

# **INTEROPERABILITY + SEMANTICS = CHECK!**

## **Smart and Cost Effective Data Modelling and Tools of the Future**

Miika Alonen, CSC – IT Center for Science

Suvi Remes, CSC – IT Center for Science

Patrik Maltusch, Aalto University

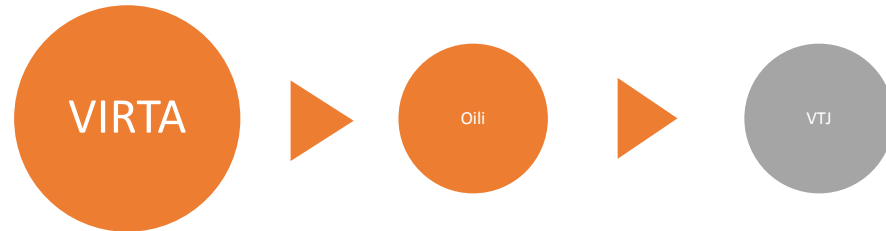
Mikael af Hällström, Finnish Tax Administration

# Summary

- Interoperability challenges
- Interoperability case: Student Transcript
- Framework for Semantic Interoperability
- Implementing the Framework:
  - Interoperability workbench

# Present – Information chaos

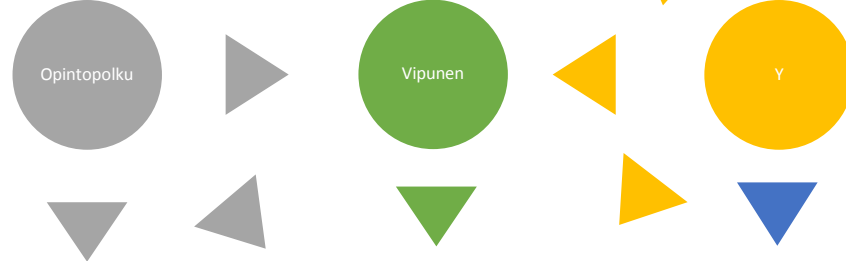
Organisation and application  
specific documentation



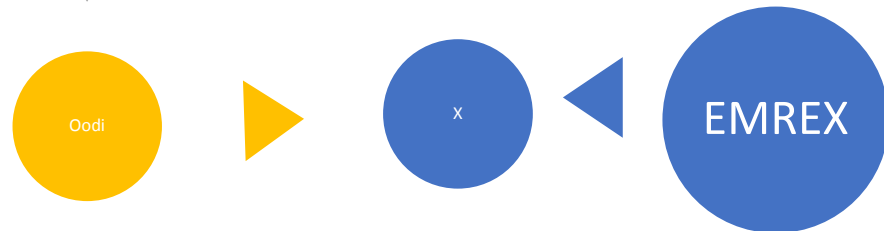
Redefining data models



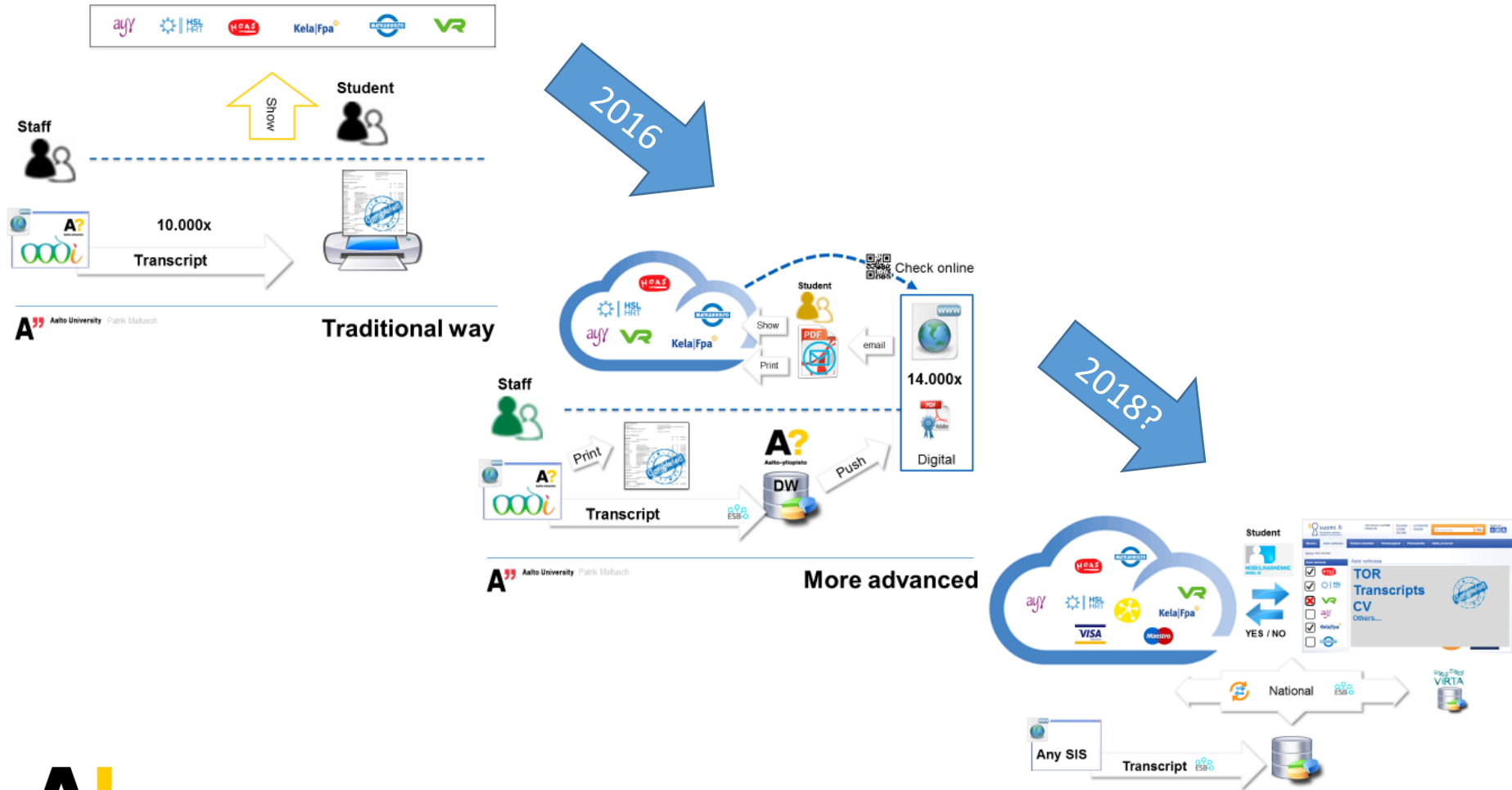
Point to point  
Integration



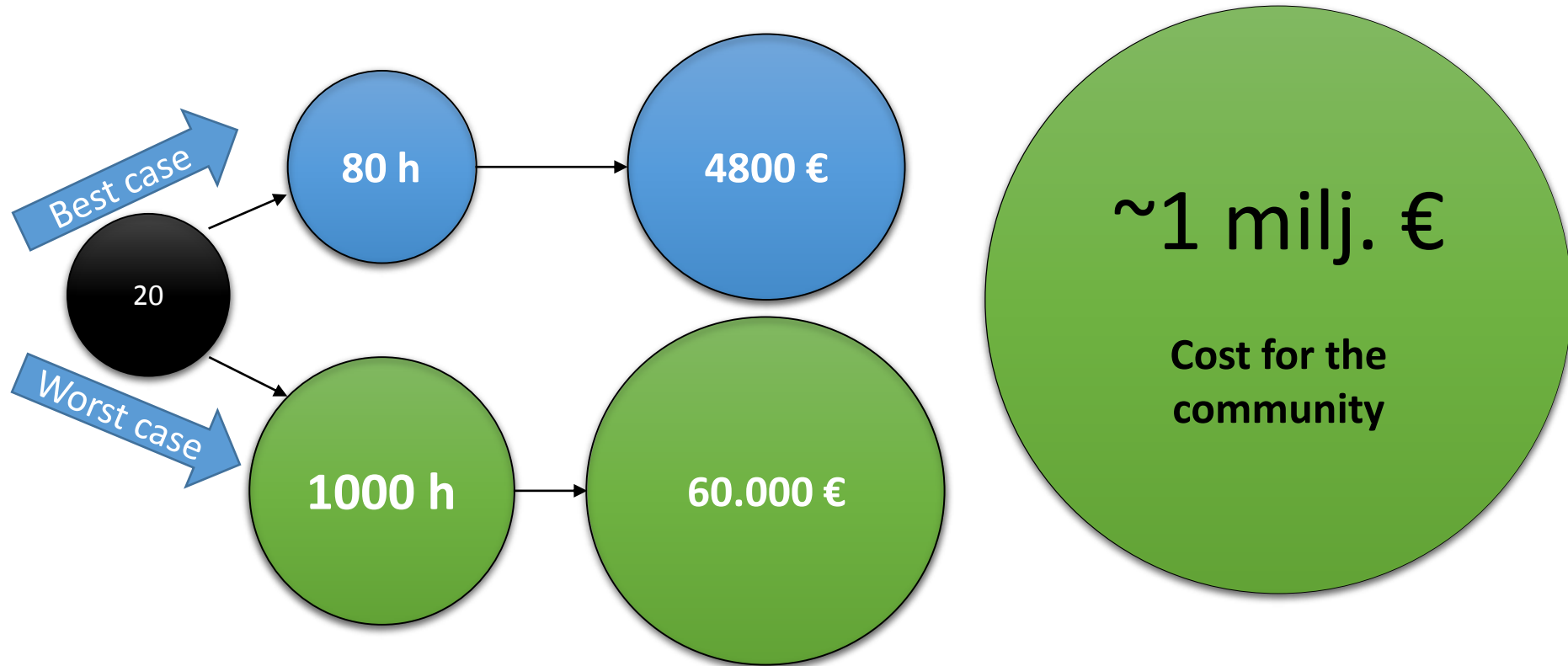
No change management  
over organisation borders



