Findable Accessible Interoperable Reusable

DIGITAL INFRASTRUCTURES for RESEARCH 2017
Serving the user base

B2SHARE-FAIR
Implementing B2SHARE as a FAIR Data Point

João Moreira[1,2], Thijs Cobben[3]
[1] Services, Cyber-security and Safety group (SCS), University of Twente, Netherlands.
[2] Dutch Techcentre for Life Sciences (DTL), Netherlands. j.luizrebelomoreira@utwente.nl
[3] SURFsara BV, Nethelands. thijs.cobben@surfsara.nl

Design problem: improve B2SHARE interoperability with the FAIR Data Point (FDP)

Goal

Enable B2SHARE to behave as a FDP: make B2SHARE compliant with the FDP specification

Requirements

FAIR DATA POINT

B2SHARE

EUDAT

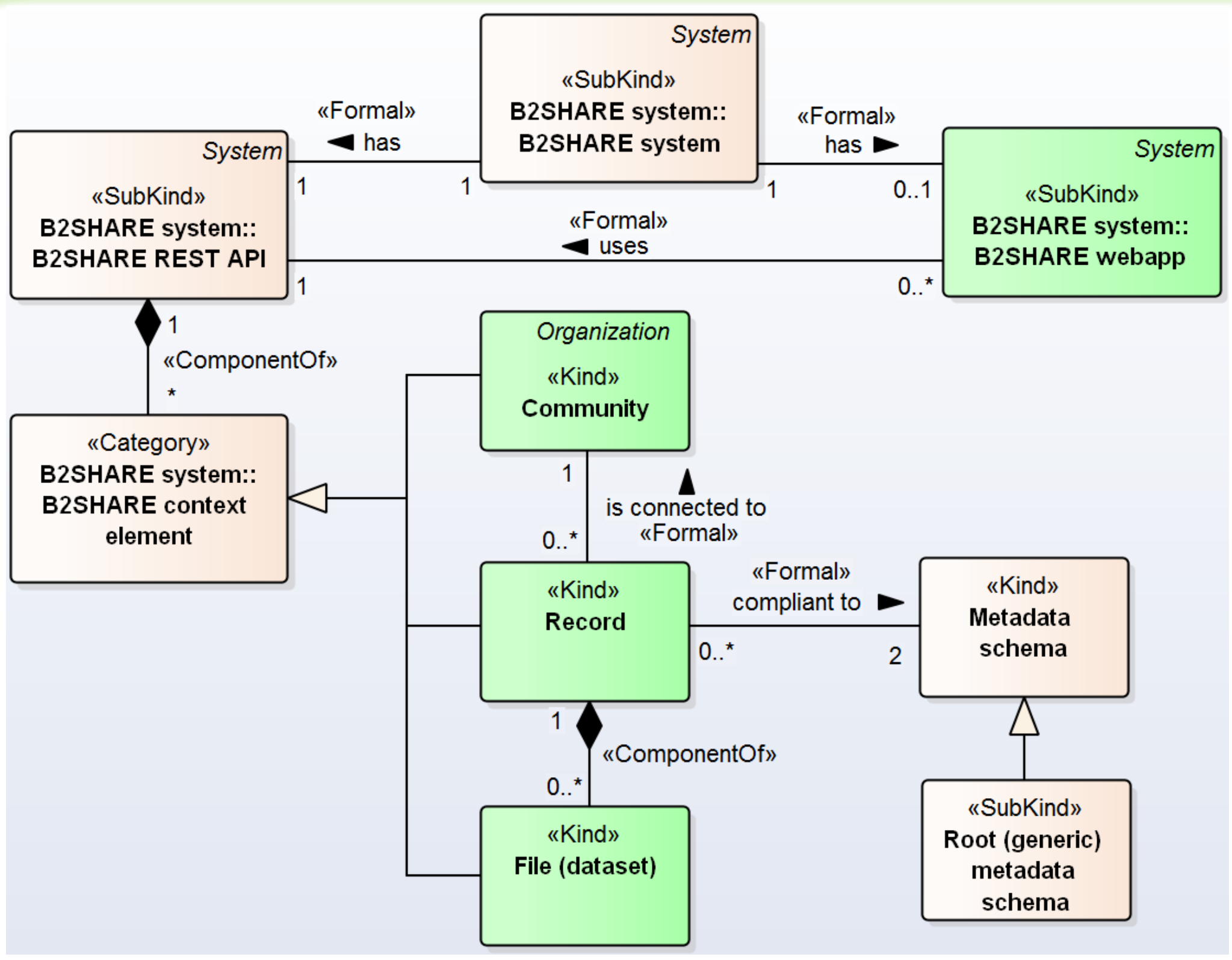
Extend B2SHARE's REST/JSON API to align with FDP metadata layers [1]:

L1. Metadata about the **data repository**: information about the FDP itself and its governing authority, based on the *RE3Data Schema* [2].

