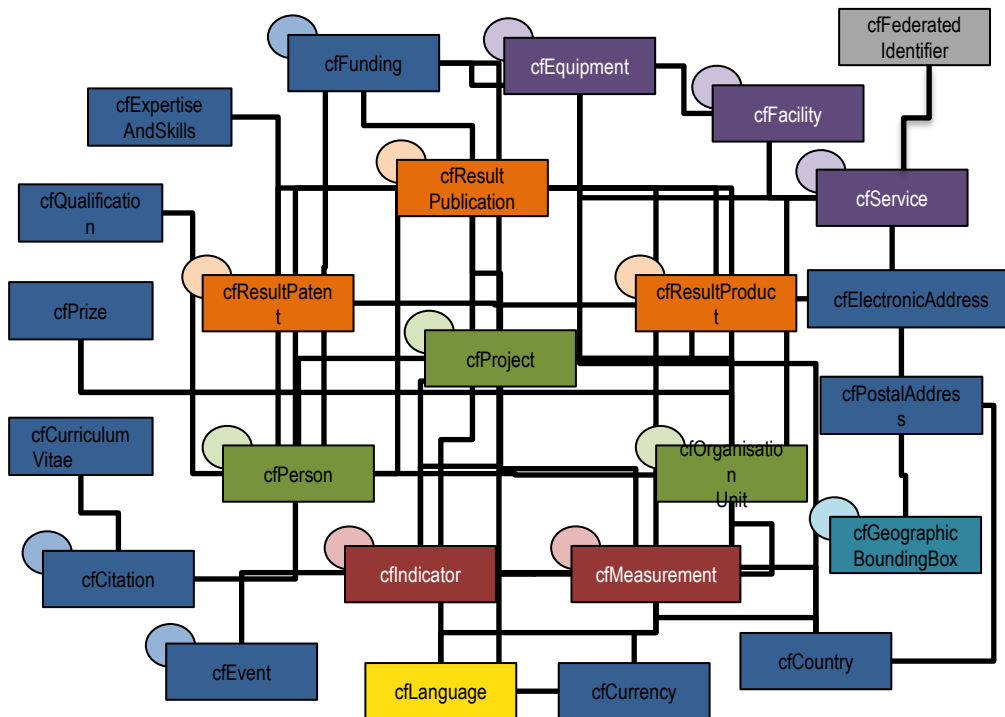


CERIF 1.5 Tutorial



November 13th, 2013
euroCRIS Membership
Meeting
Porto, Portugal

Jan Dvořák

jan.dvorak@infoscience.cz

euroCRIS

- CERIF TG Leader since 2013
- CRIS2012 (Prague, June 2012) Organizer

Charles University in Prague

- Faculty of Arts
 - Institute of Information Studies & Librarianship

InfoScience Praha

- Research & Development & Innovation Information System
(the national CRIS for CZ)

Most slides by Brigitte Jörg

What is Research Information?

Information about:

- Researchers
- Organisations (Research-performing, Funding)
- Funding Programmes, Calls
- Projects (Proposed, Ongoing, Completed)
- Publications, Patents, Data, Products
- Facilities, Equipment, Services
- Addresses, Geographic Bindings, Languages
- **And their Relationships**

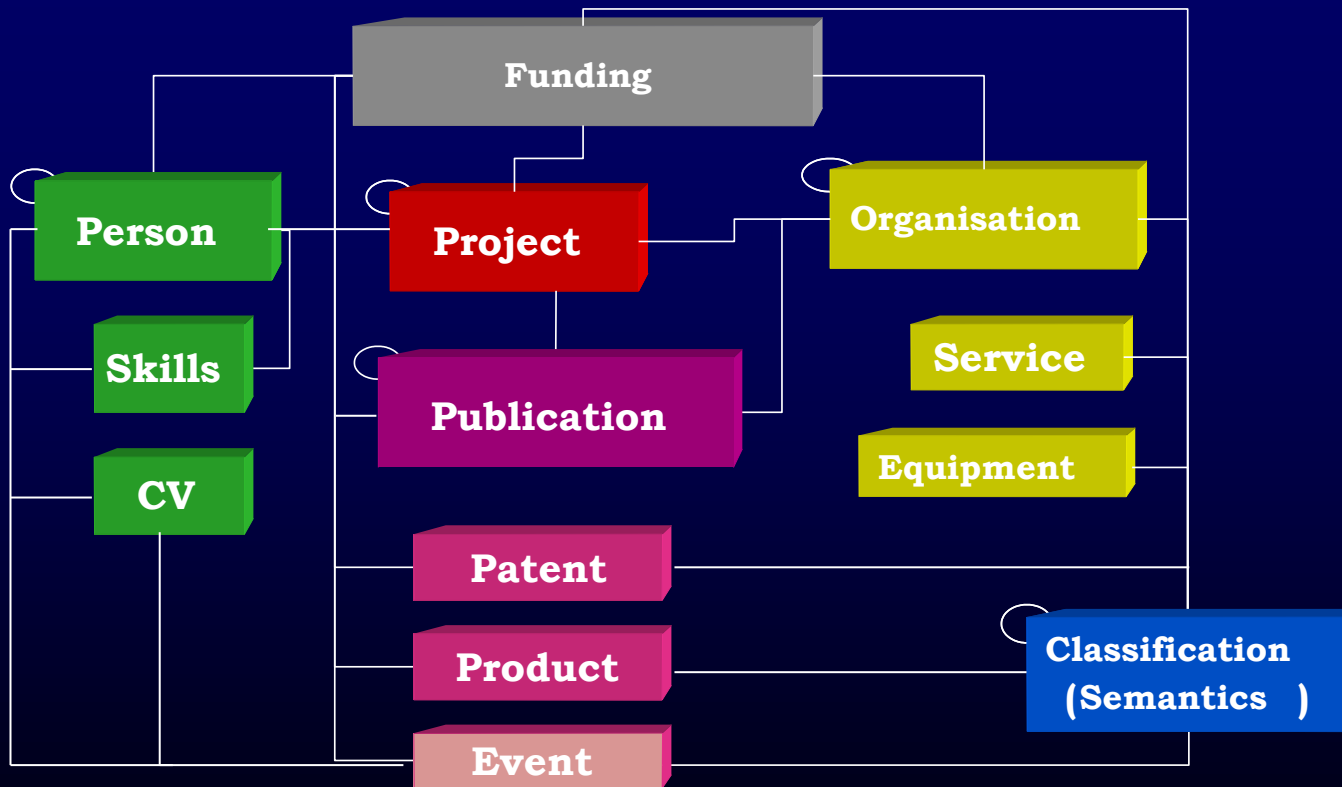
Who needs Research Information?