# Interoperability CASE: Student Transcript

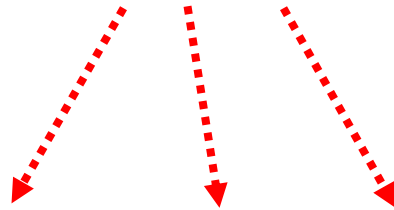


# Making data sources semantic interoperable for a DW



# Equation for Interoperability

SEMANTIC IMPACT

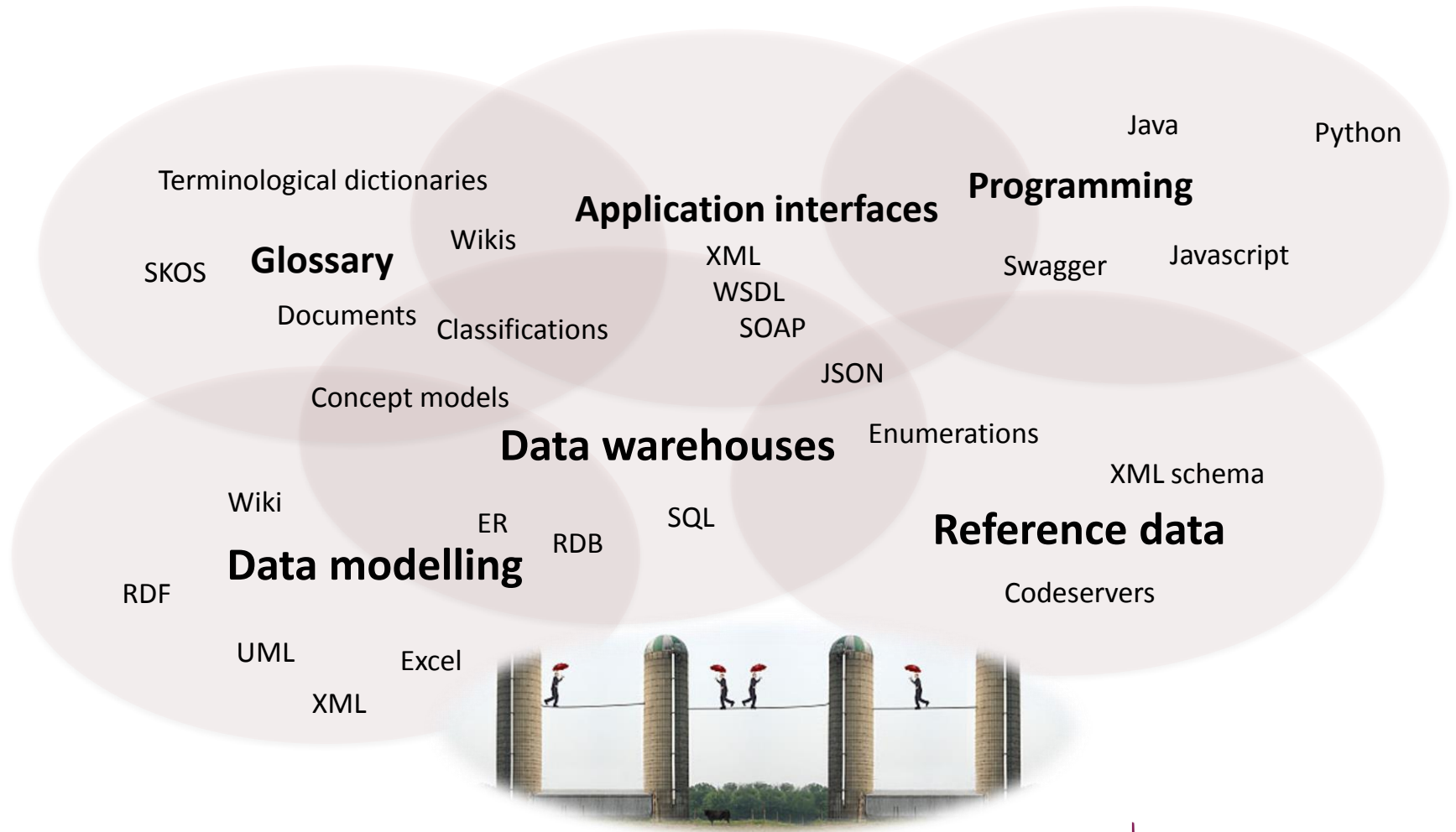


$$ICC * ESB * EIM * DW = \frac{\sum \textit{benefits}}{\sum \textit{cost}} > 1$$

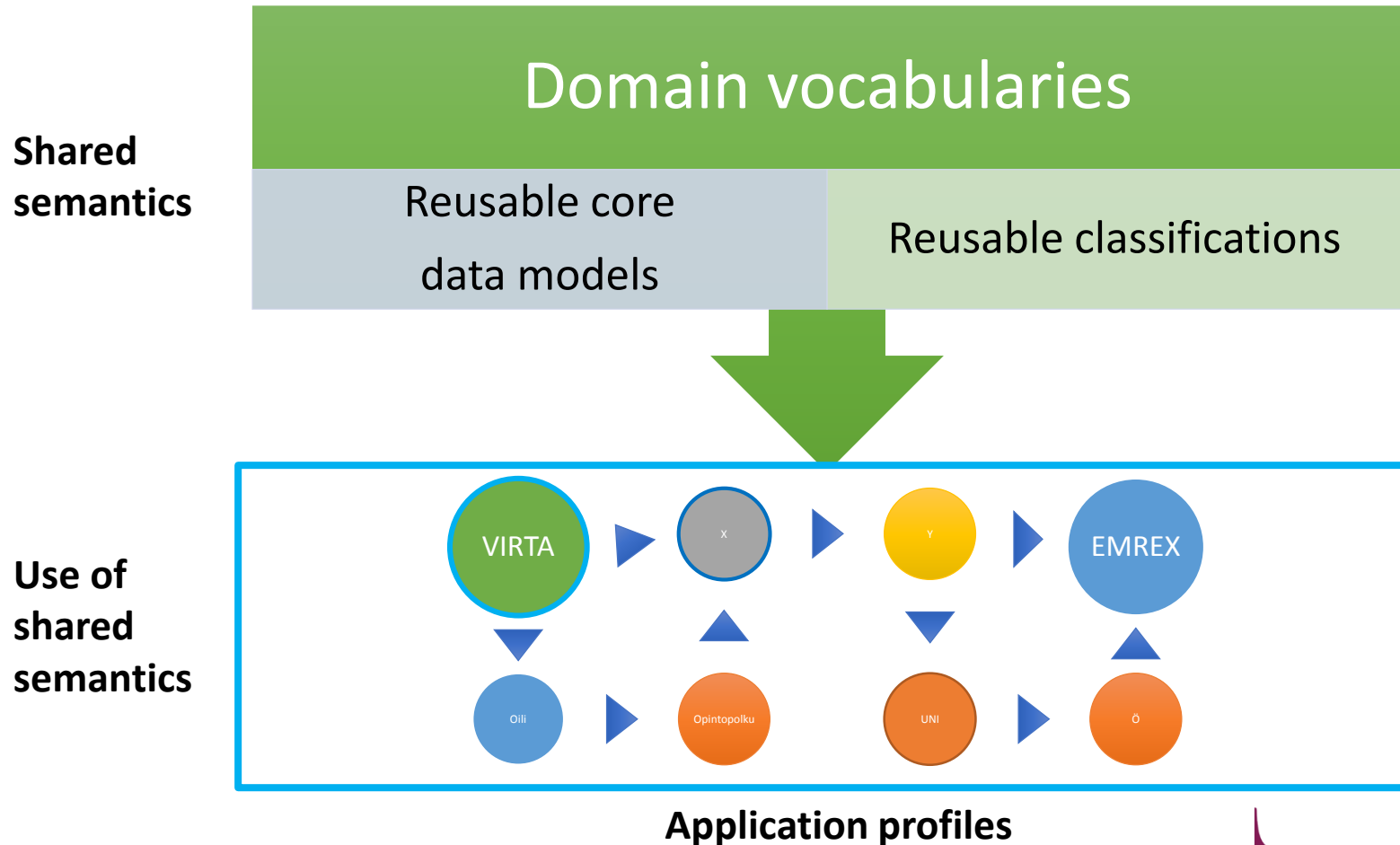
by Patrik Maltusch @aalto.fi



# Used methods for documenting APIs?



# Vision – Data model reuse and uniform documentation





# Application profiles

An Application profile seeks to address the interoperability requirements between systems by:

