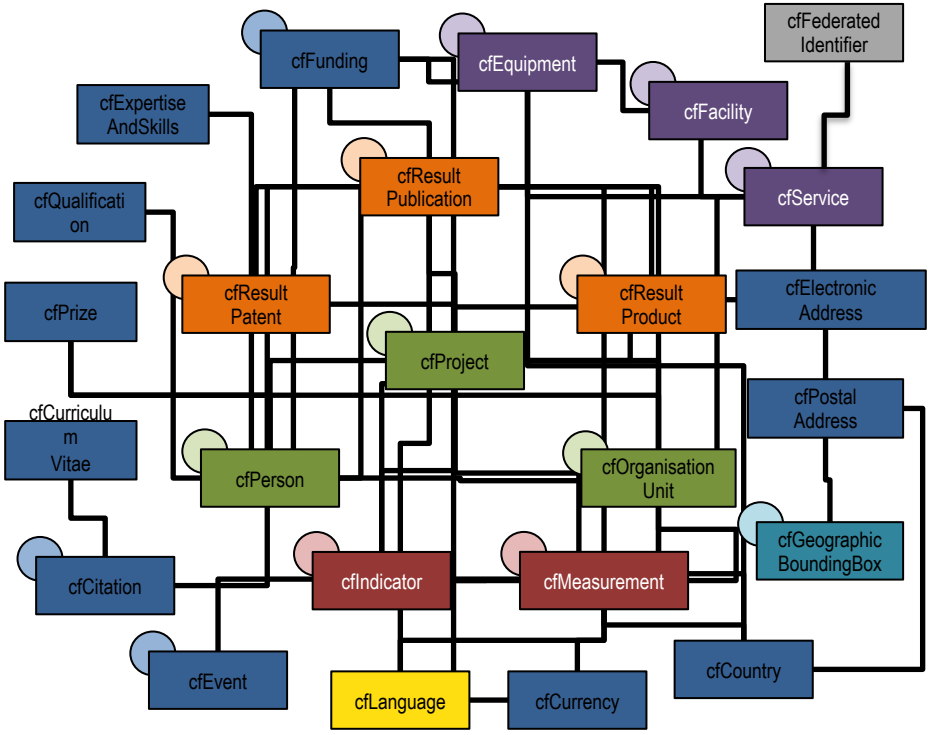
CERIF Tutorial



Jan Dvořák

May 30th, 2023

euroCRIS Membership Meeting
Brussels, Belgium

*This deck of slides is based on the CERIF Tutorial by **Brigitte Jörg**, CERIF TG Leader 2004–2012, and contains slides by **Valérie Brasse**, euroCRIS Executive for projects 2011–2015. It also contains slides co-developed by **Dragan Ivanović** for the CERIF Refactoring Project.*

Jan Dvořák https://orcid.org/0000-0001-8985-152X jan.dvorak@ff.cuni.cz

euroCRIS

- CERIF TG Leader since 2013
- President since 2023

Charles University, Institute of Information Studies and Librarianship

- Researcher & Lecturer

Czech Technical University in Prague, Computing and Information Centre

- IS Analyst: V3S+EZOP – operations and further development of the in-house built institutional CRIS

InfoScience Praha (2004–2016)

- Leader of the team: Research, Development & Innovation Information System (the national CRIS for [CZ])
- In other roles for the system since 1997

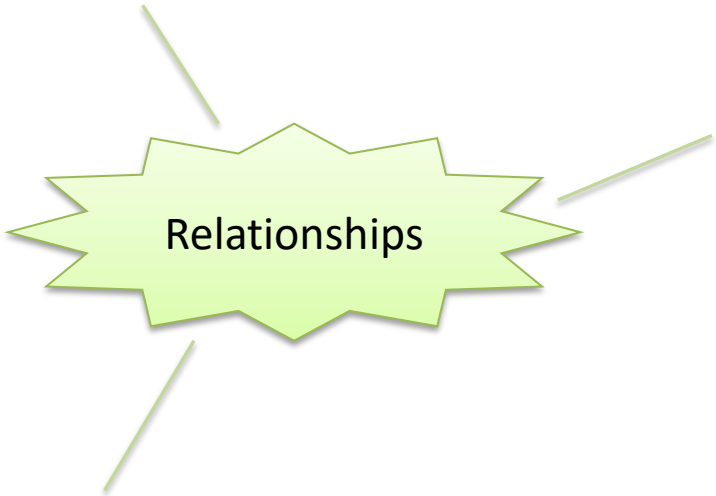
Research Information

The process of research

- Research projects
- Funding
- Research infrastructures

The research actors

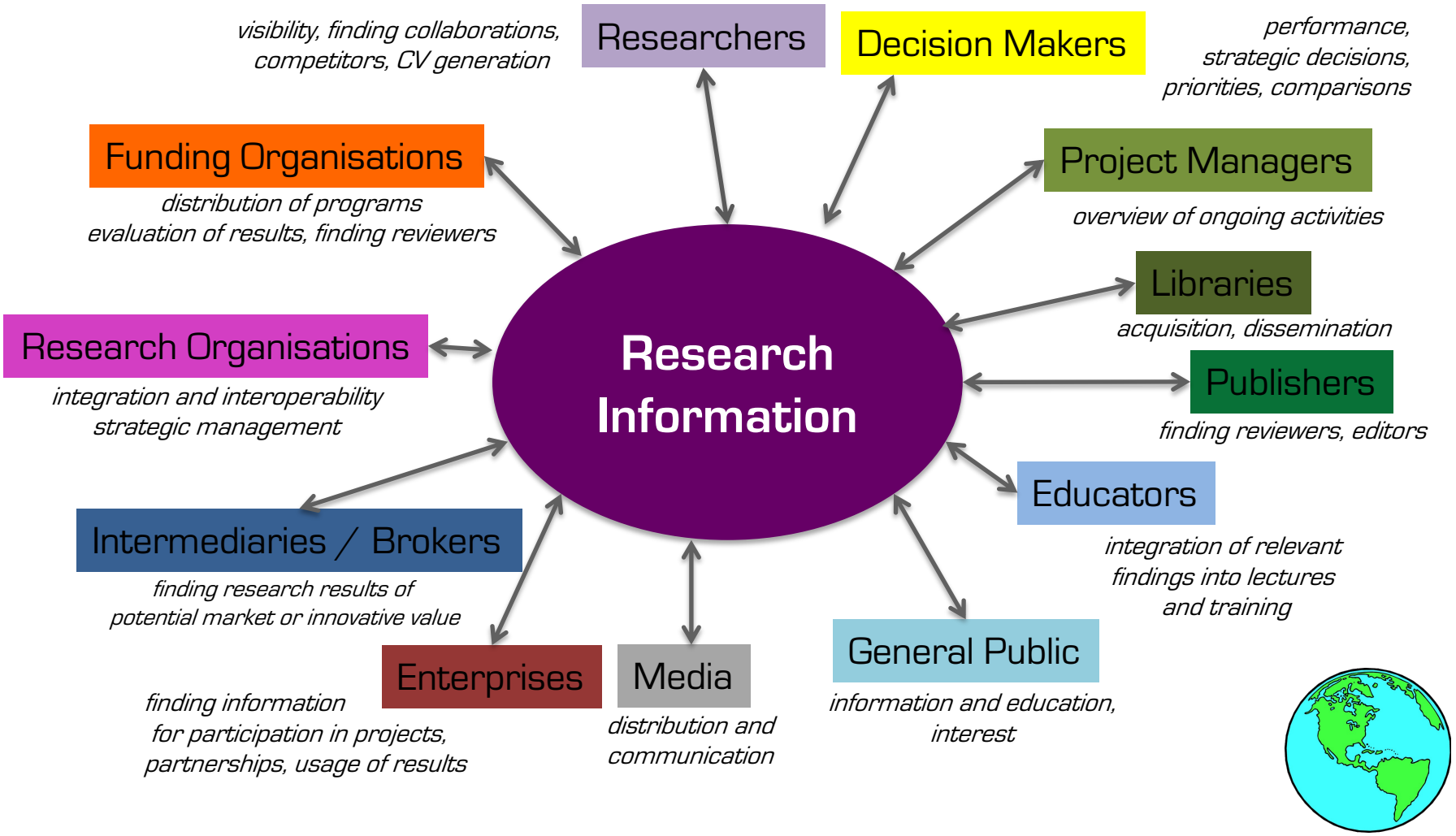
- Researchers
- Institutions
- Funders
- Publishers
- Facility operators
- Associations



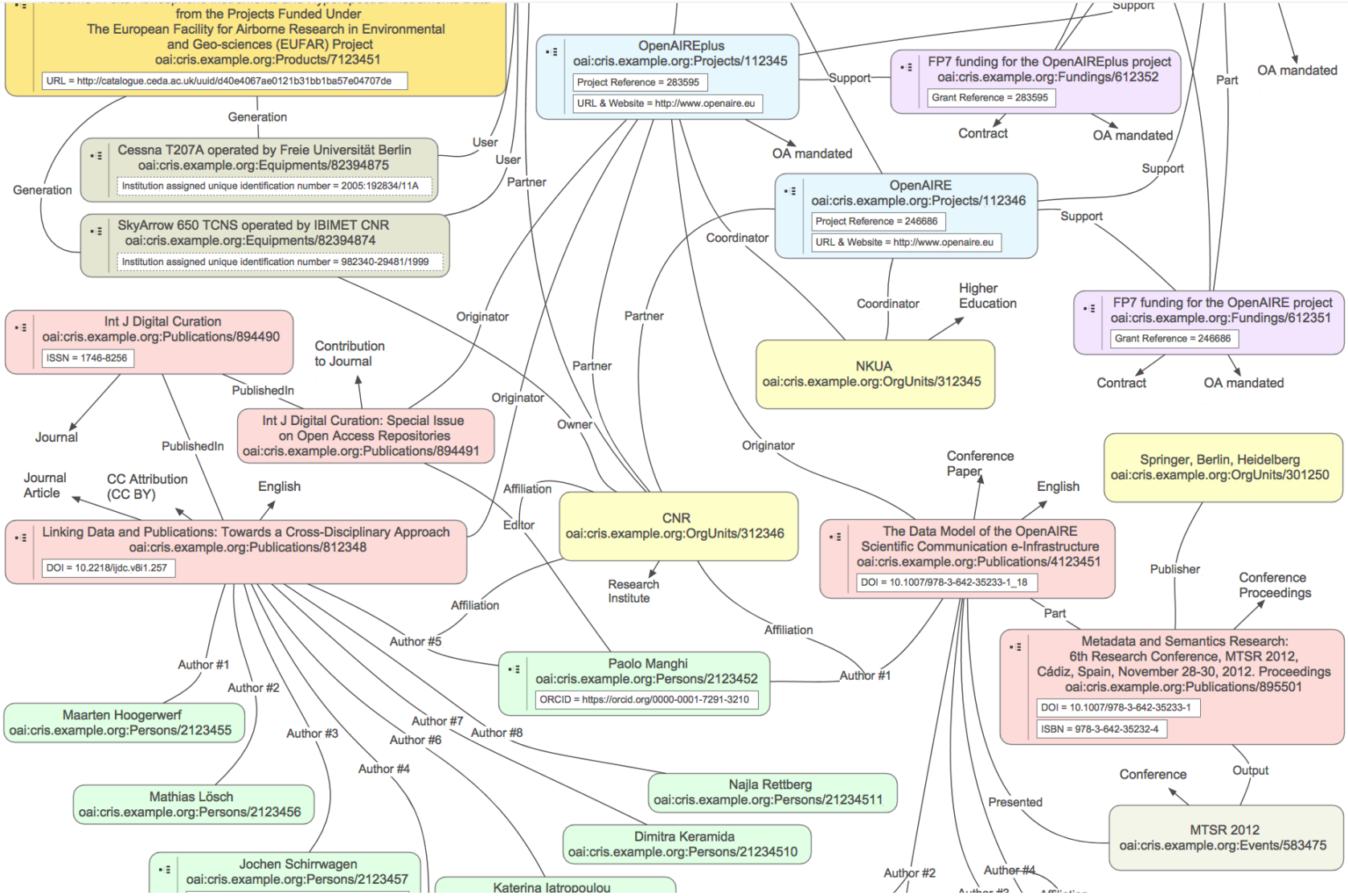
Research results

- Outputs (Publications, Research Datasets, Patents, ...)
- Outcomes, Impacts

Who needs Research Information?



Research Information is heavily interlinked



Different viewpoints on Research Information

Information perspective	Useful for
Researcher, his/her activities, results, …	Researcher profile / CV system
Research Project, its consortium, team, funding, results, …	Project webpage / report
Organisational Unit, its activities, results, ...	Unit webpage / report / evaluation report
Publication	Bibliometrics / panel evaluations
Research Facility	Tracking & reporting usage

What characterizes a research project?

An acronym

A code (identifier), for ex a Grant number

A name or title

A (planned) start date

A (planned) end date or duration

[A source of funding]

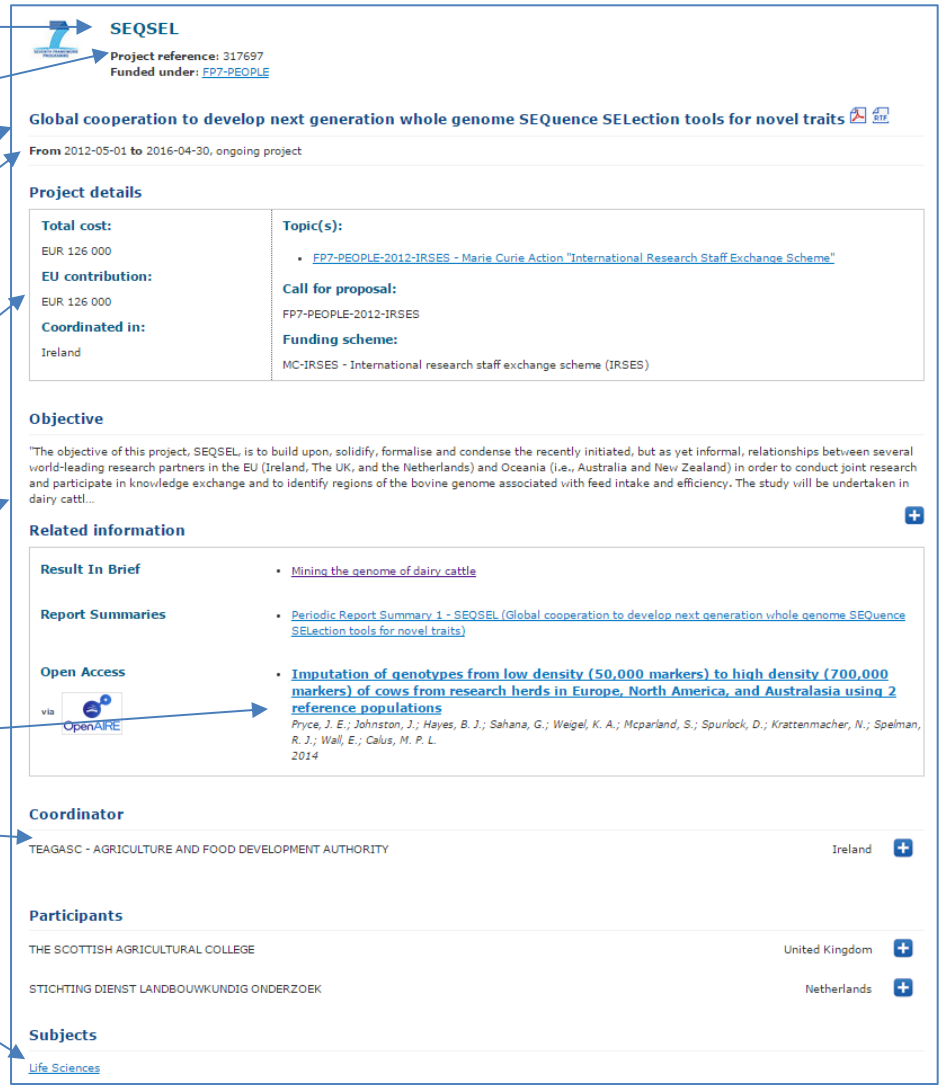
A short or long description (abstract)

[A few scientific publications]

[A project coordinator]

A web page (URI)

[A research domain]



SEQSEL

Project reference: 317697
Funded under: FP7-PEOPLE

Global cooperation to develop next generation whole genome SEQuence SELection tools for novel traits

From 2012-05-01 **to** 2016-04-30, ongoing project

Project details

Total cost:
EUR 126 000
EU contribution:
EUR 126 000
Coordinated in:
Ireland

Topic(s):
- FP7-PEOPLE-2012-IRSES - Marie Curie Action "International Research Staff Exchange Scheme"

Call for proposal:
FP7-PEOPLE-2012-IRSES
Funding scheme:
MC-IRSES - International research staff exchange scheme (IRSES)

Objective

"The objective of this project, SEQSEL, is to build upon, solidify, formalise and condense the recently initiated, but as yet informal, relationships between several world-leading research partners in the EU (Ireland, The UK, and the Netherlands) and Oceania (i.e., Australia and New Zealand) in order to conduct joint research and participate in knowledge exchange and to identify regions of the bovine genome associated with feed intake and efficiency. The study will be undertaken in dairy cattl...