Kinds of questions we want to support

- How many articles has author X published in 2011 as a first author?
- How many times have articles by author X been cited by the end of the previous year?
- Did author X publish with institutionally external authors?
- In how many FP7 projects does/ did organisation Z participate?
- How many publications have resulted from project Y?
- How many people have been employed in the course of FP7 projects from the 1st call in the New Member States?
- How many PhD students have participated in national research projects in country C? In which countries have they earned their masters degrees?
- How many women have been involved in FP7 projects?
- How often have articles in journal A been requested in 2010?
- How many articles have been published in field B?

The Ultimate Answer: Common European Research Information Format

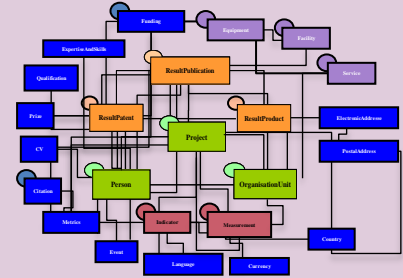
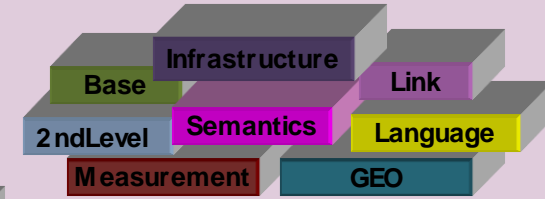
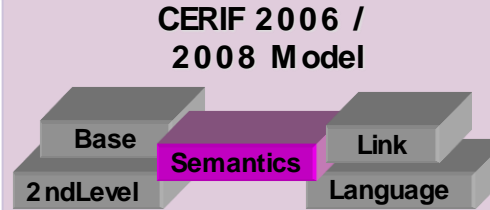


The CERIF Evolution

CERIF 1.5
CERIF 1.4 (XML)
CERIF 1.3

Similar Ideas
UN/ UNESCO
OECD
CODATA

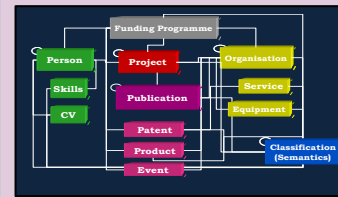
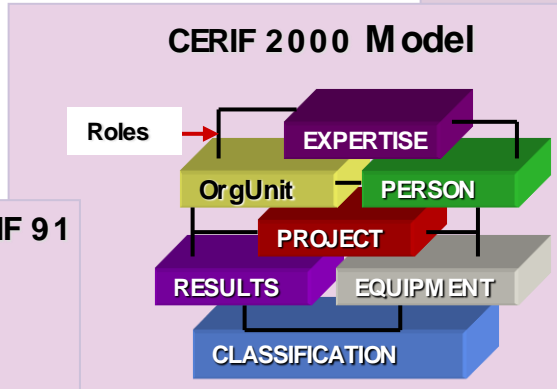
EU
Working Group
on Research
Databases
Workshop



FORMAL SEMANTICS

- Data Model
- Infrastructure
 - Facility, Equipment, Service
- Measurement & Indicator
 - Entities and Link Tables
- Geographic Bounding Box
- CERIF 1.3 Vocabulary
 - UUIDs
 - Terms
 - Schemes
- CERIF 1.4 new XML format
- CERIF 1.5 Federated Identifiers

+ Linked Data



- Data Model
- Multilinguality
- Controlled Vocabulary
- Roles / Types
- User-driven
- EC Recommendation to Member States

- Data Model
- Model Normalization
 - Robust/ Consistent Structure
 - Extensible Structure
 - Semantic Layer
- XML Exchange Specification
- Elaboration on Publication
- CERIF Core Semantics (2008 1.2)

PROJECT

Acronym: ERGO
Participant:
Keith Jeffery, Anne Asserson, many more
Organisations:

1987 1991 2000 2006 2008 2012

Common European Research Information Format

- CERIF is an EU Recommendation to Member States

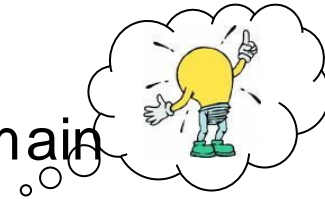
<http://cordis.europa.eu/cerif/>

- The European Commission (EC) has authorised euroCRIS to maintain and develop CERIF and its usage

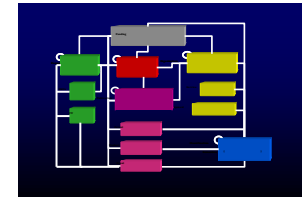
<http://www.eurocris.org/Index.php?page=CERIFreleases&t=1>

Model Levels

- **Conceptual Level** (Specification)
Concepts relevant for the research domain and their relationships



- **Logical Level** (ER Model)
Entities and their relationships



- **Physical Level** (Database Scripts)
Data Definition commands for the database

```
SQL Script
-----
CREATE Table cfPers (...);
CREATE Table cfProj (...);
CREATE Table cfOrgUnit (...);
```

- **Semantic Layer** (Declared Semantics)
A formalized controlled vocabulary describing a general contextual semantics of the research domain inline with the conceptual, logical and machine description

CERIF Model Structure (Views)

