TOWARDS USABLE ICDD CONTAINERS FOR ONTOLOGY-DRIVEN DATA LINKING AND LINK VALIDATION

Philipp Hagedorn (RUB), Madhumitha Senthilvel (RWTH), Hans Schevers (BuildingBits),

Lucas B. Verhelst (BIM-Connected)

Building Bits BIM CONNECTED RWTH AACHEN UNIVERSITY RUHR UNIVERSITÄT BOCHUM RUB

Towards usable ICDD containers

We all worked with ICDD and
were connected by **Pieter Pauwels**
after a Master Thesis Defense at TU/e.

Hans
Works together with Lucas

Lucas
Works in Eindhoven

Philipp
PhD Exchange

Madhumitha
PhD Exchange

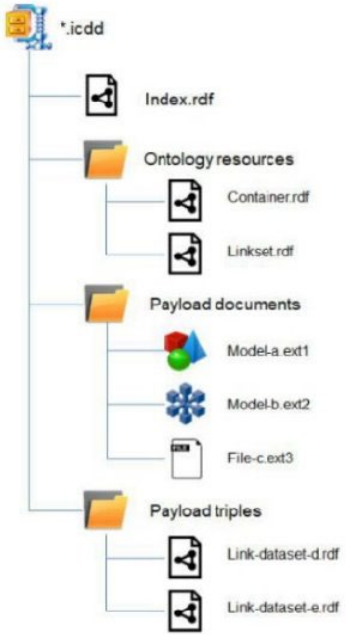
- Discussed ICDD in research and practice
- Recognized issues when ICDD is implemented
- Proceeded to write this (discussion) paper

Towards usable ICDD containers

- **Aim of ICDD**:

  Provide a vendor-neutral structure for exchanging heterogeneous distributed building data

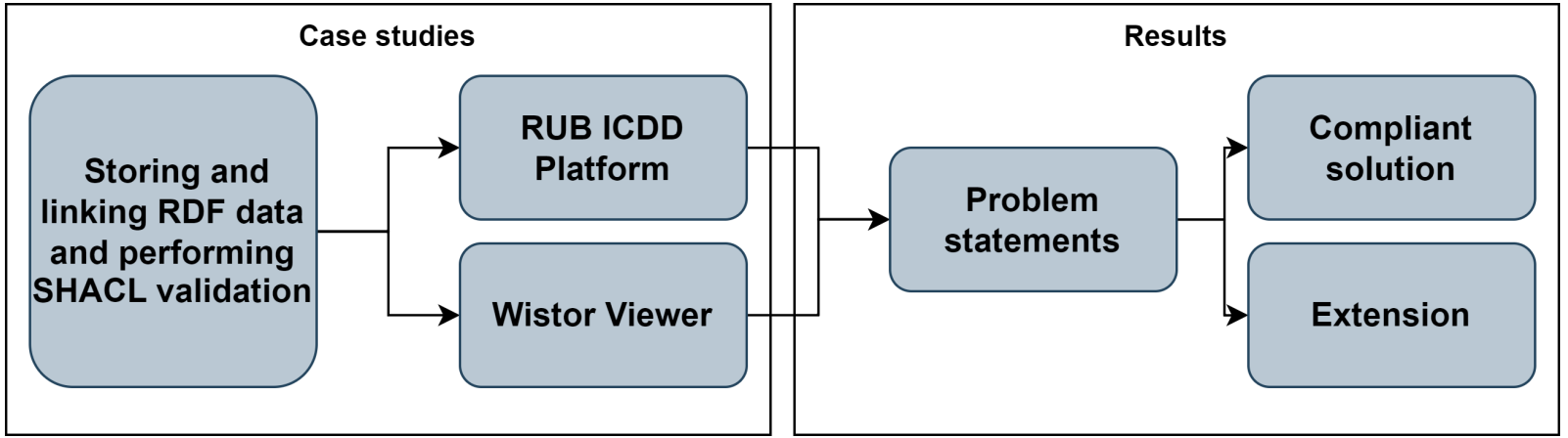
  - Part 1
    - specifies a container with a structure that holds the payload
    - Index.rdf: metadata of contained files
    - Linkset files: (deep) links between these files
  - Part 2: extended linkset
  - Encoded in Semantic Web formats, i.e., RDF and OWL



