Title: Requirements for TRUST semantic interoperability

Short description: This paper explains the IDS certification process, one reference implementation and the certification information flow based on the IDS certification criteria catalogue for components (IDS3-c-co), DAPS and IDS Info model. Finally, we identify what is required to develop in the future.

Sonia Jimenez, Senior Business Consultant and project manager, IDSA
Jörg Langkau, Researcher & Development, nicos AG
Silvia Castellvi, Senior Consultant, IDSA
Speaker (s) at the workshop: Jörg Langkau, nicos AG

Abstract. *The International Data Spaces initiative (IDS) is building an ecosystem to facilitate data exchange in a secure, trusted, and semantically interoperable way. It aims at providing a basis for smart services and cross-company business processes while at the same time guaranteeing data owners' sovereignty over their data. This paper explains the IDS certification process based on IDS Reference Architecture Model (RAM) defines static trust based on certification and the dynamic trust based on active monitoring of participants and core components. After explaining the certification process, the IDS Information Model and the IDS Certification Criteria Catalogs for Component and a machine-readable version of the criteria catalogue and test specifications. Finally, this paper identifies future developments to develop* decentralised Verifiable Credentials (VC) *in a federated ecosystem.*
ecosystem.

1. Introduction

The International Data Spaces (IDS, formerly: Industrial Data Space) was started as a research project in 2015 by the German government. As of 2022, the IDS contains +130 members from academia and private organizations. The mission of the IDS is to create “a secure data space that supports enterprises of different industries and different sizes in the autonomous management of data.”

The International Data Spaces initiative[1] targets the sovereignty of data exchange by promoting a standard for virtual data spaces for reliable data exchange among business partners. To achieve the goal of sovereignty data exchange aspects of data management, semantic data interoperability and security have to be addressed[2]. **IDS trust** for data sharing and data exchange is a fundamental requirement.

The International Data Spaces (IDS) Reference Architecture Model (RAM)[3] defines two basic types of trust: 1) Static Trust, based on the certification of participants and core technical components, and 2) Dynamic Trust, based on active monitoring of participants and core technical components. Preliminary actions and interactions are required for data exchange in an IDS data space or the IDS ecosystem. These are necessary for every participant and involve the Certification Body, Evaluation Facilities, and the Dynamic Attribute Provisioning Service (DAPS).

[1] Home - International Data Spaces
[2] Bader et al. 2020
[3] IDS Association. Reference Architecture Model (Ver. 3.0). International Data Spaces Association. Berlin, Germany, 2019; https://internationaldataspaces. org/download/16630/.

2. The state-of-the-art section

a. International Data Spaces (IDS)

The architecture of the IDS is described in the International Data Spaces (IDS) Reference Architecture Model (RAM)[4] document published by the initiative. In the IDS ecosystem, each participant is evaluated and certified before being granted access to the trusted business ecosystem. All components in the IDS relay all components of the IDS rely on state-of-the-art **Security and data sovereignty** measures; security is mainly ensured by the evaluation and certification of each technical component used in the IDS. Finally, a data owner in the IDS attached usage restriction information to their data before it is transferred to a data consumer; the data consumer must fully accept the data owner's usage policy.

The IDS connector acts as the core component in the IDS. Any organization who wants to participate in the IDS needs to set up an IDS Connector. Every Connector participating in the International Data Spaces must have a unique identifier and a valid certificate. In addition, each Connector must be able to verify the identity of other Connectors (with special conditions being applied here; e.g., security profiles).

The IDS Information Model (formerly called IDS Vocabulary) is an RDFS vocabulary and serves as a metamodel within the IDS. It provides the basis for the creation of machine-interpretable self-descriptions for the IDS Connector, datasets to be offered in the IDS, security profiles and contracts.

b. IDS Certification

Data security and data sovereignty are the fundamental value propositions of the International Data Spaces. Data sovereignty can be defined as a natural person's or legal entity's capability of being in full control of its data. To enable this control, each participant needs to follow the agreed rules for the IDS and requires reliable information about the guarantees offered by potential business partners. The certification ensures adherence to rules, which is based on defined standards.

Any organisation or individual seeking permission to operate components in the International Data Spaces needs to pass the Operational Environment Certification, ensuring secure processes and management of components. Comparably, each IDS component is expected to adhere to IDS specifications and protect the data transferred and processed. It shall allow participants to assess the possible consequences of data sharing and provide transparent information about possible guarantees with regard to Usage Control. Therefore, components need to pass a Component Certification before they may be used in the IDS. While the certification of organizations and individuals focuses on security and trust, the certification of components additionally refers to compliance with technical requirements ensuring interoperability.

The IDS uses a Certification Scheme comprising all processes, rules, and standards governing the certification process to ensure a consistent process in the certification of participants and core components. The IDS Certification Scheme follows best practices from other internationally recognised certification concepts.
The realization of the IDS Certification schema requires different roles responsible for different tasks:

- Applicants (actively submit an application to trigger the certification process).
- Evaluation Facilities, (responsible for carrying out the detailed technical and/or organizational evaluation work during a certification process), and
- one Certification Body (oversees the certification process regarding quality assurance and framework governance. It defines standard evaluation procedures and supervises the actions of the Evaluation Facilities).