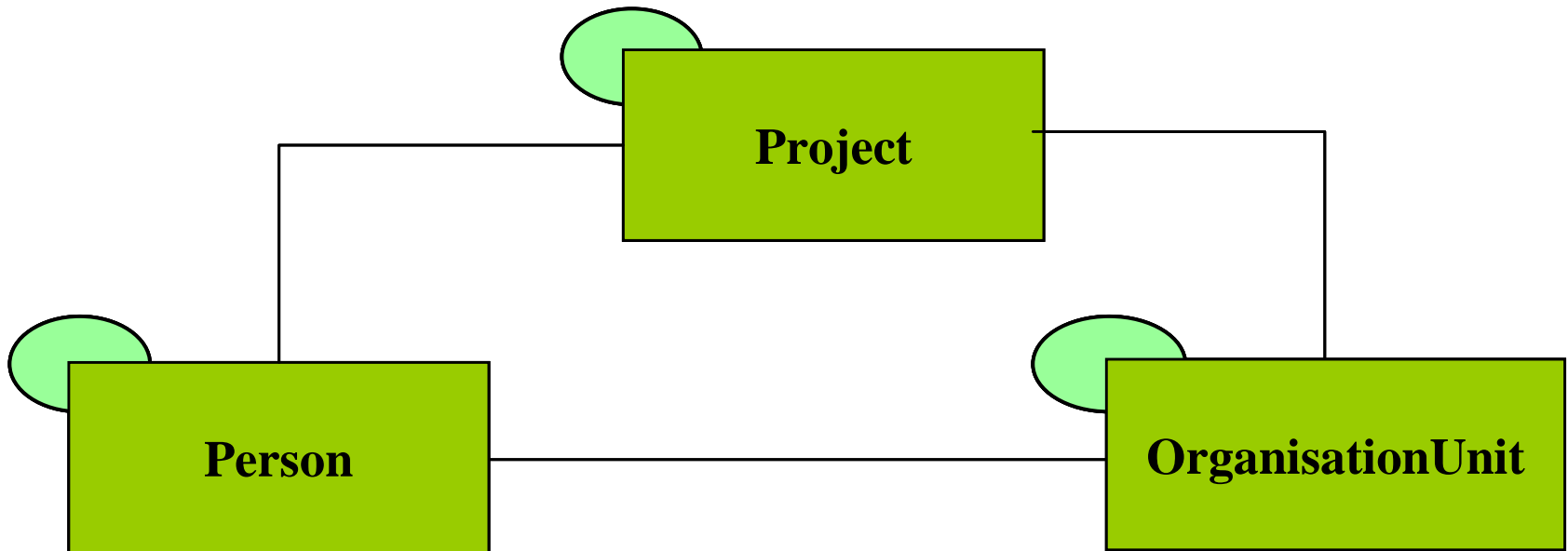
CERIF Entity Types

- Base Entities
- Result Entities
- Infrastructure Entities
- 2nd Level Entities
- Geographic Bounding Box
- Link Entities

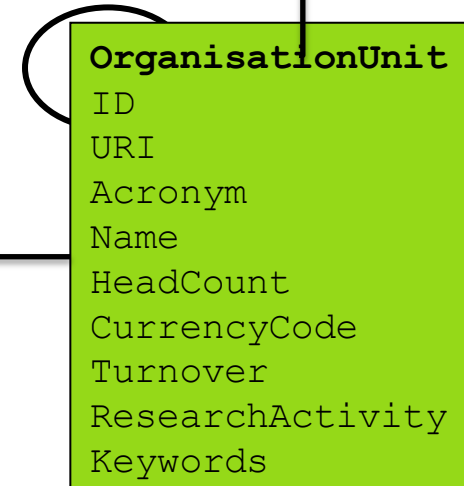
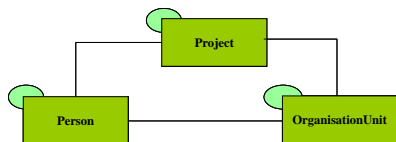
CERIF Features

- Multiple Language
- Semantics
- Measures & Indicators

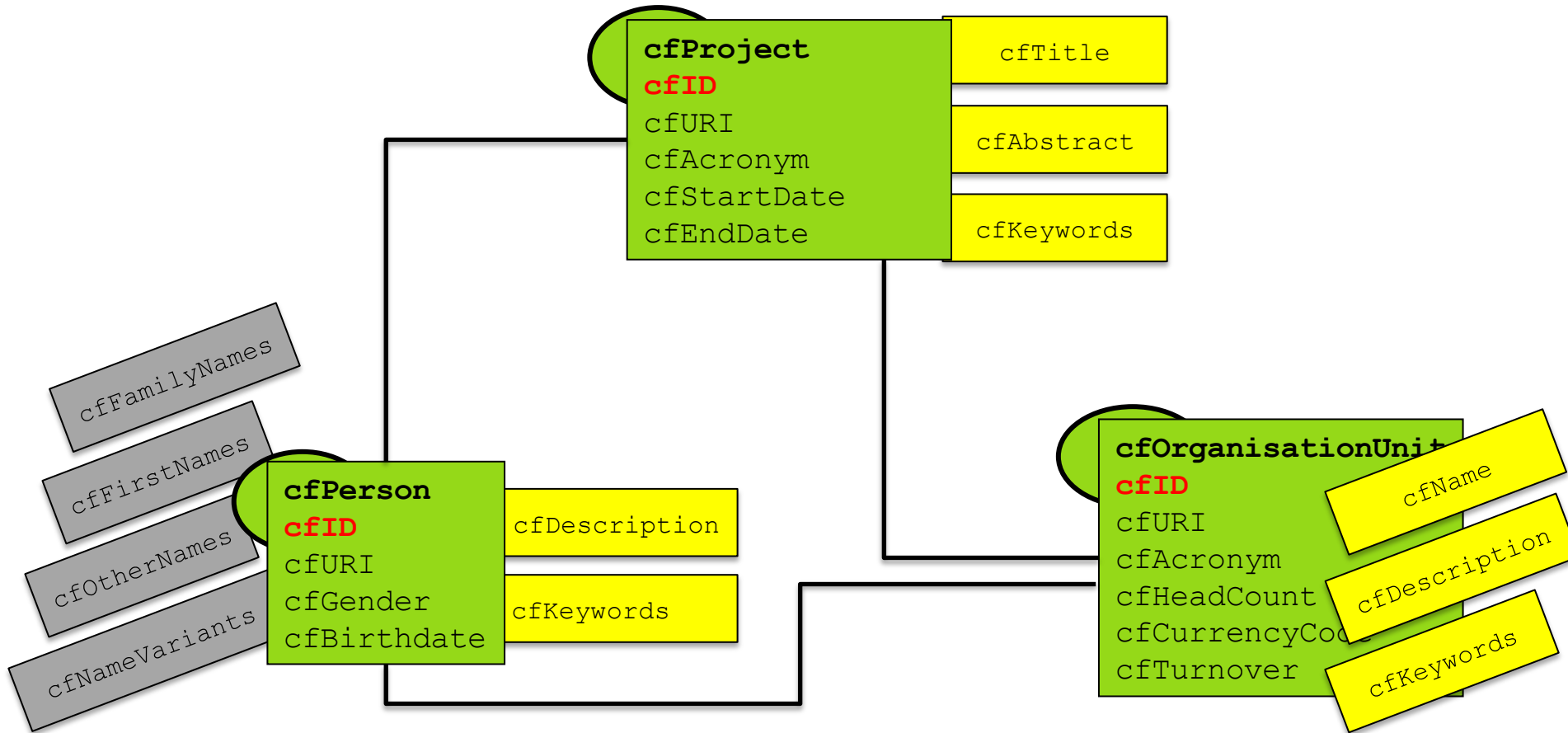
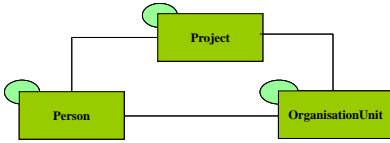
CERIF Base Entities