- retaining conformance with a base standard
- defining new requirements in an open and interoperable manner

## Examples of standard Application profiles:

**CEN/CWA 15903:** Metadata for Learning Opportunities

**CEN/CWA 16132:** European Learner Mobility Achievement Information

**EMREX AP (2016) \***: Report for the recognition of external studies

\* <https://goo.gl/fJ0021>

**2000** - Idea of Application Profiles evolved from DESIRE Registry project: “mixing and matching” metadata elements

**2003** - Guidelines for creating application profiles (CEN/CWA 14855)

**2005** – Guidelines for machine-readable representation .. (CEN/CWA 15248)

**2006** - Guidelines building application profiles in e-learning (CEN/CWA 15555)

**2008** - Guidelines for Dublin Core Application Profiles

**2011 - 2016** – MLR Framework (ISO/IEC 19788-1)

# Need for common framework

## Shared concepts with the business and IT:

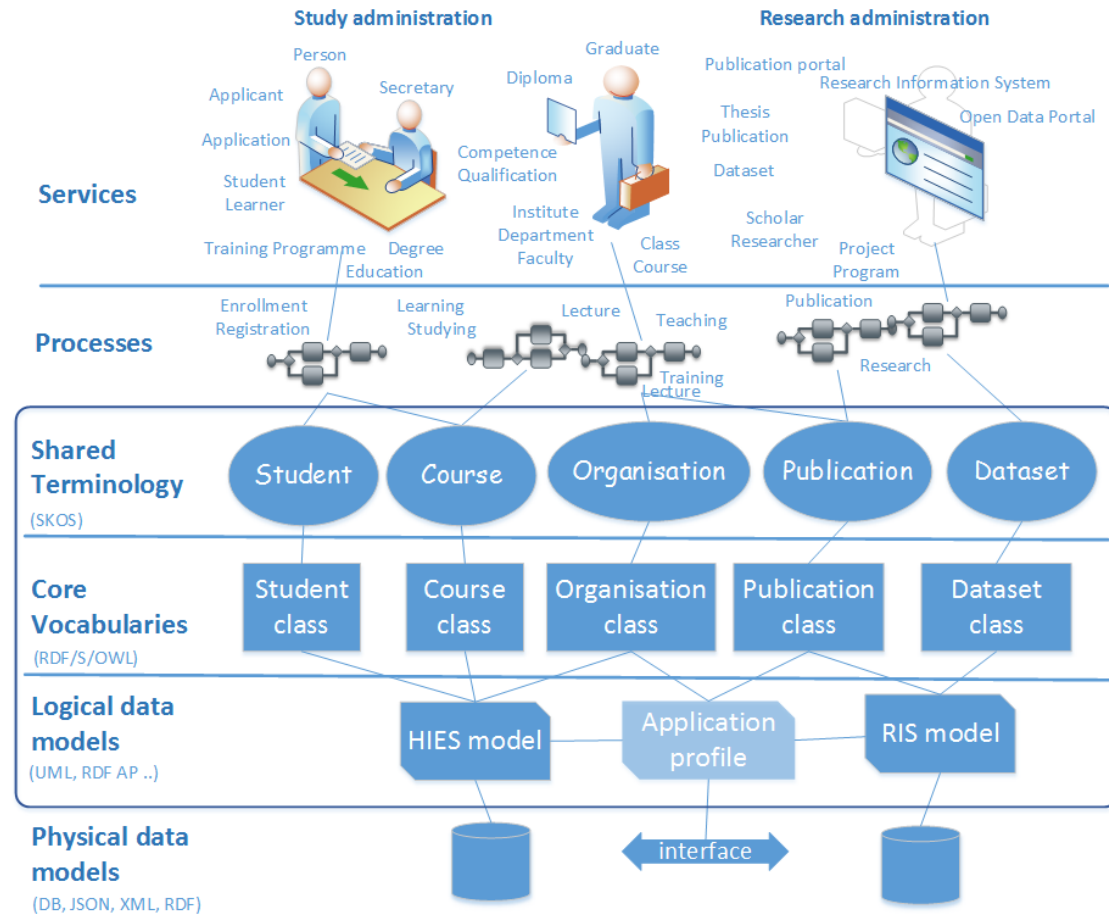
1. Well defined concepts
2. Unique identifiers
3. Machine readable format

## Service innovation and data modeling based on business needs:

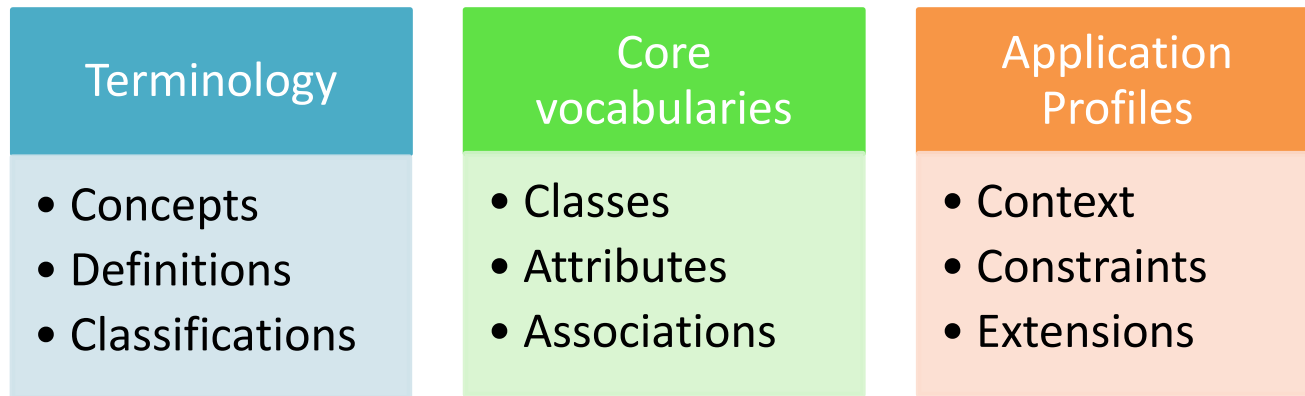
1. Reuse terms and definitions
2. Create reusable components
3. Focus on the interfaces and integration

## Framework for semantic interoperability:

- How to publish core vocabularies and application profiles?
- How to reuse standards?
- How to reuse core vocabularies in the implementations?
- How to document the metadata reuse?
- How to document application interfaces?



# Semantic interoperability framework



## Common metadata architecture for data modelling:

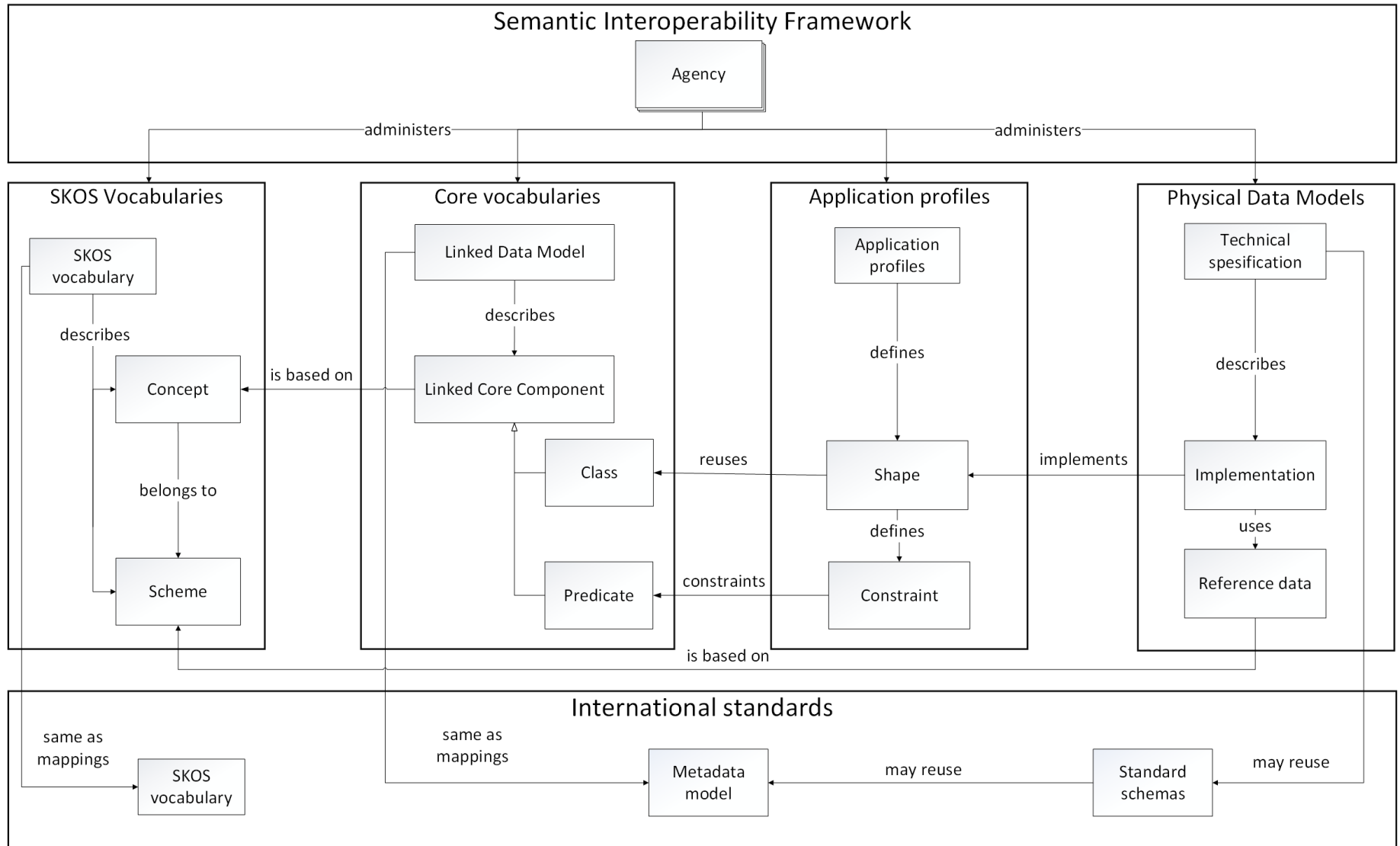
- Machine readable terminology
- Reusable core vocabularies and core components
- Documented reuse of core components