L2. Metadata about the **catalog/groups of datasets**: information about the collection of datasets, based on *W3C's Data Catalog Vocabulary* (DCAT) [3].

L3. Metadata about the **datasets**: information about an individual dataset in the collection, available for access or download in one or more formats, based on DCAT.

L4. Metadata about the dataset's **distributions**: information about concrete serialization for the datasets (e.g., XML, CSV, Excel, relational database), i.e. accessible form of a dataset, based on DCAT.



B2SHARE REST API

For interacting with B2SHARE via external services or applications. GET methods:

- List all communities: `/api/communities`
- Get community schema: `/api/communities/$COMMUNITY_ID/schemas/last`
- List all records: `/api/records`
- List records per community: `/api/records/?q=community:COMMUNITY_ID`
- Search records: `/api/records/?q=$QUERY_STRING`
- Search drafts: `/api/records/?drafts`
- Get a specific record: `/api/record/RECORD_ID`
- List the files uploaded into a record object: `/api/files/FILE_BUCKET_ID`

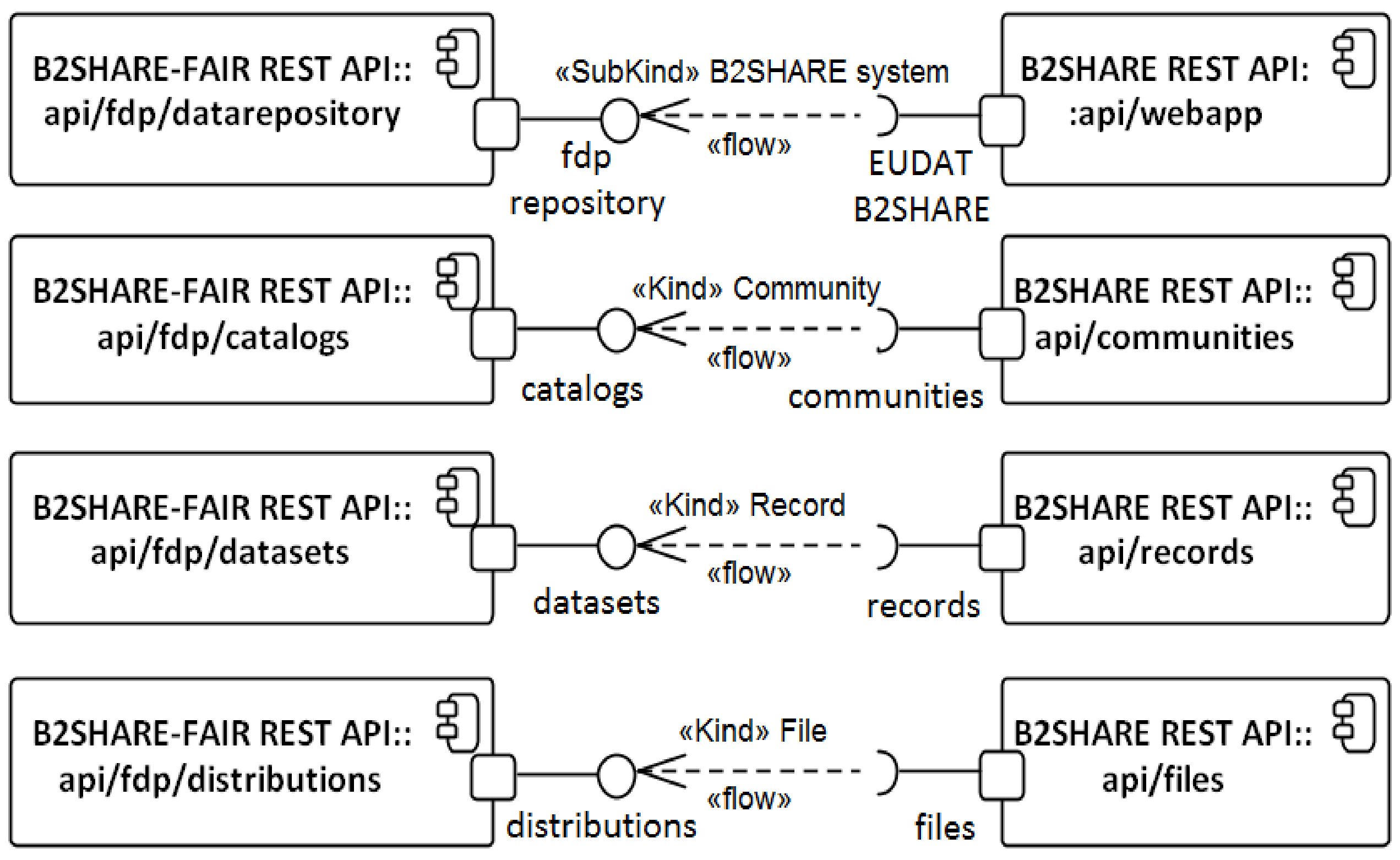
Solution: B2SHARE-FAIR

REST/JSON-LD (JSON for Linked Data) API as a proxy with metadata translations between B2SHARE and FDP

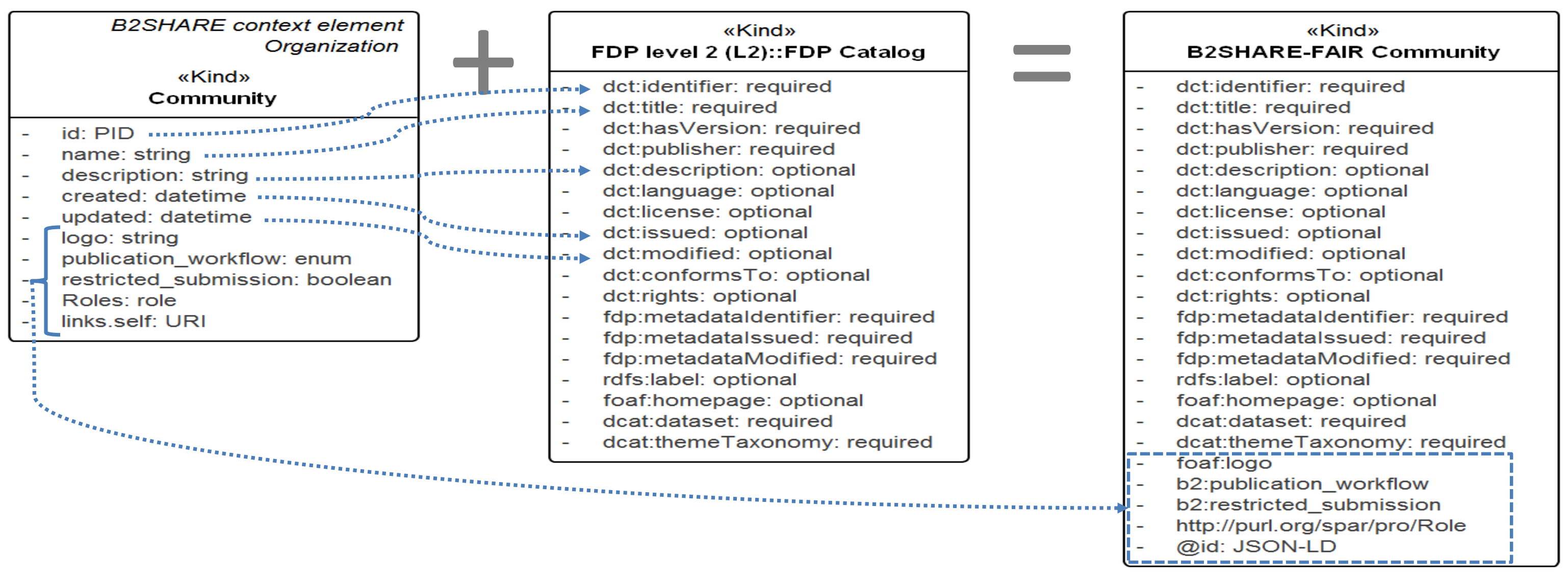
1- High-level metadata mappings

FDP metadata level	B2SHARE entity
L1: Data repository	EUDAT/B2SHARE website
L2: Group (catalog)	Community
L3: Dataset	Record
L4: Dataset distribution	File