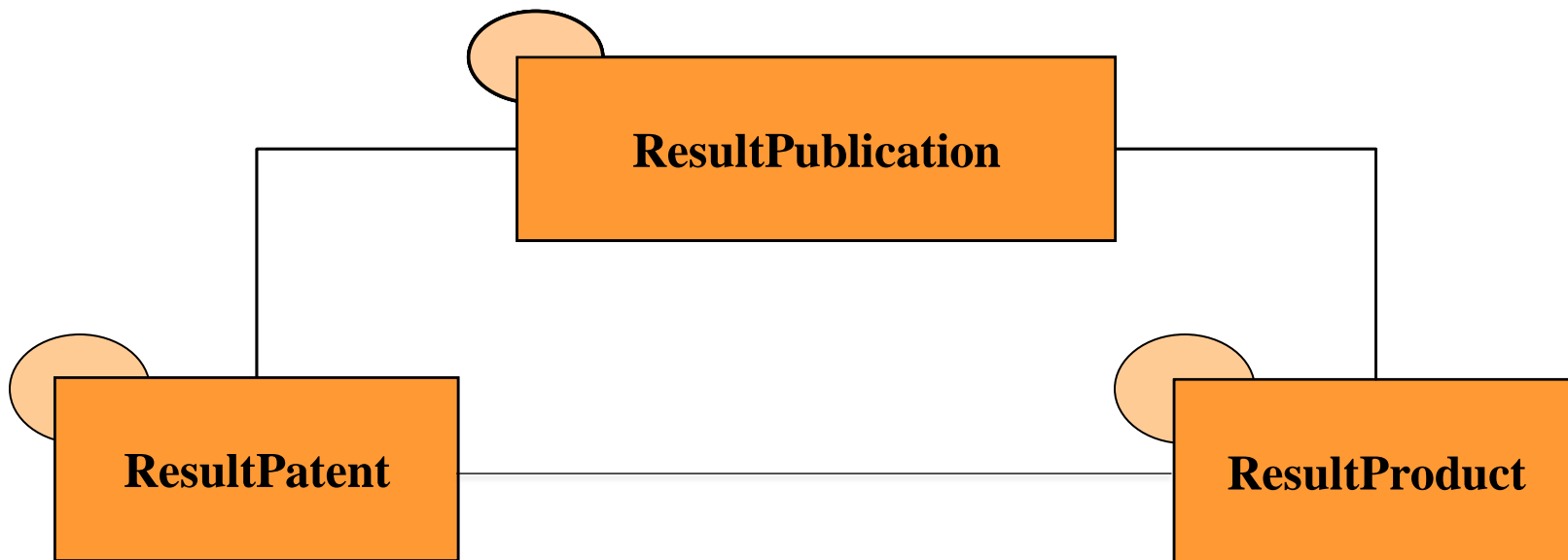
CERIF Base Entities



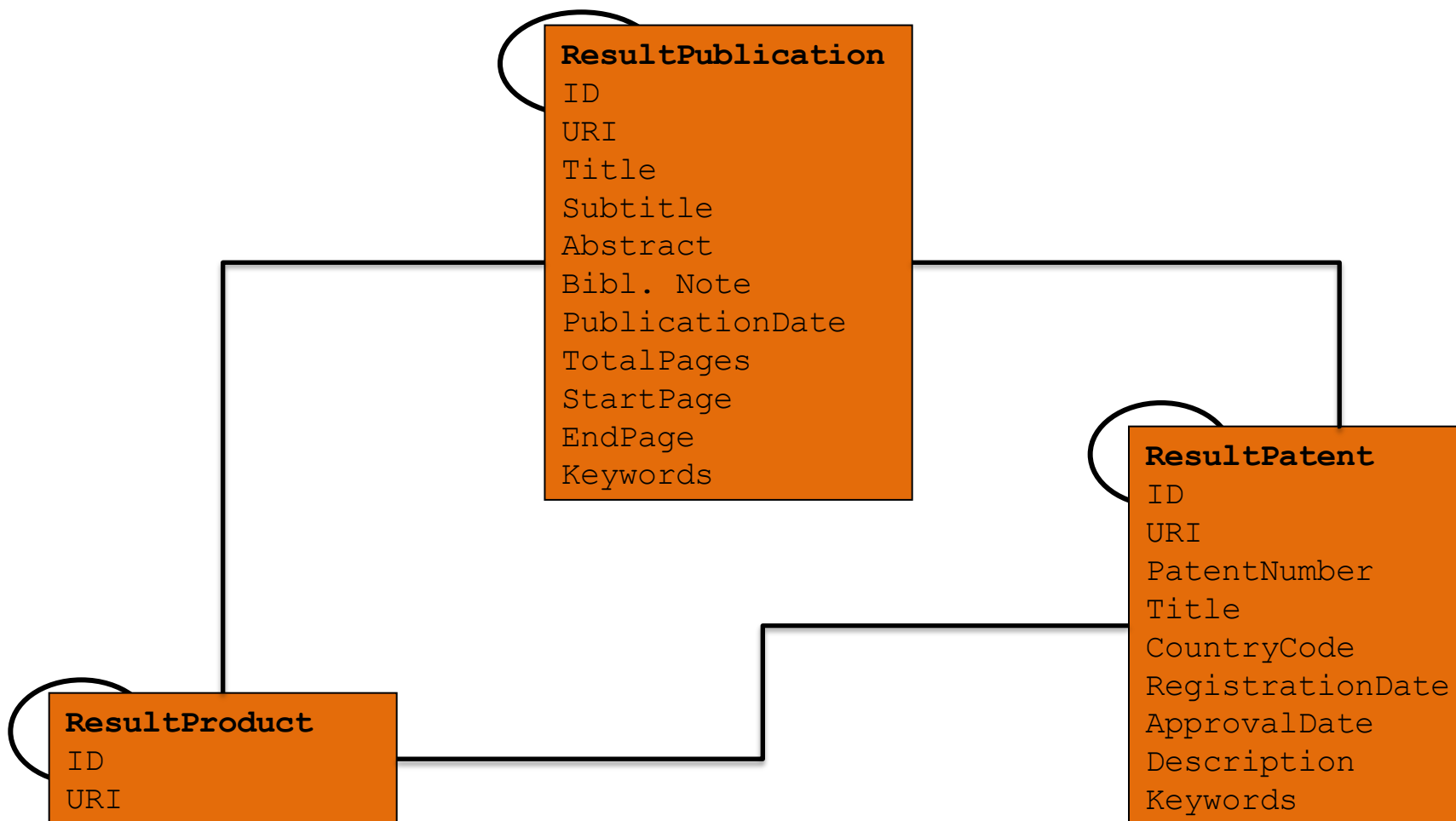
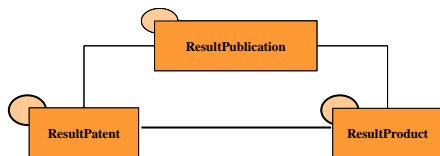
CERIF Base Entities