Related information

Result In Brief
- Mining the genome of dairy cattle

Report Summaries
- Periodic Report Summary 1 - SEQSEL (Global cooperation to develop next generation whole genome SEQuence SELection tools for novel traits)

Open Access
via OpenAIRE
- **Imputation of genotypes from low density (50,000 markers) to high density (700,000 markers) of cows from research herds in Europe, North America, and Australasia using 2 reference populations**
  Pryce, J. E.; Johnston, J.; Hayes, B. J.; Sahana, G.; Weigel, K. A.; Mcparland, S.; Spurlock, D.; Krattenmacher, N.; Spelman, R. J.; Wall, E.; Calus, M. P. L.
  2014

Coordinator

TEAGASC - AGRICULTURE AND FOOD DEVELOPMENT AUTHORITY — Ireland

Participants

THE SCOTTISH AGRICULTURAL COLLEGE — United Kingdom

STICHTING DIENST LANDBOUWKUNDIG ONDERZOEK — Netherlands

Subjects

Life Sciences

Source: http://cordis.europa.eu/project/rcn/106635_en.html

A name or title

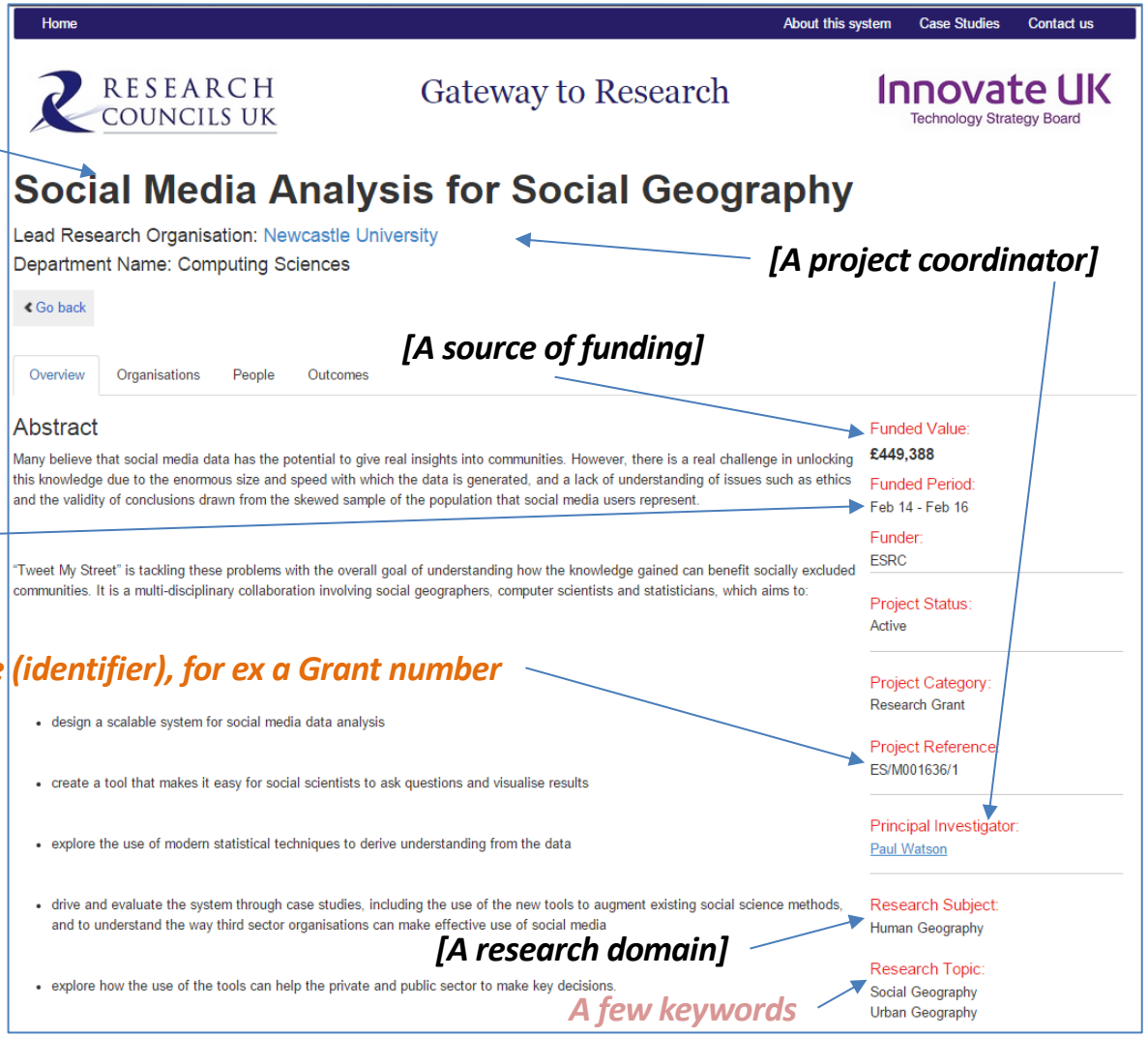
A short or long description (abstract)

A (planned) start date
A (planned) end date or duration

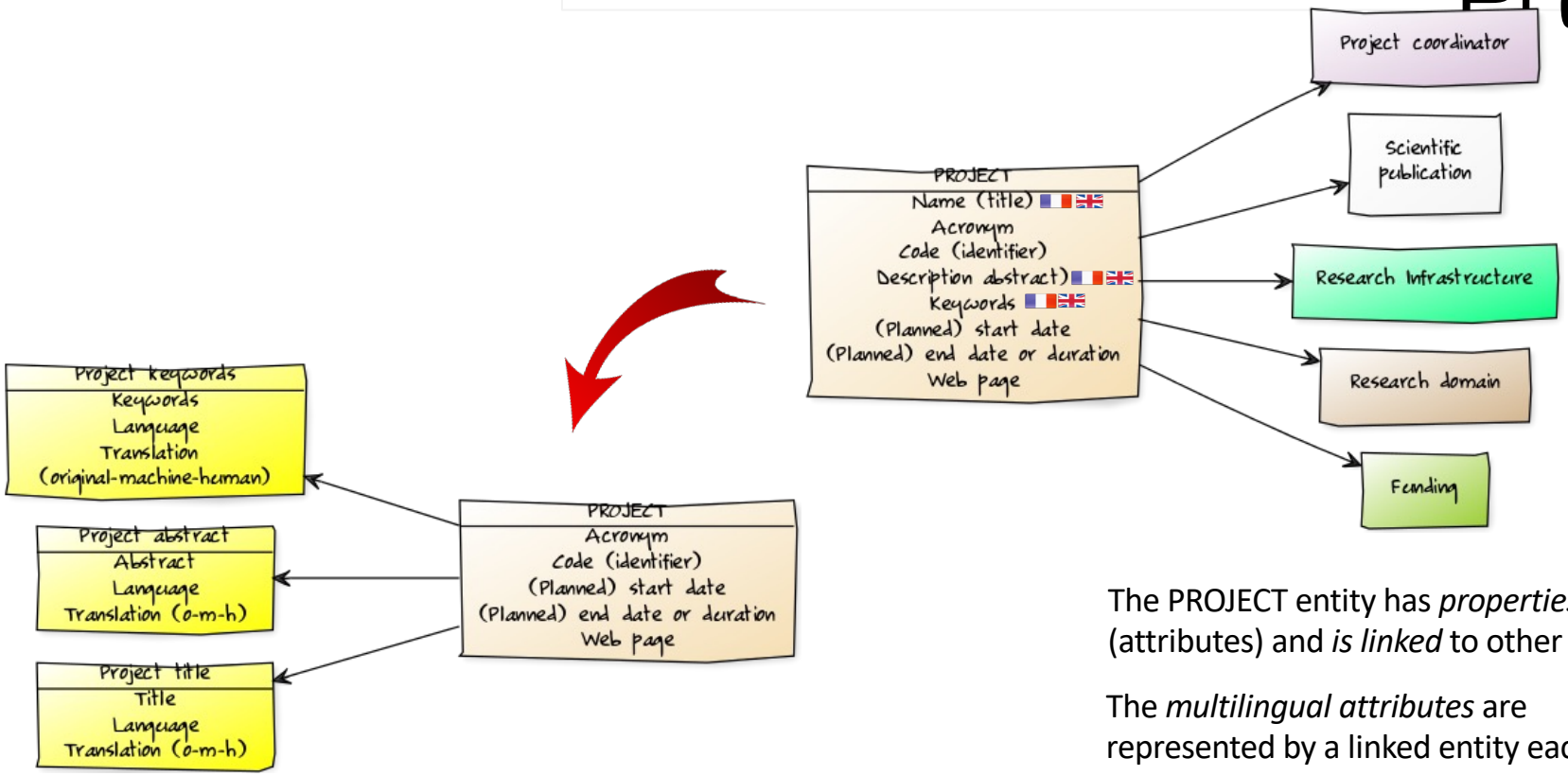
A code (identifier), for ex a Grant number

A web page (URI)

Source: http://gtr.rcuk.ac.uk/project/A49CA721-687A-4D55-8FDF-9B60375B6EA8



Metadata for a Research Project



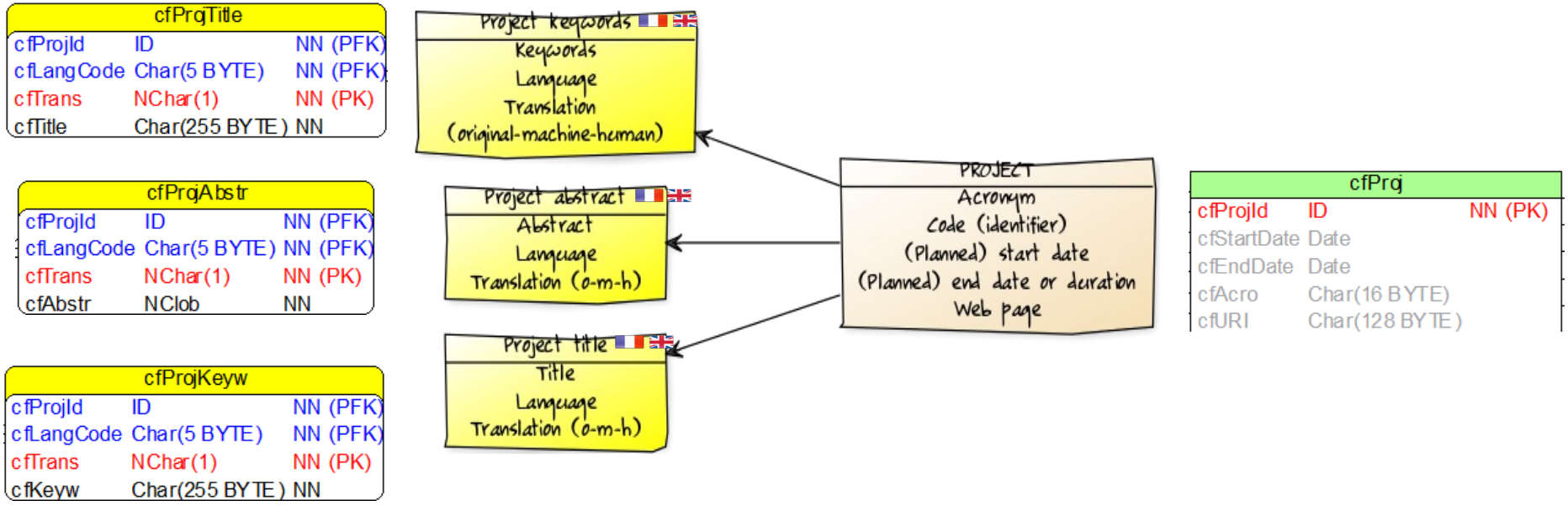
The PROJECT entity has *properties* (attributes) and *is linked* to other entities.

The *multilingual attributes* are represented by a linked entity each.

* "start date" and "end date" are deprecated in v1.6

Metadata for a Research Project

CERIF naming rule: in English, abbreviated, starting with *cf*
Example: Project title = cfProjTitle



Example in DB

CERIF Project example database record			
Data	**Attribute**	**Table**	**Type**
project-ist-world *	**cfProjId**	**cfProj**	**base**
IST World	**cfAcro**	**cfProj**	**base**
http://www.ist-world.org/	**cfURI**	**cfProj**	**base**
2005-04-01	**cfStartDate**	**cfProj**	**base**
2007-11-30	**cfEndDate**	**cfProj**	**base**
Knowledge Base for RTD Competencies in IST	**cfTitle**	**cfProjTitle**	**lang[en,o]**
Wissensbasis für RTD Kompetenzen im Bereich IST	**cfTitle**	**cfProjTitle**	**lang[de,h]**
IST, Research Information, NMS, Portal,	**cfKeyw**	**cfProjKeyw**	**lang**
The objective of the project is to set up and populate an information portal with innovative functionalities …	**cfAbstr**	**cfProjAbstr**	**lang**

cfProj

cfProjId (PK)	cfAcro	cfURI	cfStartDate	cfEndDate
project-ist-world	IST World	http://...	2005-04-01	2007-11-30

cfProjTitle, PK = cfProjID + cfLangCode + cfTrans

cfProjID (FK)	cfLangCode	cfTrans	cfTitle
project-ist-world	EN	O	Knowledge Base for RTD Competencies in IST
project-ist-world	DE	H	Wissensbasis für RTD Kompetenzen im Bereich IST

cfProjKeyw, PK = cfProjID + cfLangCode + cfTrans

cfProjID (FK)	cfLangCode	cfTrans	cfKeyw
project-ist-world	EN	O	IST, Research Information, NMS, Portal

cfProjAbstr, PK = cfProjID + cfLangCode + cfTrans

cfProjID (FK)	cfLangCode	cfTrans	cfAbstr
project-ist-world	EN	O	The objective of the project is to set…

Source: http://www.eurocris.org/Uploads/Web%20pages/CERIF-1.3/Specifications/CERIF1.3_FDM.pdf

Project definition

A temporary endeavor undertaken to create a unique product, service or result.
Source: the Project Management Institute, https://www.pmi.org/about/learn-about-pmi/what-is-project-management

In the research information domain, one typically tracks:
(1) **research** projects, where the result is an addition to the body of knowledge of the mankind,
(2) technology **development** projects, where the result is a particular technology or product,
(3) **innovation** projects, where the result is an improvement of a product or process, and
(4) projects that **create or enhance infrastructure** for research, technology development or innovation.

Depending on the scope one can also track finer levels of granularity: stages, work packages, sometimes even down to individual tasks. All such activities are also modelled using the Project entity and linked using the recursive link relationship.

The Project entity only captures details of the project scope and plan. Information about the resources needed to execute the project such as the funding (i.e., the grants received), the people and organisations involved, the supporting infrastructures, the outputs produced, etc. is contained in separate entities (the Funding entity, the Person entity, the OrgUnit entity, the infrastructure entities, the result entities respectively) and is linked to the Project.

INTERMEDIATE SUMMARY

- CERIF is:
  - A conceptual model
  - A storage format in relational database
  - A set of exchange formats (XML, Linked Data)

Common European Research Information Format

- CERIF supports multilingual data, storing the original value of a literal attribute, and for any other language, a value translated by a machine and/or a human
- So far, we have seen the CERIF Entity "PROJECT" (cfProj)

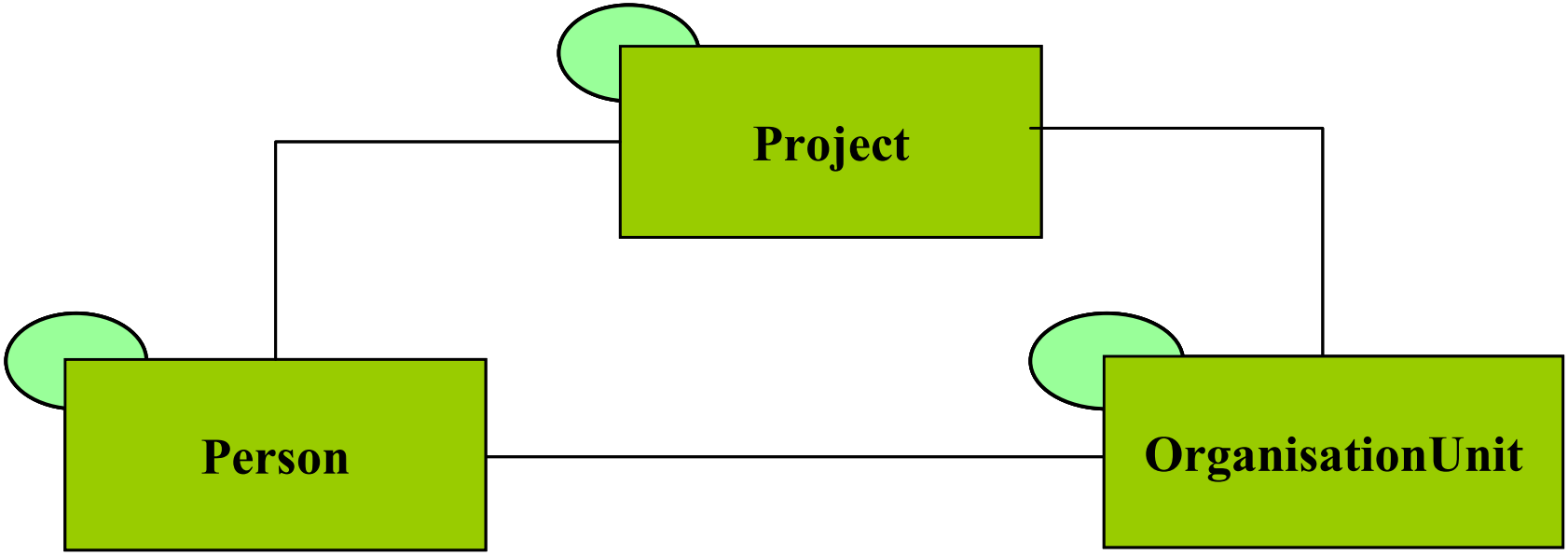
We have seen how to represent, store or exchange metadata about research projects.

Similarly:

- What characterises a person (researcher, Ph.D.,…)?
- What characterises an organisation (research laboratory, institute,…)?