# Interoperability objectives

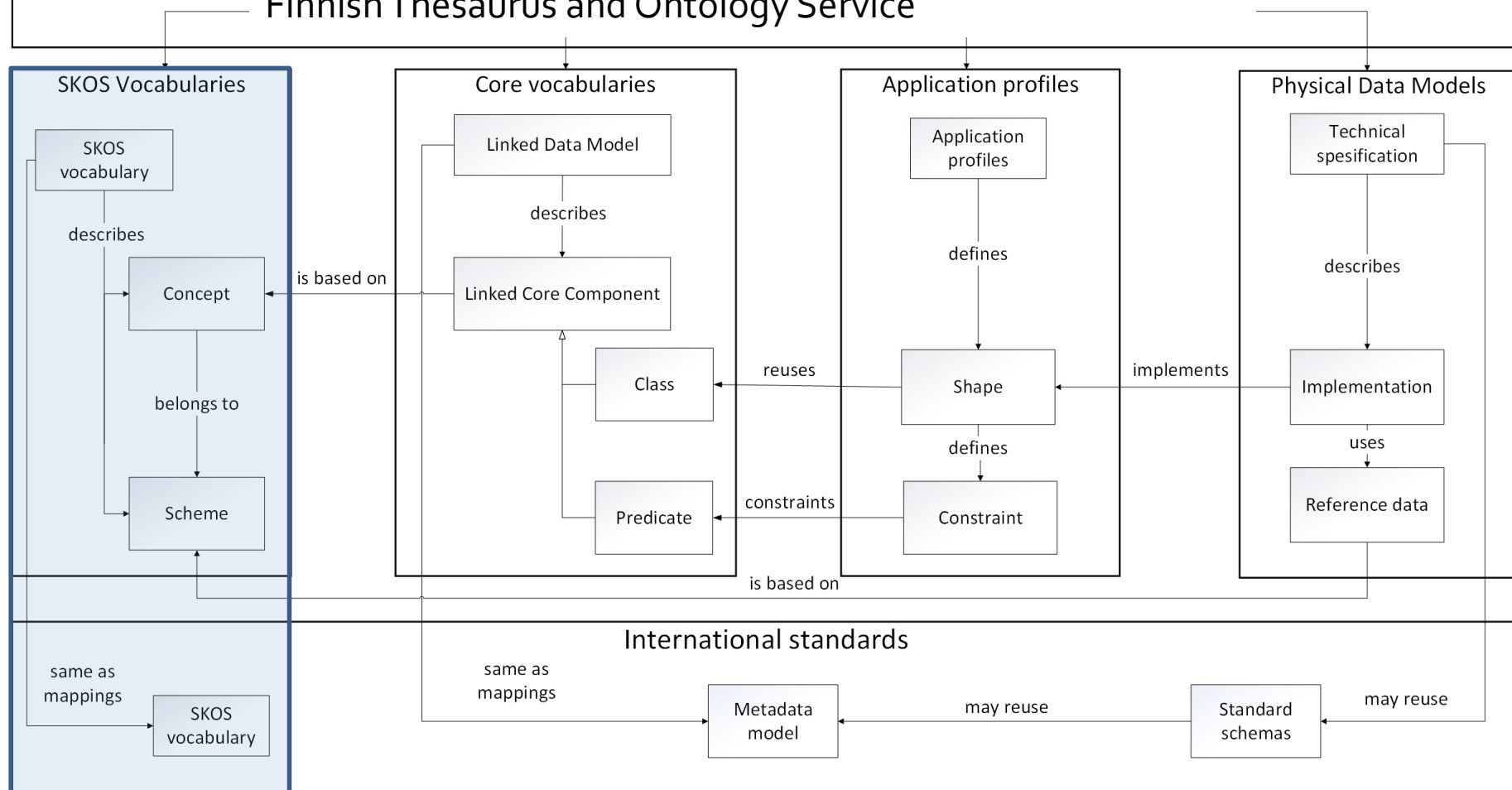
- Shared metadata models for education and research
- Modular and reusable metadata definitions
- Improve readability and understandability of data models
- Interoperability with international standards
- Promote standard reuse
- Formal and semantic mappings to CEN and ISO standards

# Interoperability benefits

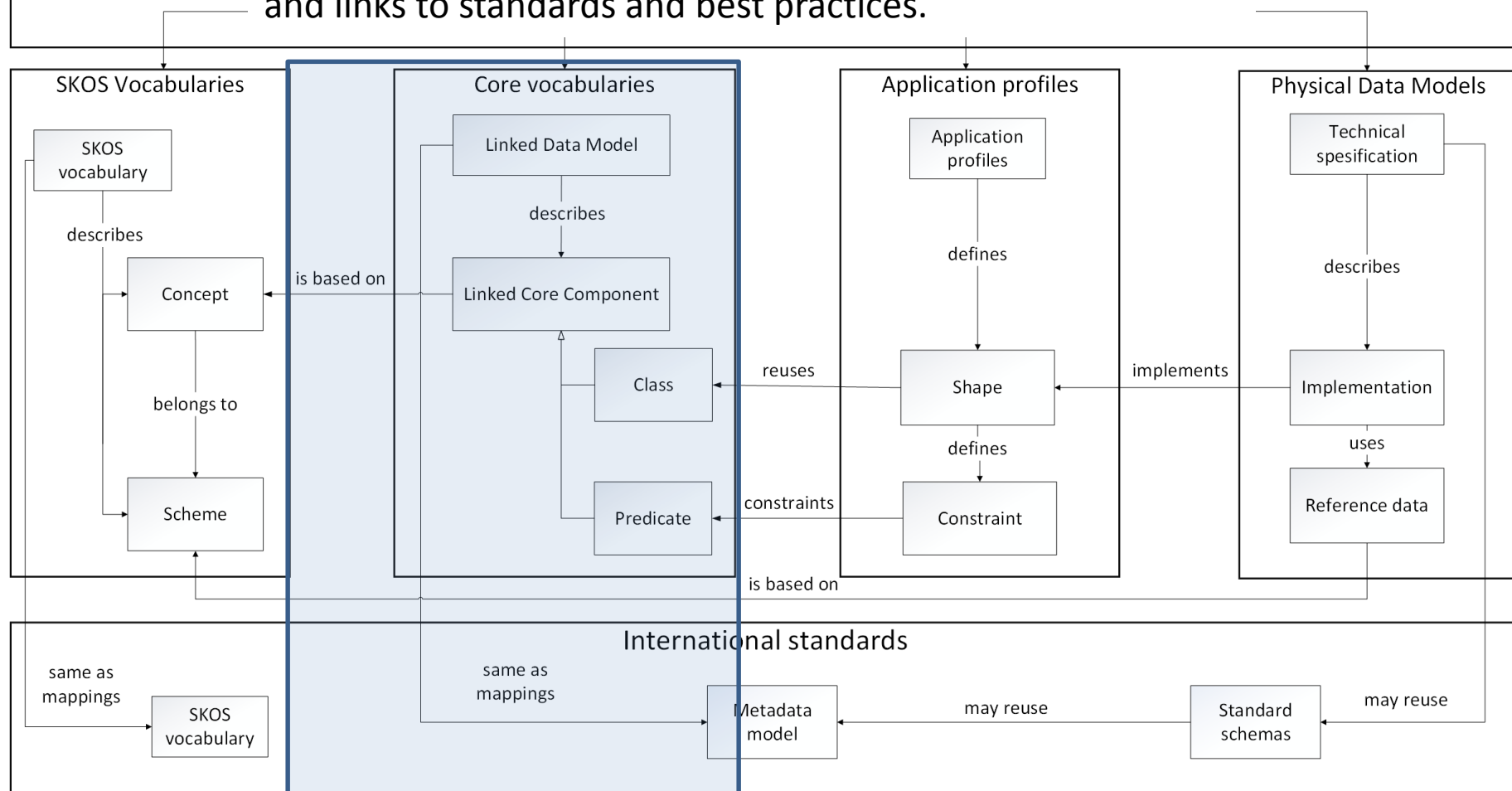
- Collaborative tool for metadata publishing
  - Benefit from external expertise
- Avoid redefinition of data models
  - Lower integration costs
- Shared terminology
  - Less confusion
- Same terms for many communications needs
  - Study administration
  - User interfaces