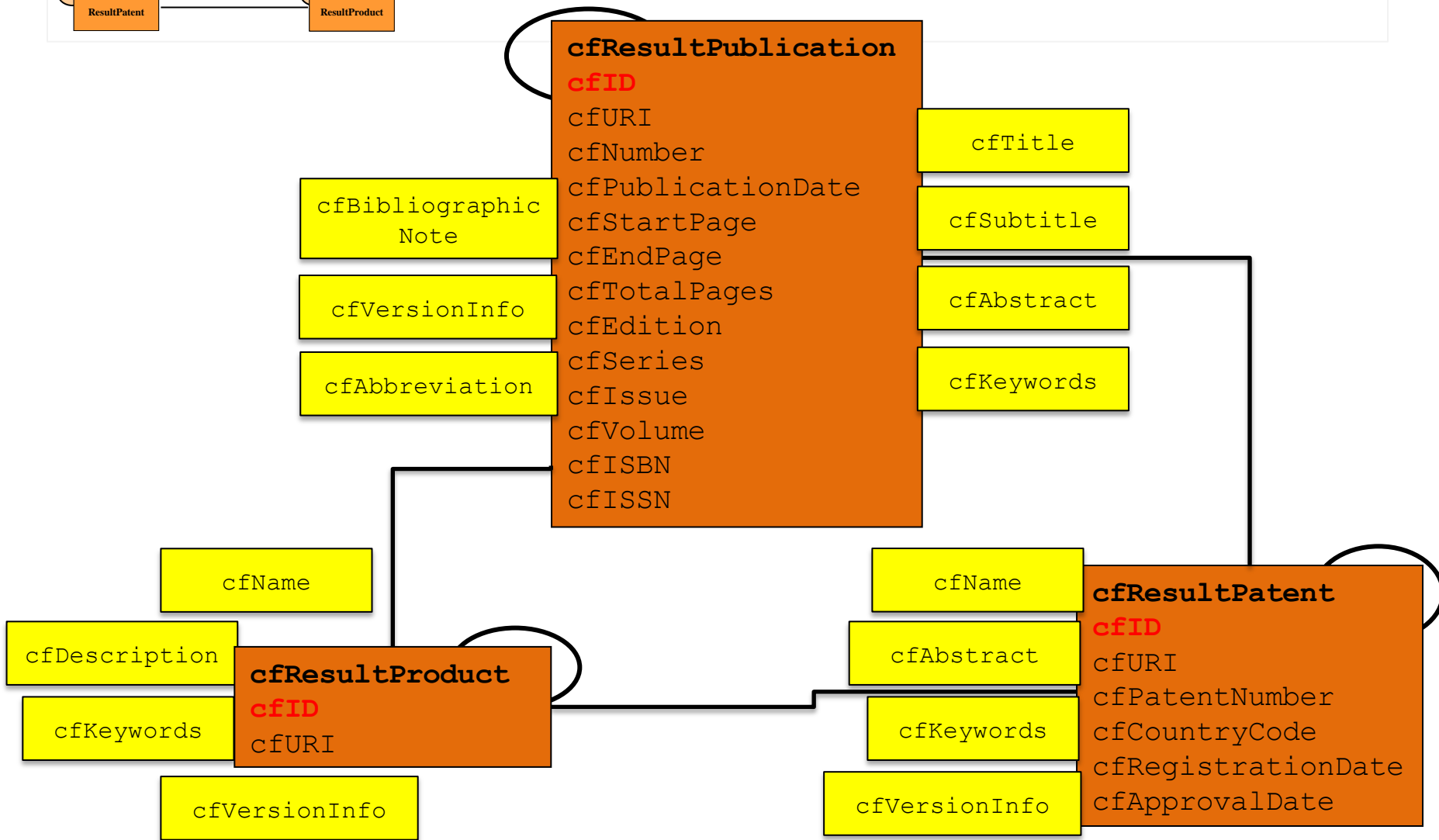
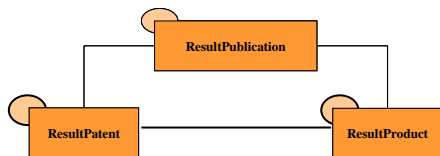
CERIF Result Entities