[4] IDS Association. Reference Architecture Model (Ver. 3.0). International Data Spaces Association. Berlin, Germany, 2019; https://internationaldataspaces. org/download/16630/.

The certification follows the same process for all certification profiles in Operational Environment and Component Certification. It consists of the following three phases:

1. Application Phase: The main goal of this stage is the successful start of the IDS evaluation and certification process.

2. Evaluation Phase: The main goal of this stage is the evaluation of an applicant or core component based on the defined evaluation criteria.

3. Certification Phase: The main goal of this stage is the examination of the evaluation report by the certification body, which issues a certificate if the result of the evaluation process is positive.

c. Verifiable Credentials (VC)/Verifiable Presentation (VP)

A verifiable credential (VC) is a tamper-evident credential that has authorship that can be cryptographically verified. Verifiable credentials can be used to build verifiable presentations, which can also be cryptographically verified. The claims in a credential can be about different subjects.

A *verifiable presentation* is a tamper-evident presentation encoded in such a way that authorship of the data can be trusted after a process of cryptographic verification. Certain types of verifiable presentations might contain data that is synthesized but do not contain the original, verifiable credentials (for example, zero-knowledge proofs).

3. IDS Certification process description.

One example of data spaces is the Mobility Data Space[5] lunched in 2020 as a multi-stakeholder project in Germany that aims data-driven mobility services and data sovereignty of the data holders and trust among all participants. Figure 1. shows the basics data and metadata flows to provide services for intermodal end-to-end services, traffic management services or services to increase road safety for individual drivers. To ensure a consistent process in the certification of participants and core components, the IDS uses a Certification Scheme comprising all processes, rules, and standards governing the certification process in addition to the federated catalogue. It connects the data providers with data users by the respective IDS Connector component.

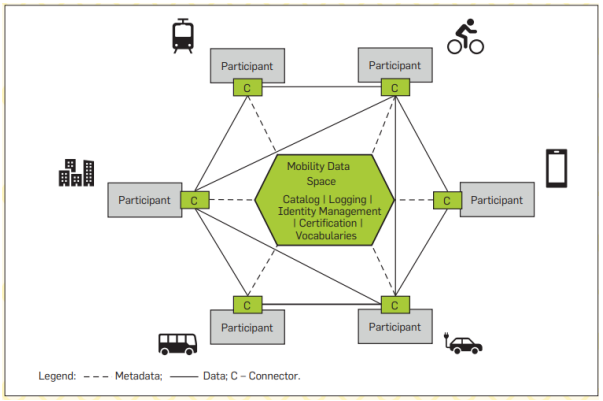


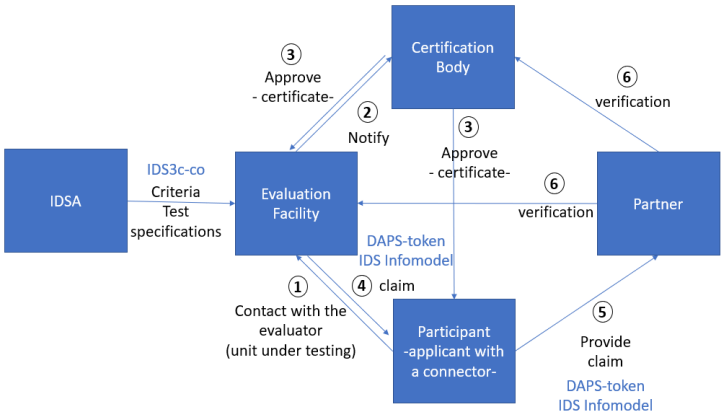
Figure 1. Mobility Data Space overview

Figure 2. illustrates the roles and interactions required for issuing a digital identity management (IdM) in the IDS. The Certification Body, together with selected Evaluation Facilities, oversees the certification of the participants and the core technical components in the IDS. These Governance Bodies make sure that only compliant organizations are granted access to the trusted business ecosystem. In this process, the Certification Body supervises the actions and decisions of the Evaluation Facilities.

To evaluate the signed certificate claims, participants and components are subject to a certification process – an additional means to establish trust within and across data spaces. A normative, tamper-proof reference of certification, security and identity attributes is maintained by IDS infrastructure components, operated as part of an Identity Management (Fig. 1).

[5] See https://mobility-dataspace.eu/

The following steps describe the certification process: 1- The applicant contracts an approved Evaluation Facility to carry out the evaluation according to the IDS certification scheme. 2- The Evaluation Facility carry out the detailed technical and organizational evaluation of the component and notify the Certification Body the evaluation report. 3- The Certification Body review, provides comments, generates the approval of application for certification and the authorization of a X.509 certificate. 4- Evaluation Facility provide a signed certified claim (DAPS-token + IDS Info model) to the participant. 5- The participant provides the claim (DAPS-token + IDS Info model) to the Partner. 6. The Partner verify the claim with the Evaluation Facility and the Certification Body.



What we already have now is the IDS3c-co, the DAPS-Token and the IDS Information.