2- Proxy requests on B2SHARE REST API



3- Enrich semantics with model-driven transformations at runtime



B2SHARE context element *Organization* «Kind» Community

- id: PID
- name: string
- description: string
- created: datetime
- updated: datetime
- logo: string
- publication_workflow: enum
- restricted_submission: boolean
- Roles: role
- links.self: URI

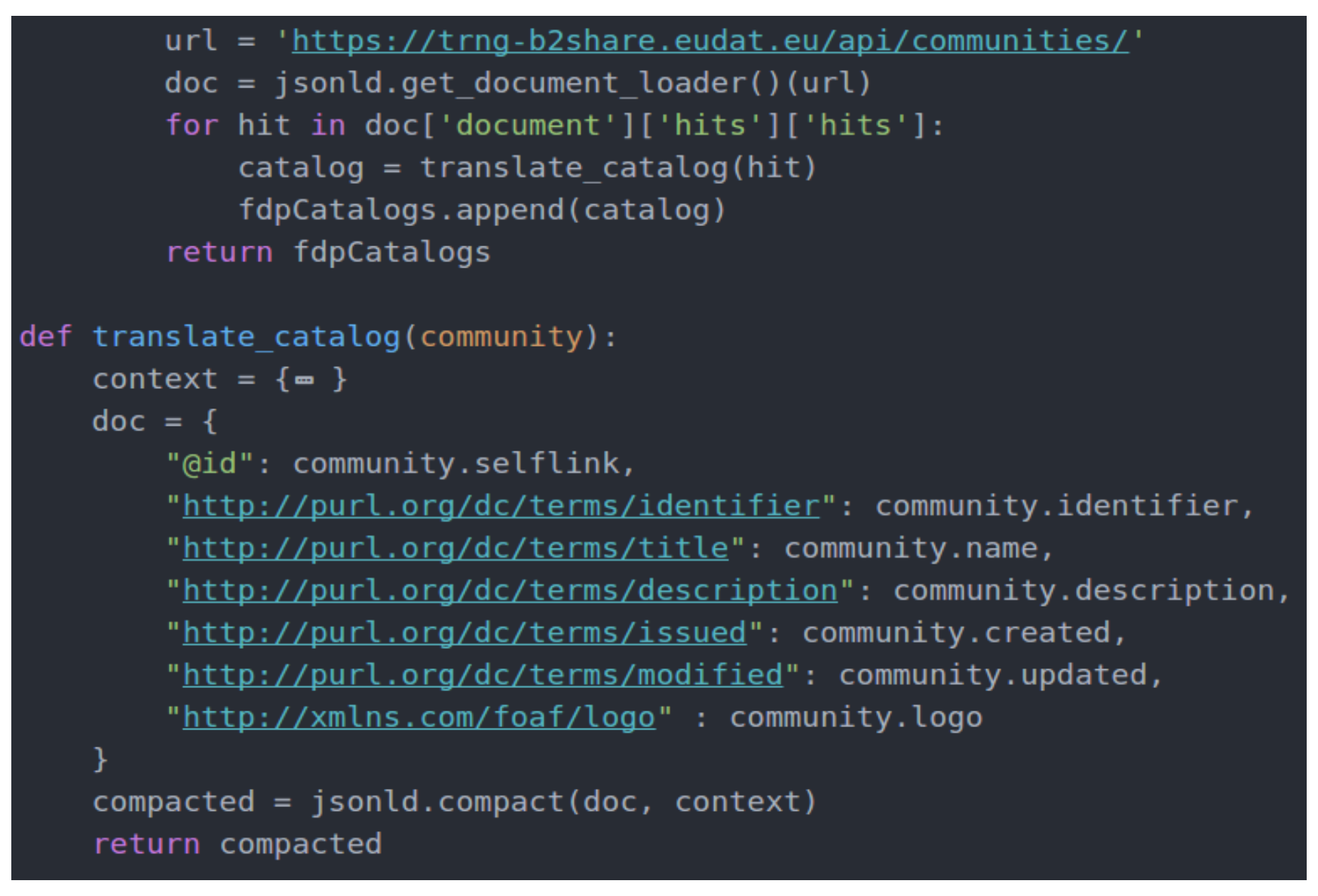
«Kind» FDP level 2 (L2)::FDP Catalog

- dct:identifier: required
- dct:title: required
- dct:hasVersion: required
- dct:publisher: required
- dct:description: optional
- dct:language: optional
- dct:license: optional
- dct:issued: optional
- dct:modified: optional
- dct:conformsTo: optional
- dct:rights: optional
- fdp:metadataIdentifier: required
- fdp:metadataIssued: required