Source: https://pixabay.com/en/chemistry-teacher-science-1027781/

CERIF Base Entities



Person definition

A human being as an individual.
Source: https://en.oxforddictionaries.com/definition/person

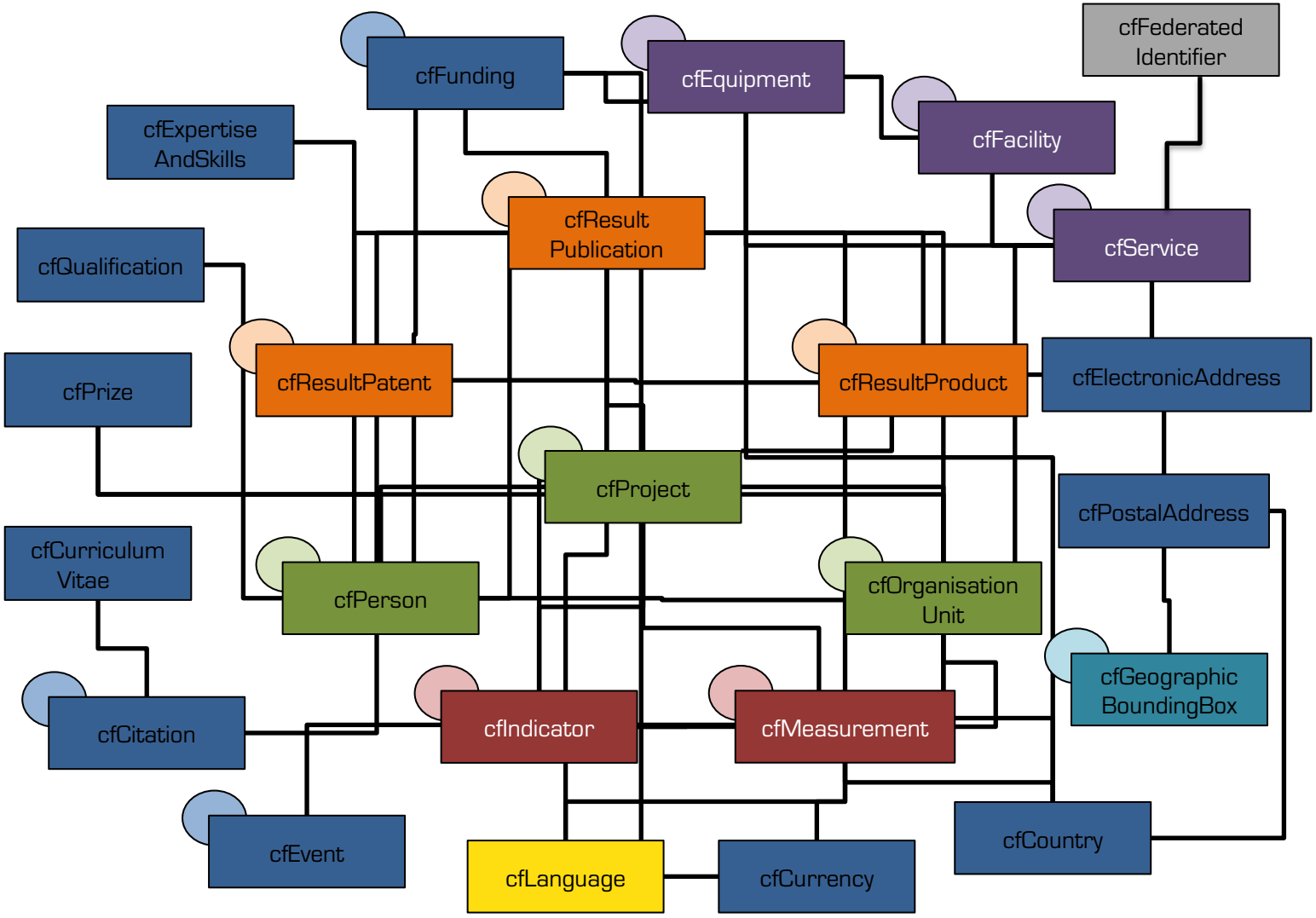
The kind of involvement of a Person in the research ecosystem is specified in the links with the organisations, the services, etc. This typically includes:
(1) **researchers** (Persons performing research in an Organisation Unit as employees or students);
(2) **authors** and **contributors** (Persons signing a publication, creators of data sets, software developers, etc.);
(3) **investigators** and project **participants** (Persons involved in a Project as principal investigators, co investigators, project managers, consultants, etc.);
(4) **management** (directors, rectors, deans, department heads, etc.);
(5) **support** staffs (technicians, responsible for Equipment, librarians and digital asset curators, administrative staff, etc.).
One Person typically has many of these relationships.

OrgUnit definition

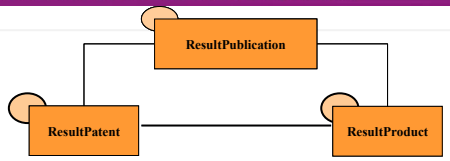
Organisation Unit: an organisation, a unit therein, a committee or any other group of people that has a collective goal. Organisation Units are not necessarily formalized as legal entities.

In the research information domain Organisation Units typically represents:
(1) organisations that **perform research** (universities, research institutes, corporations) and their subdivisions (faculties, schools, departments, research groups) and other associated bodies (boards, advisory bodies);
(2) organisations that **fund research** (funders, their divisions and evaluation panels);
(3) scientific **associations and networks**;
(4) **publishers**, facility **operators** and other **service providers** in the research space;
(5) **authorities**, such as patent offices and standardization or supervision bodies; and
(6) other bodies: editorial **boards**, evaluation **panels**, or **committees** of all kinds.

Common European Research Information Format



CERIF Result Entities

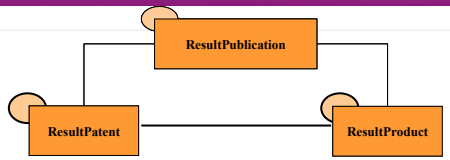


CERIF Result Entities

ResultPublication
ID
URI
Title
Subtitle
Abstract
Bibl. Note
PublicationDate
TotalPages
StartPage
EndPage
Keywords
Type
Status
Subject

ResultPatent
ID
URI
PatentNumber
Title
CountryCode
RegistrationDate
ApprovalDate
Description
Keywords
Type
Status
Subject

ResultProduct
ID
URI
Type
Status
Subject



CERIF Result Entities

cfResultPublication
cfID
cfURI
cfNumber
cfPublicationDate
cfStartPage
cfEndPage
cfTotalPages
cfEdition
cfSeries
cfIssue
cfVolume

cfTitle
cfSubtitle
cfAbstract
cfKeywords
cfBibliographic Note
cfVersionInfo
cfAbbreviation

Type
Status
Subject

cfResultProduct
cfID
cfURI

cfName
cfDescription
cfKeywords
cfVersionInfo

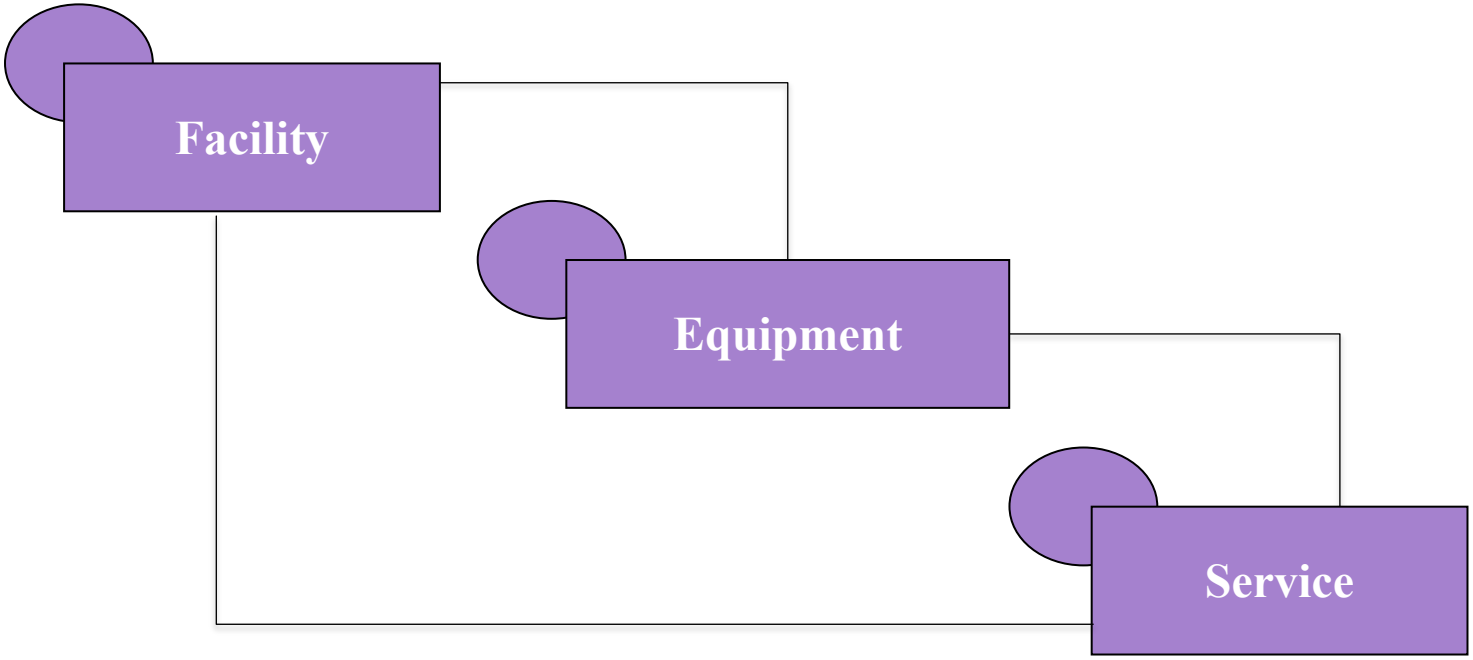
Type
Status
Subject

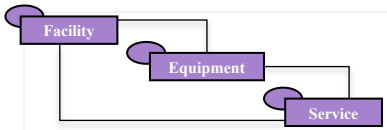
cfResultPatent
cfID
cfURI
cfPatentNumber
cfCountryCode
cfRegistrationDate
cfApprovalDate

cfName
cfAbstract
cfKeywords
cfVersionInfo

Type
Status
Subject

CERIF Infrastructure Entities



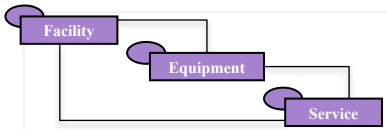


CERIF Infrastructure Entities

Facility
ID
Acronym
URI
Title
Description
Keywords
Type
Status
Subject

Equipment
ID
Acronym
URI
Title
Description
Keywords
Type
Status
Subject

Service
ID
Acronym
URI
Title
Description
Keywords
Type
Status
Subject



CERIF Infrastructure Entities

cfName
cfDescription
cfKeywords
cfFacility
cfID
cfURI
cfAcronym
Type
Status
Subject

cfName
cfDescription
cfKeywords
cfEquipment
cfID
cfURI
cfAcronym
Type
Status
Subject

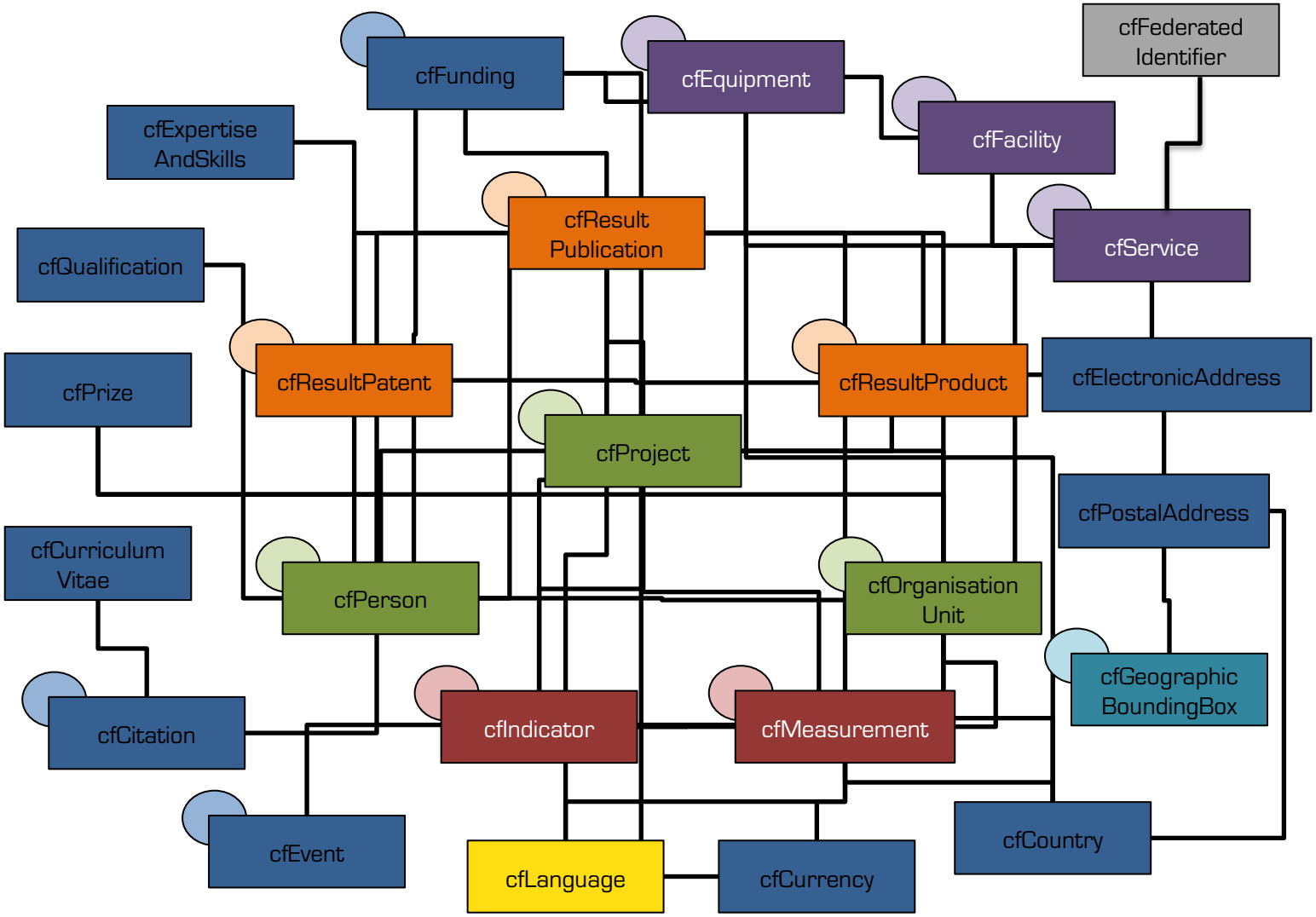
cfName
cfDescription
cfKeywords
cfService
cfID
cfURI
cfAcronym
Type
Status
Subject

CERIF General Pattern

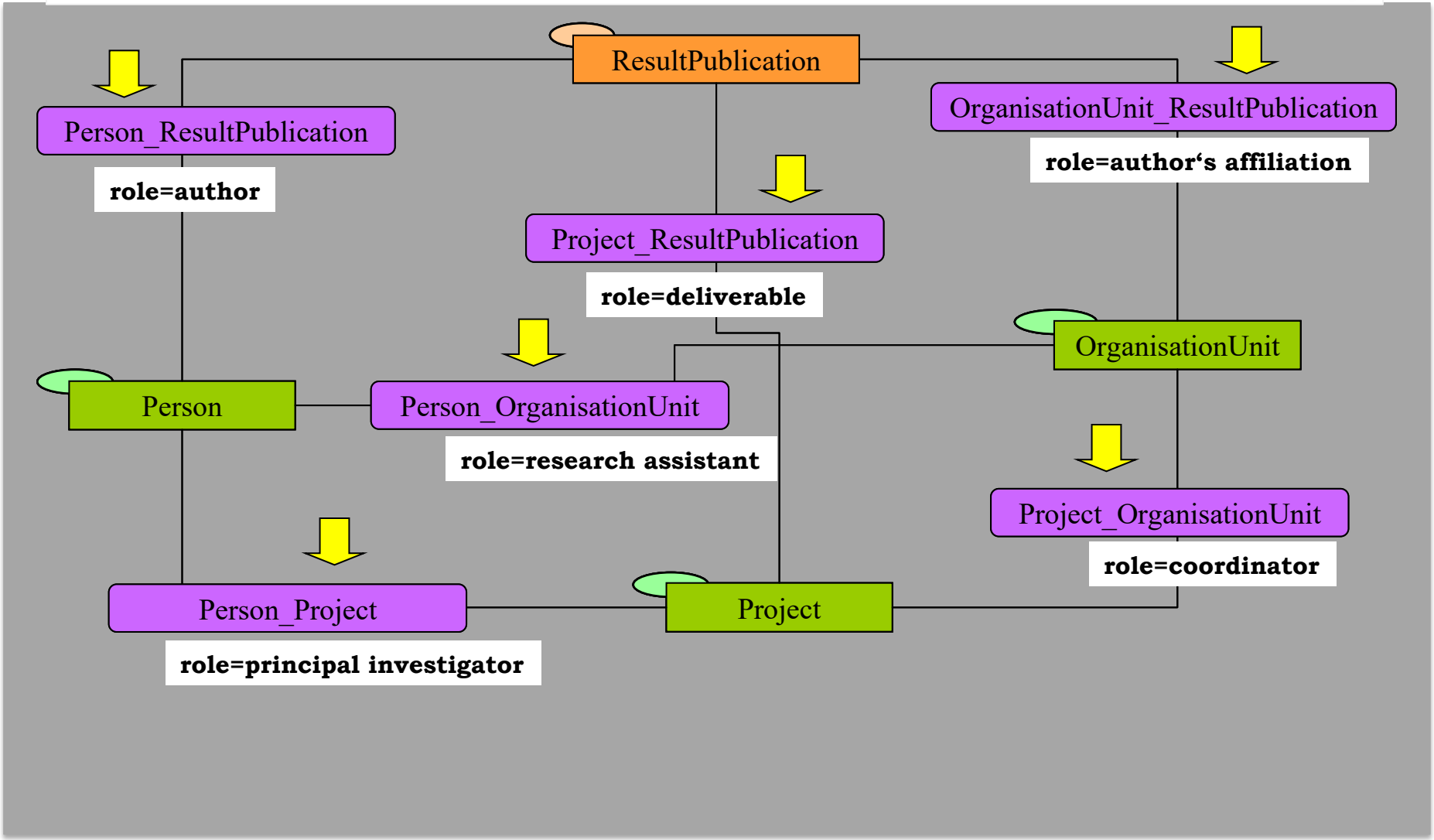
A typical CERIF entity has:

- Identifier (internal)
- Attributes
  - the basic ones
  - the multi-lingual ones
- External Identifiers
- Classifications
  - Type
  - Status
  - Subject area
  - + other
- Links
  - to other entities
  - recursive, a.k.a. self-referential

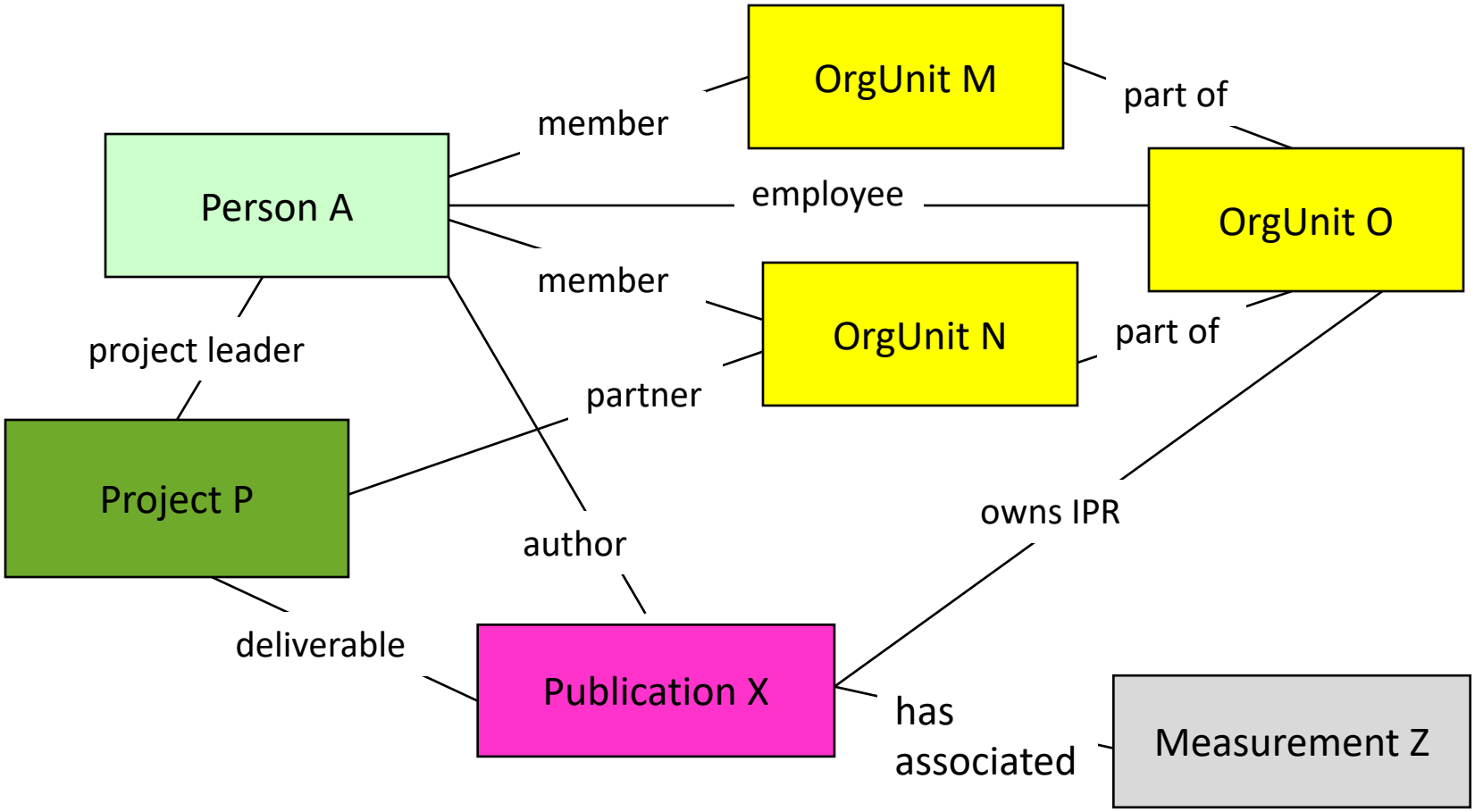
CERIF 1.6


cfFunding
cfEquipment
cfFederated Identifier
cfExpertise AndSkills
cfFacility
cfResult Publication
cfQualification
cfService
cfPrize
cfResultPatent
cfResultProduct
cfElectronicAddress
cfProject
cfPostalAddress
cfCurriculum Vitae
cfPerson
cfOrganisation Unit
cfGeographic BoundingBox
cfCitation
cfIndicator
cfMeasurement
cfEvent
cfLanguage
cfCurrency
cfCountry

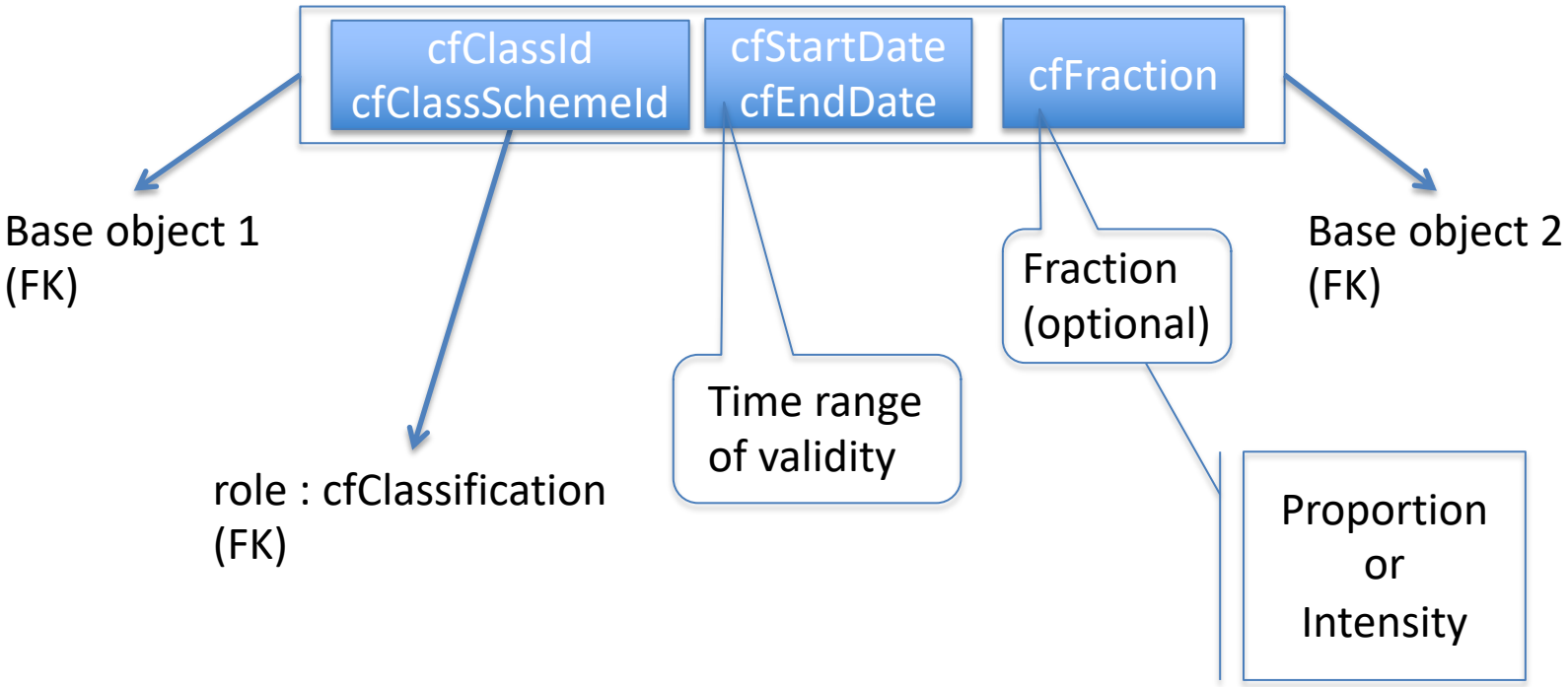
Some CERIF Link Entities



Another example
(slide by Keith Jeffery)

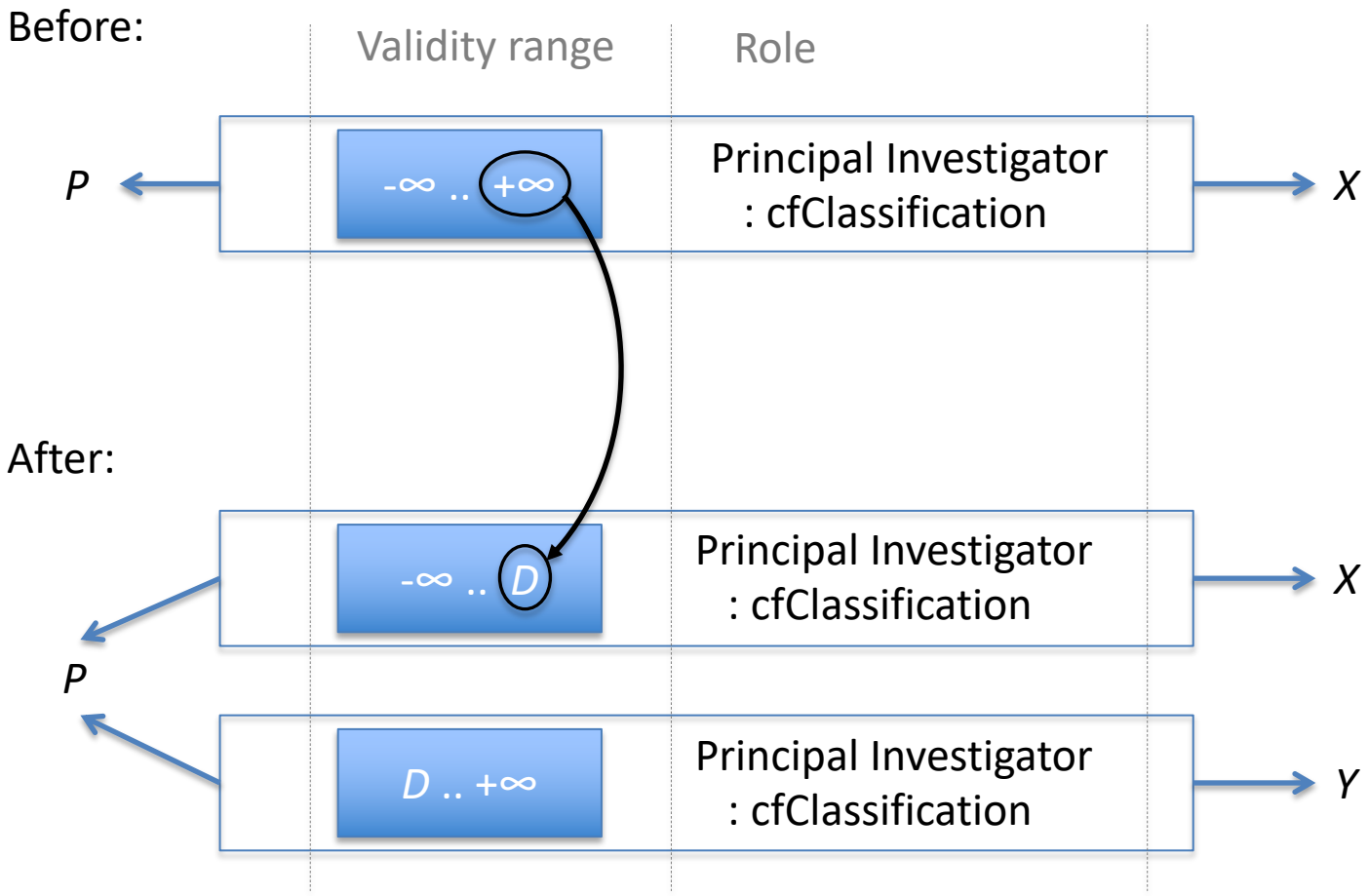


Generic Linking Entity Structure

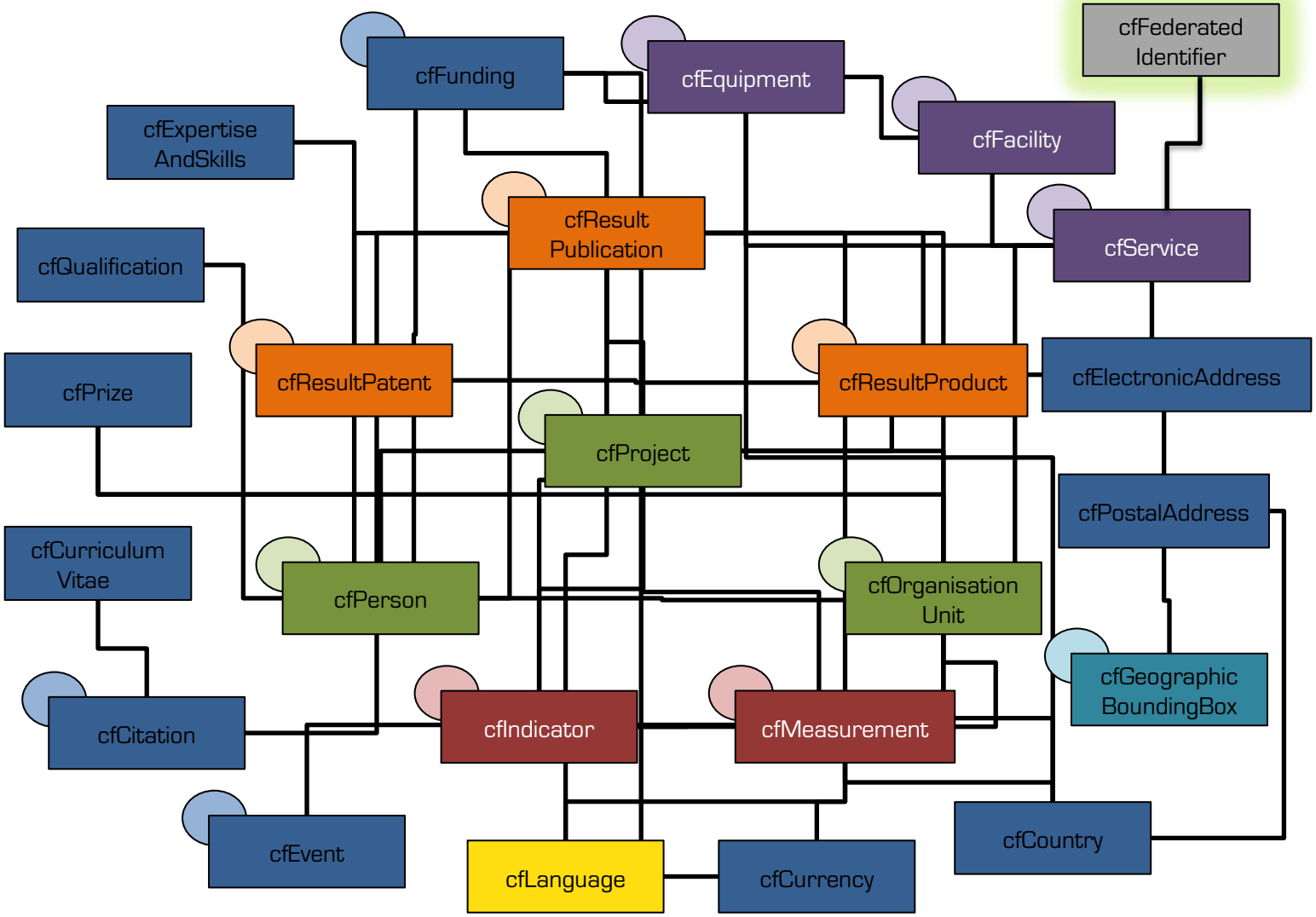


Recording Change in CERIF

Example: The Principal Investigator of project *P* changes: *X* is replaced by *Y* effective date *D*.