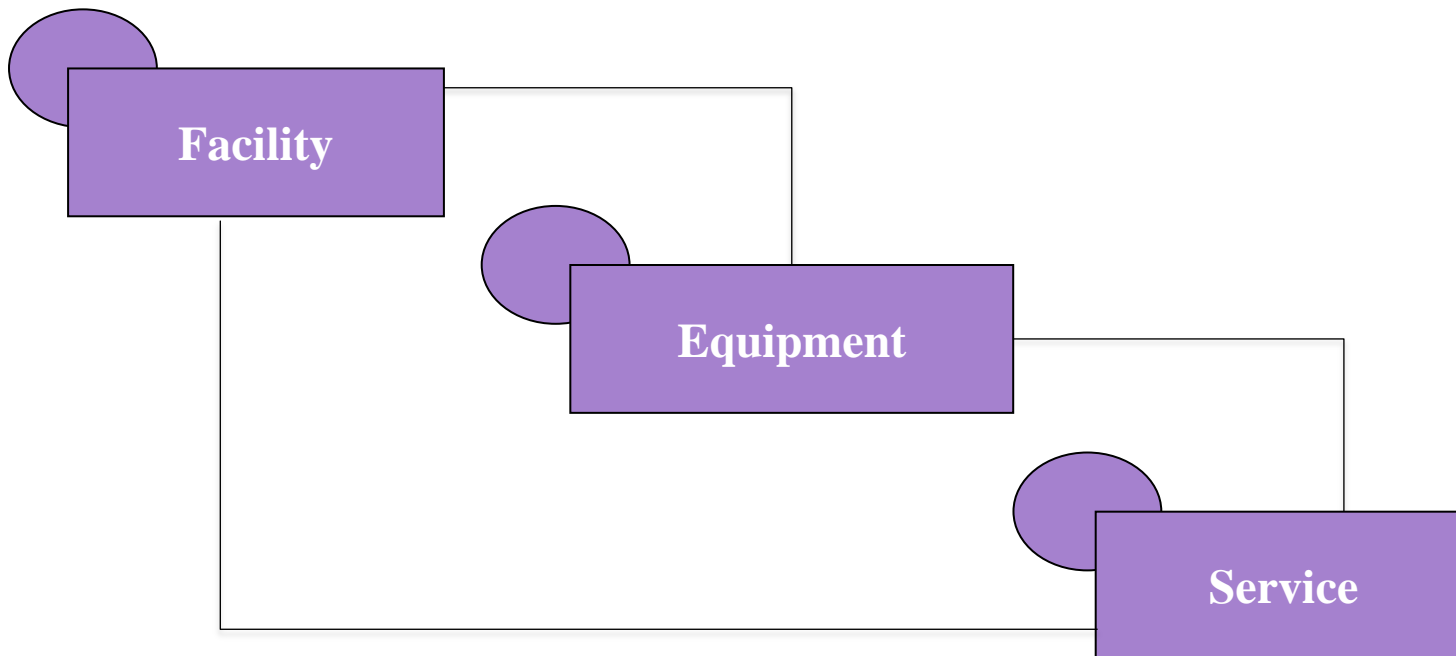
CERIF Result Entities



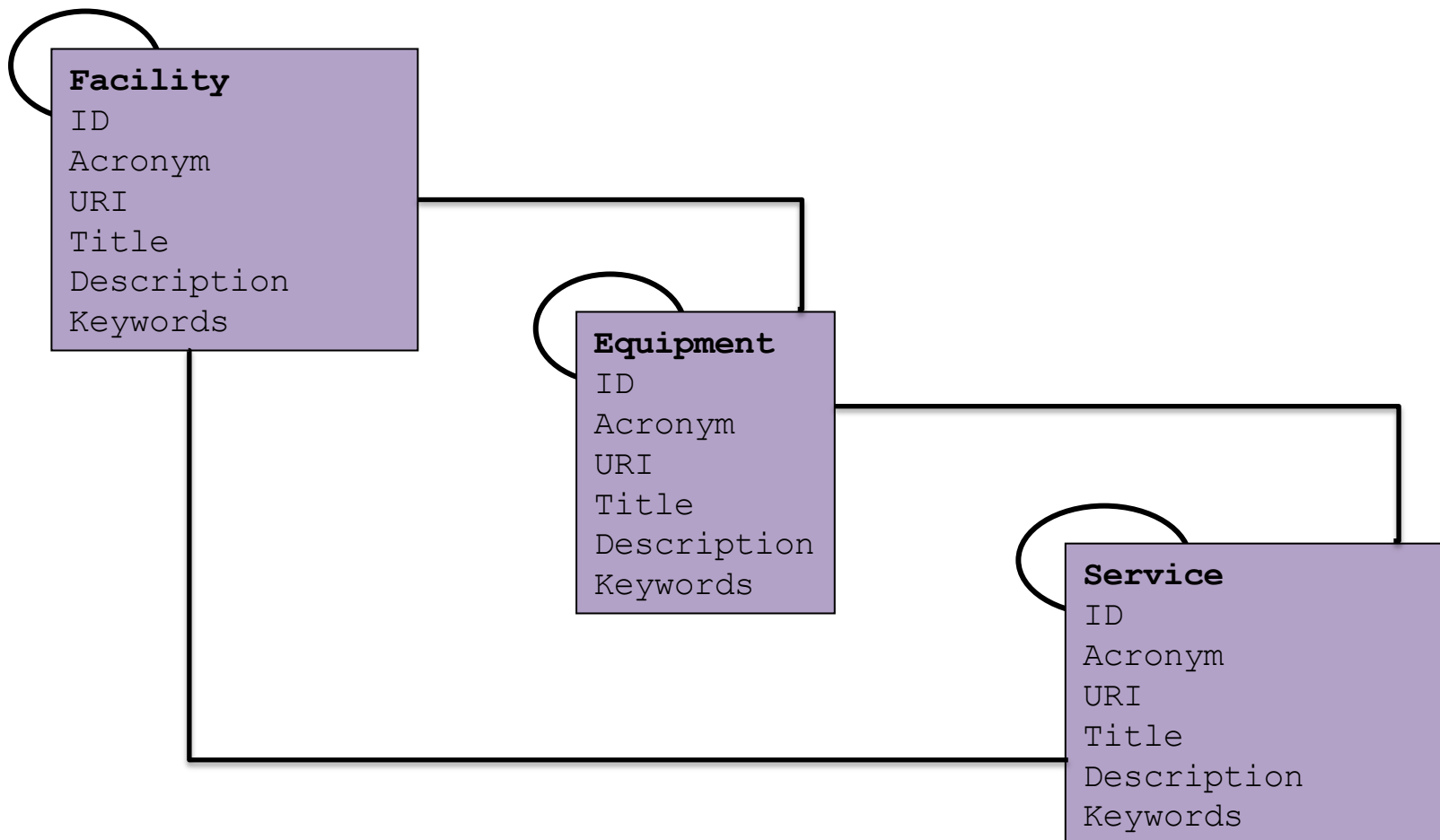
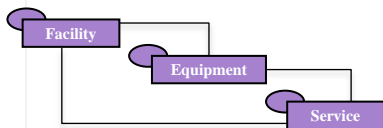
CERIF Result Entities



