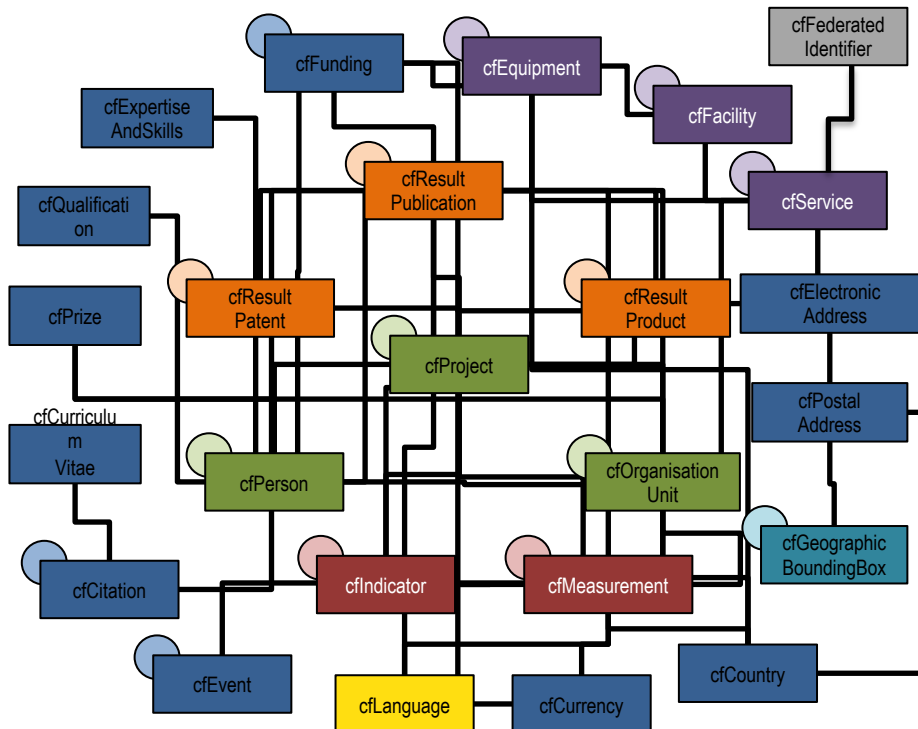


CERIF Tutorial



Jan Dvořák

May 11th, 2022

CRIS 2022 Conference
Dubrovnik, Croatia

Jan Dvořák

 <https://orcid.org/0000-0001-8985-152X>

 jan.dvorak@ff.cuni.cz

euroCRIS

- CERIF TG Leader since 2013

Charles University, Institute of Information Studies and Librarianship

- Researcher & Lecturer

Czech Technical University in Prague, Computing and Information Centre

- IS Analyst: V3S+EZOP – 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

*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.*

Research Information

= Information about research

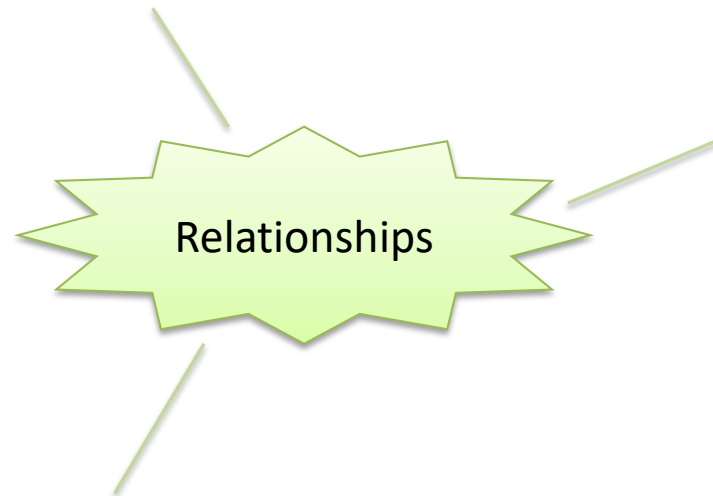
Research Information

The process of research

- Research projects
- Funding
- Research infrastructures

The research actors

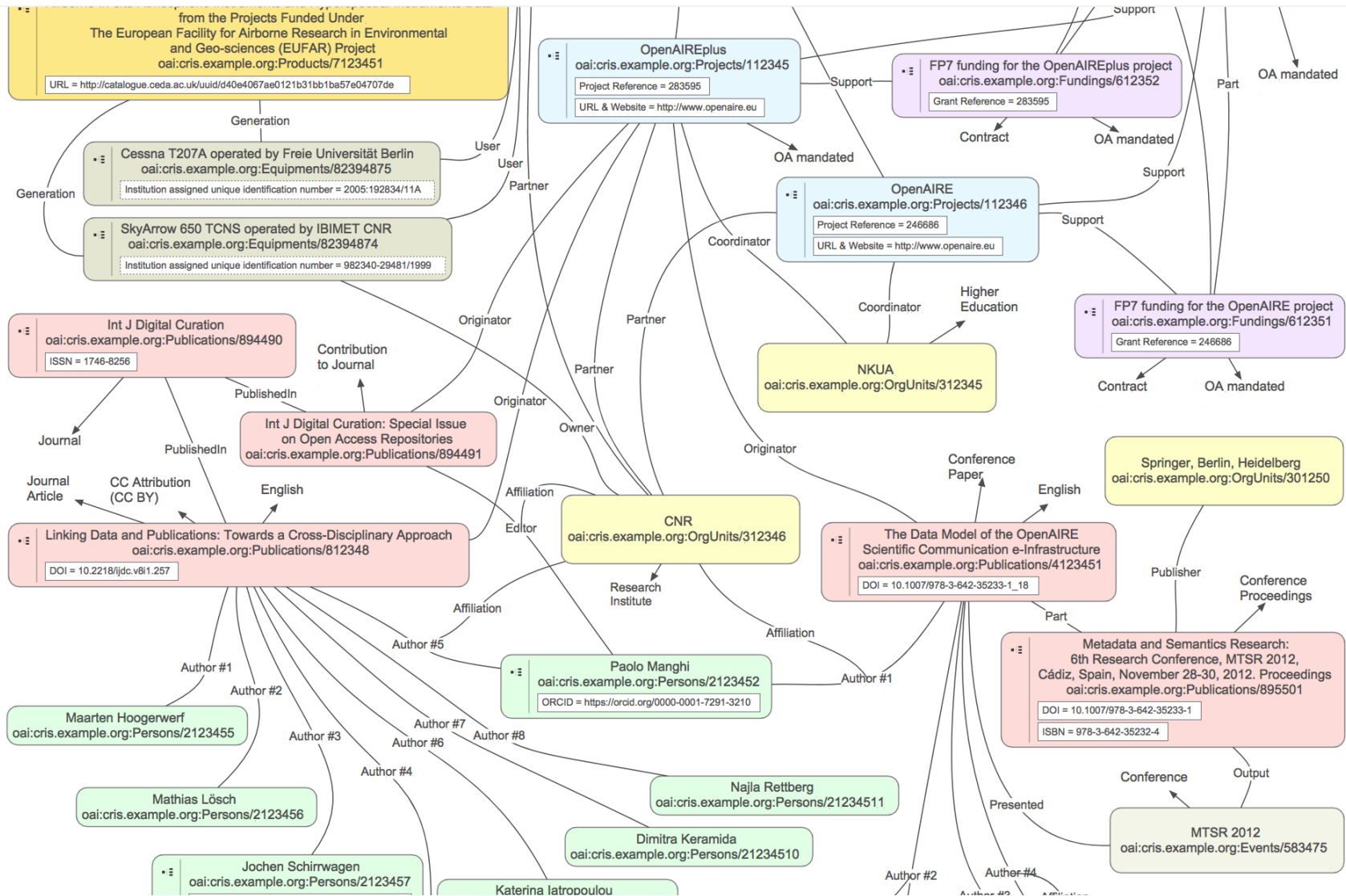
- Researchers
- Institutions
- Funders
- Publishers
- Facility operators
- Associations



Research results

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

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 characterises 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 research domain]

A web page (URI)

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

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	Topic(s): FP7-PEOPLE-2012-IRSES - Marie Curie Action "International Research Staff Exchange Scheme"
EU contribution: EUR 126 000	Call for proposal: FP7-PEOPLE-2012-IRSES
Coordinated in: Ireland	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 cattle..."

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 

- [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](#)

Home About this system Case Studies Contact us

RESEARCH COUNCILS UK Gateway to Research Innovate UK Technology Strategy Board

Social Media Analysis for Social Geography

Lead Research Organisation: [Newcastle University](#)
Department Name: Computing Sciences

[Go back](#)

Overview Organisations People Outcomes

Abstract

Many believe that social media data has the potential to give real insights into communities. However, there is a real challenge in unlocking this knowledge due to the enormous size and speed with which the data is generated, and a lack of understanding of issues such as ethics and the validity of conclusions drawn from the skewed sample of the population that social media users represent.

"Tweet My Street" is tackling these problems with the overall goal of understanding how the knowledge gained can benefit socially excluded communities. It is a multi-disciplinary collaboration involving social geographers, computer scientists and statisticians, which aims to:

- design a scalable system for social media data analysis
- create a tool that makes it easy for social scientists to ask questions and visualise results
- explore the use of modern statistical techniques to derive understanding from the data
- drive and evaluate the system through case studies, including the use of the new tools to augment existing social science methods, and to understand the way third sector organisations can make effective use of social media
- explore how the use of the tools can help the private and public sector to make key decisions.

Funded Value: £449,388
Funded Period: Feb 14 - Feb 16
Funder: ESRC
Project Status: Active
Project Category: Research Grant
Project Reference: ES/M001636/1
Principal Investigator: [Paul Watson](#)
Research Subject: Human Geography
Research Topic: Social Geography Urban Geography

A name or title

A short or long description (abstract)

[A project coordinator]

[A source of funding]

A (planned) start date

A (planned) end date or duration

A code (identifier), for ex a Grant number

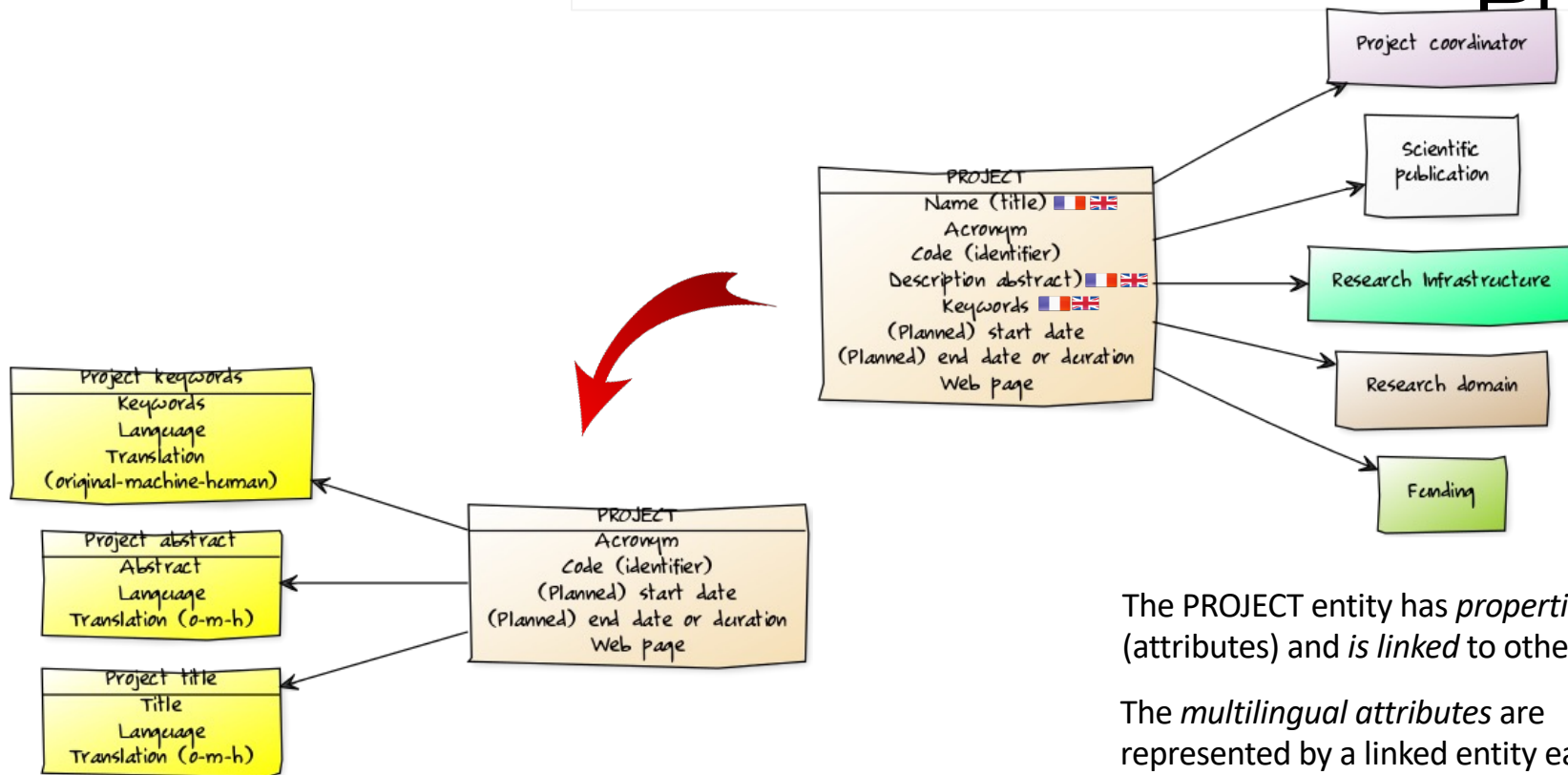
A web page (URI)