CERIF 1.6


cfFederated Identifier
cfFunding
cfEquipment
cfExpertise AndSkills
cfFacility
cfResult Publication
cfQualification
cfService
cfPrize
cfResultPatent
cfResultProduct
cfElectronicAddress
cfProject
cfPostalAddress
cfCurriculum Vitae
cfPerson
cfOrganisation Unit
cfGeographic BoundingBox
cfCitation
cfIndicator
cfMeasurement
cfEvent
cfCountry
cfLanguage
cfCurrency

CERIF Federated Identifiers

- Records the "tag" by which an object is known elsewhere
- For any CERIF research entity
- "Identifier Types" classification scheme
- (optionally) Connects to a Service representing the issuer of the identifier
    - Usually an information system

CERIF Federated Identifiers: Examples

Publication

- DOI
- PubMed Central ID
- WoS UT, Scopus EID
- ISBN, ISSN

Person

- ORCID iD
- ISNI
- ResearcherID, Scopus Author ID
- IdRef, DAI, Número Lattes
- Social Security Number or another administrative identifier
- Staff Id in an HR system

Project/Grant

- Funder's reference number
- Organisation's reference number

Organisation

- ROR ID
- FundRefID
- Wikidata reference
- ISNI
- RingGold number
- Org Enhanced (WoS)
- VAT Identification Number or another administrative identifier

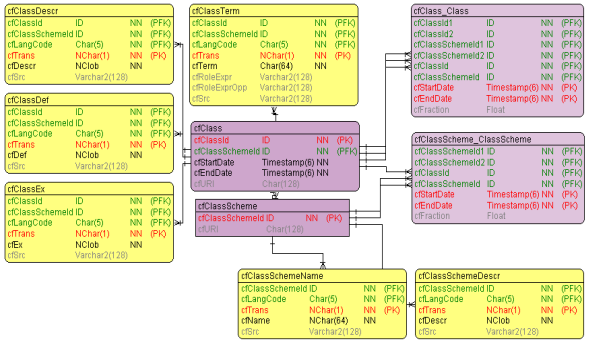
Classifications: External Codes

CERIF Semantic Layer

Central place to store declared semantic classifications

Allows to capture any Schema or Structure

- Flat Lists
- Thesauri
- Classification Systems (e.g. SKOS, ...)
- Taxonomies
- Ontologies



Open / Extensible in all directions

- New Schemas
- New Concepts / Terms
- New Relationships

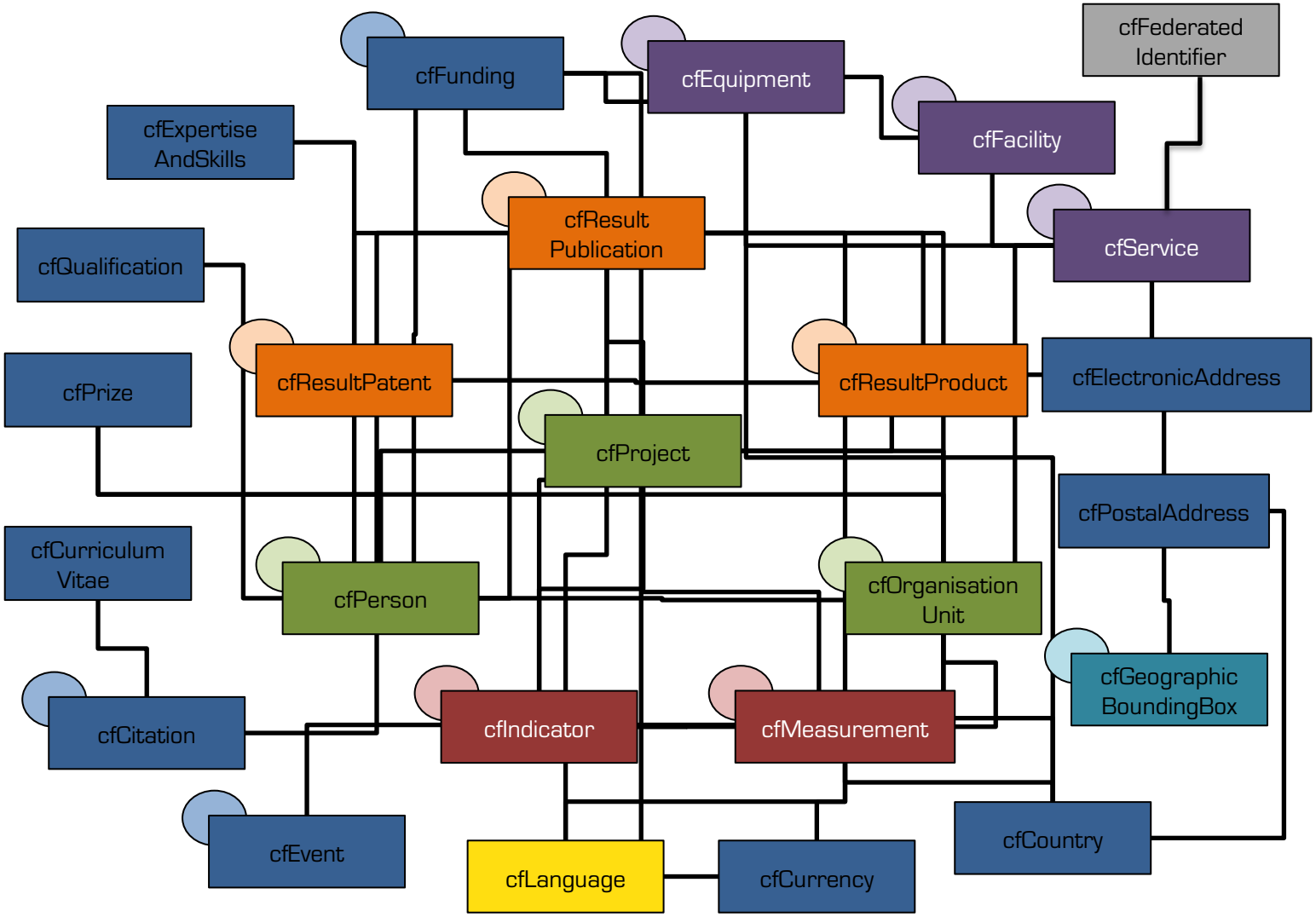
Enables to manage

- Roles, Types, Statuses, ... Semantics
- Subject Headings
- Versioning and archiving (start+end dates)

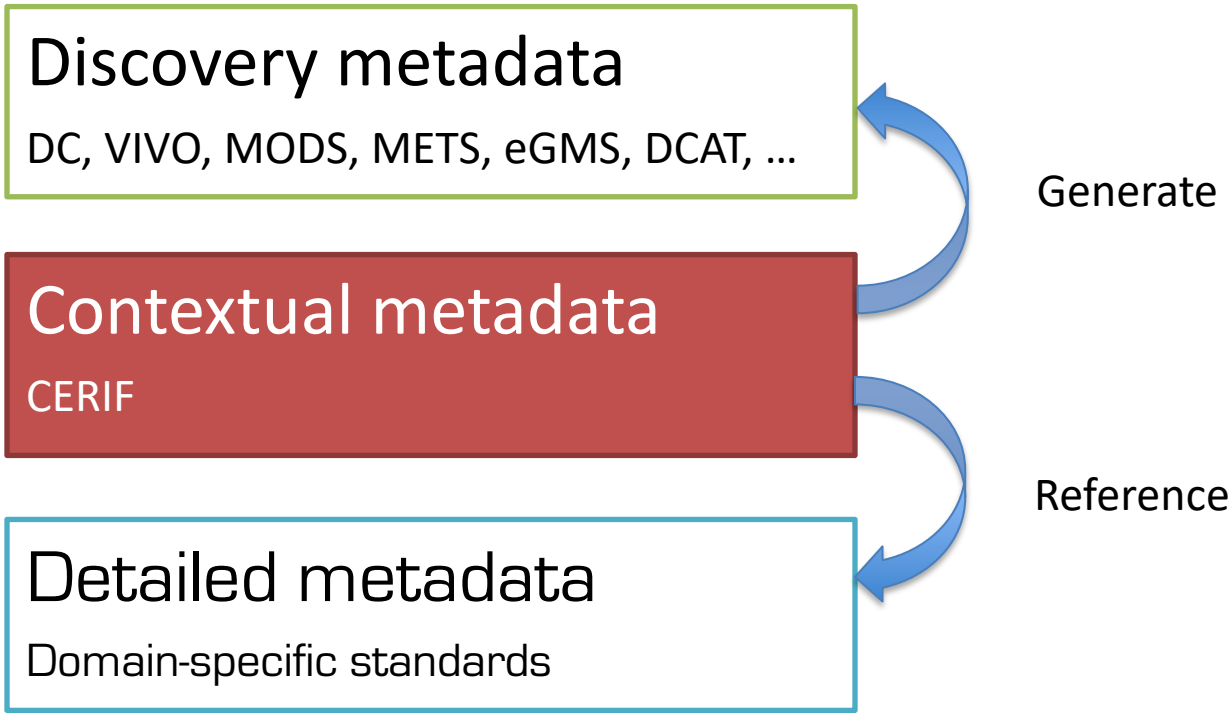
Allows for Mappings between Schemes

- skos:closeMatch, skos:exactMatch or any other mapping you need

CERIF 1.6


cfFunding
cfEquipment
cfFederated Identifier
cfExpertise AndSkills
cfFacility
cfResult Publication
cfQualification
cfService
cfPrize
cfResultPatent
cfResultProduct
cfElectronicAddress
cfProject
cfPostalAddress
cfCurriculum Vitae
cfPerson
cfOrganisation Unit
cfGeographic BoundingBox
cfCitation
cfIndicator
cfMeasurement
cfEvent
cfLanguage
cfCurrency
cfCountry

Metadata Layers



Current Research Information System

- Ongoing
- Past, of current interest
- Currently planned or decided upon

CERIF development

By the CERIF Task Group of euroCRIS

Adopting open-source software projects
tools & best practices:

→ https://github.com/EuroCRIS/CERIF-DataModel

→ https://github.com/EuroCRIS/CERIF-Vocabularies

→ CC BY license

Two branches:

- master: latest official release (1.6.1)
- develop: on-going development

Model Info

Model Info | ER Diagram | Entities | Attributes | Keys | Relationships | Domains

Basic Information

Project	CERIF: the Common European Research Information Format
Model	CERIF Data Model
Version	CERIF current development DM
Company	euroCRIS
Author	CERIF Task Group
Created	30.6.2014
Last Modified	22.12.2017

Statistic Information

Entities	295				
Attributes	1829				
Domains	1				
User Data Types	0				
Relationships	671	Identifying	638	Non-identifying	33
Indexes	0				
Keys	298	Primary Keys	295	Alternate Keys	3
Views	0				
View Relationships	0				
Procedures	0				
Defaults	0				
Schemas	0				
Users	0				
User Groups	0				
Notes	0				

Common European Research Information Format

- CERIF is an EU Recommendation to Member States
- The European Commission (EC) has authorised euroCRIS to maintain and develop CERIF and its usage

http://cordis.europa.eu/cerif/

INTEROPERABILITY

Interoperability

1. Agreed Semantics
2. Agreed Format
3. Agreed Protocol

CERIF-XML exchange formats (based on XML Schema)

Original

1:1 with the ER structure

Only uses embedding for multilingual texts

→ Many foreign key relationships

→ Takes several API requests to get a presentable form of an object

Ex.: OpenAIRE Guidelines for CRIS Managers 1.0.
DOI 10.5281/zenodo.17065

2nd Generation

Template XML Schema → adaptation

Profiles: Useful subsets of CERIF for specific research information exchange scenarios

1. Specify a subset of CERIF entities & attributes
2. Fix semantic vocabularies to use
3. Add integrity constraints

Ex.: OpenAIRE Guidelines for CRIS Managers 1.1.
DOI 10.5281/zenodo.1298649

OpenAIRE Guidelines for CRIS Managers 1.1

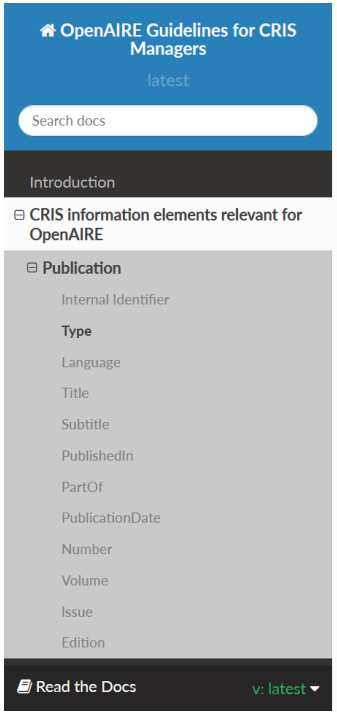
- Dvořák, Jan, Bollini, Andrea, Rémy, Laurent, & Schirrwagen, Jochen. (2018). OpenAIRE Guidelines for CRIS Managers 1.1 (Version 1.1.1). Zenodo. DOI 10.5281/zenodo.2316420

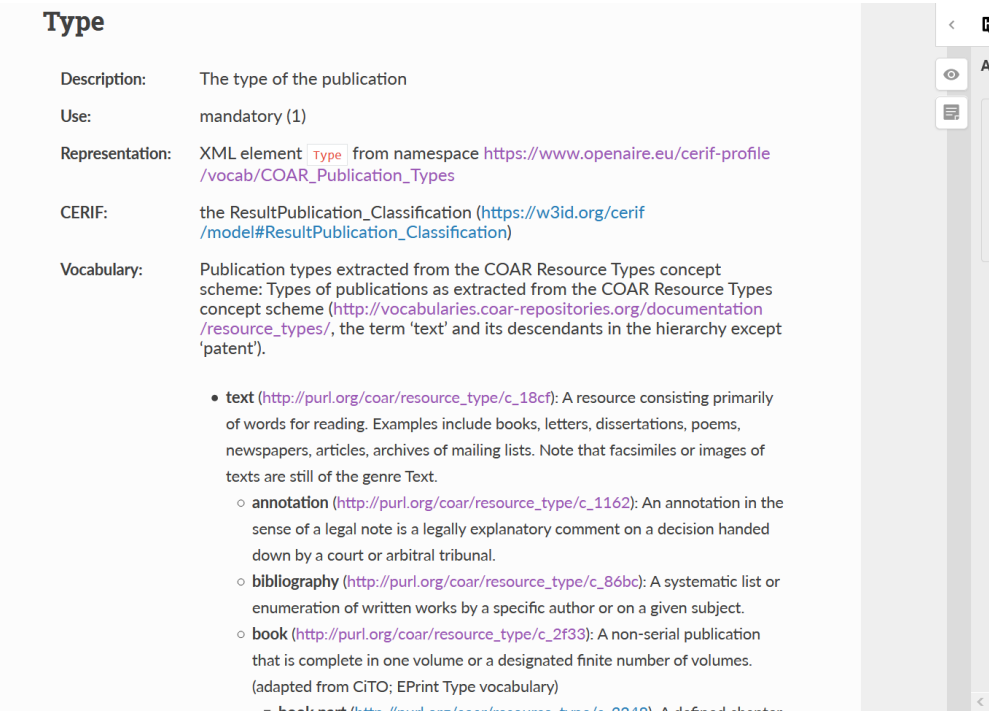
- Document on the web https://openaire-guidelines-for-cris-managers.readthedocs.io/en/latest/index.html
  - Introduction
  - CRIS Information elements relevant for OpenAIRE
    - CERIF, CERIF XML
  - Technical implementation guidelines
    - CERIF XML as payload of OAI-PMH 2.0

- Version 1.1.2 being finalized https://github.com/openaire/guidelines-cris-managers with Andreas Czerniak and Dragan Ivanović

Version 1.1: Aligned vocabularies

- COAR Resource Types (v. 3.1)
- COAR Access Rights
- ISSN Media List





Type

Description: The type of the publication

Use: mandatory (1)

Representation: XML element `Type` from namespace https://www.openaire.eu/cerif-profile/vocab/COAR_Publication_Types

CERIF: the ResultPublication_Classification (https://w3id.org/cerif/model#ResultPublication_Classification)

Vocabulary: Publication types extracted from the COAR Resource Types concept scheme: Types of publications as extracted from the COAR Resource Types concept scheme (http://vocabularies.coar-repositories.org/documentation/resource_types/, the term 'text' and its descendants in the hierarchy except 'patent').

- **text** (http://purl.org/coar/resource_type/c_18cf): A resource consisting primarily of words for reading. Examples include books, letters, dissertations, poems, newspapers, articles, archives of mailing lists. Note that facsimiles or images of texts are still of the genre Text.
  - **annotation** (http://purl.org/coar/resource_type/c_1162): An annotation in the sense of a legal note is a legally explanatory comment on a decision handed down by a court or arbitral tribunal.
  - **bibliography** (http://purl.org/coar/resource_type/c_86bc): A systematic list or enumeration of written works by a specific author or on a given subject.
  - **book** (http://purl.org/coar/resource_type/c_2f33): A non-serial publication that is complete in one volume or a designated finite number of volumes. (adapted from CiTO; EPrint Type vocabulary)

Information scope of the OpenAIRE Guidelines for CRIS Managers 1.1

The Core:

- The CRIS as a Service
- Organisations
- Researchers

Activities:

- Projects
- Funding
- Events

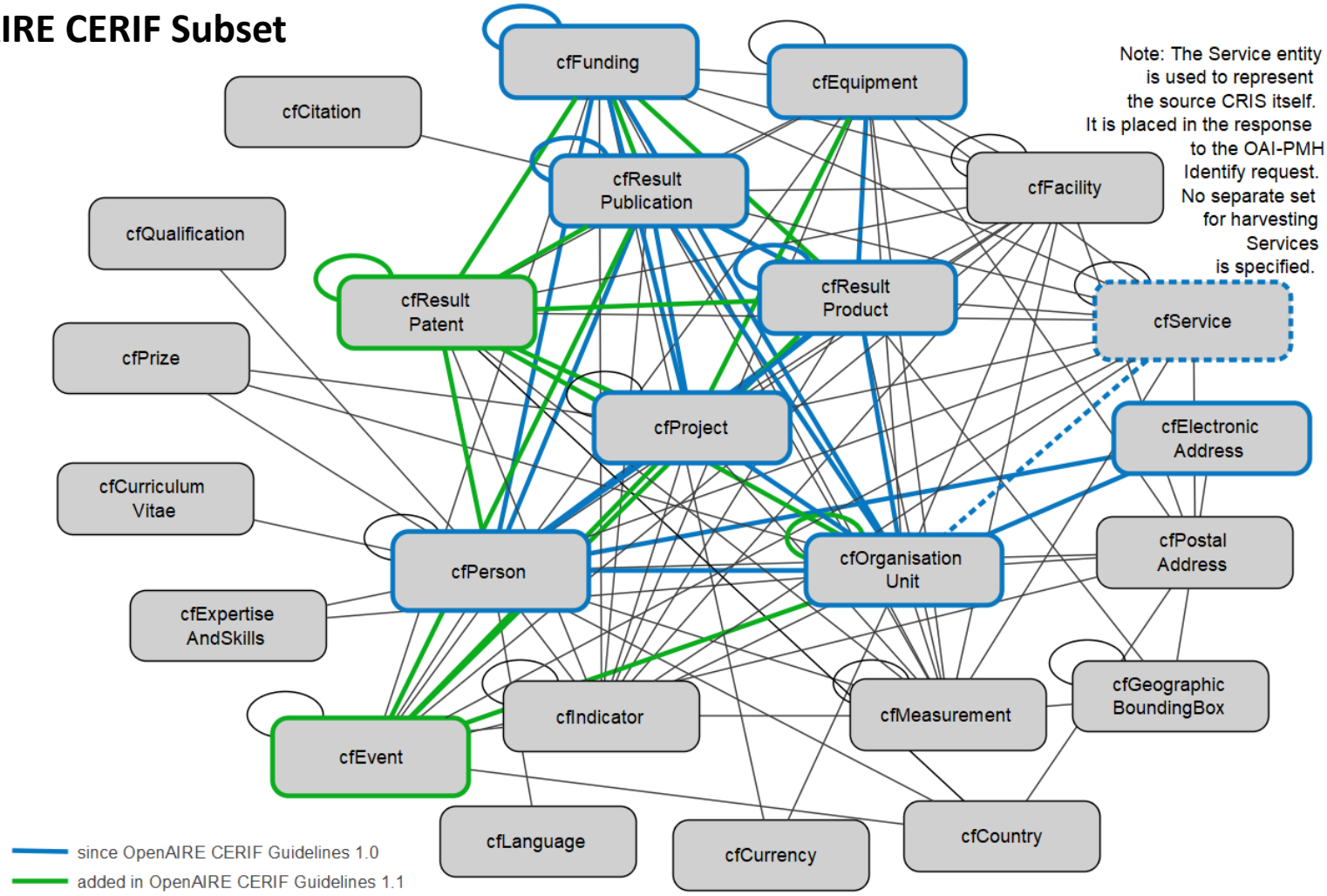
Outputs:

- Publications
- Datasets, Software & other Products
- Patents

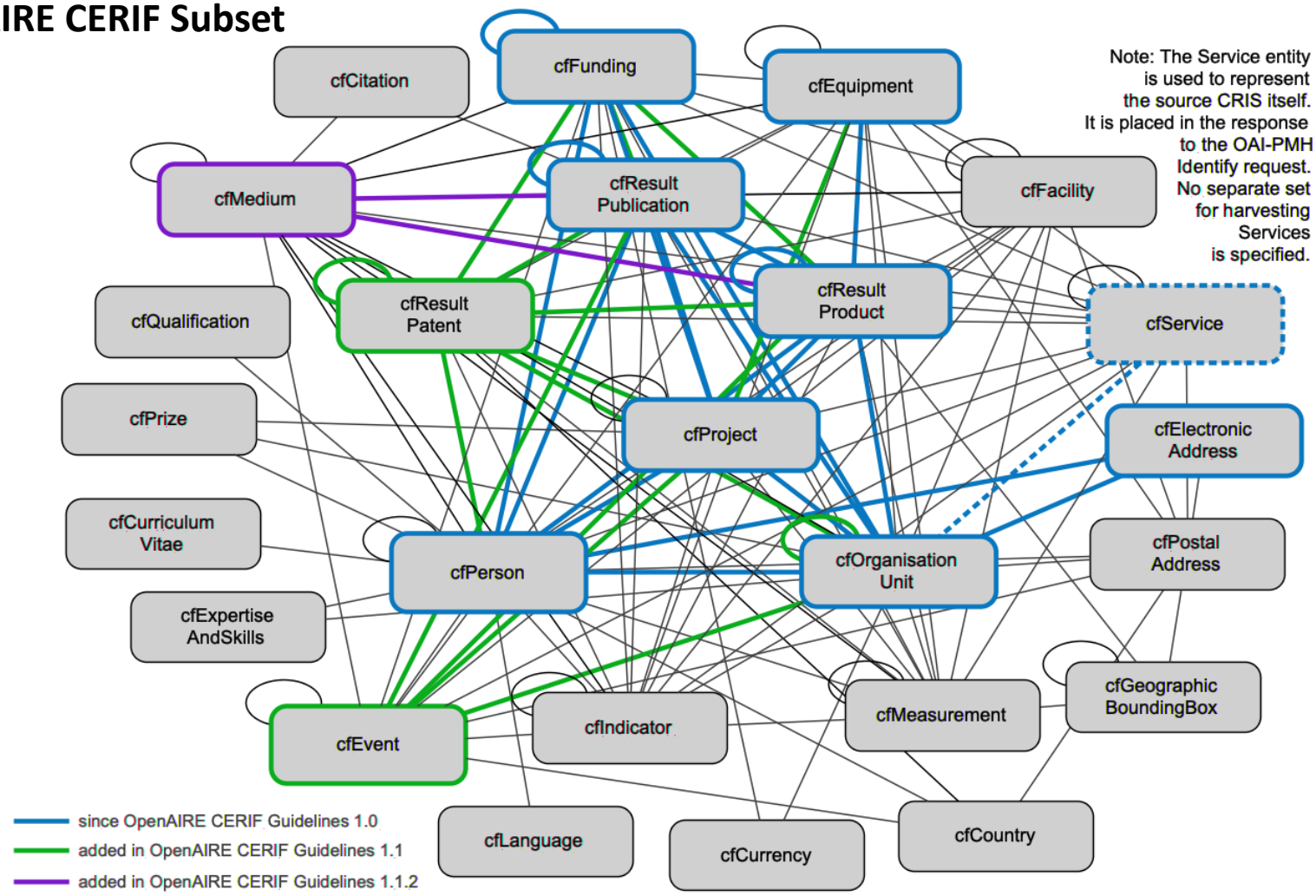
Infrastructure:

- Equipment

OpenAIRE CERIF Subset



OpenAIRE CERIF Subset



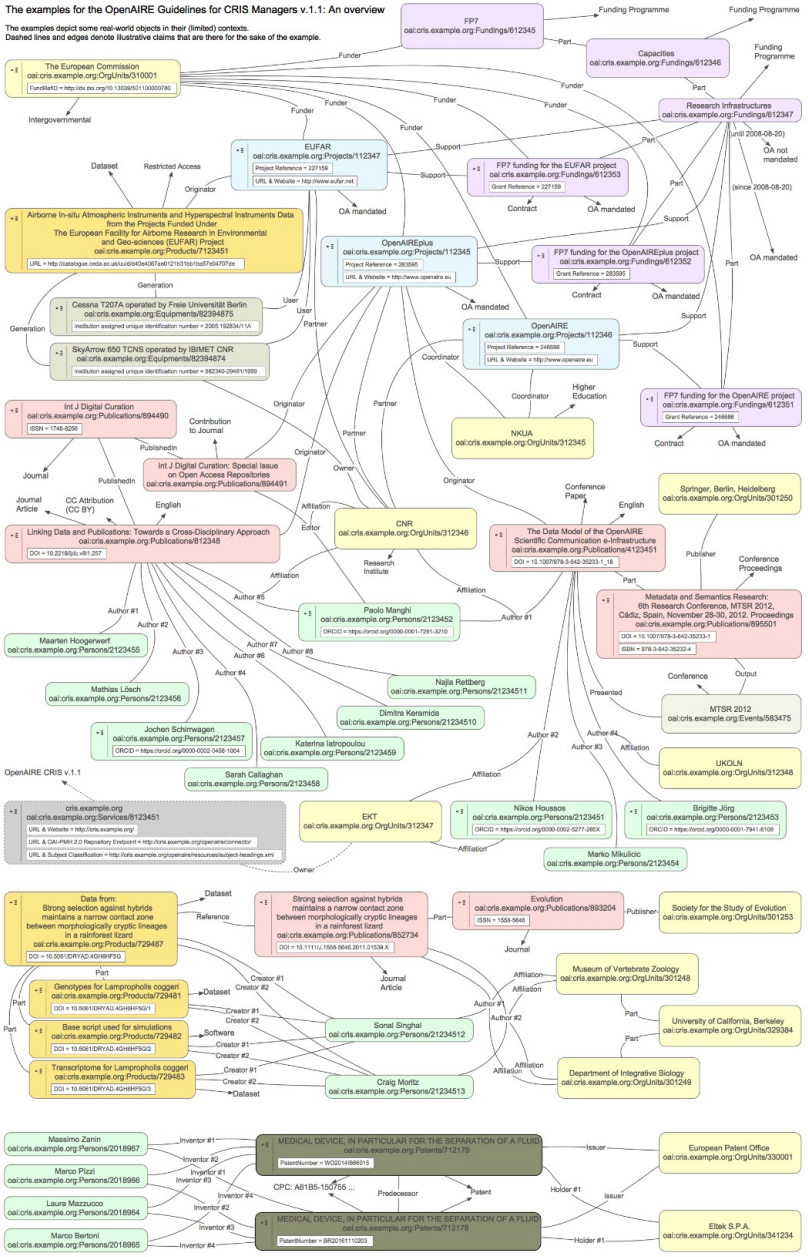
Identifiers

- Publication:
  - DOI
  - Handle
  - PMCID
  - ISI Number (WoS Accession Number)
  - SCP Number
  - ISSN
  - ISBN
  - URL
  - URN
- Persons:
  - ORCID iD
  - ISNI
  - ResearcherID
  - Scopus Author ID
- Products:
  - DOI
  - ARK
  - Handle
  - URL
  - URN

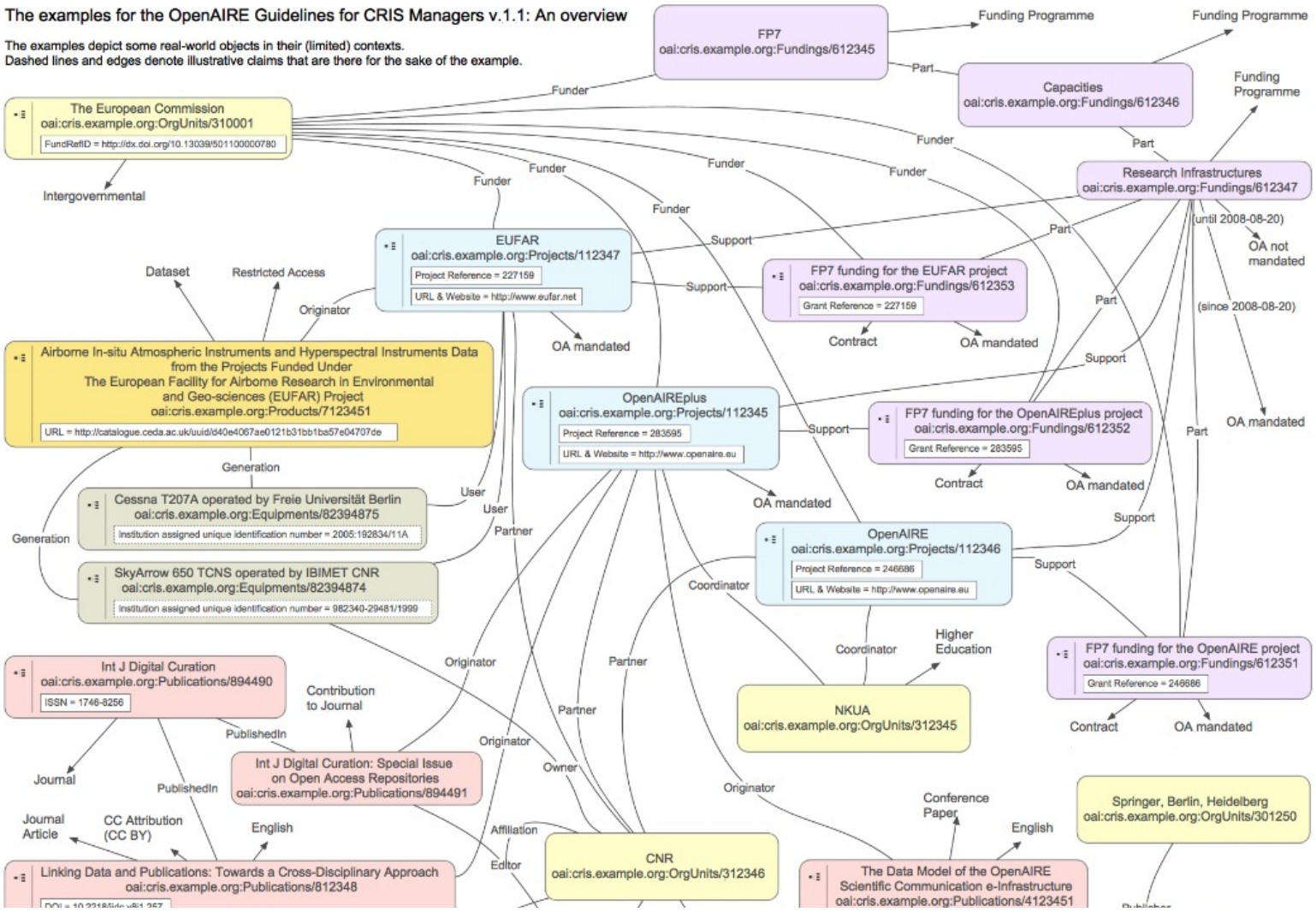
Plus a generic option

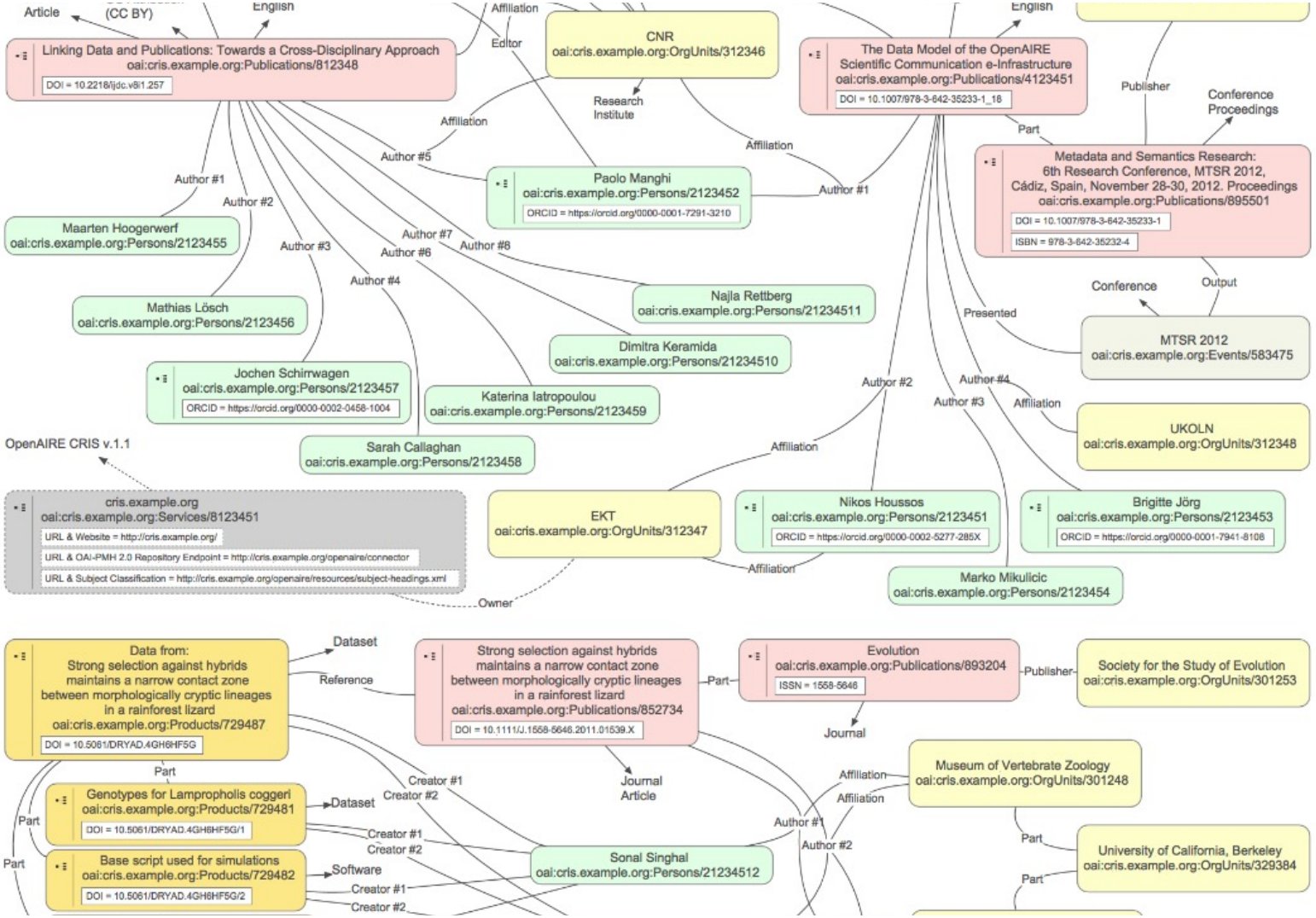
Supporting artifacts

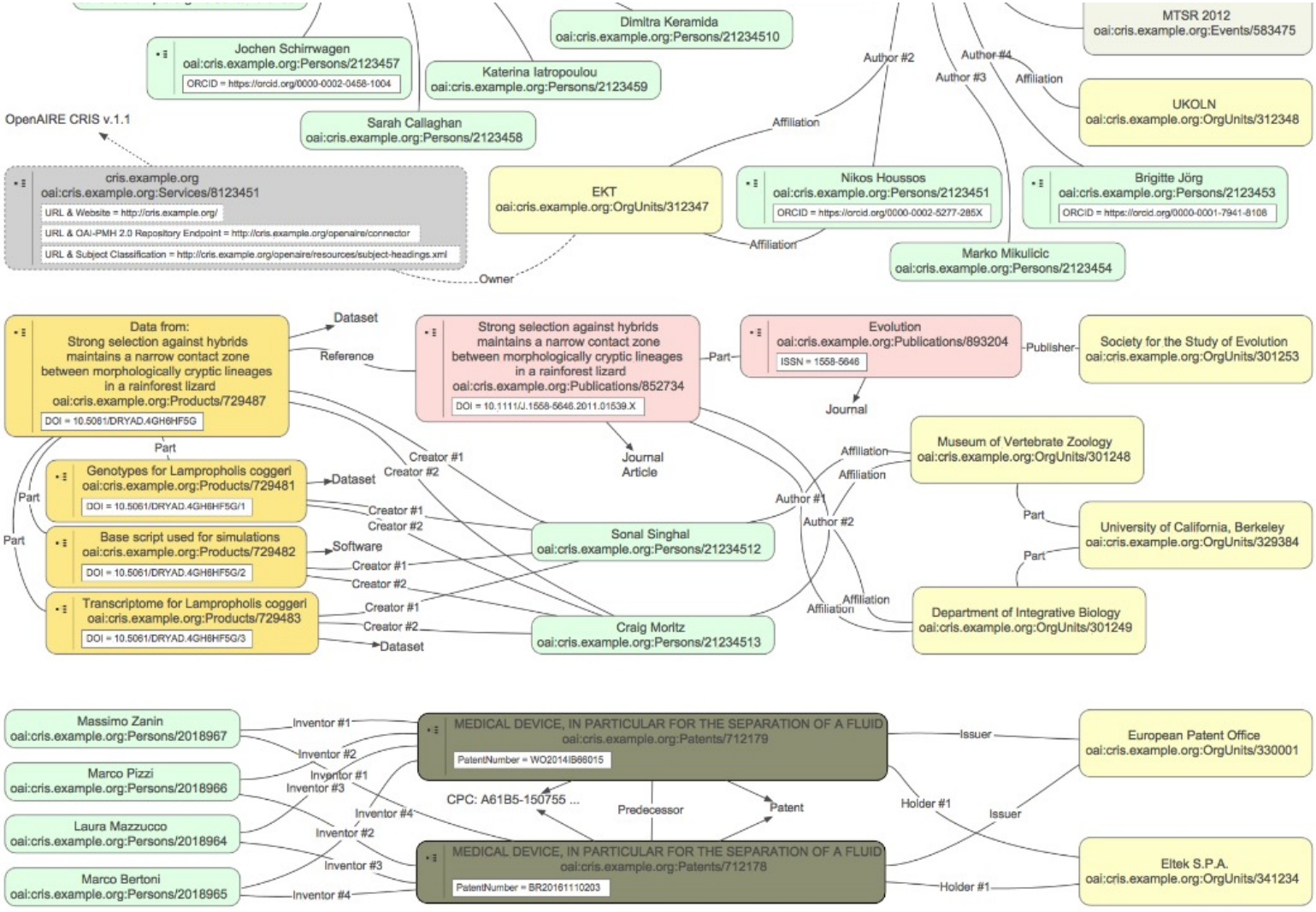
- XML Schema
  - Namespace https://www.openaire.eu/cerif-profile/1.1/
  - A few Schematron integrity rules embedded
- A comprehensive set of examples
- Prototype validator