Source: ISO 21597-1

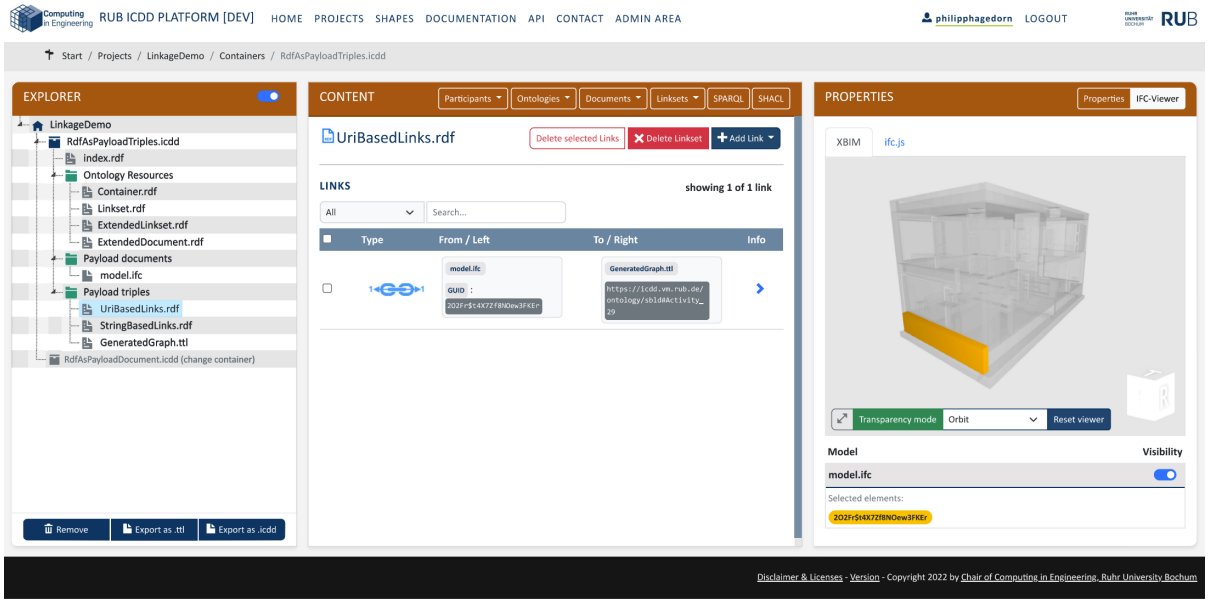
Towards usable ICDD containers

How can we make linking and validation of heterogeneous data in ICDD containers more usable?