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

[A research domain]

A few keywords

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*

cfProjTitle			
cfProjId	ID	NN (PFK)	
cfLangCode	Char(5 BYTE)	NN (PFK)	
cfTrans	NChar(1)	NN (PK)	
cfTitle	Char(255 BYTE)	NN	

cfProjAbstr			
cfProjId	ID	NN (PFK)	
cfLangCode	Char(5 BYTE)	NN (PFK)	
cfTrans	NChar(1)	NN (PK)	
cfAbstr	NClob	NN	

cfProjKeyw			
cfProjId	ID	NN (PFK)	
cfLangCode	Char(5 BYTE)	NN (PFK)	
cfTrans	NChar(1)	NN (PK)	
cfKeyw	Char(255 BYTE)	NN	

Project keywords
Keywords
Language
Translation (original-machine-human)

Project abstract
Abstract
Language
Translation (o-m-h)

Project title
Title
Language
Translation (o-m-h)

PROJECT
Acronym
Code (identifier)
(Planned) start date
(Planned) end date or duration
Web page

cfPrj			
cfProjId	ID	NN (PK)	
cfStartDate	Date		
cfEndDate	Date		
cfAcro	Char(16 BYTE)		
cfURI	Char(128 BYTE)		

Representation in Database

Format, Unicity, Not-null, Foreign Key (FK), composed Primary Key (set of PFK and PK)

cfPrj		
cfProjId	ID	NN (PK)
cfStartDate	Date	
cfEndDate	Date	
cfAcro	Char(16 BYTE)	
cfURI	Char(128 BYTE)	

cfPrjAbstr		
cfProjId	ID	NN (PFK)
cfLangCode	Char(5 BYTE)	NN (PFK)
cfTrans	NChar(1)	NN (PK)
cfAbstr	NClob	NN

Key	Attribute Name	Domain	Data Type	Not Null	Unique
PFK	cfProjId	ID	Char(128 BYTE)	YES	NO
PFK	cfLangCode		Char(5 BYTE)	YES	NO
PK	cfTrans		NChar(1)	YES	NO
	cfAbstr		NClob	YES	NO

Key Type	Constraint Name	Attributes
PK	pk_cfProjAbstr	cfProjId, cfLangCode, cfTrans

Key	Attribute Name	Domain	Data Type	Not Null	Unique	Check	Default	Comments
PK	cfProjId	ID	Char(128 BYTE)	YES	NO	NO		The project identifier (cfProjId) propagates to system-internal project-related entities: cfProjectTitle cfProjectAbstract cfProjectKeywords cfProject_Service cfProject_Event cfProject_Facility cfProject_Equipment cfProject_Classification cfProject_ResultPatent cfProject_ResultProduct cfProject_ResultPublication cfProject_Project cfProject_Person cfProject_PrizeAward cfProject_Funding cfProject_OrganisationUnit cfProject_Medium cfProject_Indicator cfProject_Measurement
	cfStartDate		Date	NO	NO	NO		{@deprecated 1.6} use the cfStartDate of an appropriate unary classification of this cfProject or an appropriate link between this cfProject and the project participant (cfOrganisationalUnit or cfPerson)
	cfEndDate		Date	NO	NO	NO		{@deprecated 1.6} use the cfEndDate of an appropriate unary classification of this cfProject or an appropriate link between this cfProject and the project participant (cfOrganisationalUnit or cfPerson)
	cfAcro		Char(16 BYTE)	NO	NO	NO		
	cfURI		Char(128 BYTE)	NO	NO	NO		

Example in DB

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...

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

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

Representation in XML

Enclosing XML element = CERIF entity physical name (cfProj)

Enclosed XML elements = CERIF entity's attributes (cfProjId, cfAcro,...)

```
<cfProj>
  <cfProjId>system-internal-project-identifier</cfProjId>
  <cfAcro>project acronym</cfAcro>
  <cfTitle cfLang="en" cfTrans="o">original project title in English</cfTitle>
  <cfTitle cfLang="fr" cfTrans="h">le titre du projet en français, traduction humaine</cfTitle>
  <cfTitle cfLang="de" cfTrans="m">der Titel des Projekts auf Deutsch, maschinelle Übersetzung</cfTitle>
</cfProj>
```

XML attributes are used for multilingual CERIF attributes

cfLang, cfTrans:

- o for original language
- h for human translation
- m for machine translation

Representation and example in Linked Data

CERIF entity

Attributes

Multilingual attributes

```
@prefix xsd:      <http://www.w3.org/2001/XMLSchema#> .
@prefix rdfs:     <http://www.w3.org/2000/01/rdf-schema#> .
@prefix foaf:     <http://xmlns.com/foaf/0.1/> .
@prefix cerif:    <http://eurocris.org/cerif#> .
@prefix dc:       <http://purl.org/dc/elements/1.1/> .
@prefix dcterms:  <http://purl.org/dc/terms/> .

<http://cris.myOrganization.org/resource/projects/VOA3R>
a      cerif:Project ;

rdfs:label "Repositorio de Agricultura y Acuicultura de acceso abierto virtual"@es-es ,
"Virtual Open Access Agriculture & Aquaculture Repository"@en-uk ;

dc:title "Repositorio de Agricultura y Acuicultura de acceso abierto virtual"@es-es ,
"Virtual Open Access Agriculture & Aquaculture Repository"@en-uk ;

cerif:title "Repositorio de Agricultura y Acuicultura de acceso abierto virtual"@es-es ,
"Virtual Open Access Agriculture & Aquaculture Repository"@en-uk ;

foaf:homepage <http://voa3r.eu/> ;
cerif:uri <http://voa3r.eu/> ;

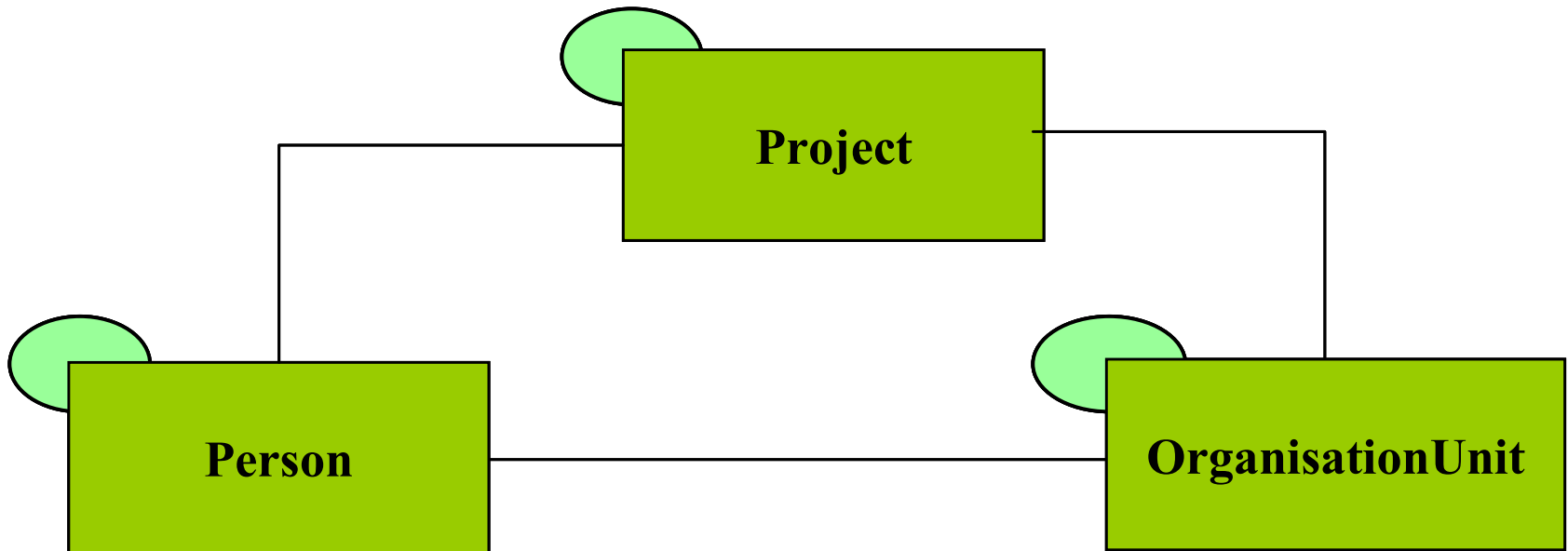
cerif:abstract "The general objective of the VOA3R project is to improve the spread of
European agriculture and aquaculture research results by using an innovative approach to
sharing open access research products. "@en-uk , "El objetivo general del proyecto VOA3R
es mejorar la difusión de los resultados de investigación en la agricultura y
acuicultura europea, mediante el uso de un enfoque innovador para el intercambio de
productos de investigación de acceso abierto. "@es-es ;

dcterms:abstract "The general objective of the VOA3R project is to improve the spread of
European agriculture and aquaculture research results by using an innovative approach to
sharing open access research products. "@en-uk , "El objetivo general del proyecto VOA3R
es mejorar la difusión de los resultados de investigación en la agricultura y
acuicultura europea, mediante el uso de un enfoque innovador para el intercambio de
productos de investigación de acceso abierto. "@es-es ;

cerif:internalIdentifier "ff8080812ddb916a012ddb9170b60001" ;
cerif:acronym "VOA3R" ;
cerif:startDate "2010-06-01"^^xsd:date ;
cerif:endDate "2013-05-31"^^xsd:date ;
```

Source in 2016: <http://cerif-linked-data.googlecode.com/files/Proposal%20of%20Recommendations%20-%20Report.docx>

CERIF Base Entities



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.

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.

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.

INTERMEDIATE SUMMARY

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



Common
European
Research
Information
Format

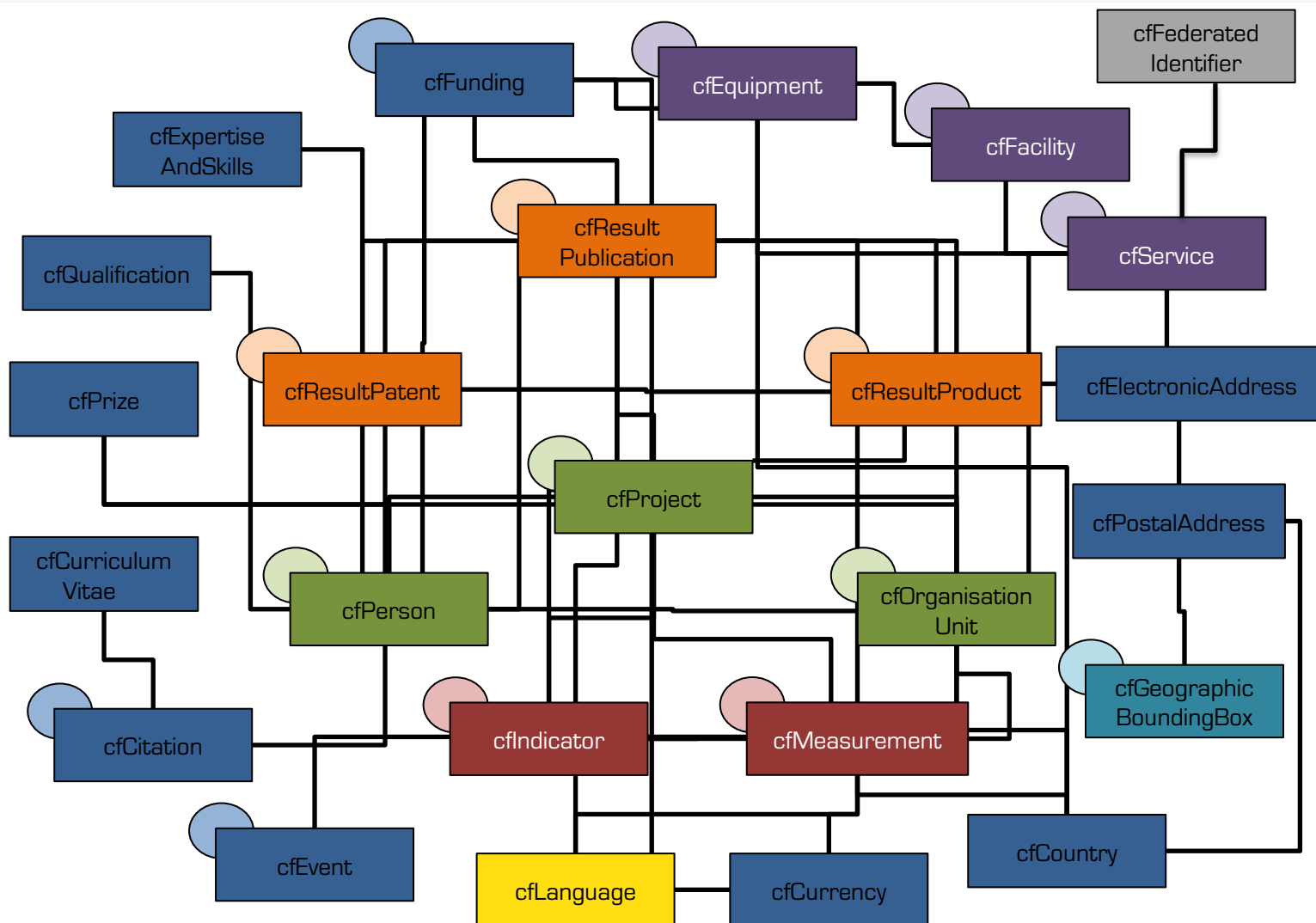
We have seen how to represent, store or exchange metadata about research projects.