A Research Graph:
The example set from the OpenAIRE Guidelines for CRIS Managers 1.1


The examples for the OpenAIRE Guidelines for CRIS Managers v.1.1: An overview
The examples depict some real-world objects in their (limited) contexts.
Dashed lines and edges denote illustrative claims that are there for the sake of the example.
FP7
oai:cris.example.org:Fundings/612345
Funding Programme
Funding Programme
Part
Capacities
oai:cris.example.org:Fundings/612346
Funding Programme
Part
Research Infrastructures
oai:cris.example.org:Fundings/612347
(until 2008-08-20)
OA not mandated
(since 2008-08-20)
OA mandated
Part
The European Commission
oai:cris.example.org:OrgUnits/310001
FundRefID = http://dx.doi.org/10.13039/501100000780
Intergovernmental
Funder
Funder
Funder
Funder
Funder
Funder
Funder
EUFAR
oai:cris.example.org:Projects/112347
Project Reference = 227159
URL & Website = http://www.eufar.net
OA mandated
Support
Support
FP7 funding for the EUFAR project
oai:cris.example.org:Fundings/612353
Grant Reference = 227159
Contract
OA mandated
Dataset
Restricted Access
Originator
Airborne In-situ Atmospheric Instruments and Hyperspectral Instruments Data from the Projects Funded Under The European Facility for Airborne Research in Environmental and Geo-sciences (EUFAR) Project
oai:cris.example.org:Products/7123451
URL = http://catalogue.ceda.ac.uk/uuid/d40e4067ae0121b31bb1ba57e04707de
Generation
Generation
Cessna T207A operated by Freie Universität Berlin
oai:cris.example.org:Equipments/82394875
Institution assigned unique identification number = 2005:192834/11A
SkyArrow 650 TCNS operated by IBIMET CNR
oai:cris.example.org:Equipments/82394874
Institution assigned unique identification number = 982340-29481/1999
User
User
Partner
OpenAIREplus
oai:cris.example.org:Projects/112345
Project Reference = 283595
URL & Website = http://www.openaire.eu
OA mandated
Support
FP7 funding for the OpenAIREplus project
oai:cris.example.org:Fundings/612352
Grant Reference = 283595
Contract
OA mandated
Part
Support
OpenAIRE
oai:cris.example.org:Projects/112346
Project Reference = 246686
URL & Website = http://www.openaire.eu
Support
Coordinator
Coordinator
Higher Education
NKUA
oai:cris.example.org:OrgUnits/312345
FP7 funding for the OpenAIRE project
oai:cris.example.org:Fundings/612351
Grant Reference = 246686
Contract
OA mandated
Int J Digital Curation
oai:cris.example.org:Publications/894490
ISSN = 1746-8256
Journal
PublishedIn
Contribution to Journal
Int J Digital Curation: Special Issue on Open Access Repositories
oai:cris.example.org:Publications/894491
PublishedIn
Journal Article
CC Attribution (CC BY)
English
Linking Data and Publications: Towards a Cross-Disciplinary Approach
oai:cris.example.org:Publications/812348
Originator
Partner
Partner
Originator
Owner
Affiliation
Editor
CNR
oai:cris.example.org:OrgUnits/312346
Originator
Conference Paper
English
The Data Model of the OpenAIRE Scientific Communication e-Infrastructure
oai:cris.example.org:Publications/4123451
Springer, Berlin, Heidelberg
oai:cris.example.org:OrgUnits/301250


Article
(CC BY)
English
Linking Data and Publications: Towards a Cross-Disciplinary Approach
oai:cris.example.org:Publications/812348
DOI = 10.2218/ijdc.v8i1.257
Affiliation
Editor
CNR
oai:cris.example.org:OrgUnits/312346
Research Institute
Affiliation
Author #5
Paolo Manghi
oai:cris.example.org:Persons/2123452
ORCID = https://orcid.org/0000-0001-7291-3210
Author #1
Author #1
Author #2
Maarten Hoogerwerf
oai:cris.example.org:Persons/2123455
Author #3
Author #7
Author #6
Author #8
Author #4
Najla Rettberg
oai:cris.example.org:Persons/21234511
Mathias Lösch
oai:cris.example.org:Persons/2123456
Dimitra Keramida
oai:cris.example.org:Persons/21234510
Jochen Schirrwagen
oai:cris.example.org:Persons/2123457
ORCID = https://orcid.org/0000-0002-0458-1004
Katerina Iatropoulou
oai:cris.example.org:Persons/2123459
Sarah Callaghan
oai:cris.example.org:Persons/2123458
OpenAIRE CRIS v.1.1
cris.example.org
oai:cris.example.org:Services/8123451
URL & Website = http://cris.example.org/
URL & OAI-PMH 2.0 Repository Endpoint = http://cris.example.org/openaire/connector
URL & Subject Classification = http://cris.example.org/openaire/resources/subject-headings.xml
Owner
EKT
oai:cris.example.org:OrgUnits/312347
Affiliation
Author #2
Author #3
Author #4
Nikos Houssos
oai:cris.example.org:Persons/2123451
ORCID = https://orcid.org/0000-0002-5277-285X
Brigitte Jörg
oai:cris.example.org:Persons/2123453
ORCID = https://orcid.org/0000-0001-7941-8108
Affiliation
Marko Mikulicic
oai:cris.example.org:Persons/2123454
English
The Data Model of the OpenAIRE Scientific Communication e-Infrastructure
oai:cris.example.org:Publications/4123451
DOI = 10.1007/978-3-642-35233-1_18
Publisher
Conference Proceedings
Part
Metadata and Semantics Research: 6th Research Conference, MTSR 2012, Cádiz, Spain, November 28-30, 2012. Proceedings
oai:cris.example.org:Publications/895501
DOI = 10.1007/978-3-642-35233-1
ISBN = 978-3-642-35232-4
Conference
Output
Presented
MTSR 2012
oai:cris.example.org:Events/583475
Affiliation
UKOLN
oai:cris.example.org:OrgUnits/312348
Data from: Strong selection against hybrids maintains a narrow contact zone between morphologically cryptic lineages in a rainforest lizard
oai:cris.example.org:Products/729487
DOI = 10.5061/DRYAD.4GH6HF5G
Dataset
Reference
Strong selection against hybrids maintains a narrow contact zone between morphologically cryptic lineages in a rainforest lizard
oai:cris.example.org:Publications/852734
DOI = 10.1111/J.1558-5646.2011.01539.X
Part
Evolution
oai:cris.example.org:Publications/893204
ISSN = 1558-5646
Publisher
Society for the Study of Evolution
oai:cris.example.org:OrgUnits/301253
Journal
Part
Genotypes for Lampropholis coggeri
oai:cris.example.org:Products/729481
DOI = 10.5061/DRYAD.4GH6HF5G/1
Dataset
Creator #1
Creator #2
Part
Base script used for simulations
oai:cris.example.org:Products/729482
DOI = 10.5061/DRYAD.4GH6HF5G/2
Software
Creator #1
Creator #2
Journal Article
Sonal Singhal
oai:cris.example.org:Persons/21234512
Author #1
Author #2
Affiliation
Affiliation
Museum of Vertebrate Zoology
oai:cris.example.org:OrgUnits/301248
Part
University of California, Berkeley
oai:cris.example.org:OrgUnits/329384
Part


Dimitra Keramida
oai:cris.example.org:Persons/21234510
MTSR 2012
oai:cris.example.org:Events/583475
Jochen Schirrwagen
oai:cris.example.org:Persons/2123457
ORCID = https://orcid.org/0000-0002-0458-1004
Katerina Iatropoulou
oai:cris.example.org:Persons/2123459
Author #2
Author #4
Author #3
Affiliation
UKOLN
oai:cris.example.org:OrgUnits/312348
OpenAIRE CRIS v.1.1
Sarah Callaghan
oai:cris.example.org:Persons/2123458
Affiliation
cris.example.org
oai:cris.example.org:Services/8123451
URL & Website = http://cris.example.org/
URL & OAI-PMH 2.0 Repository Endpoint = http://cris.example.org/openaire/connector
URL & Subject Classification = http://cris.example.org/openaire/resources/subject-headings.xml
EKT
oai:cris.example.org:OrgUnits/312347
Nikos Houssos
oai:cris.example.org:Persons/2123451
ORCID = https://orcid.org/0000-0002-5277-285X
Brigitte Jörg
oai:cris.example.org:Persons/2123453
ORCID = https://orcid.org/0000-0001-7941-8108
Affiliation
Marko Mikulicic
oai:cris.example.org:Persons/2123454
Owner
Data from:
Strong selection against hybrids maintains a narrow contact zone between morphologically cryptic lineages in a rainforest lizard
oai:cris.example.org:Products/729487
DOI = 10.5061/DRYAD.4GH8HF5G
Dataset
Reference
Strong selection against hybrids maintains a narrow contact zone between morphologically cryptic lineages in a rainforest lizard
oai:cris.example.org:Publications/852734
DOI = 10.1111/J.1558-5646.2011.01539.X
Part
Evolution
oai:cris.example.org:Publications/893204
ISSN = 1558-5646
Publisher
Society for the Study of Evolution
oai:cris.example.org:OrgUnits/301253
Journal
Journal Article
Part
Genotypes for Lampropholis coggeri
oai:cris.example.org:Products/729481
DOI = 10.5061/DRYAD.4GH8HF5G/1
Dataset
Base script used for simulations
oai:cris.example.org:Products/729482
DOI = 10.5061/DRYAD.4GH8HF5G/2
Software
Transcriptome for Lampropholis coggeri
oai:cris.example.org:Products/729483
DOI = 10.5061/DRYAD.4GH8HF5G/3
Dataset
Creator #1
Creator #2
Sonal Singhal
oai:cris.example.org:Persons/21234512
Craig Moritz
oai:cris.example.org:Persons/21234513
Author #1
Author #2
Affiliation
Museum of Vertebrate Zoology
oai:cris.example.org:OrgUnits/301248
University of California, Berkeley
oai:cris.example.org:OrgUnits/329384
Department of Integrative Biology
oai:cris.example.org:OrgUnits/301249
Part
Massimo Zanin
oai:cris.example.org:Persons/2018967
Marco Pizzi
oai:cris.example.org:Persons/2018966
Laura Mazzucco
oai:cris.example.org:Persons/2018964
Marco Bertoni
oai:cris.example.org:Persons/2018965
Inventor #1
Inventor #2
Inventor #3
Inventor #4
MEDICAL DEVICE, IN PARTICULAR FOR THE SEPARATION OF A FLUID
oai:cris.example.org:Patents/712179
PatentNumber = WO2014IB66015
CPC: A61B5-150755 ...
Predecessor
Patent
MEDICAL DEVICE, IN PARTICULAR FOR THE SEPARATION OF A FLUID
oai:cris.example.org:Patents/712178
PatentNumber = BR20161110203
Issuer
Holder #1
European Patent Office
oai:cris.example.org:OrgUnits/330001
Eltek S.P.A.
oai:cris.example.org:OrgUnits/341234

<Publication xmlns="*https://www.openaire.eu/cerif-profile/1.1/*" id="***852734***">
<Type xmlns="*https://www.openaire.eu/cerif-profile/vocab/COAR_Publication_Types*">**http://purl.org/coar/resource_type/c_6501**<!-- journal article --></Type>
<Title xml:lang="***en***">**Strong selection against hybrids maintains a narrow contact zone between morphologically cryptic lineages in a rainforest lizard**</Title>
<PublishedIn>
<Publication id="***893204***">
<Type scheme="*https://www.openaire.eu/cerif-profile/vocab/COAR_Publication_Types*">**http://purl.org/coar/resource_type/c_0640**<!-- journal --></Type>
<Title xml:lang="***en***">**Evolution**</Title>
<ISSN medium="http://issn.org/vocabularies/Medium#Online">**1558-5646**</ISSN>
<Publishers>
<!-- […] -->
</Publishers>
</Publication>
</PublishedIn>
<Volume>**66**</Volume>
<Issue>**5**</Issue>
<StartPage>**1474**</StartPage>
<EndPage>**1489**</EndPage>
<DOI>**10.1111/J.1558-5646.2011.01539.X**</DOI>
<!-- […] -->
</Publication>

```
<Publication xmlns="https://www.openaire.eu/cerif-profile/1.1/" id="852734">
    <!-- [ … ] -->
    <Authors>
        <Author>
            <Person>
                <PersonName>
                    <FamilyNames>Singhal</FamilyNames>
                    <FirstNames>Sonal</FirstNames>
                </PersonName>
            </Person>
            <Affiliation>
                <OrgUnit id="301248">
                    <Name xml:lang="en">Museum of Vertebrate Zoology</Name>
                    <PartOf>
                        <OrgUnit id="329384">
                            <Name xml:lang="en">University of California, Berkeley</Name>
                        </OrgUnit>
                    </PartOf>
                </OrgUnit>
            </Affiliation>  <!-- [ … ] -->
        </Author>  <!-- [ … ] -->
    </Authors>
    <!-- [ … ] -->
</Publication>
```

<Publication xmlns="*https://www.openaire.eu/cerif-profile/1.1/*" id="***852734***">

<!-- […] -->

<Keyword xml:lang="***en***">**hybrid zones**</Keyword>

<Keyword xml:lang="***en***">**phylogeography**</Keyword>

<Abstract xml:lang="***en***">**Phenotypically cryptic lineages comprise an important yet understudied part of biodiversity;**
<!-- […] --></Abstract>

<References>

<Product id="***729487***">

<Type xmlns="*https://www.openaire.eu/cerif-profile/vocab/COAR_Product_Types*">**http://purl.org/coar/resource_type/c_ddb1**<!-- dataset --></Type>

<Name xml:lang="***en***">**Data from: Strong selection against hybrids maintains a narrow contact zone between morphologically cryptic lineages in a rainforest lizard**</Name>

<VersionInfo xml:lang="***en***">**1**</VersionInfo>

<DOI>**10.5061/DRYAD.4GH6HF5G**</DOI>

<!-- […] -->

</Product>

</References>

</Publication>

CERIF Profiles

Producers know what to include

Consumers know what to expect

CERIF Profiles

Useful subsets of CERIF
for specific research information
interchange scenarios

- Entities & attributes:
  - Profile ⊂ CERIF
- Semantic vocabularies:
  - Profile – specific choices
  - Sources: CERIF & beyond
- Integrity constraints:
  - Profile ⊃ CERIF

Profile data is CERIF

CERIF Refactoring

- Address shortcomings of current CERIF
- Modernize CERIF
  - Take away the perceived complexity of CERIF
  - Involve the community in further development of CERIF
  - New modelling approach (and tool), modularity, modern serialization formats
- Ongoing project started 2021

CERIF REFACTORING

CERIF Refactoring Goals

1. Keep the strong features of current CERIF
   a. Multilinguality
2. Take away the perceived complexity of CERIF
   a. Emphasize the conceptual model
   b. Improve documentation
3. Modernize CERIF
   a. Change the modeling notation
   b. Adapt CERIF for usage in APIs and for Linked Open Data
   c. Modern serialization formats
   d. Allow for systematic provenance tracking and verifiable credentials
4. Involve the community in further development of CERIF
   a. Modularity
   b. Open source sw development practices