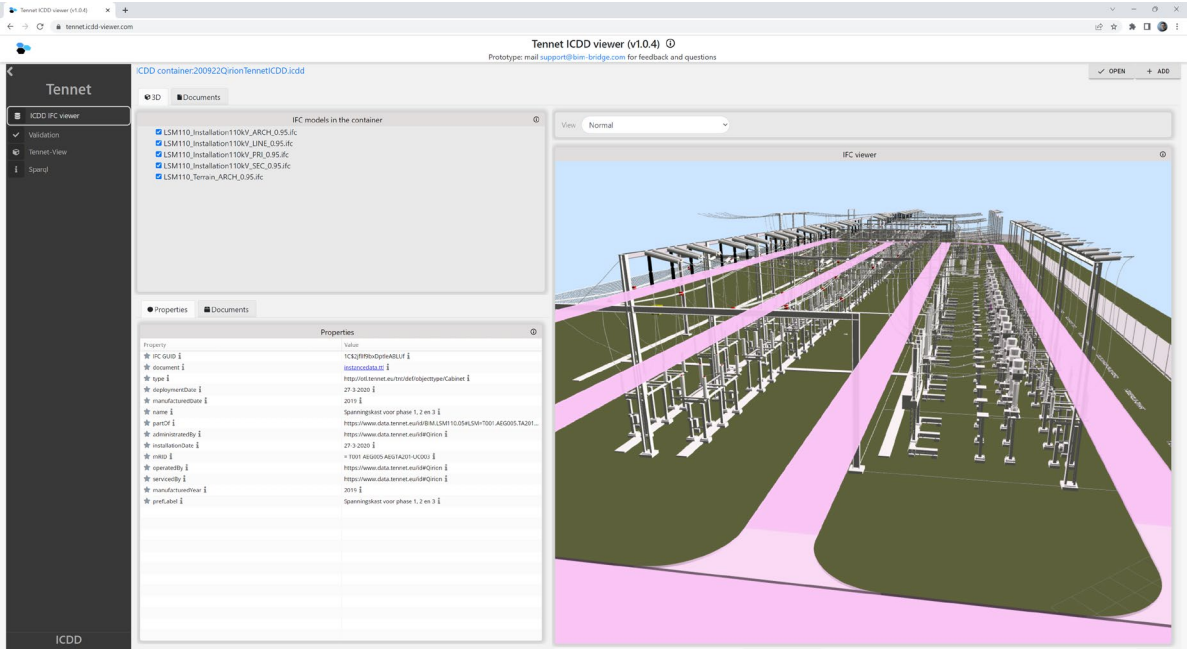
RUB ICDD Platform

- Collaboration platform based on ICDD containers
- Option to store **RDF files** either in payload documents or in payload triples
- Registers **RDF files** from payload triples additionally using an extension of the external document class
- Enables **SHACL exchange requirements** as payload triples or remotely as external documents



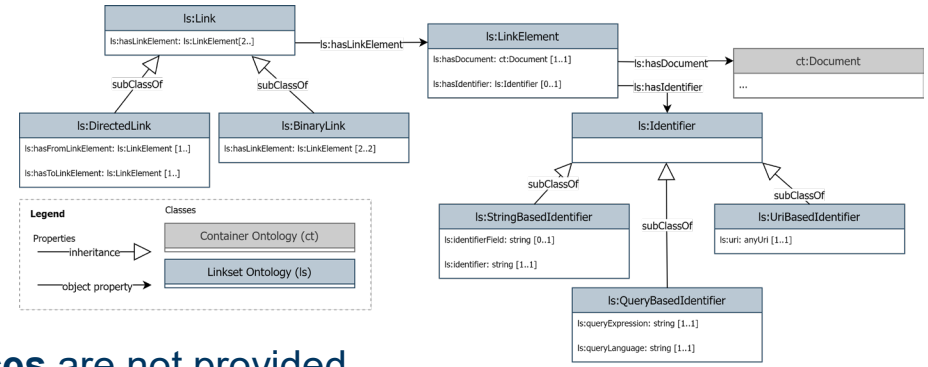
Wistor ICDD Viewer

- Based on **low code platform**
- **Opening, validating and viewing** ICDD containers
- **Validation** of ICDD container, SHACL restrictions and IFC files
- **Viewing validation results** in an integrated way from different perspectives
- **Dashboards** with validation results and statistics about ICDD container



Towards usable ICDD containers

- **Problem statements**
  1. Linking mechanism of ICDD is unnecessarily **cumbersome** for linking RDF data
  2. **Overhead** of the linking structure in ICDD may lead to redundancy
  3. **Storing and/or registering RDF data in ICDD** is not fully resolved in the standard[1]
  4. **SHACL exchange requirements for ICDD containers** only provided for conformance not for content validation
  5. **Extensions** of the ontologies are **not allowed** by the standard
  6. Standard **formats for deep link references** are not provided