Similarly:

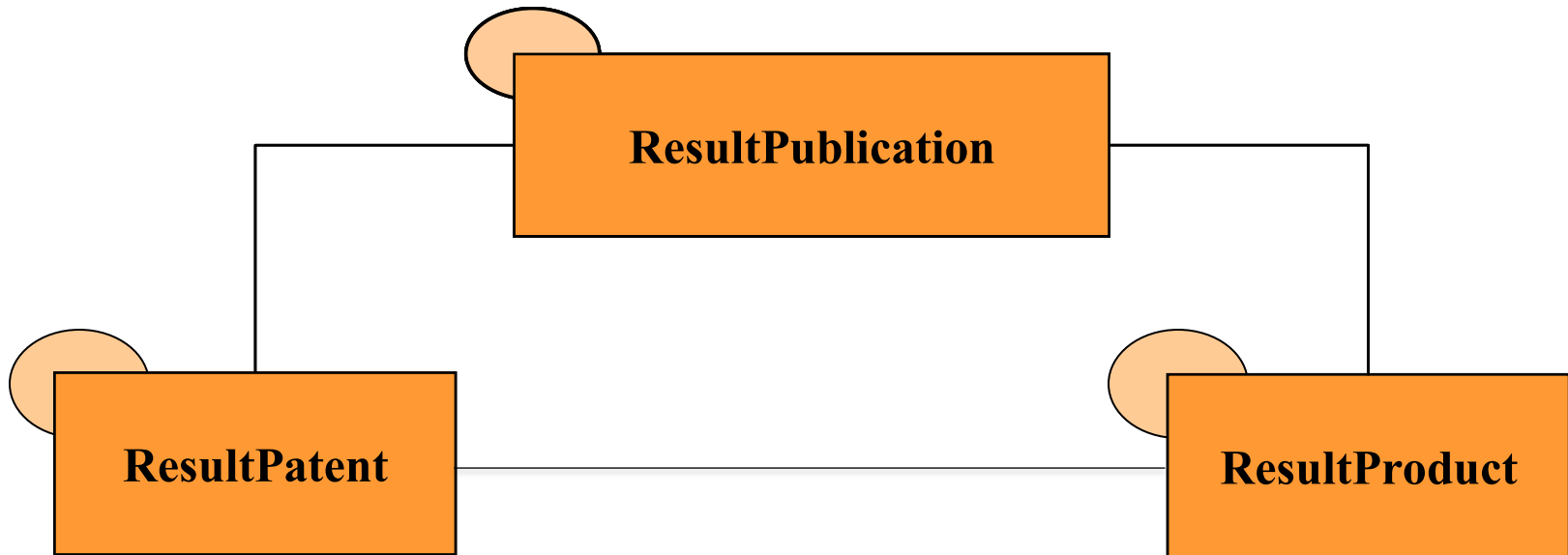
- What characterises a person (researcher, Ph.D.,...)?
- What characterises an organisation (research laboratory, institute,...)?