Domain specific vocabularies are administered by different agencies and published in the standard SKOS model in the Finnish Thesaurus and Ontology Service

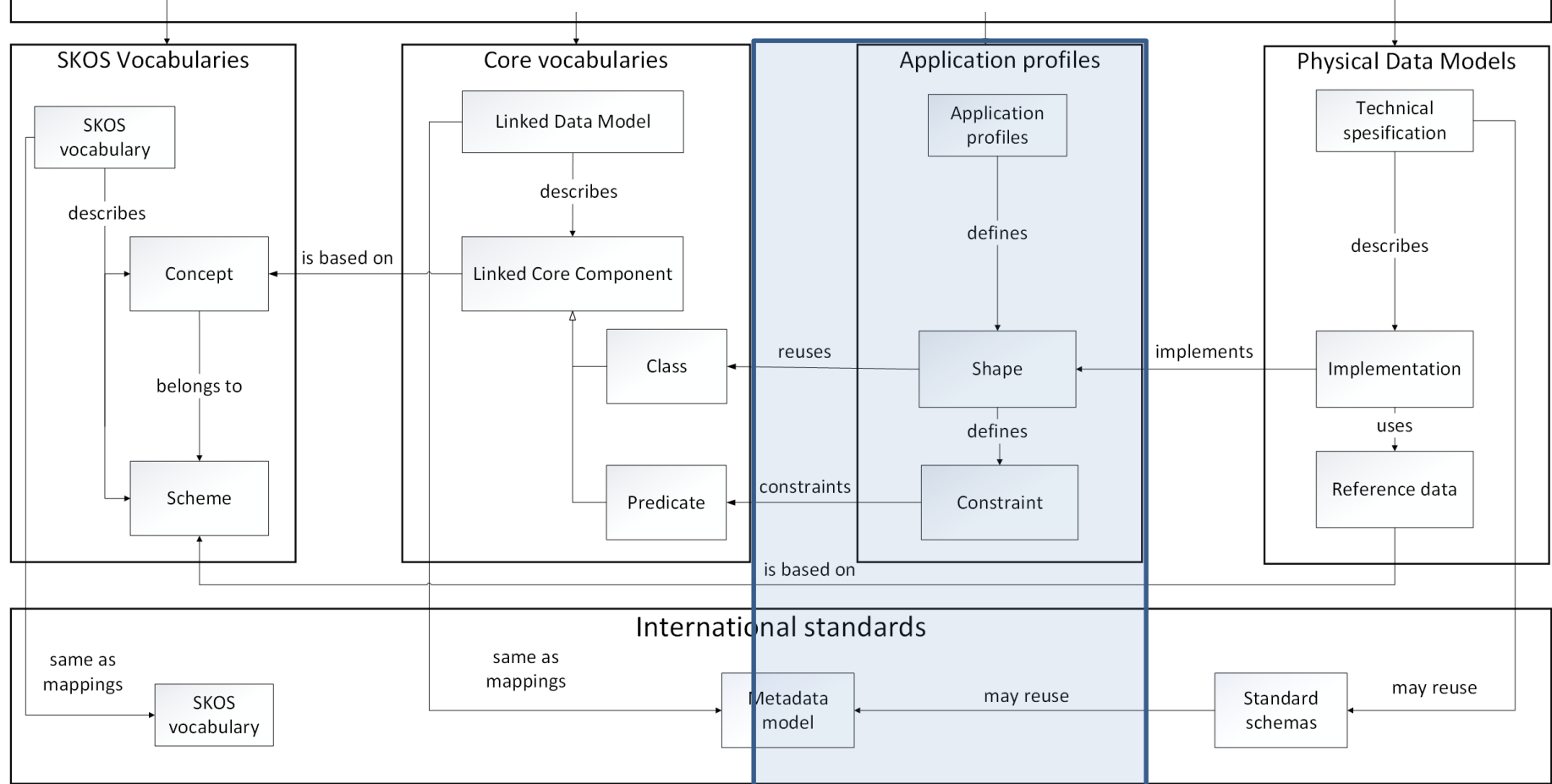


Core vocabularies are published as Linked Data that defines re-usable classes and properties based on shared concepts and links to standards and best practices.