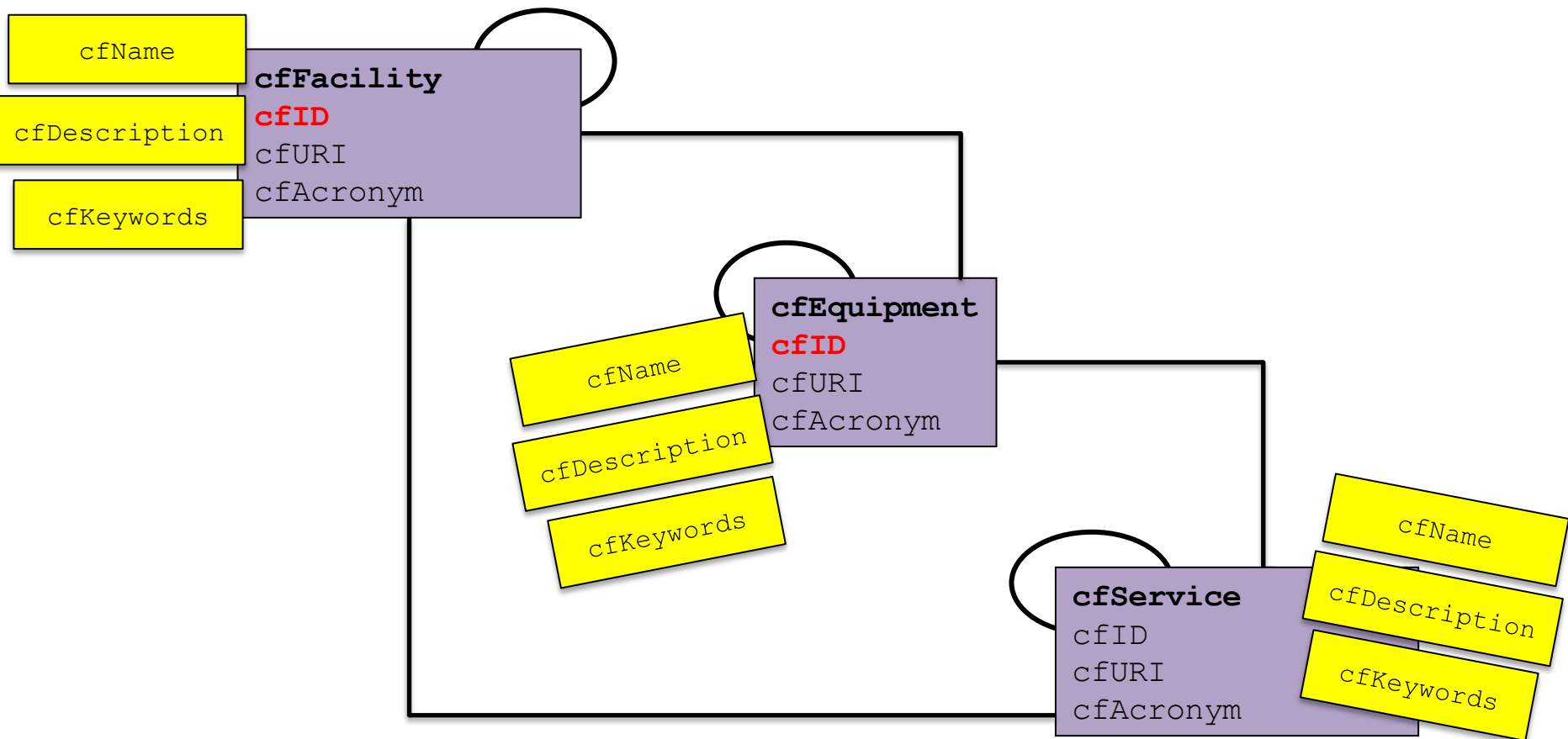
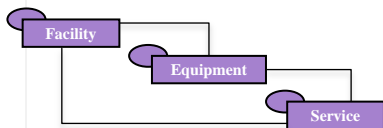
CERIF Infrastructure Entities