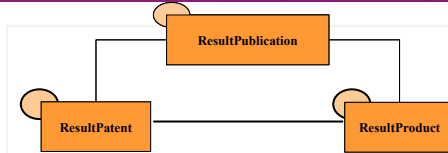


Common European Research Information Format

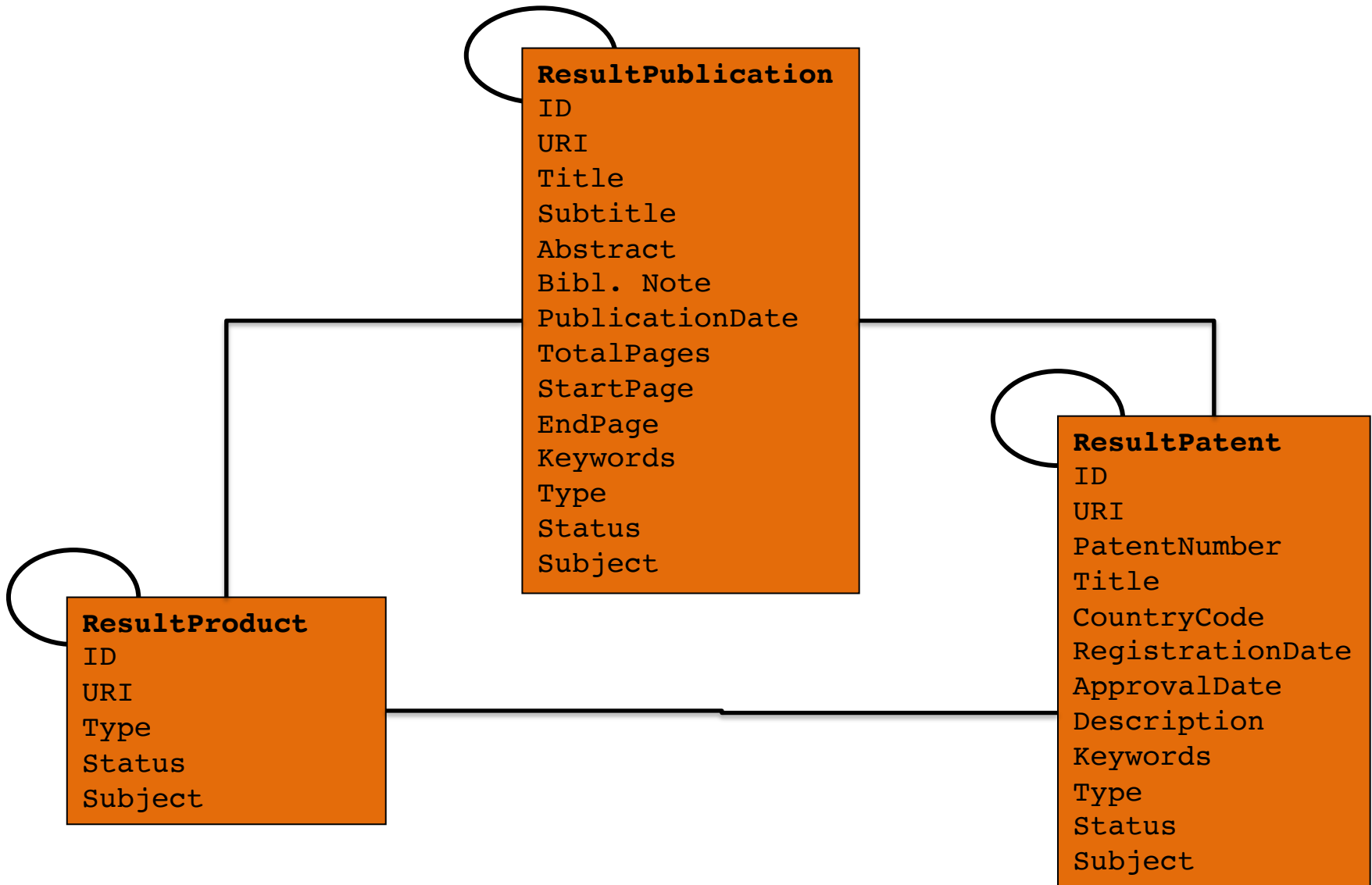


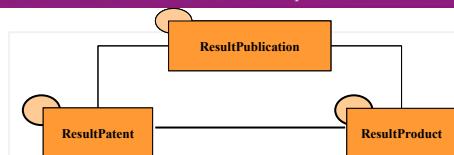
CERIF Result Entities



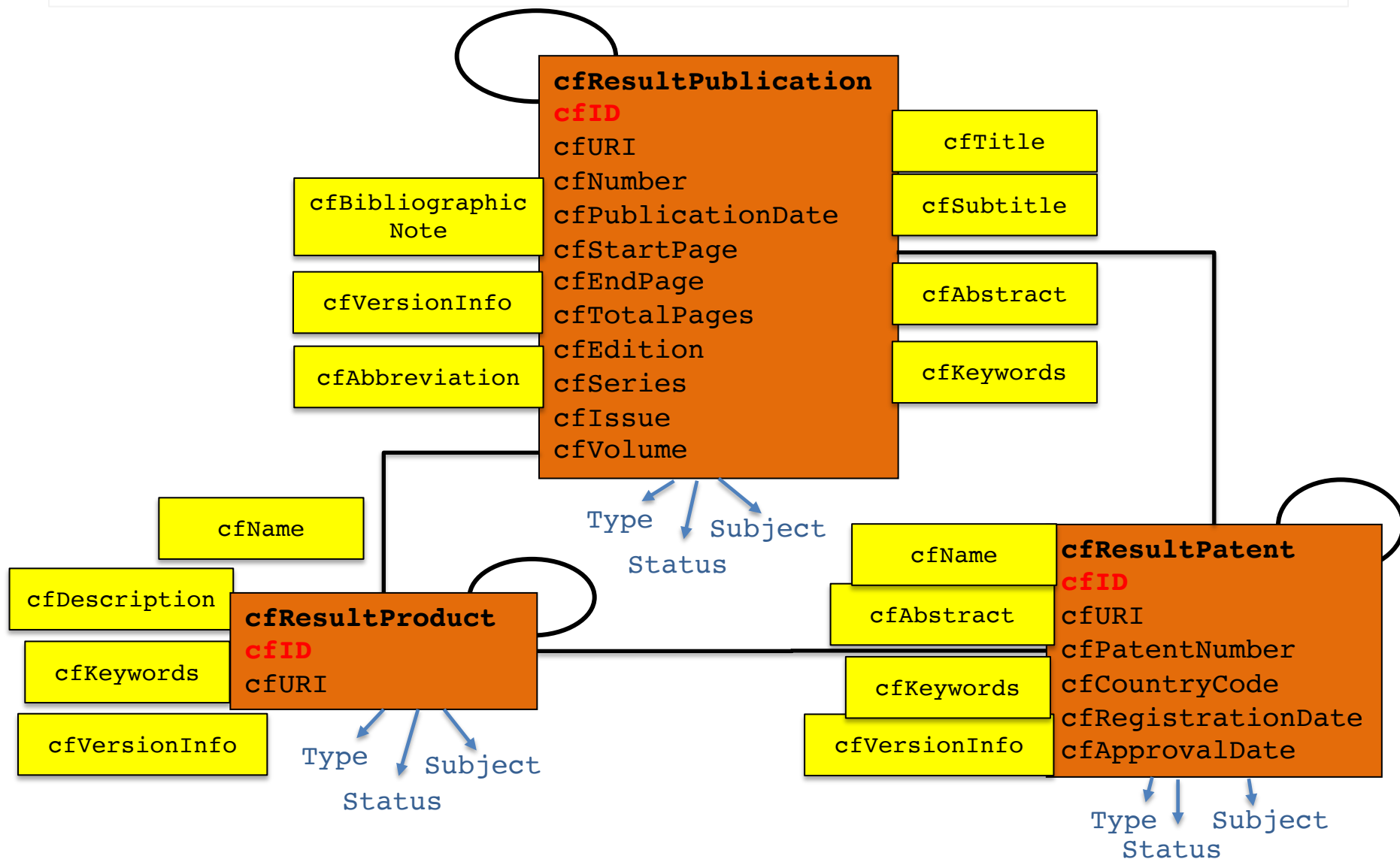


CERIF Result Entities

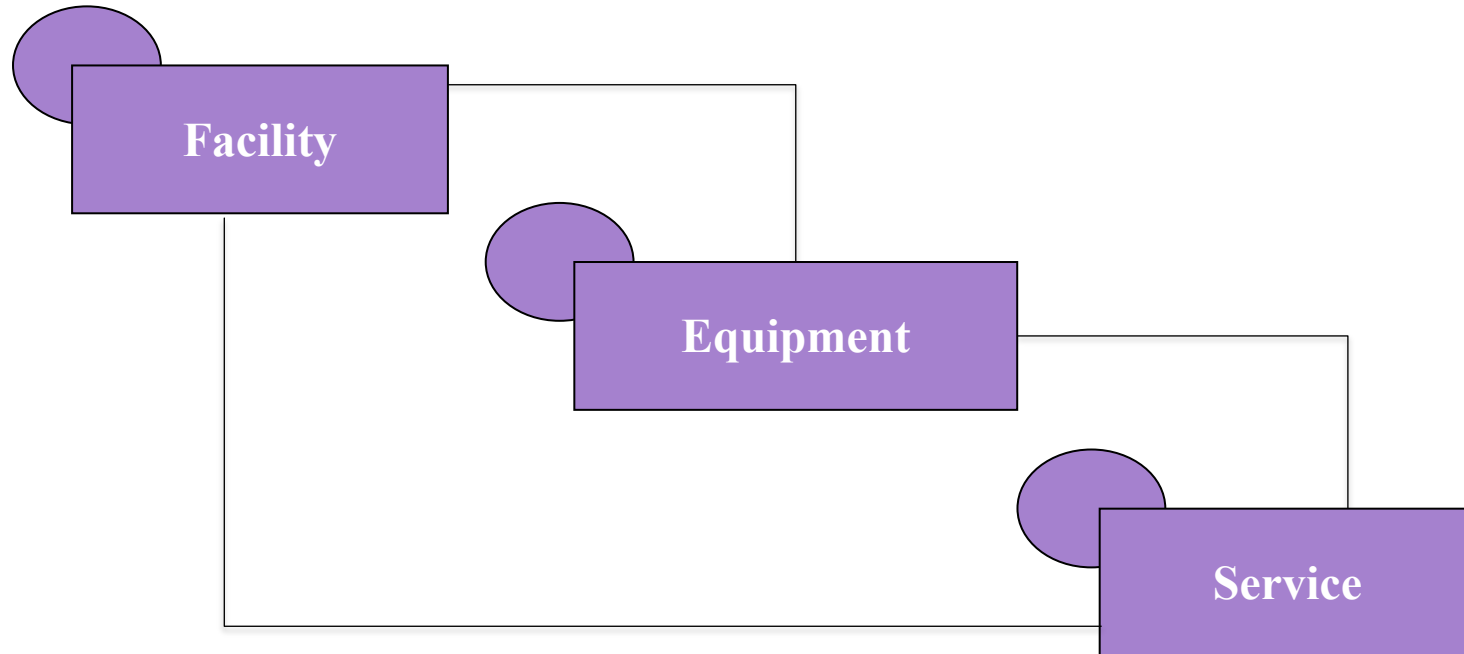


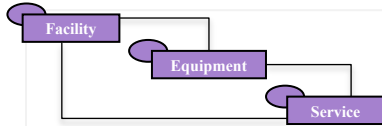


CERIF Result Entities

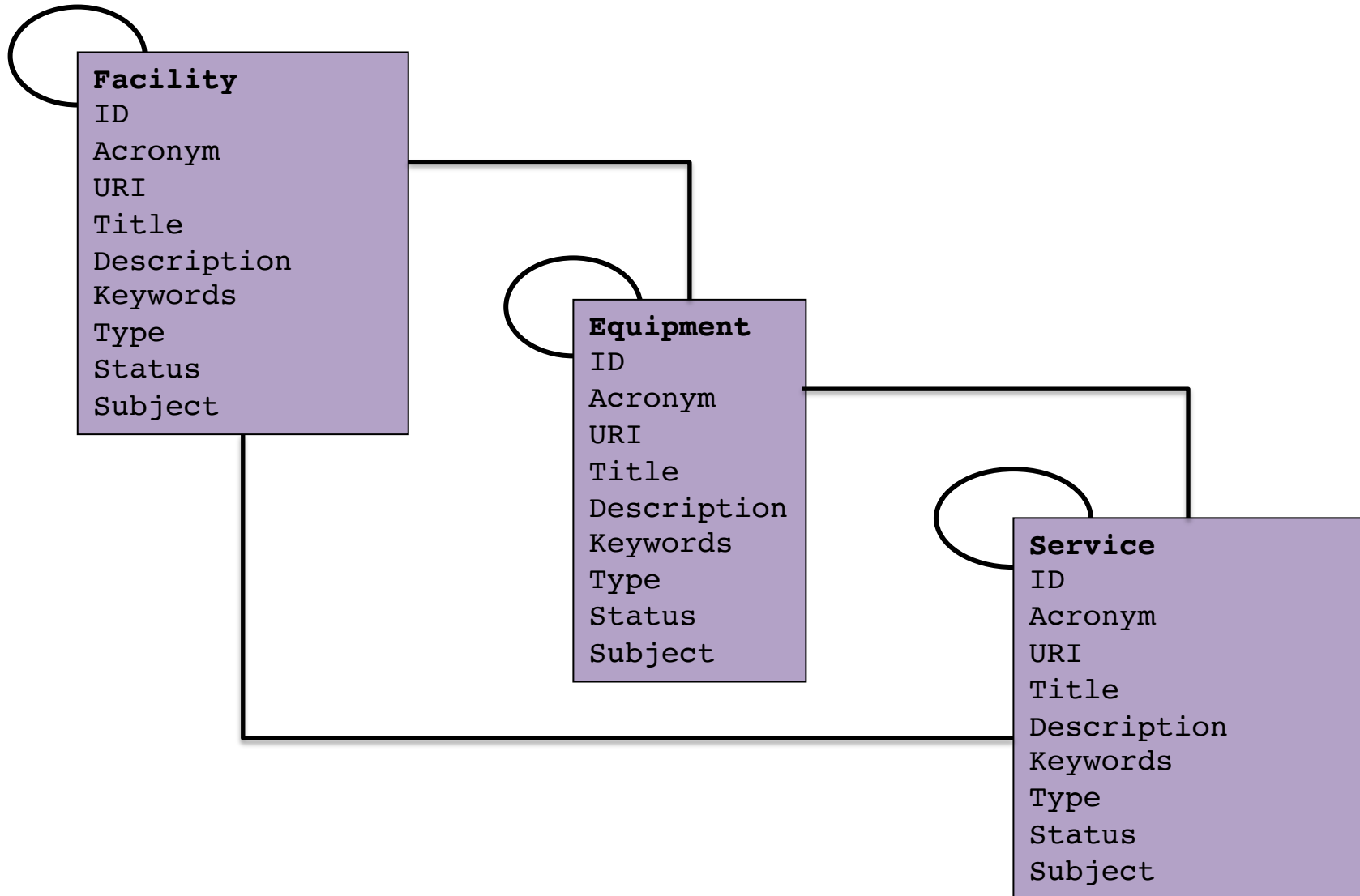


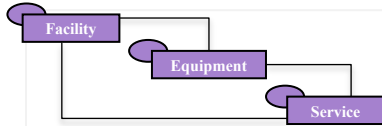
CERIF Infrastructure Entities



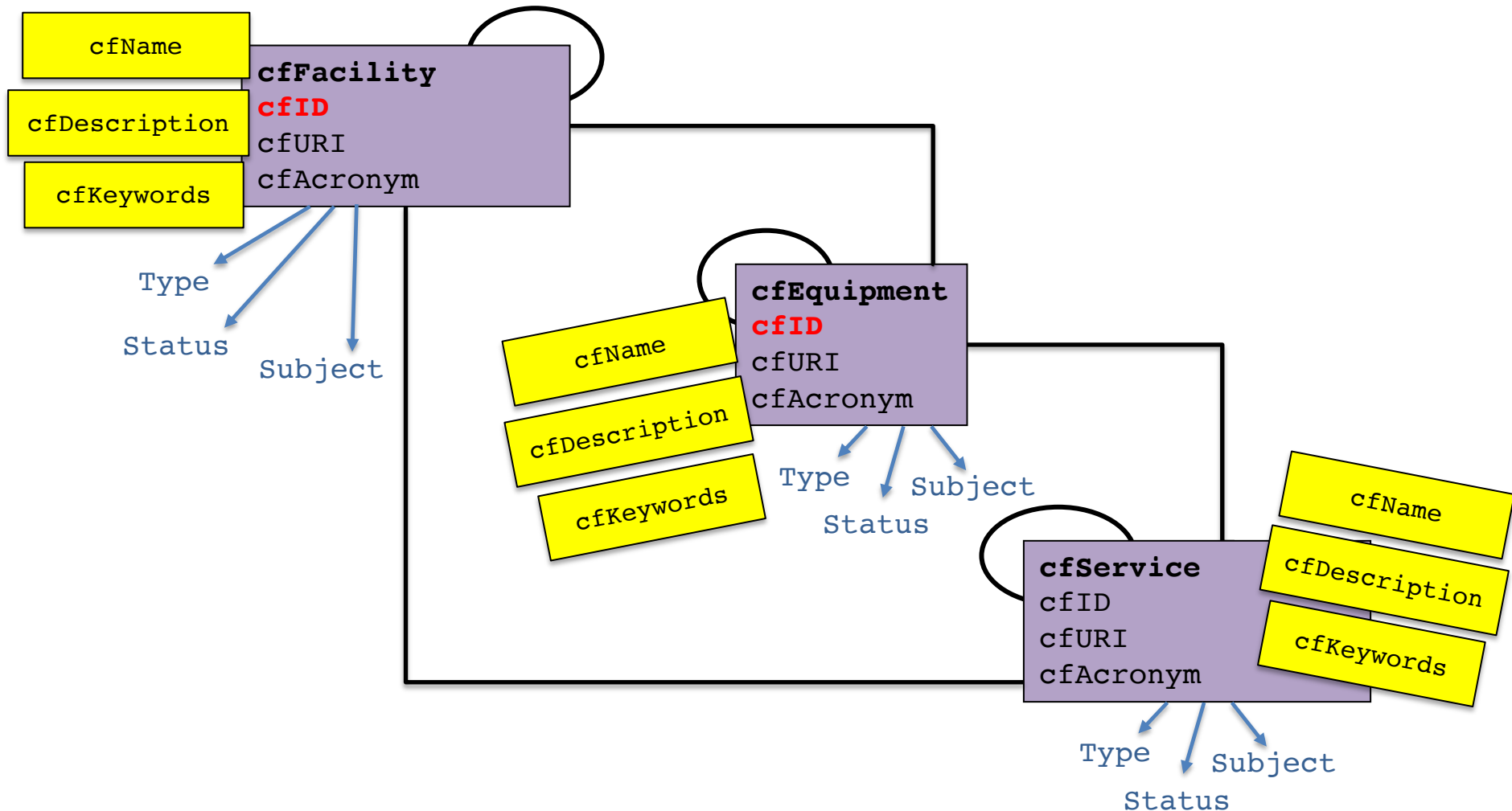


CERIF Infrastructure Entities





CERIF Infrastructure Entities

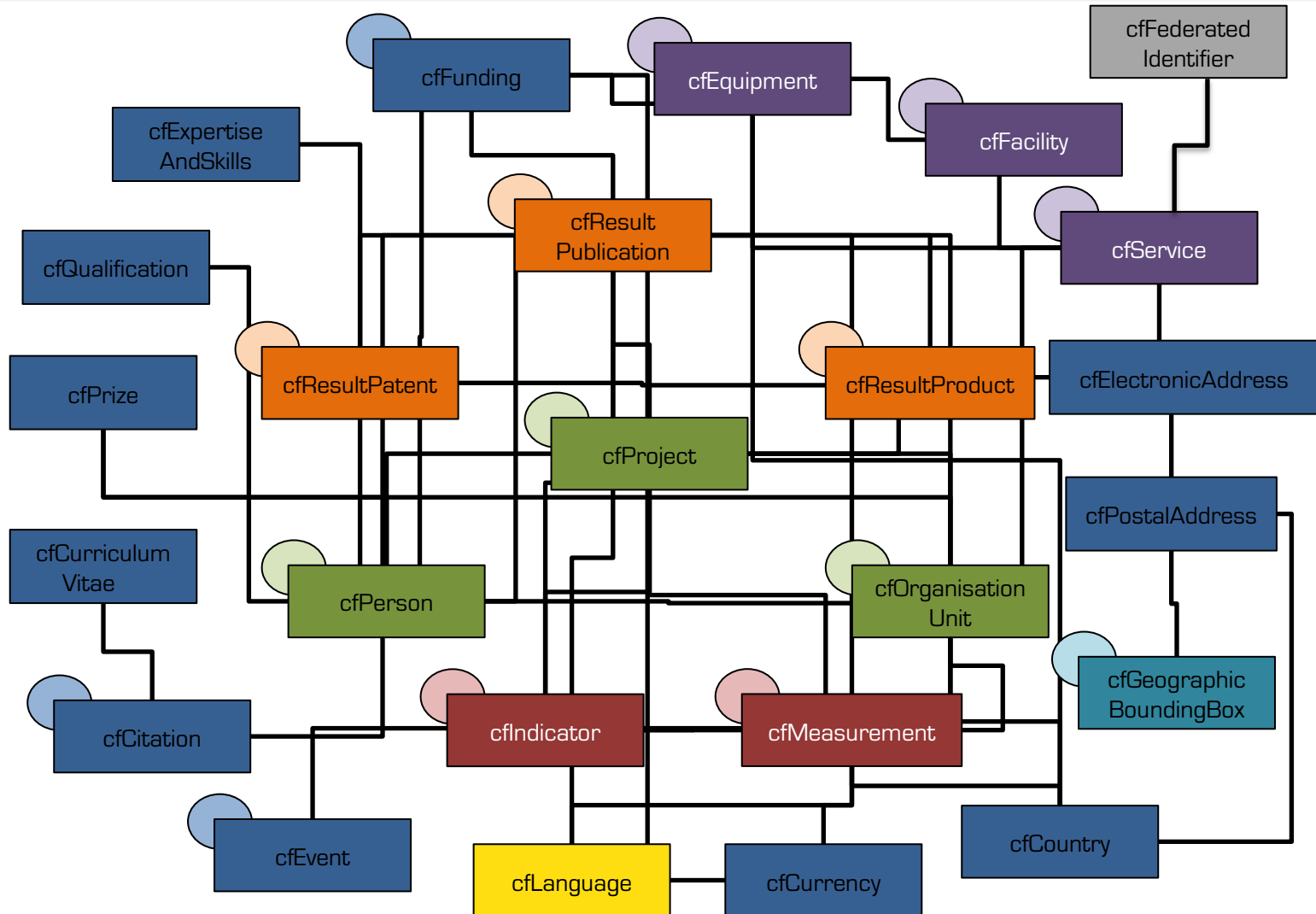


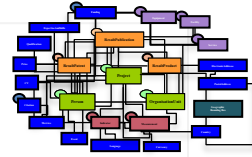
CERIF General Pattern

A typical CERIF entity has:

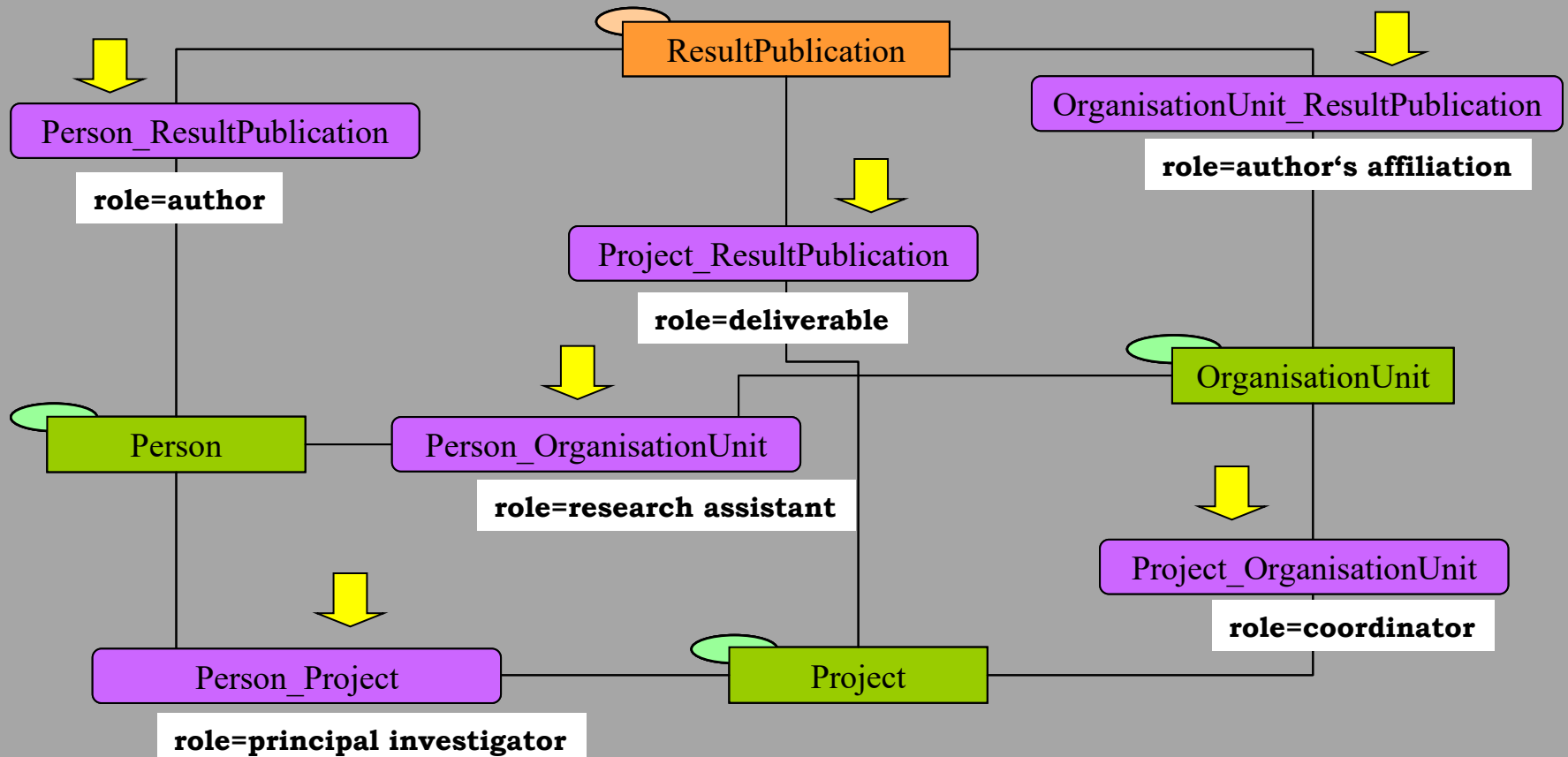
- Identifier (internal)
- Attributes
 - the basic ones
 - the multi-lingual ones
- External Identifiers
- Classifications
 - Type
 - Status
 - Subject area
 - + other
- Links
 - to other entities
 - recursive, a.k.a. self-referential

CERIF 1.6

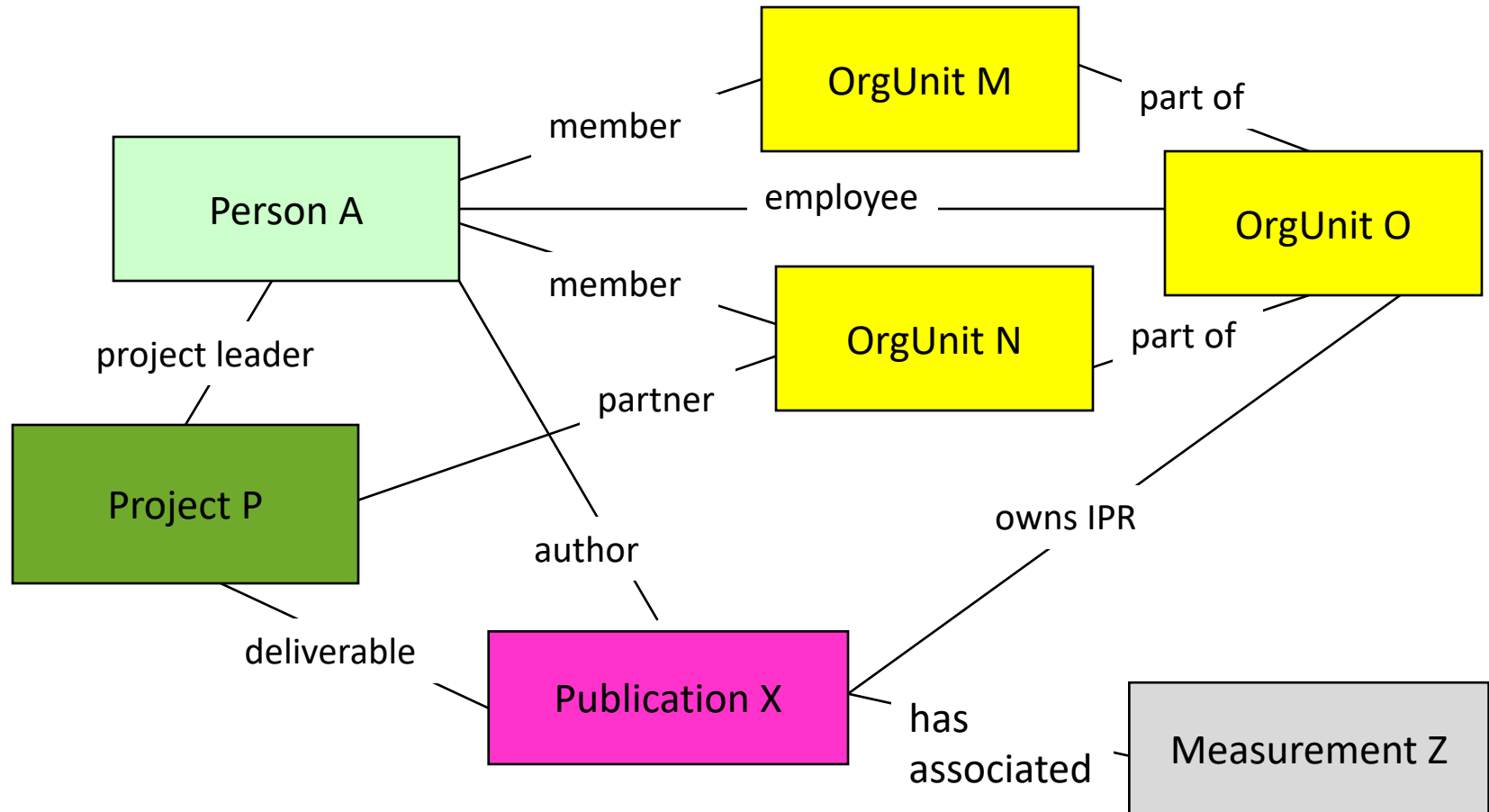




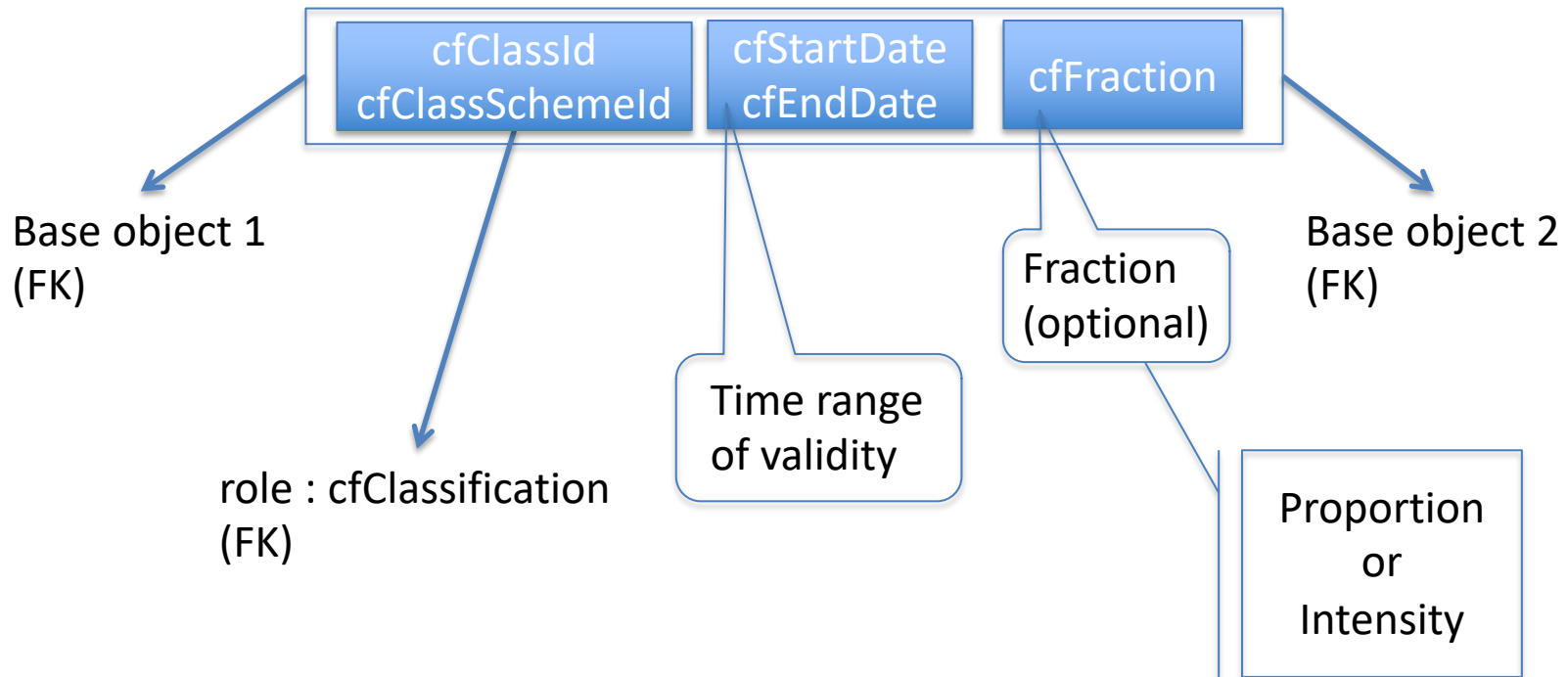
Some CERIF Link Entities



Another example
(slide by Keith Jeffery)