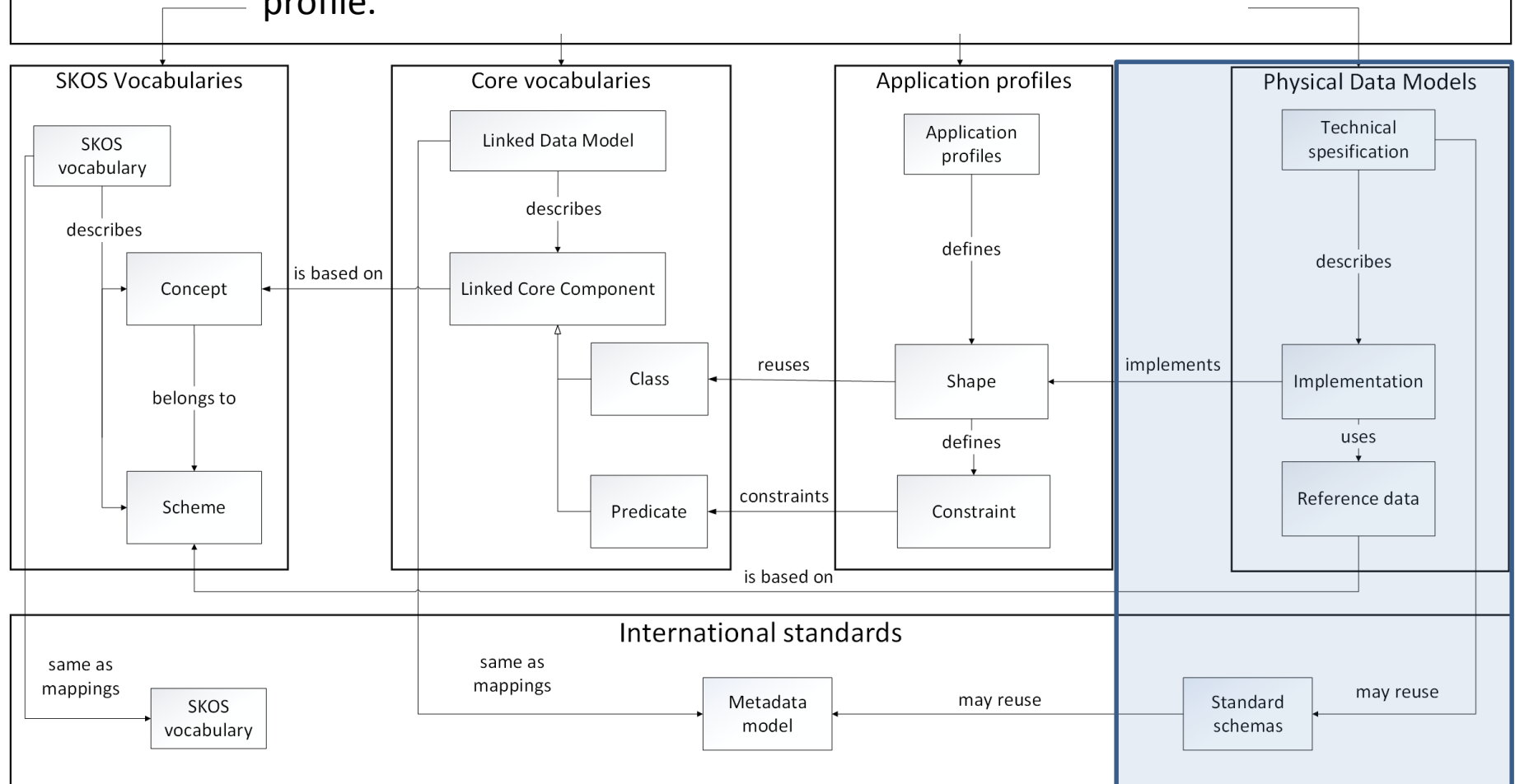




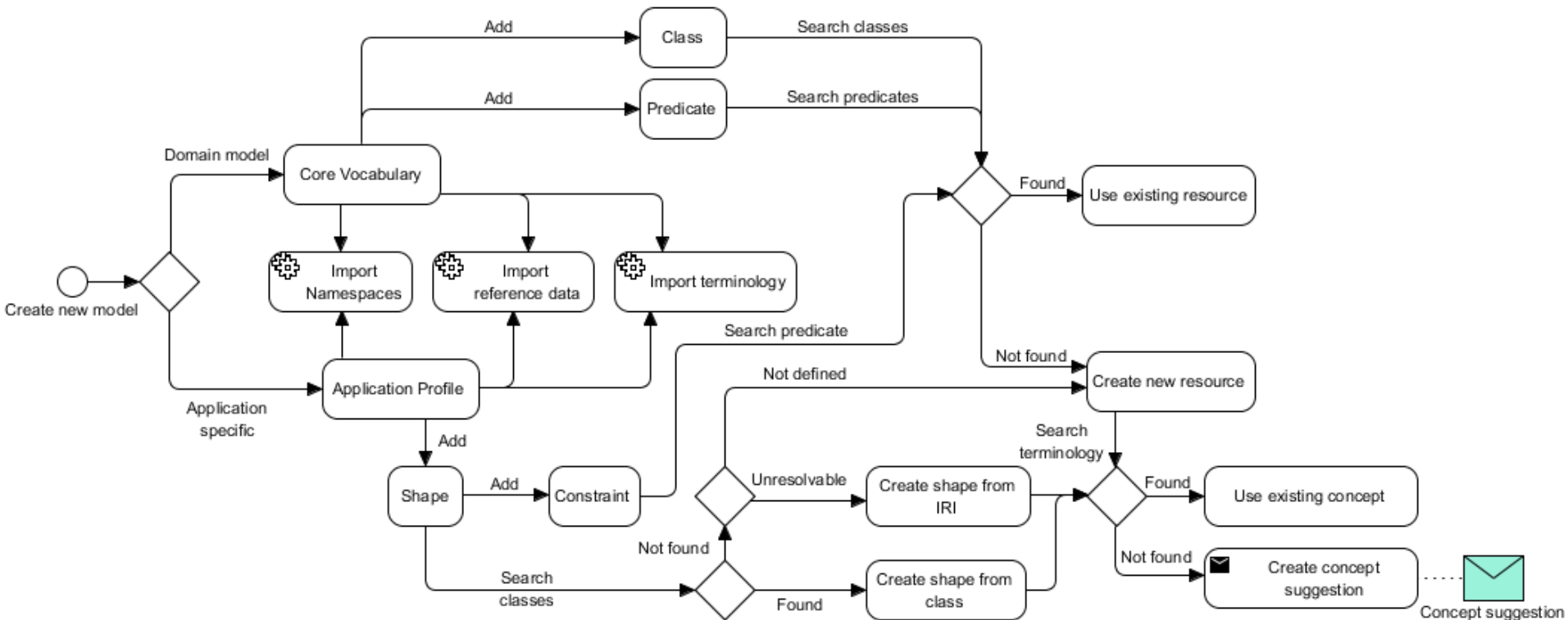
Domain specific data models and interfaces are documented as Application profiles that re-use the classes and properties from the Core Vocabularies



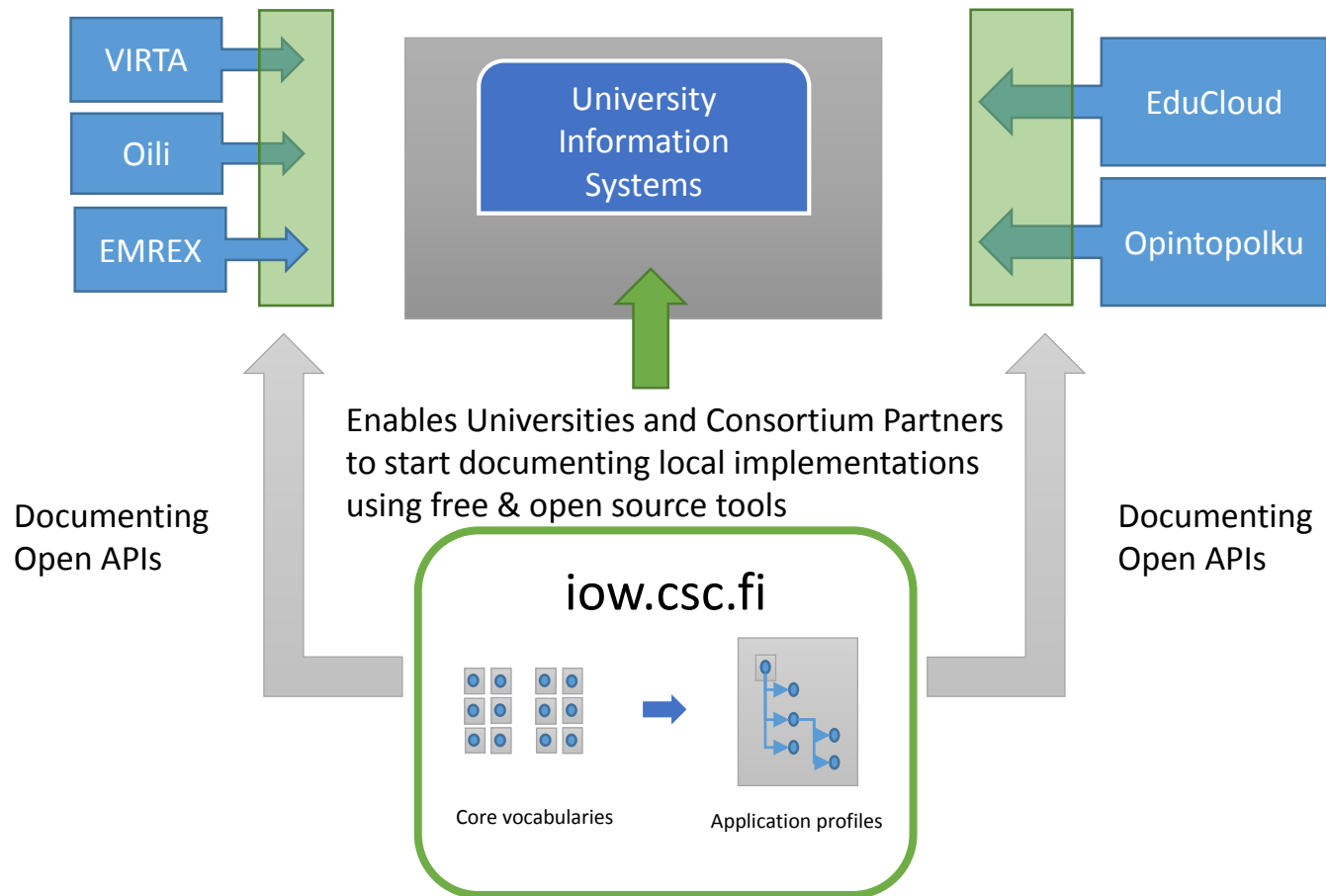
Data models are implemented with languages that best suit the given architecture by implementing the application profile.



# Simplified process for describing Core vocabularies and Application profiles



# Piloting Interoperability Framework



# Interoperability workbench

- Collaborative online tool for creating Core Vocabularies and Application Profiles:

The screenshot shows the Interoperability workbench website. The header is blue with the text 'Interoperability descriptions' and links for 'Suomeksi', 'In English', and 'Login'. The main content area is white. On the left, there is a sidebar with a 'Interoperability method' section containing three paragraphs of text. The first paragraph asks 'What are the descriptions for interoperability?' and explains that interoperability descriptions are a set of highly reusable metadata. The second paragraph asks 'What is the interoperability method?' and describes it as a technique to describe metadata in a technology-independent way. The third paragraph asks 'What can the interoperability workbench do?' and states that it is a set of tools for developing, documenting, and publishing interoperability descriptions. On the right, there is a 'Search' section with a search bar containing the word 'Studen' and a magnifying glass icon. Below the search bar, there is an 'Advanced search' section with four results: 'Electronic Mobility Report', 'Learner', 'Learning Opportunity Instance', and 'Report'. Each result has a small icon and a brief description. Below the search section, there is a 'Browse' section with two links: 'Common vocabularies' and 'Learning, Education and Culture'. At the bottom of the page, there is a footer with contact information for CSC - IT Center for Science Ltd and a version string: 'Version 17.05.2016 12:22 Source code: Frontend, Backend, Database licensed under the European Union Public Licence'.

Interoperability descriptions

Suomeksi In English Login

FRONT PAGE

## Interoperability method

What are the descriptions for interoperability?

Interoperability descriptions are set of highly reusable metadata that enables harmonization of data models and creation of interoperable services. The more we rely on digitalisation in our transaction and businesses, the more we need to ensure that both cross-human and machine-to-machine communication is accurate and able to preserve the meaning of the data – without constant assistance in the interpretation. Interoperability descriptions consist of concepts defined according to the terminological theory as the controlled terminology, reusable information components based on the shared concept definitions and application profiles specifying the use of the information components.