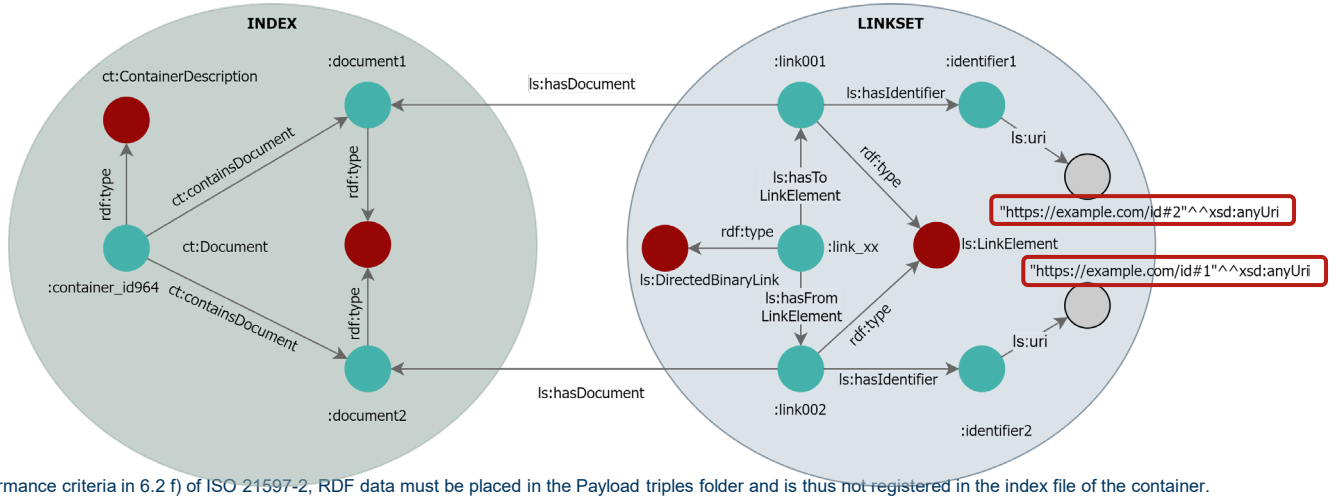


1) According to the conformance criteria in 6.2 f) of ISO 21597-2, RDF data must be placed in the Payload triples folder and is thus not registered in the index file of the container.

Compliant solution (stay within the standard)

Store RDF data and SHACL files in Payload documents folder - without having them treated specially

- Link to RDF data done via `ls:URIBasedIdentifier` as facilitated by current standard
- Enables resolvable links when shared with other stakeholders

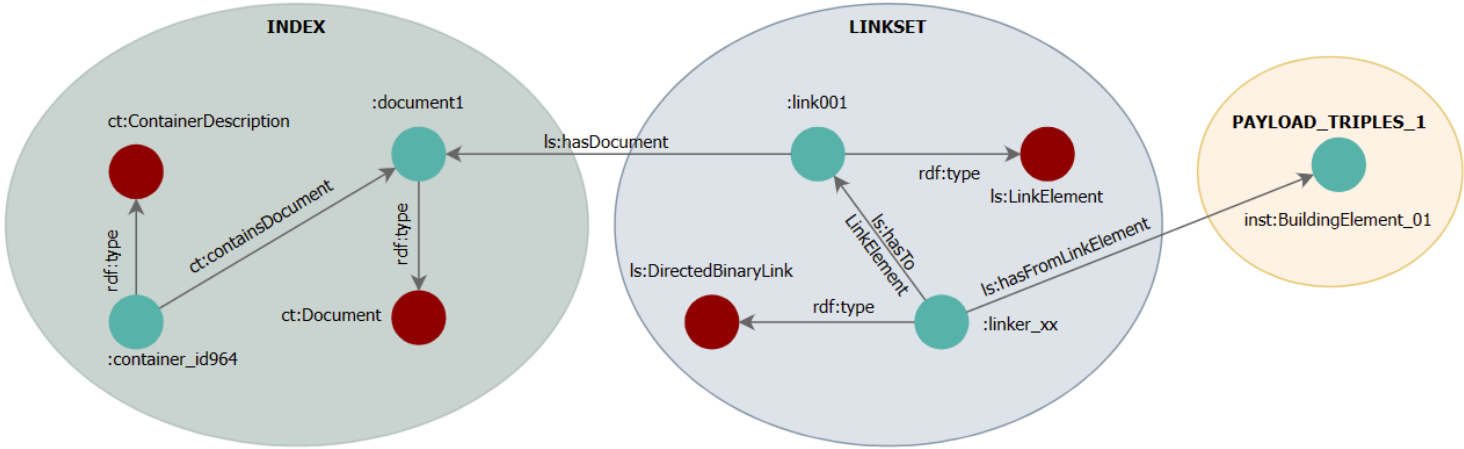


1) According to the conformance criteria in 6.2 f) of ISO 21597-2, RDF data must be placed in the Payload triples folder and is thus not registered in the index file of the container.