- fdp:metadataModified: required
- rdfs:label: optional
- foaf:homepage: optional
- dcat:dataset: required
- dcat:themeTaxonomy: required

«Kind» B2SHARE-FAIR Community

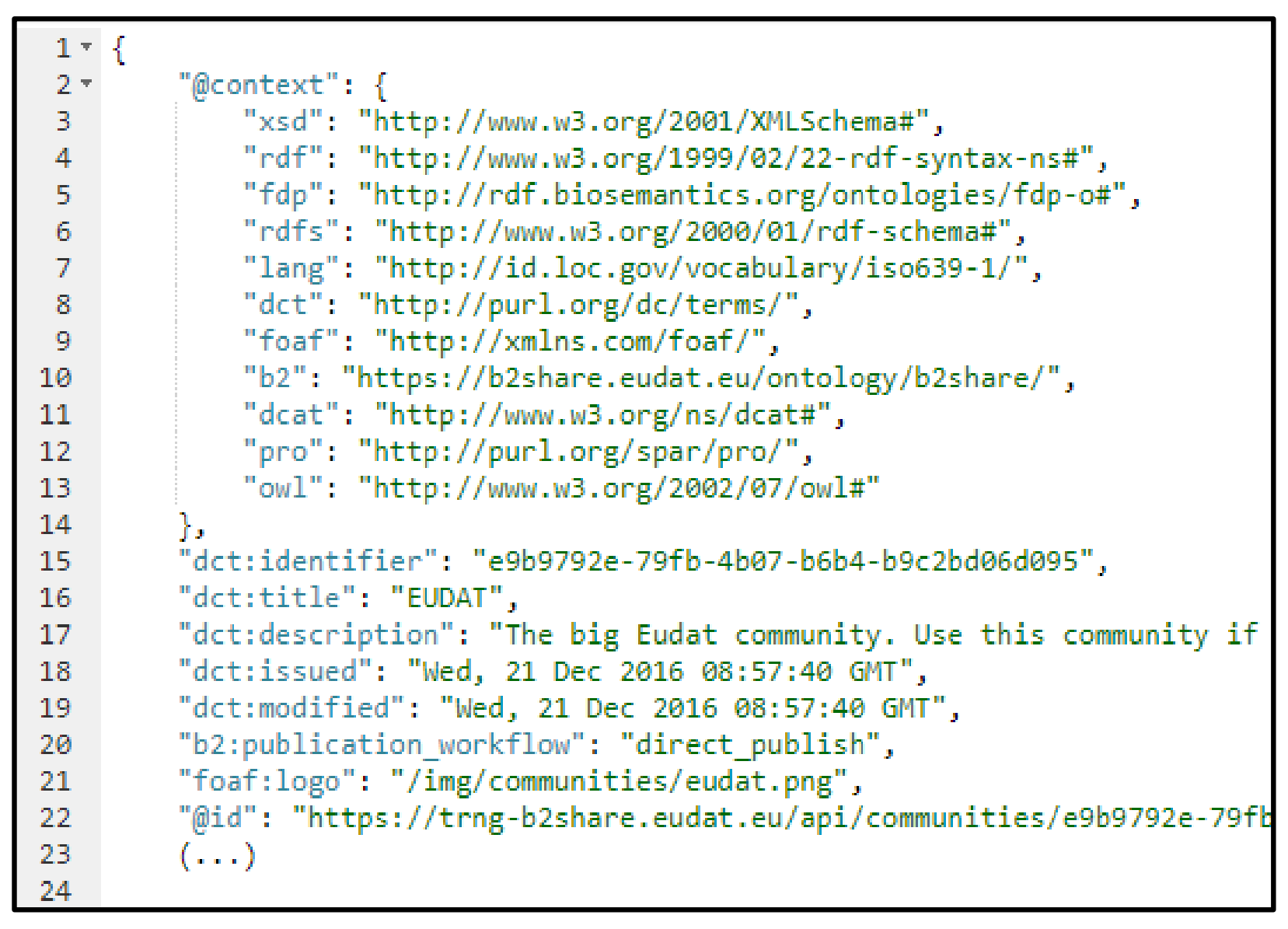
- dct:identifier: required
- dct:title: required
- dct:hasVersion: required
- dct:publisher: required
- dct:description: optional
- dct:language: optional
- dct:license: optional
- dct:issued: optional
- dct:modified: optional
- dct:conformsTo: optional
- dct:rights: optional
- fdp:metadataIdentifier: required
- fdp:metadataIssued: required
- fdp:metadataModified: required
- rdfs:label: optional
- foaf:homepage: optional
- dcat:dataset: required
- dcat:themeTaxonomy: required
- foaf:logo
- b2:publication_workflow
- b2:restricted_submission
- http://purl.org/spar/pro/Role
- @id: JSON-LD