What is the interoperability method?

Interoperability Method is a technique to describe the metadata of information systems in a technology independent way. It forms a structured, common architecture connecting terminological work and data modelling and reuse of codelists and classifications. The method ensures that shared definitions are applied in a systematic way and the semantics is passed to every implementation that re-uses the information components. The method is based on international standards. Please read more [here](#) about the Interoperability Method and the development of tools supporting it.

What can the interoperability workbench do?

Interoperability Workbench is a set of tools designed to be used in developing, documenting and publishing interoperability descriptions such as re-usable information components and application profiles. It offers a one-stop service for openly published data descriptions and vocabularies. These semantically

## Search

Studen

Advanced search

- [Electronic Mobility Report](#)  
Exchange document containing multiple student achievement reports
- [Learner](#)  
Student engaged in learning process, who is also part of the exchange
- [Learning Opportunity Instance](#)  
LO offered to students restricted to time, place, pace and form
- [Report](#)  
Description of the institution and subset of courses completed by the student

## Browse

- [Common vocabularies](#)
- [Learning, Education and Culture](#)

Contact info CSC - IT Center for Science Ltd

Version 17.05.2016 12:22 Source code: Frontend, Backend, Database licensed under the European Union Public Licence

# Interoperability workbench

- Tool for defining resolvable and machine readable data models
- Document the use of data models, standards and best practices

The screenshot displays the 'Interoperability descriptions' web application. The top navigation bar includes 'Suomeksi', 'In English', and 'Login'. The breadcrumb trail shows 'FRONT PAGE / LEARNING, EDUCATION AND CULTURE / ELECTRONIC MOBILITY REPORT / ELECTRONIC MOBILITY REPORT'. The main heading is 'Electronic Mobility Report' with a right-pointing arrow.

On the left, a sidebar lists various categories: Class, Attribute, and Association. Under 'Class', several items are listed, with 'Electronic Mobility Report' selected. The main content area is titled 'Electronic Mobility Report' and contains the following information:

- Data in english** (dropdown menu)
- Show history** and **Export** buttons
- Class information**
  - Class label**: Electronic Mobility Report
  - Class id**: emrex:Elmo
  - Description**: Exchange document containing multiple student achievement reports
  - Defined by profile**: Electronic Mobility Report
  - Status**: Unstable
  - Concept suggestion**
    - Concept id**: urn:uuid:da3845e9-e48a-41d9-a0d4-425efdd12bae
    - Concept label**: Electronic Mobility Report (fi)
    - Definition**: Exchange document containing multiple student achievement reports (fi)
- Class properties**
  - Report** (indicated by a right-pointing arrow)
  - Generation date** (indicated by a right-pointing arrow)

On the right, the 'Visualization' panel shows a graph of relationships between classes. The graph includes nodes for 'Electronic Mobility Report', 'Report', 'Learner', 'Teacher', 'Learning Opportunity Specification', 'Learning Opportunity Instance', 'Report Distribution', and 'Shortened Grading Table'. Relationships are indicated by arrows with labels such as 'Report', 'Learner', 'Teacher', 'Learning Opportunity Specification', 'Learning Opportunity Instance', 'Report Distribution', and 'Shortened Grading Table'.

# Integration to controlled vocabularies

- Link controlled vocabularies to created model
- Create classes and properties based on existing concepts



## Controlled vocabularies

Browse concepts

+ Add vocabulary

| Identifier | Vocabulary name                    |  |
|------------|------------------------------------|--|
| eos        | Elinikäisen oppimisen sanasto (fi) |  |
| jhsmeta    | JHMeta (fi)                        |  |

## Reference data

None added

## Imported namespaces

| Prefix | Namespace label                      |
|--------|--------------------------------------|
| adms   | adms                                 |
| cm     | Educational Credit Information Model |

+ Add reference data

Määrittele käsite uudelle luokalle

Hae ensin tarvitsemaasi käsitettä sanastoista. Jos et löydä tarpeisiisi sopivaa, voit luoda uuden käsite-ehdotuksen.

RAJAA HAKUTULOKSIA

Sanasto

Kaikki sanastot

+ ehdota 'opisk' sanastoon

opiskelija  
Elinikäisen oppimisen sanasto

opiskelija  
JHMeta

opiskelijaksi ilmoittautuminen  
Elinikäisen oppimisen sanasto

opiskeluvoimavarojen tila  
JHMeta

opiskeluvoimavarojen  
JHMeta

**Käsitteen nimi**  
opiskelija

**Käsitteen tunniste**  
<http://low.csc.fi/skos/eos#tmpOKSAID227>

**Määritelmä**  
henkilö, joka opiskelee muun kuin esi- tai perusopetuksen piirissä

Peruuta Luo uusi luokka

# Integration to classification schemes

- Link to existing reference data from code service
- Restrict allowed values by using existing reference data

ja maiden englanninkielisiin nimiin. Valtioiden ja maiden nimistä käytetään ns. lyhyttä muotoa (vrt. Suomi, Suomen tasavalta). Suomenkieliset nimet perustuvat SFS-ISO 3166 -standardiin sekä Kotimaisten kielten tutkimuskeskuksen ohjeisiin. Ruotsinkieliset nimet perustuvat Kotimaisten kielten tutkimuskeskuksen ohjeisiin ja englanninkieliset ISO 3166-standardiin sekä Kotimaisten kielten tutkimuskeskuksen ohjeisiin.

Kaikki koodistot, Alueet

Yhteisten tutkinnon osien osa-alueet

Yhteisten tutkinnon osien osa-alueet

Koulutuskoodistot

Yhteystietojen alkuperä

Yhteystietojen alkuperätiedot

Alueet

Yhteystietotyytit

Koodistolla luokitellaan henkilöiden yhteystietoja.

| Code | Name          |
|------|---------------|
| 548  | Vanuatu       |
| 554  | Uusi-Seelanti |
| 558  | Nicaragua     |

Data in english

Shape properties

Image

status

firstName

Citizenship (fi)

| Class property label | Description      |
|----------------------|------------------|
| Citizenship (fi)     | Citizenship (fi) |