Figure 2. IDS roles and interactions of certification process

How is the information flow for certification approval? How can we semantically describe the criteria, the test certification, the claims and the claims proof? How can we describe it in a semantic way? These are the challenge that we have currently.

We have most the required artefacts for certification process implementation, but we need to develop approve claim, partners verification by evaluation facility and verifiable credentials and presentation.

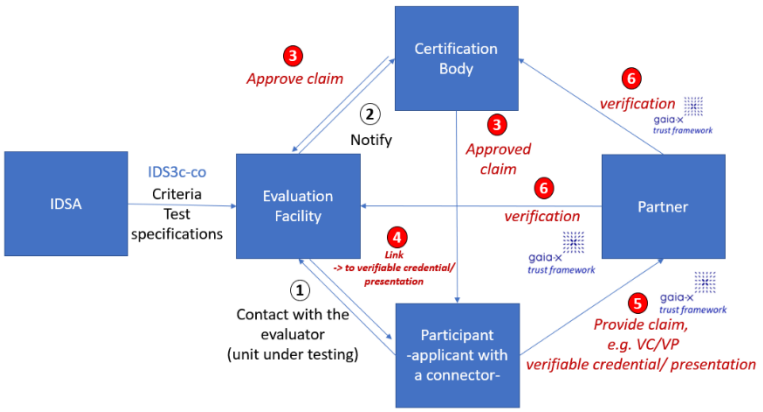


Figure 2. explains the information flow (*black*), what already has been developed and can be improved (*blue*) and figure 3. shows what is missing and needs to be developed (*red*). Also, we can see how we can integrate and use the Gaia-X trust framework.

Figure 3. Future work on IDS provides claims, VC/VP model.

What is missing are the approved claim, claim evidence proof and verifiable credentials and verifiable presentations that can replace the DAPS-tokens.

4. The IDS3c-co

The International Data Spaces Certification Criteria Catalog for Components (ids3c-co) describes the minimum requested requirements to be met by the IDS' components/roles supporting data exchange transactions. To secure the intended cross-industrial and cross-company information exchange, the Industrial Data Space core components must provide the required functionality and an appropriate level of security. As such, the core component certification is interoperability- and security-focused, while aiming to strengthen the development and maintenance process of these components.

There is a handbook of IDS Certification Criteria Catalogs for Component and a machine-readable version of the criteria catalog and test specifications to support with long-term management of these documents.
The IDS Criteria Catalog for Components tooling requirements are (R1) a machine-readable and human-readable that (R2) make possible mapping between different parts (e.g., criteria->test design spec->test cases; test case->test design spec,….), (R3) render the content into different formats and (R3) check for inconsistencies. In order to fulfil these requirements, it was decided to use a defined ontology for the entire content which helps structuring content by assigning artifact types and allowed formats to different aspects/objects. To be aligned with the utilized IDS Info Model, this ontology was defined in Turtle and is provided in the Repo. Turtle provides a structure for defining content based on the defined schema. It supports automated checks (R$) with so-called SHACL-Shapes and content can be rendered into different output formats. A version management for the specifications can be realized with versions of the utilized ontology and versioning in the Git Repo.

5. Claim creation and IDS Info Model

The Information Model is an RDFS/OWL-ontology covering the fundamental concepts of the International Data Spaces (IDS), i.e. the types of *digital contents* that are exchanged by *participants* by means of the IDS *infrastructure components*. The ontology and its documentation are published at https://w3id.org/idsa/core. [6] The IDS Information Model (IDS IM) defines the general concepts depicted in Figure. 2 along with roles required to describe actors, components, roles and interactions in a data space. This ontology serves two purposes, (1) as a catalogue of machine-readable terms and data schema for IDS components and (2) as a shared language for all stakeholders.

To evaluate these claims, participants and components are subject to a certification process – an additional means to establish trust within and across data spaces. Organisational structures, methodologies, and standards underlying that process are detailed in the IDS. Whitepaper Certification[7]. A normative, tamper-proof reference of certification, security and identity attributes is maintained by IDS infrastructure components, operated as part of an Identity Provider (Figure. 2).

6. Conclusions and future work.

IDSA uses privacy-preserving, efficient and self-sovereign identity approach in a federate identity management (FIM) solution based on DAPS and IDS Information Model. Today's federated identity management (FIM) systems have some weaknesses, as they are putting Identity Providers (IdP) at the centre of the identity ecosystem. This architectural constraint makes them a great target to attack and dependence of a trust intermediary. To avoid such security threats, IdPs have to follow heavy security controls and apply different security mechanisms to fulfil their requirements with privacy laws. In that context, different security experts came up with an alternative solution, where decentralized Verifiable Credentials (VC) would be an alternative and would shift the burden of securing privacy data from those IdPs to its real owner.

[6] https://www.researchgate.net/publication/346501057_The_International_Data_Spaces_Information_Model_-_An_Ontology_for_Sovereign_Exchange_of_Digital_Content

[7] IDSA: Whitepaper certification. Technical report, IDSA (2018), https://www.internationaldataspaces.org/publications/whitepaper-certification/