4- Implementation with Python library for JSON-LD (PyLD) [4]



```
    url = 'https://trng-b2share.eudat.eu/api/communities/'
    doc = jsonld.get_document_loader()(url)
    for hit in doc['document']['hits']['hits']:
        catalog = translate_catalog(hit)
        fdpCatalogs.append(catalog)
    return fdpCatalogs

def translate_catalog(community):
    context = {⋯ }
    doc = {
        "@id": community.selflink,
        "http://purl.org/dc/terms/identifier": community.identifier,
        "http://purl.org/dc/terms/title": community.name,
        "http://purl.org/dc/terms/description": community.description,
        "http://purl.org/dc/terms/issued": community.created,
        "http://purl.org/dc/terms/modified": community.updated,
        "http://xmlns.com/foaf/logo" : community.logo
    }
    compacted = jsonld.compact(doc, context)
    return compacted
```



```
{
    "@context": {
        "xsd": "http://www.w3.org/2001/XMLSchema#",
        "rdf": "http://www.w3.org/1999/02/22-rdf-syntax-ns#",
        "fdp": "http://rdf.biosemantics.org/ontologies/fdp-o#",
        "rdfs": "http://www.w3.org/2000/01/rdf-schema#",
        "lang": "http://id.loc.gov/vocabulary/iso639-1/",
        "dct": "http://purl.org/dc/terms/",
        "foaf": "http://xmlns.com/foaf/",
        "b2": "https://b2share.eudat.eu/ontology/b2share/",
        "dcat": "http://www.w3.org/ns/dcat#",
        "pro": "http://purl.org/spar/pro/",
        "owl": "http://www.w3.org/2002/07/owl#"
    },
    "dct:identifier": "e9b9792e-79fb-4b07-b6b4-b9c2bd06d095",
    "dct:title": "EUDAT",
    "dct:description": "The big Eudat community. Use this community if
    "dct:issued": "Wed, 21 Dec 2016 08:57:40 GMT",
    "dct:modified": "Wed, 21 Dec 2016 08:57:40 GMT",
    "b2:publication_workflow": "direct_publish",
    "foaf:logo": "/img/communities/eudat.png",
    "@id": "https://trng-b2share.eudat.eu/api/communities/e9b9792e-79fb
    (...)

```

[1] FAIR Data Point Specification. https://github.com/DTL-FAIRData/FAIRDataPoint/wiki/FAIR-Data-Point-Specification
[2] RE3Data Schema. http://www.re3data.org/schema/3-0
[3] W3C's Data Catalog Vocabulary. http://www.w3.org/ns/dcat
[4] Python library for JSON-LD (https://json-ld.org/). https://github.com/digitalbazaar/pyld

DTL DUTCH TECHCENTRE FOR LIFE SCIENCES   SURF SARA   UNIVERSITEIT TWENTE.

Contact info
João Moreira
7522 NB Enschede, Zilverling 4074
+31 53 489 4337
j.luizrebelomoreira@utwente.nl

Contact info
Thijs Cobben
1098XG Amsterdam
+31-887873000
thijs.cobben@surfsara.nl