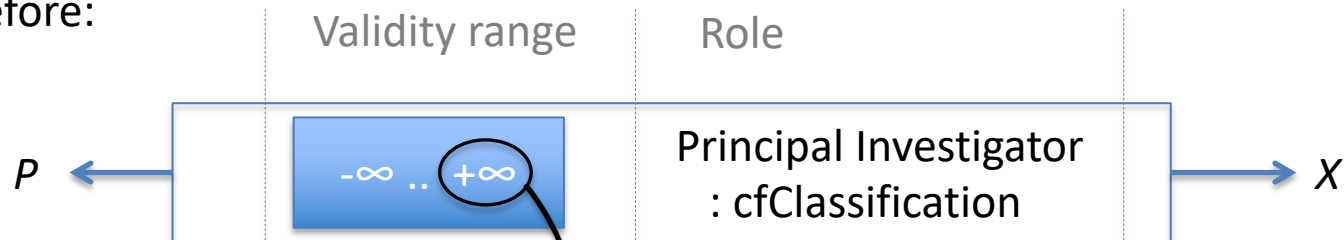
Generic Linking Entity Structure



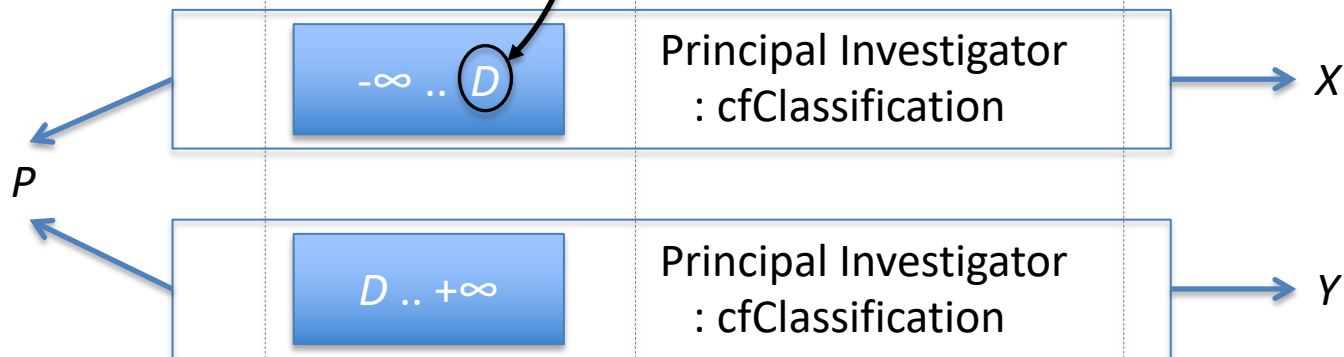
Recording Change in CERIF

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

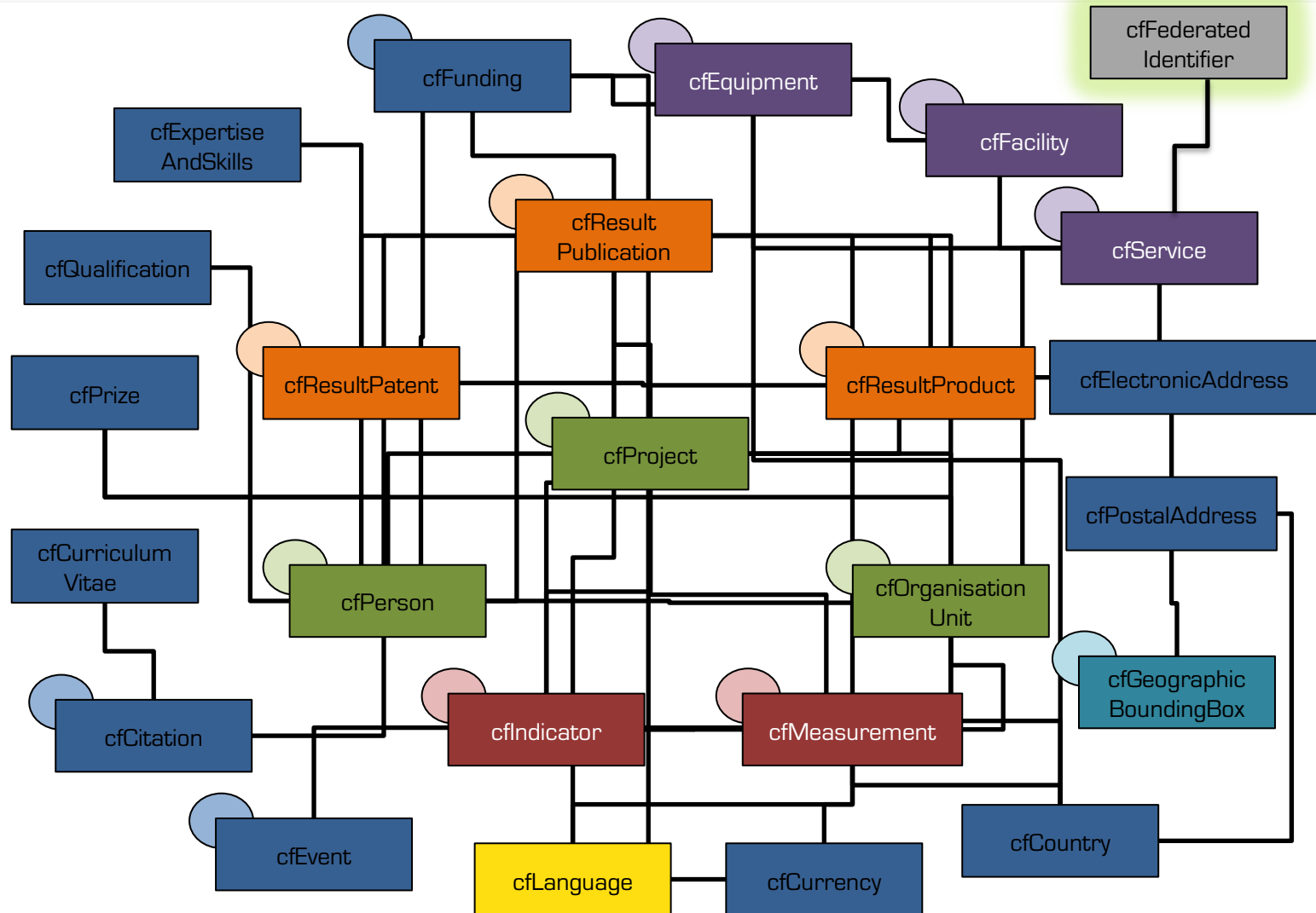
Before:



After:



CERIF 1.6



CERIF Federated Identifiers

Publication

- ISBN
- ISSN
- DOI
- WoS Accession Number
- Scopus EID
- PubMed Central ID

Person

- Social Security Number
- ORCID
- Staff Id in HR system
- Author identifier
 - IdRef, DAI, Número Lattes,
 - ResearcherID, Scopus Author ID

Project/Grant

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

Organisation

- VAT Identification Number
- FundRefID
- GridID
- OrgID

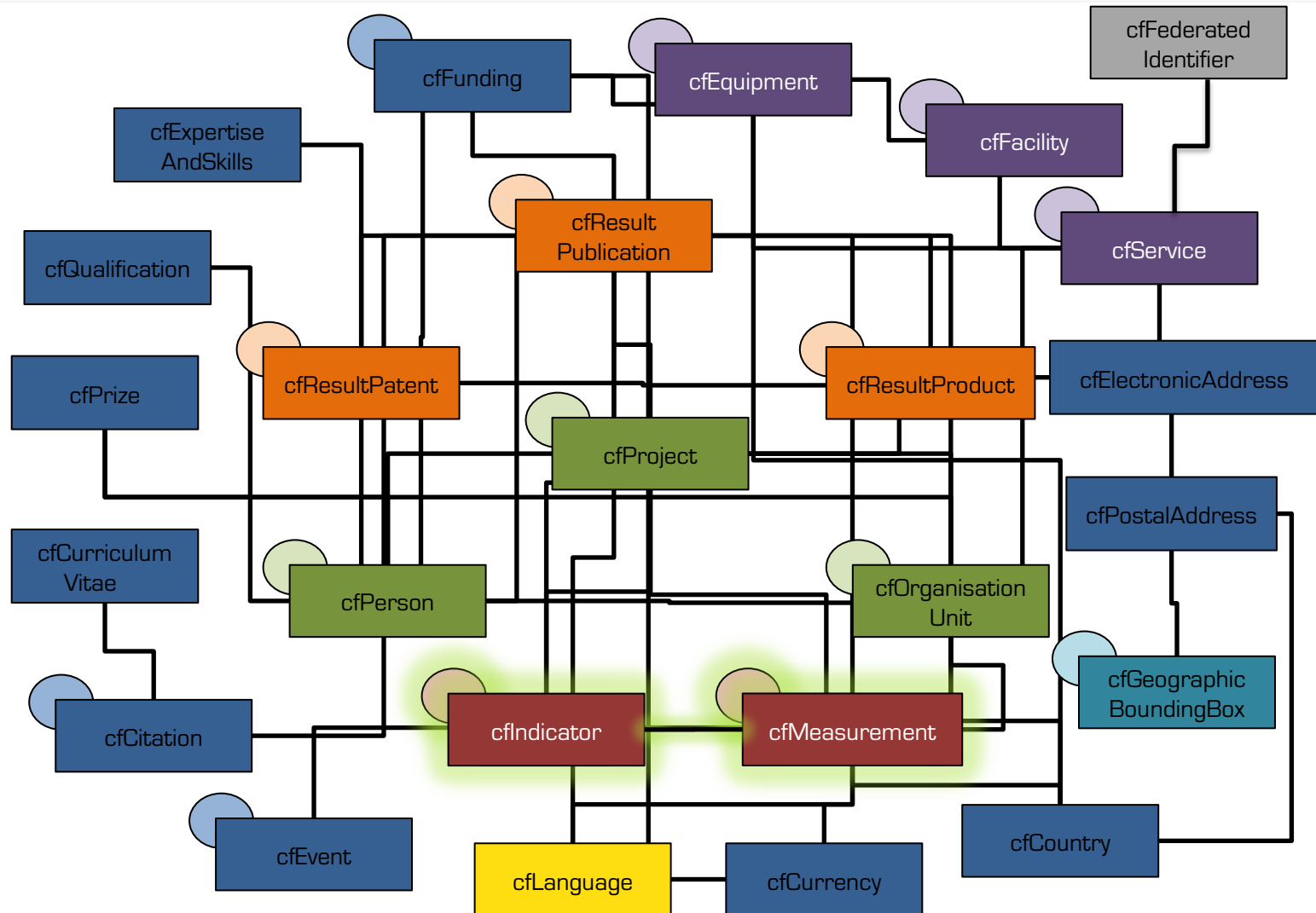
Classification

- External Code

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 1.6



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

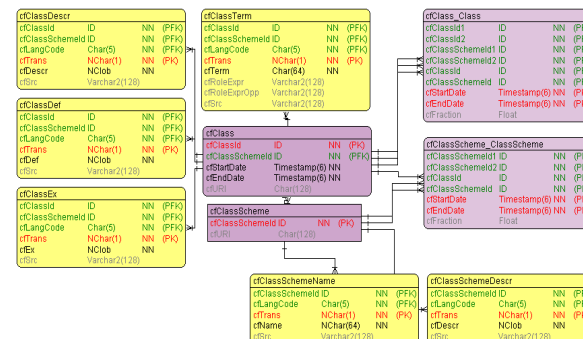
- New Schemas
- New Concepts / Terms
- New Relationships

Enables to manage

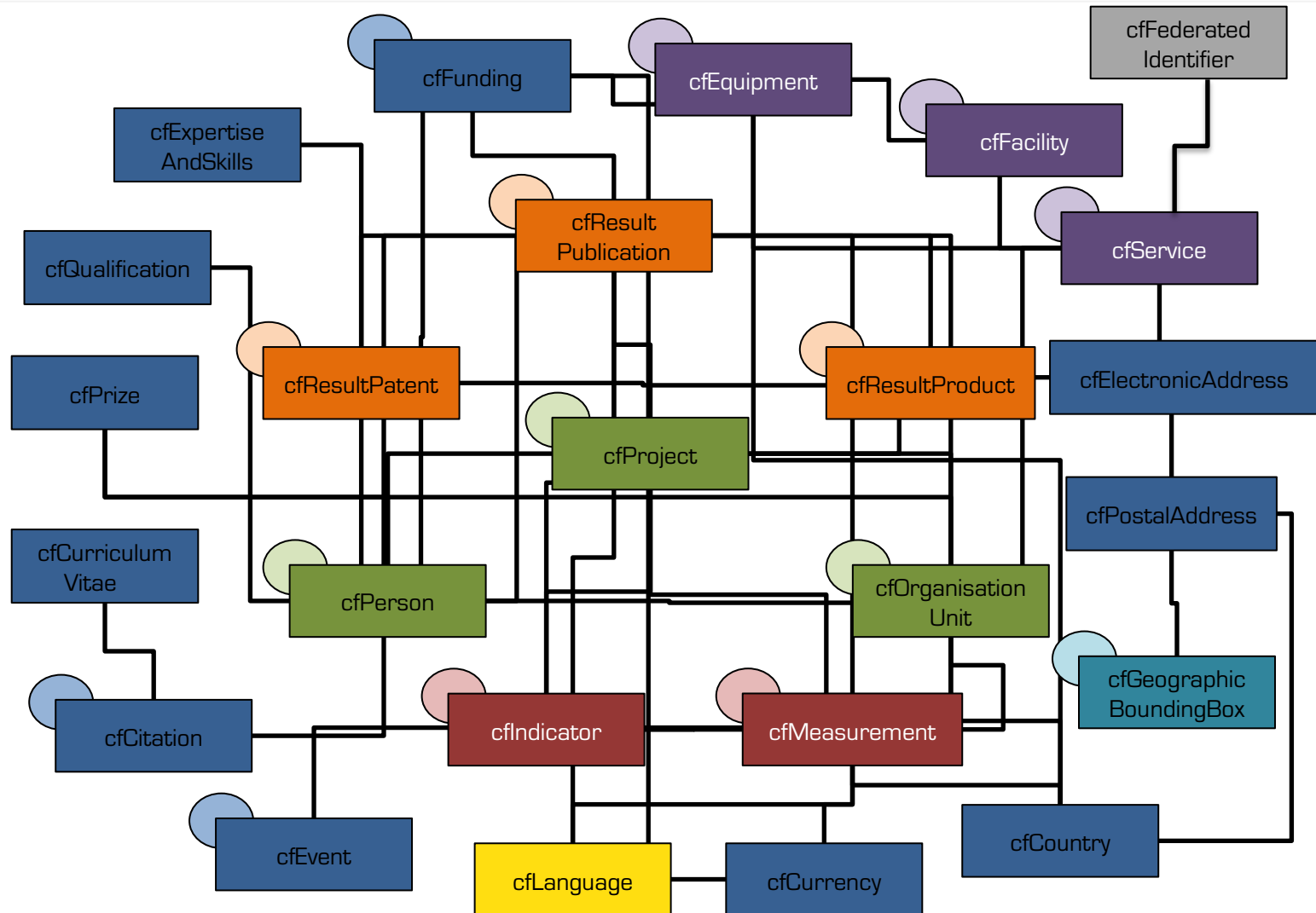
- Roles, Types, Statuses, ... Semantics
- Subject Headings
- Versioning and archiving (start+end dates)