CERIF Refactoring Means

1. New modelling approach (and tool)
   a. Improved documentation
   b. Element addressability

2. Modularity (⇒ scalability of further CERIF development)
   a. The Core and pluggable modules
   b. Datatypes to represent reusable functional groups
   c. Developed in a decentralized way

3. Modern serialization formats (with appropriate schema expressions)
   a. JSON-LD (JSON Schema, SHACL) for APIs
   b. RDF (RDF Schema, SHACL) for LOD
   c. previously: XML (XML Schema, Schematron)

CERIF Core

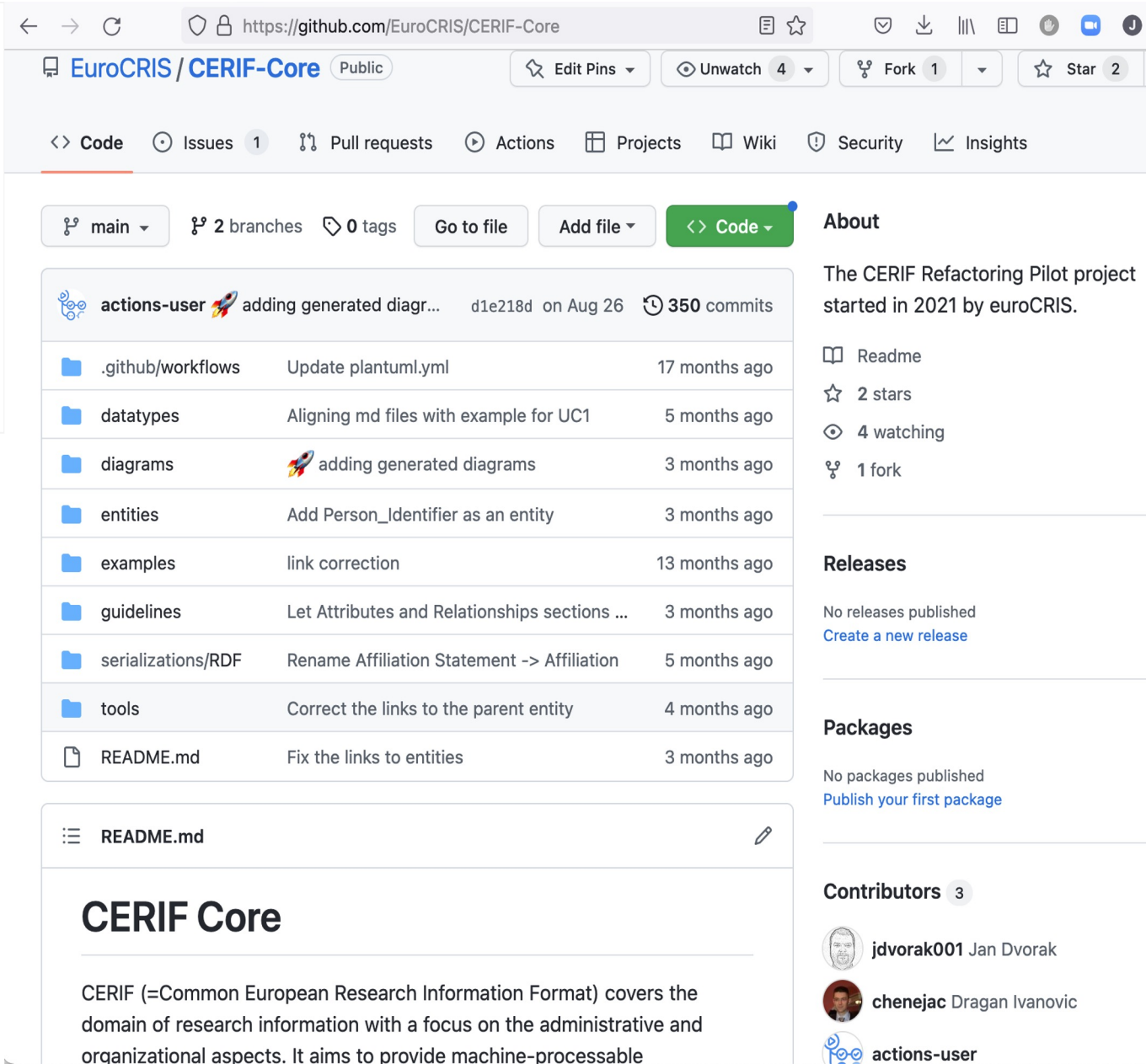
Contains:

- common data types
- abstract entities
- commonly used entities
- common constructs

The basis for:

- Particular research information interchange scenarios
- Extensions in terms of coverage

A prototype CERIF Core repository at https://github.com/EuroCRIS/CERIF-Core



CERIF Agent

Def: Anything that has the ability to perform actions.

Generalizes the former cfPerson + cfOrgUnit

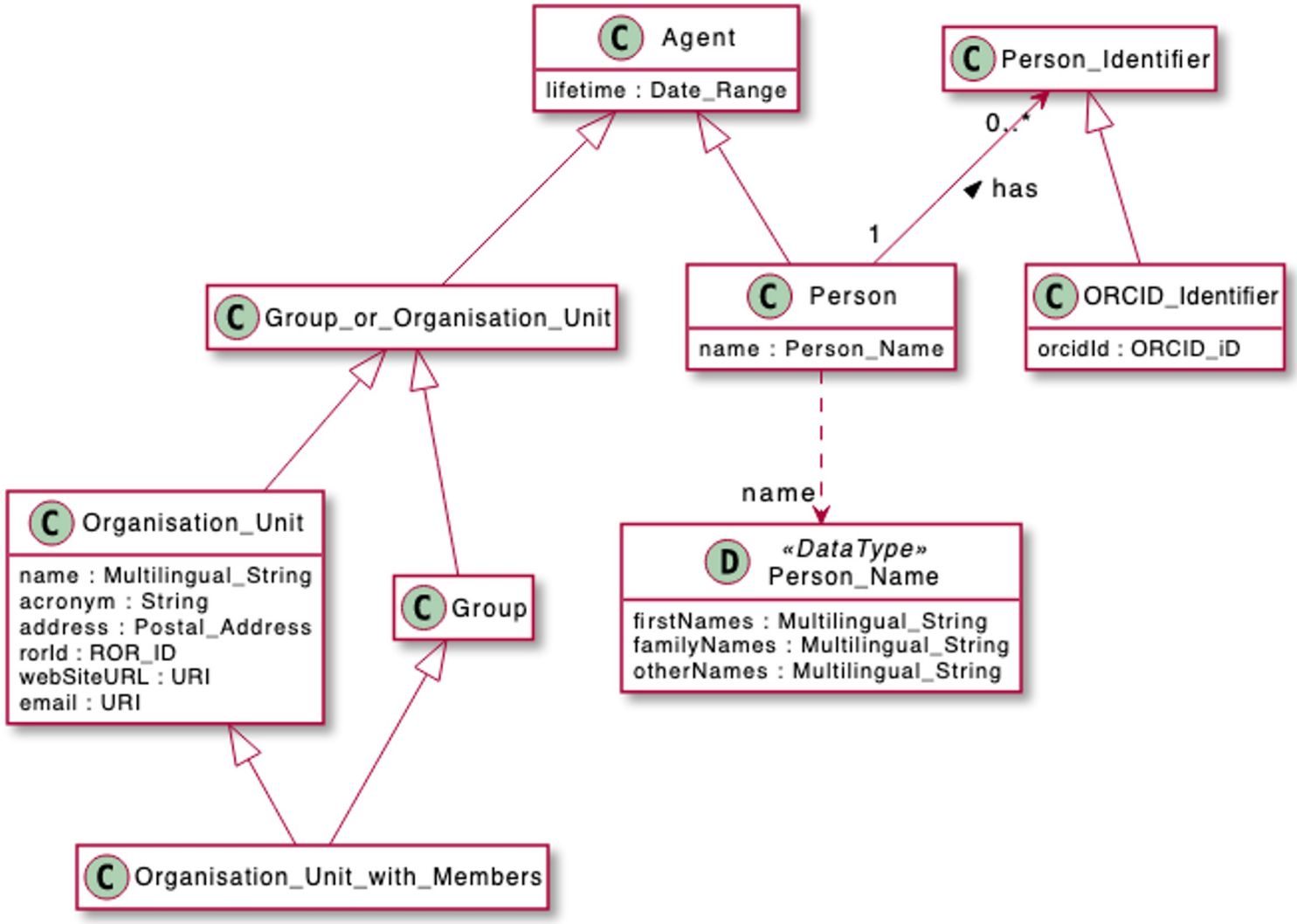
Subclassed as:

- Person
- Group_or_Organisation_Unit
  - Organisation_Unit (formalized)
  - Group (has members)

Organisation_Unit_with_Members

(following FOAF)

The Agent Hierarch



CERIF Activity

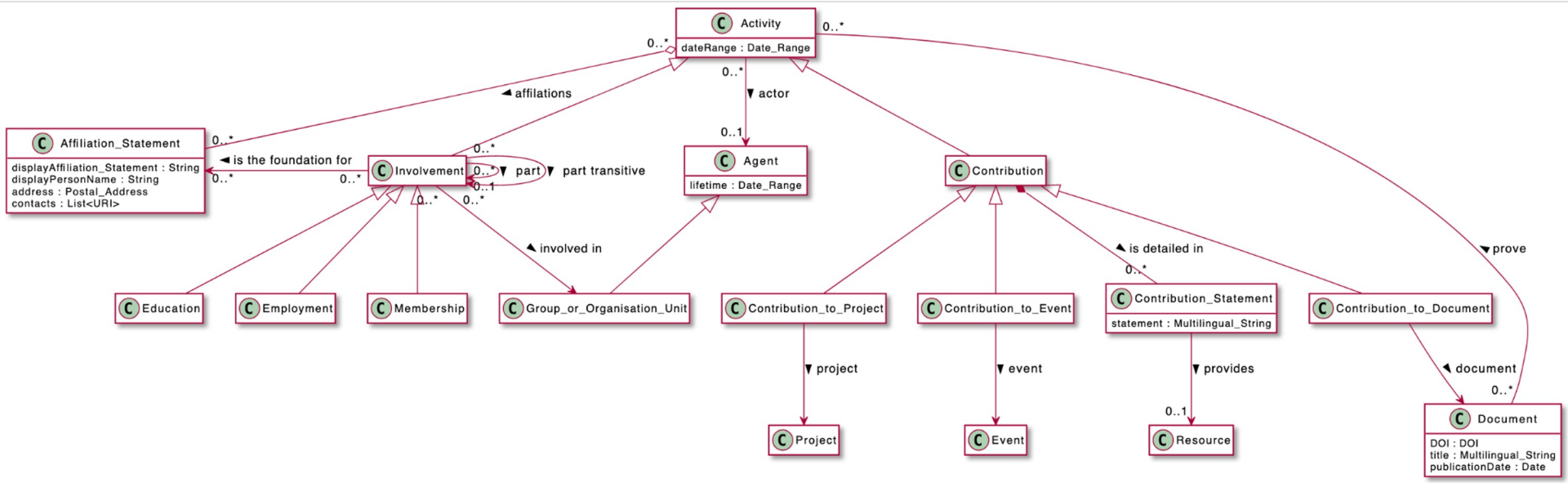
Def: Anything an Agent does.

Subclassed as:

- Involvement (Membership, Employment or Education) in a Group or Organisation Unit
- Contribution to a Document, a Project, an Event, …

This is a refinement of the Linking Entities of current CERIF starting from cfPerson or cfOrgUnit
(dateRange is an attribute of Activity)

Activity hierarchy



CERIF AffiliationStatement

To give context to the Agent in his/her/its Activities.

Person is member

Person with affiliation to OrgUnit is member

OrgUnit is member, reprepresented by Person

OrgUnit is member, reprepresented by Position

OrgUnit is member

Where does the person come from? Why are they there?

Who represents the orgunit? Who should we talk to?

CERIF Project & Funding

Project: A temporary endeavor undertaken to create a unique product, service or result.
(Source: the Project Management Institute)

Funding: Money provided by an organization or government for a particular purpose.
(Source: the Oxford Learner's Dictionary of Academic English)

Common confusion: funders, institutions and researchers alike may say "Project" when they mean "Project with Funding" or "Funding".

Funding: a specific type of Resource

It's not just Funding that enables research. There is also:

- The expertise, time & effort of Researchers
- Prior knowledge (from Documents)
- Housing the research, the administrative backing by Institutions
- Services offered by Facilities:
  - Laboratories
  - Supercomputing centres
  - Data archives
  - Libraries
  - ...
- Other Services
  - Publishing

Most of these can be had in exchange for some Funding

Resources exist in different modes

Planned

Sought for

Constructed / Secured

Offered / Invitation to use

Considered

Applied for / Negotiated

Approved / Agreed upon

Contracted

Used

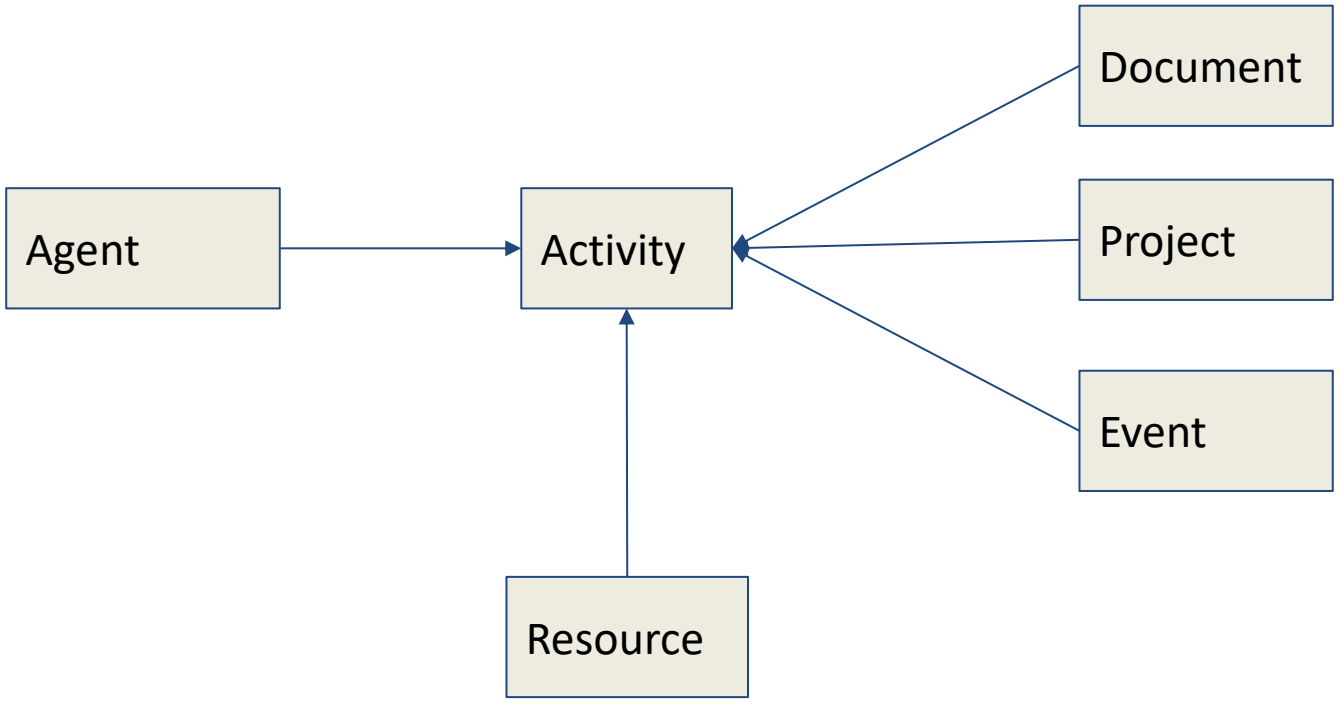
CERIF Document

Def: A document is a bounded representation of a body of information designed with the capacity (and usually intent) to communicate. A document may manifest symbolic, diagrammatic or sensory-representational information.

(Source: The Bibo Ontology, http://purl.org/ontology/bibo/Document)

Accommodates current cfResultPublication, cfResultPatent and some of cfResultProduct

CERIF Core “View from the Top”



CERIF Modules

Add functionality

- E.g. the Scholarly publication module

Customize the model for a specific scenario

- E.g. institutions to report publications produced with Horizon Europe funding to the EC (tentative)

Contributed by another group of people

Reuse & Extend stuff from the Core and other modules

Candidate modules:

- Dataset metadata
- Research software metadata
- Bibliometric indicators
- Project reporting
- Research Infrastructure description

CERIF Scholarly Publication Module

1st module started along with the Core

Scope:

- Journal Article
- Journal
- Monograph
- (Book Chapter)
- (Proceedings Paper)

https://github.com/EuroCRIS/CERIF-ScholarlyPublicationsModule

EuroCRIS / CERIF-ScholarlyPublicationsModule Public

Edit Pins | Unwatch 3 | Fork 0 | Star 0

Code | Issues | Pull requests | Actions | Projects | Wiki | Security | Insights | Settings

main | 1 branch | 0 tags | Go to file | Add file | Code

jdvorak001 Add human-readable anchors to the journal-articles relationship — 7fe3b7a on 10 Mar — 87 commits

.github/workflows	Update plantuml.yml	6 months ago
datatypes	The initial creation of datatypes ISBN and ISSN, and update of Editor...	10 months ago
diagrams	🚀 adding generated diagrams	2 months ago
entities	Add human-readable anchors to the journal-articles relationship	2 months ago
examples	link correction	6 months ago
serializations/RDF	adding examples directory	6 months ago
README.md	The Development section now references the Core	6 months ago

README.md

CERIF Scholarly Publications Module

A module to represent scholarly publications in the refactored CERIF. Its secondary purpose is to demonstrate the liason between the CERIF Core and a module.

Status

About

No description, website, or topics provided.

Readme
0 stars
3 watching
0 forks

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Contributors 3

chenejac Dragan Ivanovic
jdvorak001 Jan Dvorak

Thank you!