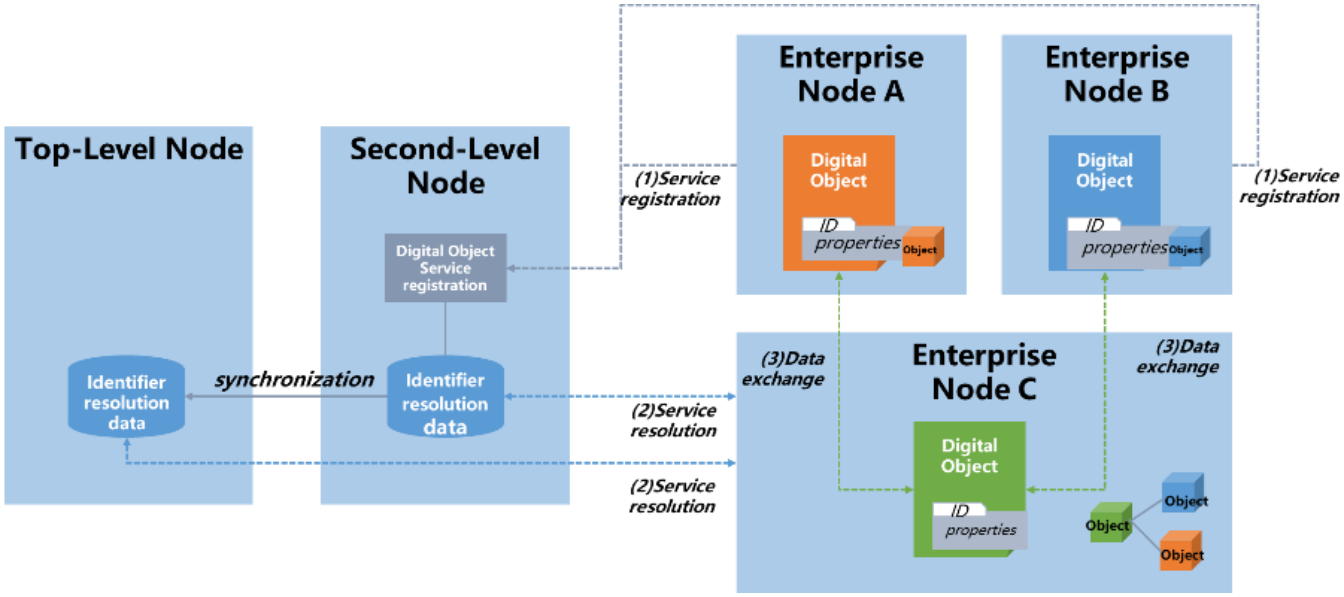

Figure 9-3. The Architecture of Enterprise Data Interaction.

10 Observations and Conclusions

Many different identification systems for the entities in an IoT system have emerged in recent years. The primary motivation for identifiers is to define precisely which entities are being referred to when exchanging data between systems. The entities needing an ID range from simple devices and single variables to complex systems of systems. Entities also include concept descriptions with further, often essential, information about other entities.

All concepts, especially those corresponding to real-world phenomena, need an ID to ensure that their semantics are unambiguous and can be correctly interpreted. Identifiers are essential for sustainable and interoperable systems. Identifiers from one ecosystem should be reusable in or easily translatable to other ecosystems. Accordingly, the scope, richness and maturity of identifier systems has grown continually.

Experience has shown the need to cover three activity fields for identifiers: registration, resolution and discovery. There is a trend towards distributed governance of identifiers, either with a trusted central ‘federal’ authority at the top or no central authority at all. The organization(s) governing the identification system and the underlying rules and standards need to be independent of companies and governments to achieve international acceptance and widespread usage.

Resolution is important to integrate applications efficiently. There will most likely be no universal solution and new solutions will emerge. However, it is to be hoped that the set of core basic standards will be extended over time based on the lessons learned in the current IoT initiatives. For example, DID could well be used more widely and be considered as a basic standard in the

future. Consortia-related projects will review and refine the existing concepts when designing new systems, ideally leading to a convergence of ideas.

IIC as an independent organization will continue playing an important role in this process by evaluating, consolidating and evolving identifier systems for consideration by other initiatives and SDOs. Identifiers will remain a key prerequisite for interoperability in IoT systems.

11 REFERENCES

- [ODI] Leigh Dodds, Libby Young "Using identifiers" guide. Open Data Institute *https://theodi.org/article/using-identifiers/*
- [RAMI4.0-2016] DIN SPEC 91345:2016-04 "Reference Architecture Model Industrie 4.0 (RAMI4.0)", April 2016. *https://www.beuth.de/en/technical-rule/din-spec-91345-en/250940128*
- [AAS-2020] German Federal Ministry for Economic Affairs and Energy (BMWi), "Details of the Asset Administration Shell –Part 1: The exchange of information between partners in the value chain of Industrie 4.0 (Version 3.0RC012.0)", 2020-November, *https://www.plattform-i40.de/PI40/Redaktion/EN/Downloads/Publikation/Details_of_the_Asset_Administration_Shell_Part1_V3.html*
- [ECLASS-2020] EClass Whitepaper, Distributed Ledger-based Infrastructure for Industrial Digital Twins, 2020-December, Version 1.1, *https://www.eclass.eu/en/association/digitalization-expert-group/eclass-and-the-digital-twin.html*
- [GAIA-X-2021a] GAIA-X Architecture Document, 21.12 Release, *https://gaia-x.eu/sites/default/files/2022-01/Gaia-X_Architecture_Document_2112.pdf*
- [GAIA-X-2021b] GAIA-X Federation Services, GAIA-X ecosystem Kickstarter, 2021-12-01, *https://gaia-x.eu/sites/default/files/2021-12/GXFS_1.pdf*
- [Gaia-X-2021c] GAIA-X IAM Framework, Version 1.2, accessed on 2021-12-13, *https://community.gaia-x.eu/s/P23ZJNLyjf7n7Zp?path=%2FReleases*
- [IDSA-2019] Reference Architecture Model of the International Data Spaces Association, Version 3.0, April 2019, accessed on 2022-01-07. *https://internationaldataspaces.org/wp-content/uploads/IDS-Reference-Architecture-Model-3.0-2019.pdf*
- [IDSA-2021] GAIA-X and IDS, Position paper, Version 1.0, January 2021, accessed on 2021-12-13, *https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/IDSA-Position-Paper-GAIA-X-and-IDS.pdf*

- [ISO10646] International Organization for Standardization, "ISO/IEC 10646:2003: Information Technology - Universal Multiple-Octet Coded Character Set (UCS)", ISO Standard 10646, December 2003.

12 AUTHORS & LEGAL NOTICE

This document is a work product of the Industry IoT Consortium Technology Working Group.

Technical Editor: Stephen Mellor (IIC staff) oversaw the process of organizing the contributions of the above Authors and Contributors into an integrated document.