Allows for Mappings between Schemes

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



CERIF 1.6



CERIF highlights

- Right level of abstraction
- Normalized model
 - Record information only once
 - Reference rather than copy
- Versatile Semantic Layer
- Time-based relationships
- Clean design, regular structure

Metadata Layers

Discovery metadata

DC, VIVO, MODS, METS, eGMS, DCAT, ...

Contextual metadata

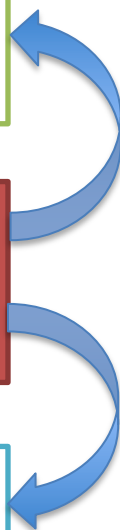
CERIF

Detailed metadata

Domain-specific standards

Generate

Reference



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

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

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 1.0.

DOI [10.5281/zenodo.17065](https://doi.org/10.5281/zenodo.17065)

Ex.: OpenAIRE Guidelines for CRIS Managers 1.1.

DOI [10.5281/zenodo.1298649](https://doi.org/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](https://doi.org/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
- Ongoing: <https://github.com/openaire/guidelines-cris-managers> for version 1.2 with Andreas Czerniak and Dragan Ivanović

Version 1.1: Aligned vocabularies

- COAR Resource Types
- COAR Access Rights
- ISSN Media List

OpenAIRE Guidelines for CRIS Managers
latest

Search docs

Introduction

CRIS information elements relevant for OpenAIRE

Publication

- Internal Identifier
- Type
- Language
- Title
- Subtitle
- PublishedIn
- PartOf
- PublicationDate
- Number
- Volume
- Issue
- Edition

Read the Docs v: latest

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

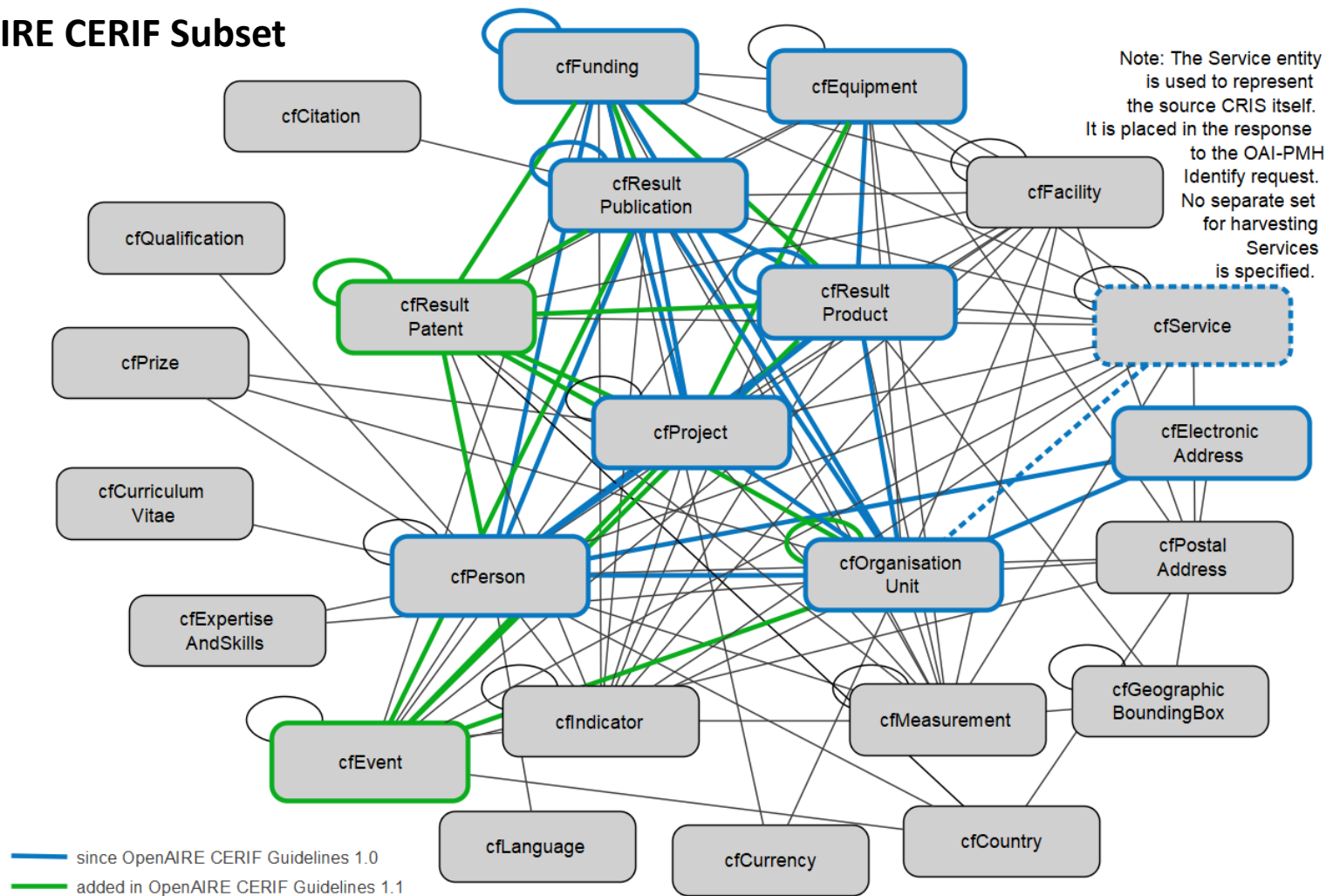
Outputs:

- Publications
- Datasets, Software & other Products
- Patents

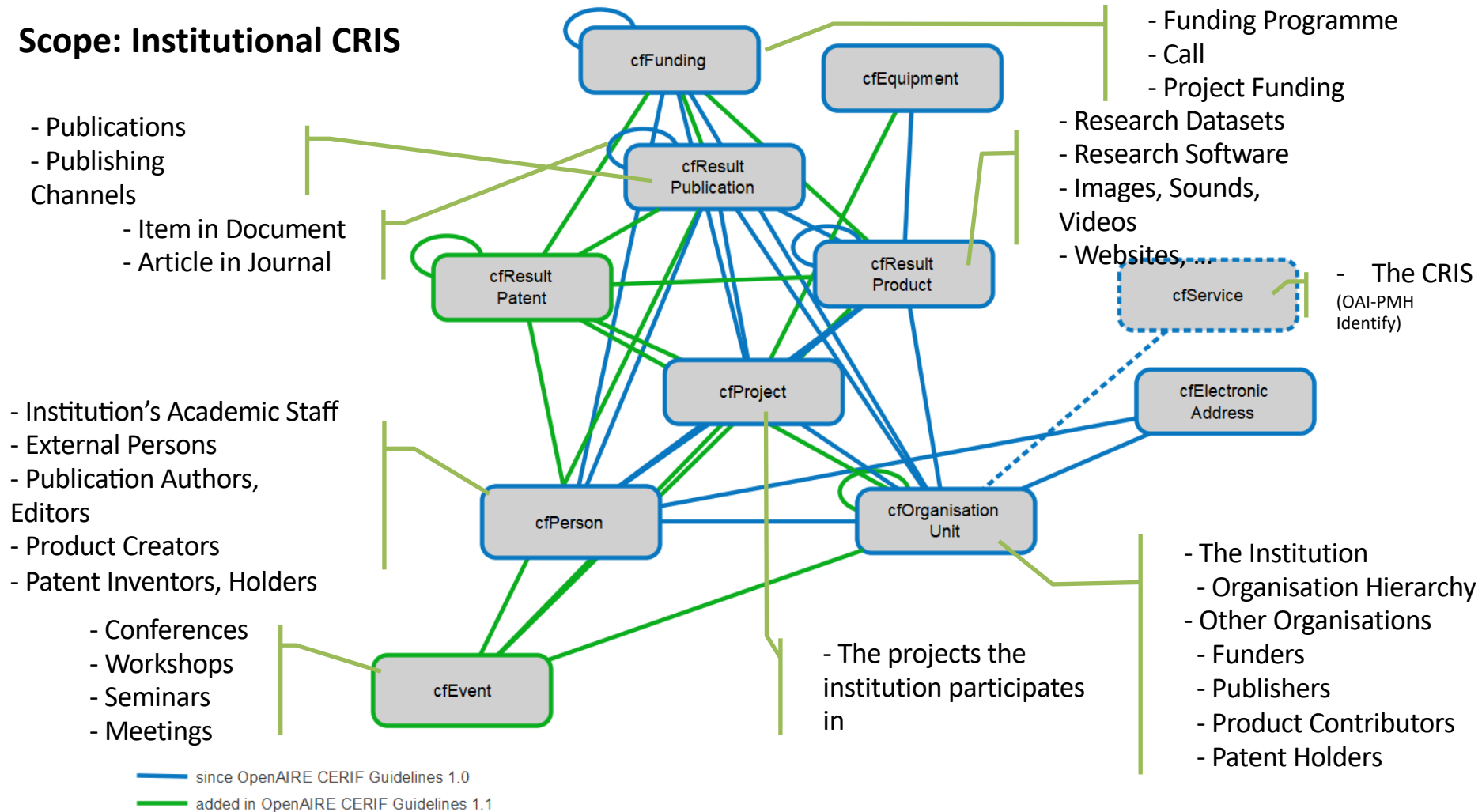
Infrastructure:

- Equipment

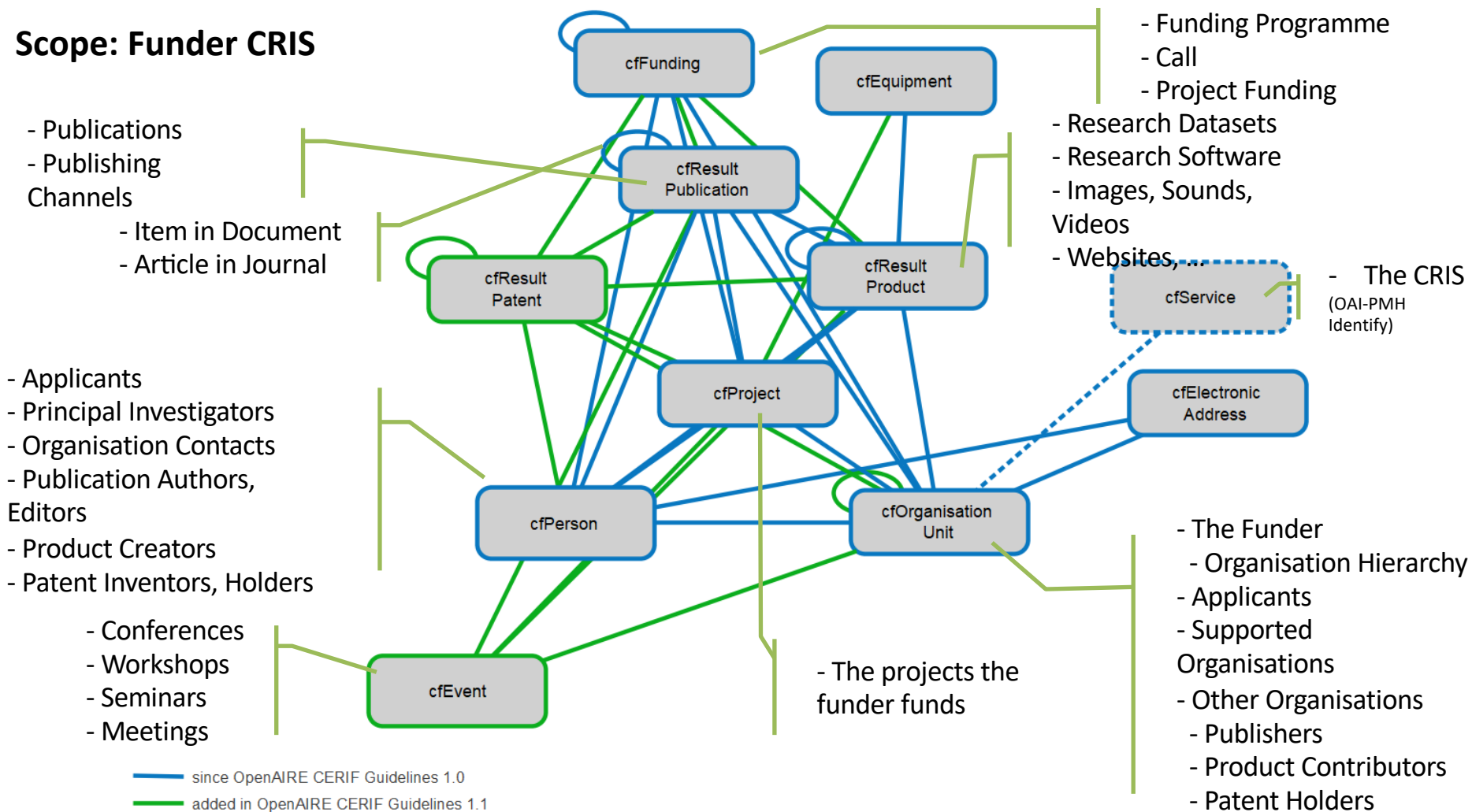
OpenAIRE CERIF Subset



Scope: Institutional CRIS



Scope: Funder CRIS



Identifiers

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

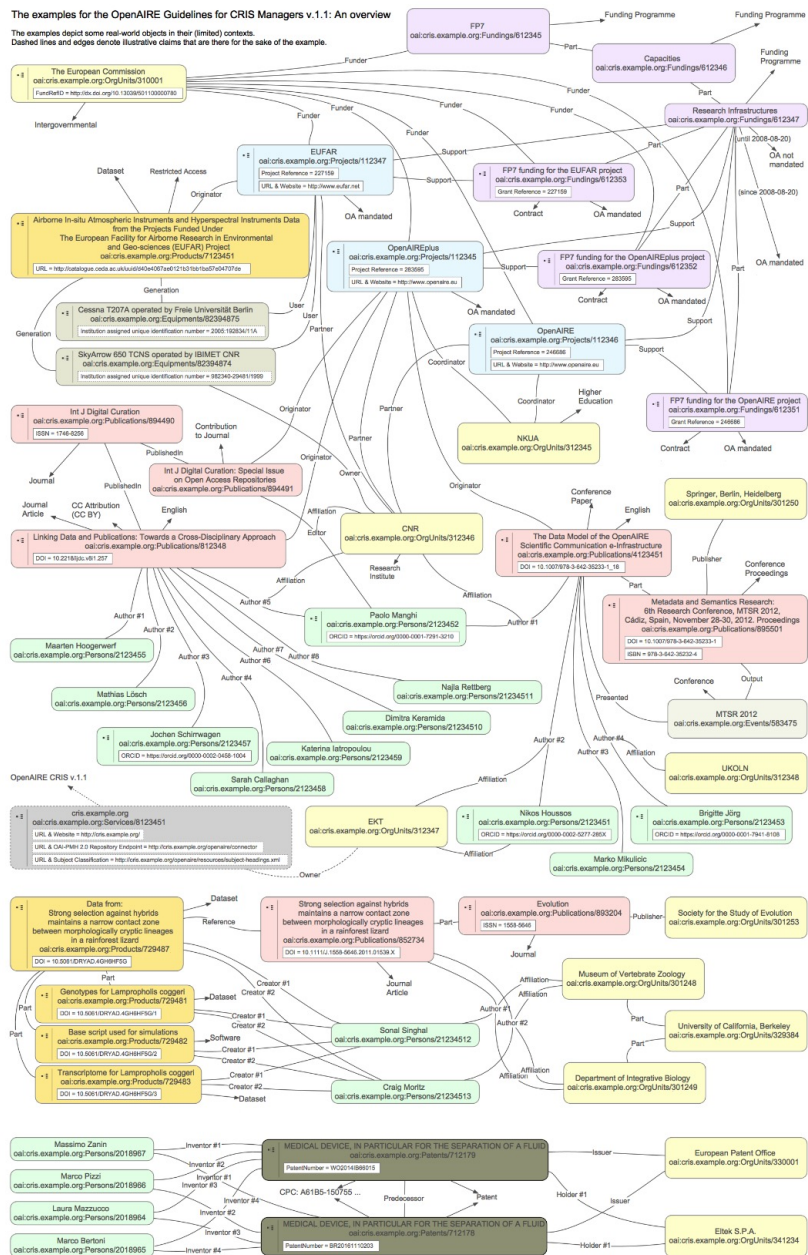
Elsewhere generic

Supporting artifacts

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

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.



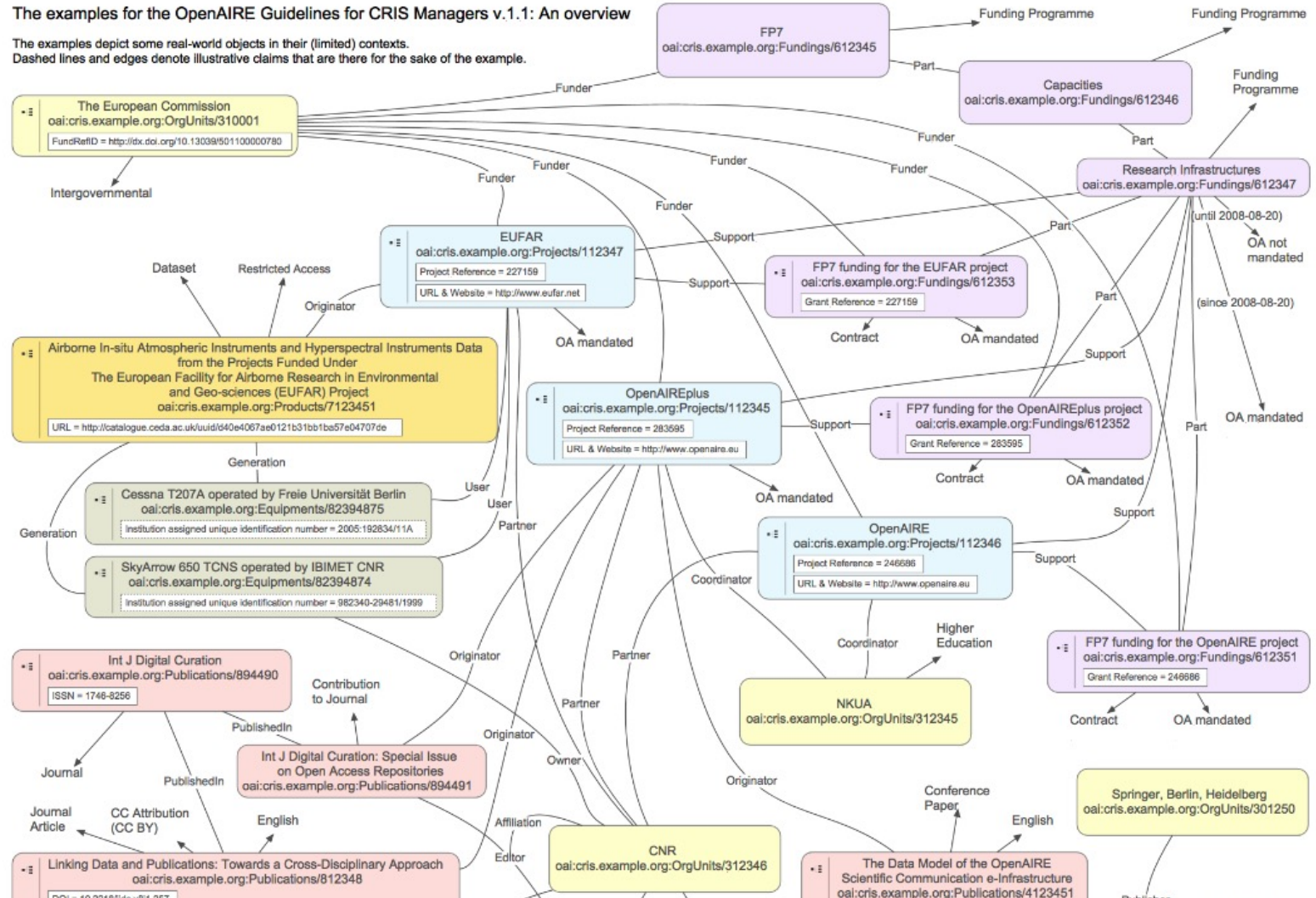
A Research Graph: The example set from the OpenAIRE Guidelines for CRIS Managers 1.1

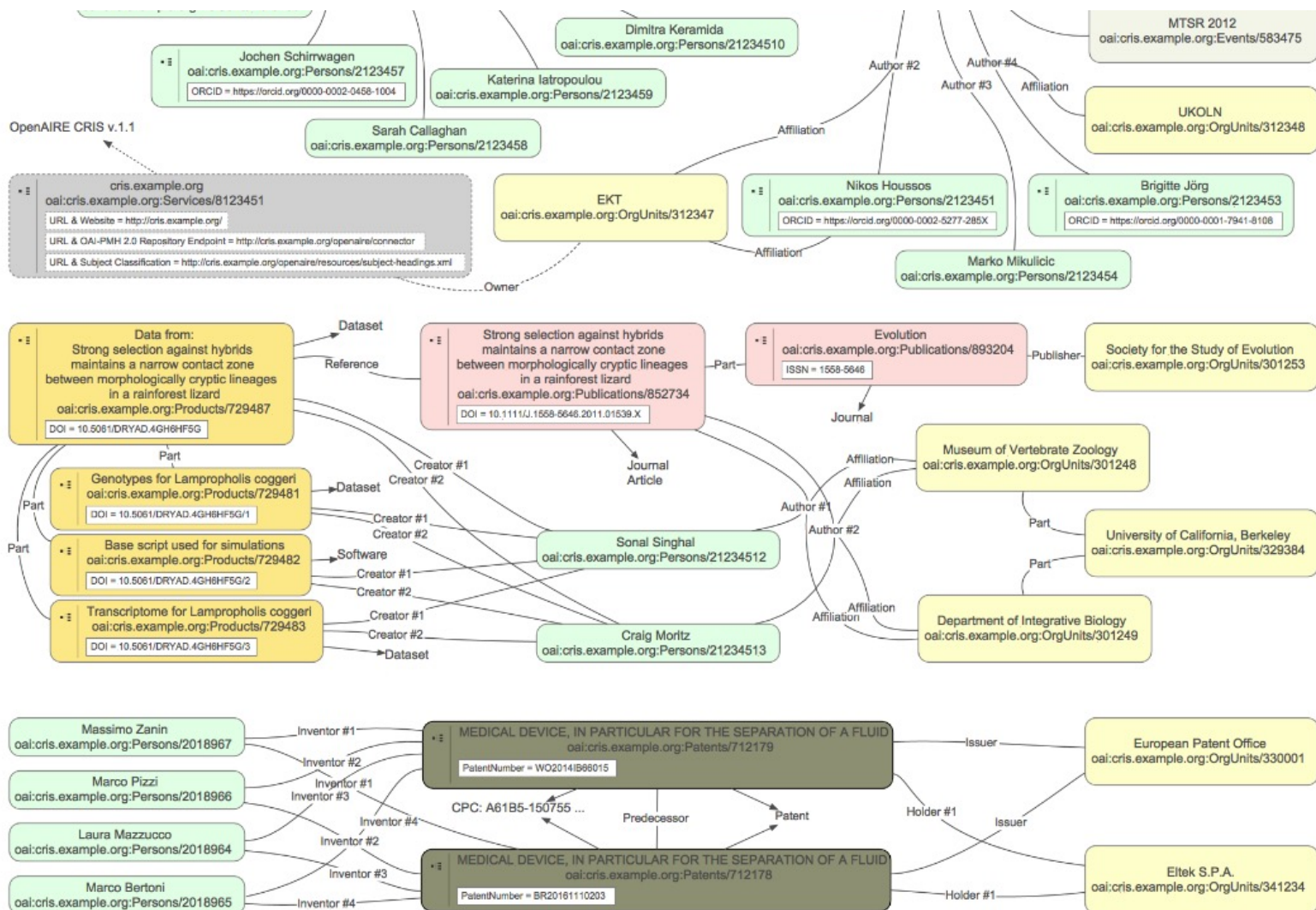
OpenAIRE Technical Meeting @ CERN

Funding Programme Funding Programme

oai:cris.example.org:Fundings/612345

Part





```
<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 in general

Useful subsets of CERIF
for specific research information
interchange scenarios

- Entities & attributes:
 - Profile $\subset$ CERIF
- Semantic vocabularies:
 - Profile – specific choices
 - Sources: CERIF & beyond
- Integrity constraints:
 - Profile $\supset$ CERIF

Profile data is
CERIF

CERIF Profiles in general

Producers know
what to include



Consumers know
what to expect

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

👉 CRIS 2022: more to be presented on Friday at 16:30 👉