| Range      | Status   |
|------------|----------|
| xsd:string | Unstable |

Optional metadata

Class property id

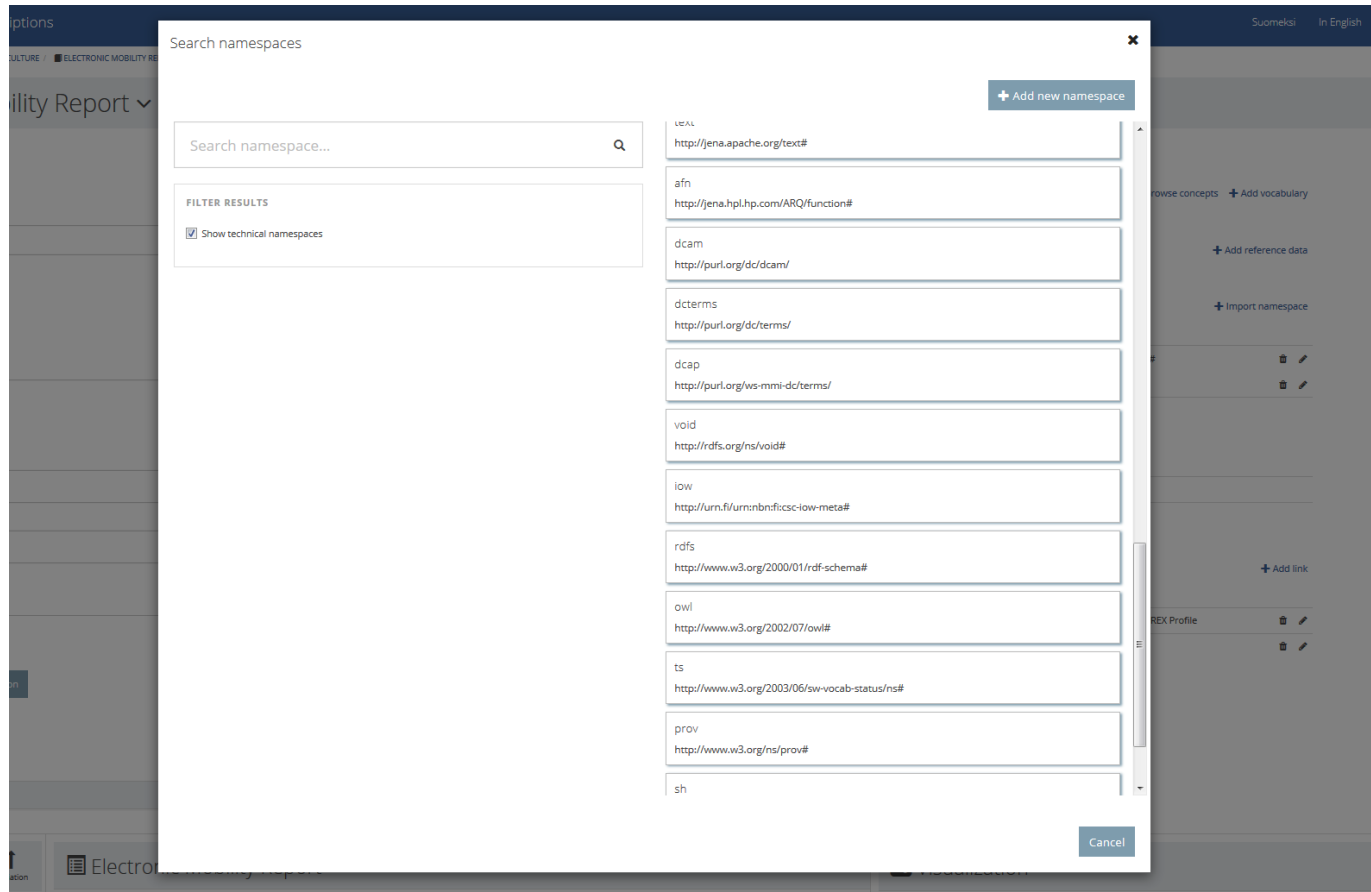
citizenship

| Reference data name    | Description                                                                                                                                                                                                                         |
|------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Valtiot ja maat 2 (fi) | Valtiot ja maat -luokituksella käytetään mm. väestötilastoissa henkilöiden kansallisuuden luokittamiseen. Valtioiden ja maiden tunnukset perustuvat kansainväliseen ISO 3166 -standardiin (International Standard ISO 3166-1. Codes |



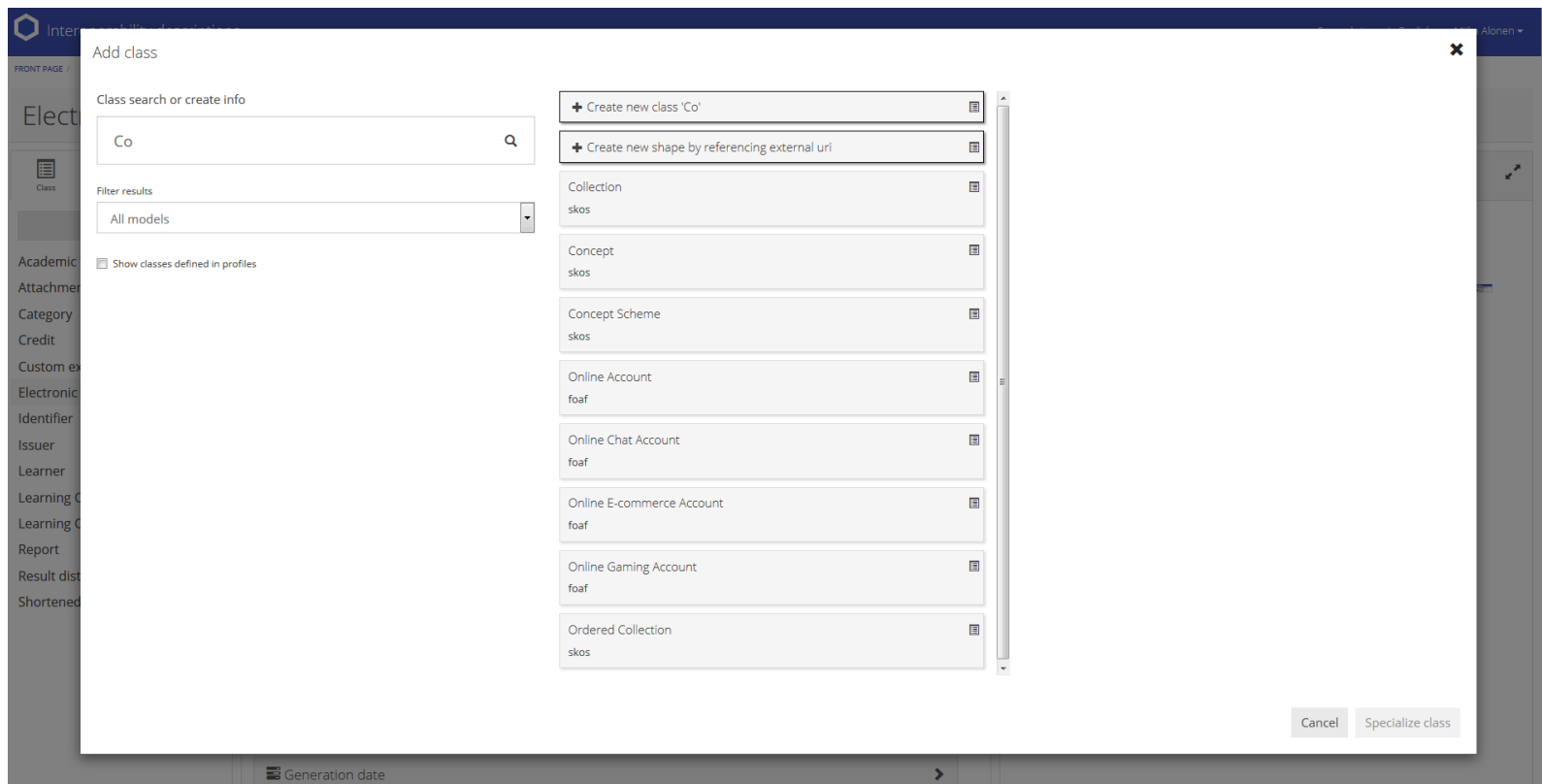
# Interoperability workbench

- Import existing models from local models and external namespaces



# Interoperability workbench

- Include and search metadata from imported standards
- Create new domain models as highly reusable metadata



# Interoperability workbench

- Map new classes and class usage to relevant standards

The screenshot displays the 'Interoperability workbench' interface. On the left is a sidebar with a list of classes: Academic Term, Attachment, Category, Custom extension container, Electronic Mobility Report (highlighted), Identifier, Issuer, Learner, Learning Opportunity Instance, Learning Opportunity Specification, Report, Result distribution, and Shortened grading table. Above this list are buttons for 'Class', 'Attribute', and 'Association', along with an '+ Add class' button.

The main area is titled 'Class information' and contains several fields and sections:

- Data in english**: A dropdown menu showing 'Electronic Mobility Report'.
- Class id**: A text field containing 'Elmo'.
- Superclass**: A text field with a placeholder 'Write identifier...' and a 'Choose class' button below it.
- Equivalent class**: A text field with a placeholder 'Write identifier...' and a list of classes below it. The list includes 'skos:Collection', 'skos:Concept', 'skos:ConceptScheme', 'skos:OrderedCollection', 'adms:Asset' (highlighted), 'adms:AssetDistribution', 'adms:AssetRepository', 'adms:Identifier', 'foaf:Agent', 'foaf:Document', and 'Generation date'.
- Description**: A text area containing 'Exchange document containing multiple student achievement reports'.
- Status**: A dropdown menu showing 'Keskenärsäinen (Unstable)'.
- Definition**: A text area containing 'Report (fi) Exchange document containing multiple student achievement reports (fi)'.
- Class property label**: A text field containing 'Generation date'.
- Range**: A dropdown menu showing 'Ajanhetki (xsd:dateTime)'.

At the bottom right, there is a 'Description' section with a text area containing 'The datetime when the file was generated' and a 'Status' dropdown menu.

# Interoperability workbench

- Export schemas in multiple formats
- Enforces Naming practices

[illegible]

# RDF

```

    "items": [
      {
        "type": "string",
      },
    ],
    "required": [
      "label",
      "count"
    ],
  },
  "Email": {
    "title": "Electronic Mobility Report&en",
    "description": "Exchange document containing multiple student achievement report&en",
    "properties": {
      "attachment": {
        "title": "Attachment&en",
        "description": "Attachments to the report&en",
        "type": "array",
        "items": [
          {
            "type": "object",
            "$ref": "#/definitions/sDocument"
          }
        ],
      },
      "date": {
        "title": "Generation date&en",
        "description": "The datetime when the file was generated&en",
        "type": "string",
        "format": "date-time"
      },
    },
    "readOnly": {
      "title": "Learner&en",
      "description": "A student whose achievements are described in the report&en",
      "type": "object",
      "$ref": "#/definitions/ExternalLearner"
    },
  },
  "report": {
    "title": "Report&en",
    "description": "Student achievement reports associated with the Learner&en",
    "minItems": 1,
    "type": "array",
    "items": [
      {
        "type": "object",
        "$ref": "#/definitions/sReport"
      }
    ],
  },
  "signature": {
    "title": "Signature&en",
    "description": "Digital signature, xmdsig-core2, for demonstrating the authenticity of EXSEX document&en",
    "type": "object",
    "$ref": "#/definitions/Signature"
  },
  "required": [
    "date",
    "report",
    "signature",
    "date",
    "report",
    "signature",
    "date",
    "report",
    "signature",
    "date",
    "report"
  ],
}

```

# JSON Schema

## JSON Schema

## XML Schema

tbd ...

# Thanks!

- Questions
- Interoperability  
workbench
- <http://iow.csc.fi>