Solution 1 for extending the standard

Referring to an RDF entity in the payload triples directly

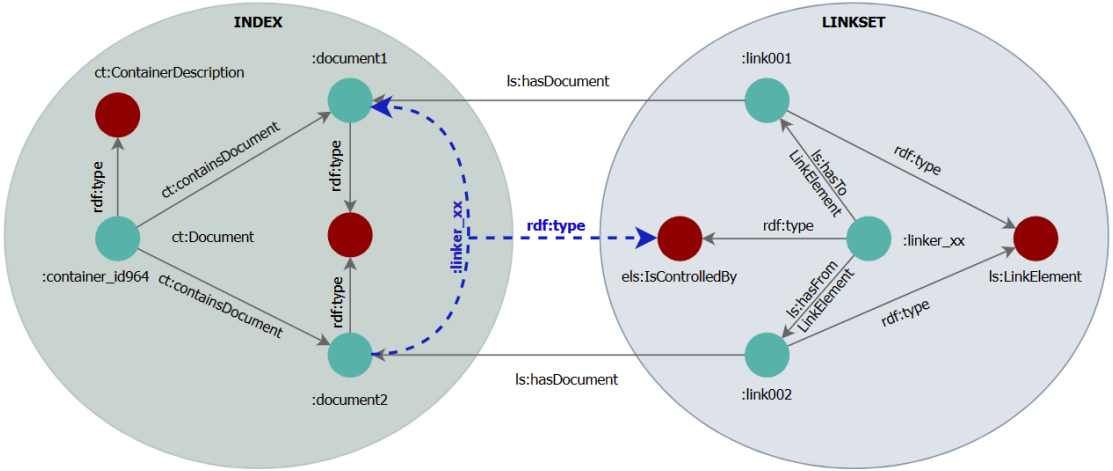
- Referencing the RDF node directly from any `ls:Link` instance
- Possible with weaker restrictions on `ls:hasFromLinkElement`, `ls:hasToLinkElement` and `ls:hasLinkElement` (introduce as a super-property)



Solution 2 for extending the standard

Direct linking mechanism between documents using the extended linkset vocabulary

- Retain original folder-based structure of ICDD and omit mandatory declaration of both `ls:Link` and `ls:LinkElement`
- Unrestricted usage of linktypes – assign a link class as a property value directly – inspired by Noy et al.[2]



2) N. Noy, M. Uschold, C. Welty, Representing Classes As Property Values on the Semantic Web, 2005. URL: https://www.w3.org/TR/2005/NOTE-swbp-classes-as-values-20050405/.

Conclusions and Recommendations

1. **The benefits of LD are not realized** by treating the datasets as normal documents.
2. To realize the benefits of LD in containers, the **extensions of classes and properties of the ICDD ontologies should be allowed**.
3. The storage of internal and external **RDF files should be registered in the index file**.
4. Encoding data requirements as SHACL shapes in the container
   → **registered inside the index file wit (e.g.) hasRequirements property**
5. Supplemental agreements:
   - **Structuring link elements**
   - **Identifiers for specific file types**, e.g., IFC, GIS, or spreadsheets
     → help interpret the links inside ICDD in different systems uniquely

TOWARDS USABLE ICDD CONTAINERS FOR ONTOLOGY-DRIVEN DATA LINKING AND LINK VALIDATION

Philipp Hagedorn (RUB), Madhumitha Senthilvel (RWTH), Hans Schevers (BuildingBits),

Lucas B. Verhelst (BIM-Connected)

Building Bits BIM CONNECTED RWTH AACHEN UNIVERSITY RUHR UNIVERSITÄT BOCHUM RUB

Compliant solution (stay within the standard)

For unique identification of entities in other file types (IFC models, GIS data, or calculation spreadsheets)

- Establish common understanding on the use of `ls:StringBasedIdentifier`

Known Issues – two prototypical implementations show:

- For IFC documents, different key-value pairs for referring to the IFC GUID can be used for linking, e.g., when utilizing different keys: "GUID", "globalID", "IFC GUID".

Hard for software systems to implement ICDD in an interoperable way.