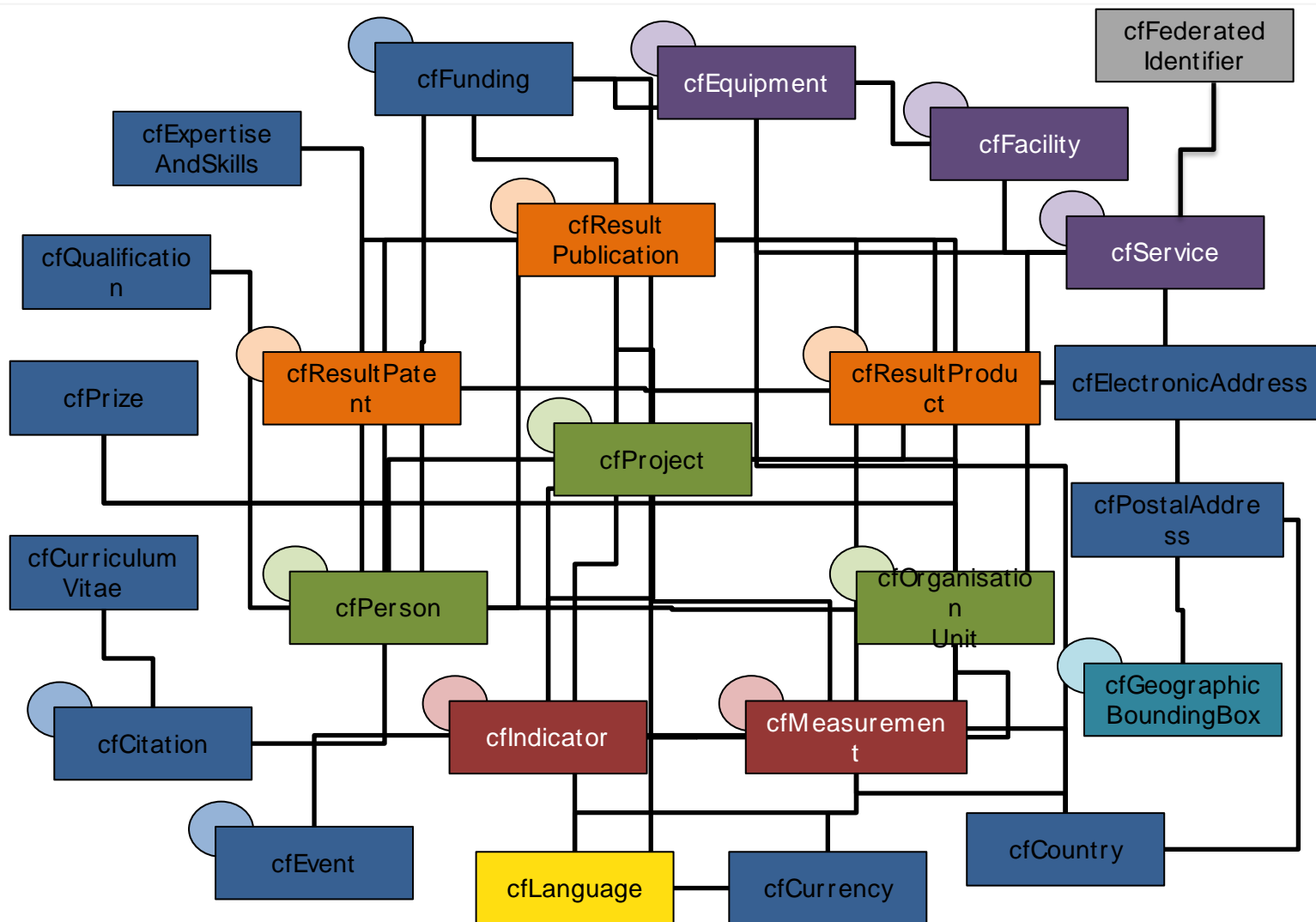
CERIF Infrastructure Entities



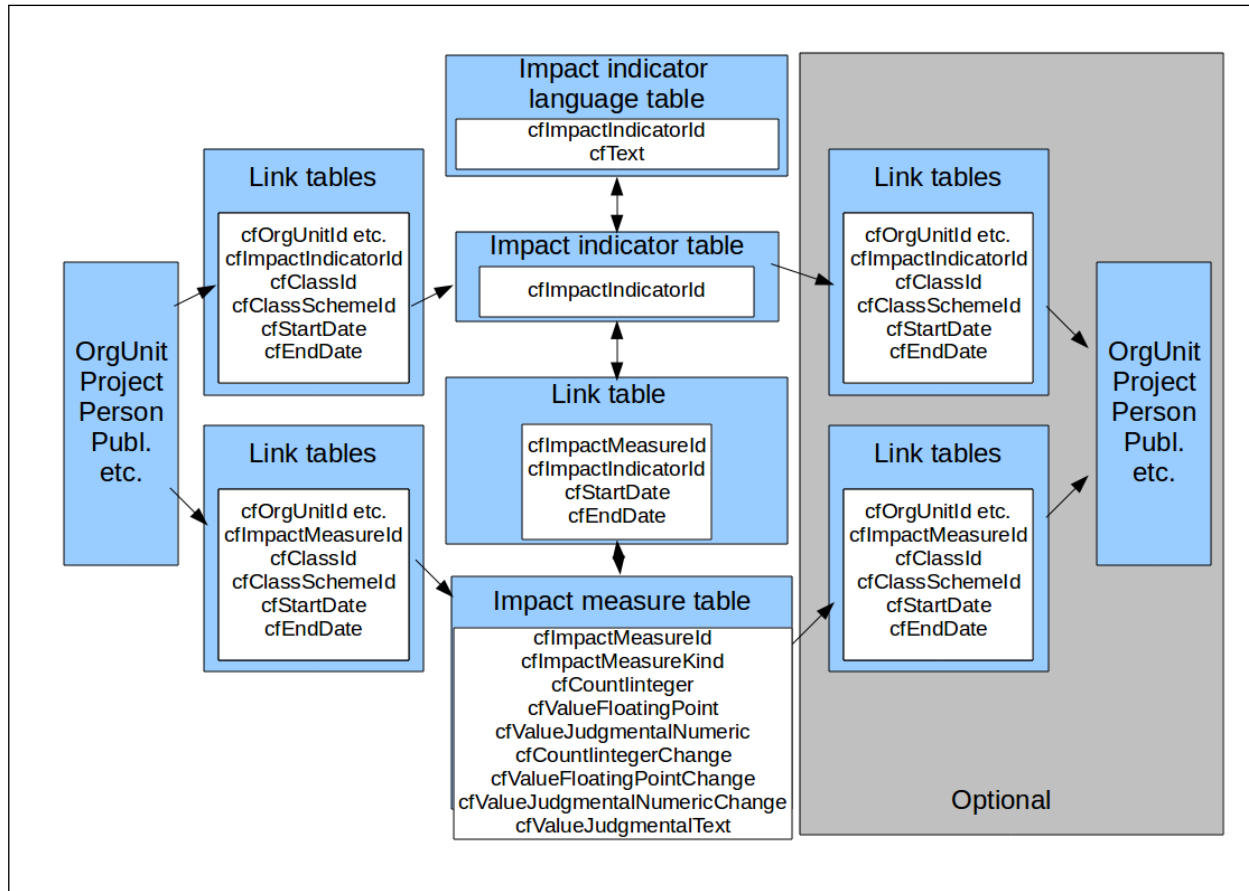
CERIF Infrastructure Entities



CERIF 1.5



Measuring Impact in CERIF (MICE)



MICE, a JISC-funded Project coordinated by Richard Gartner, Kings College, London, UK

CERIF Measurement & Indicator



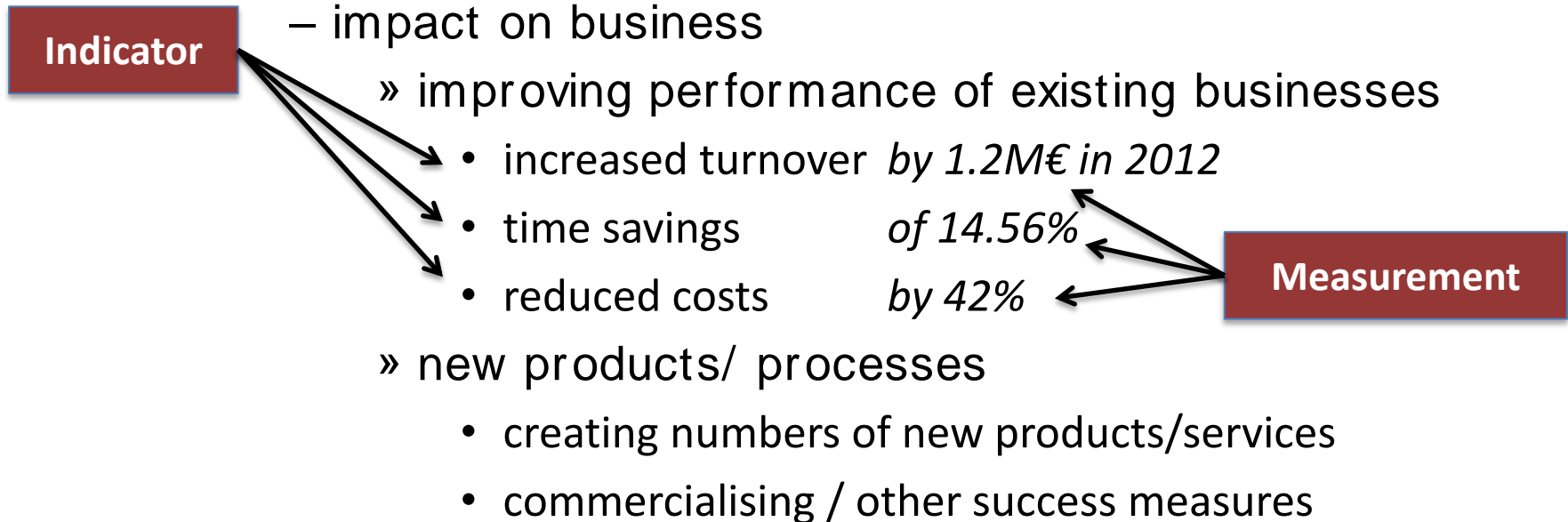
Is an Aggregation Entity

- cfMeasurementIdentifier
- cfCountInteger
- cfCountIntegerChange
- cfValueFloatingPoint
- cfCountFloatingPointChange
- cfValueJudgementalNumeric
- cfValueJudgementalNumericChange
- cfValueJudgementalText
- cfValueJudgementalTextChange
- cfURI

Measurement & Indicator (some examples)

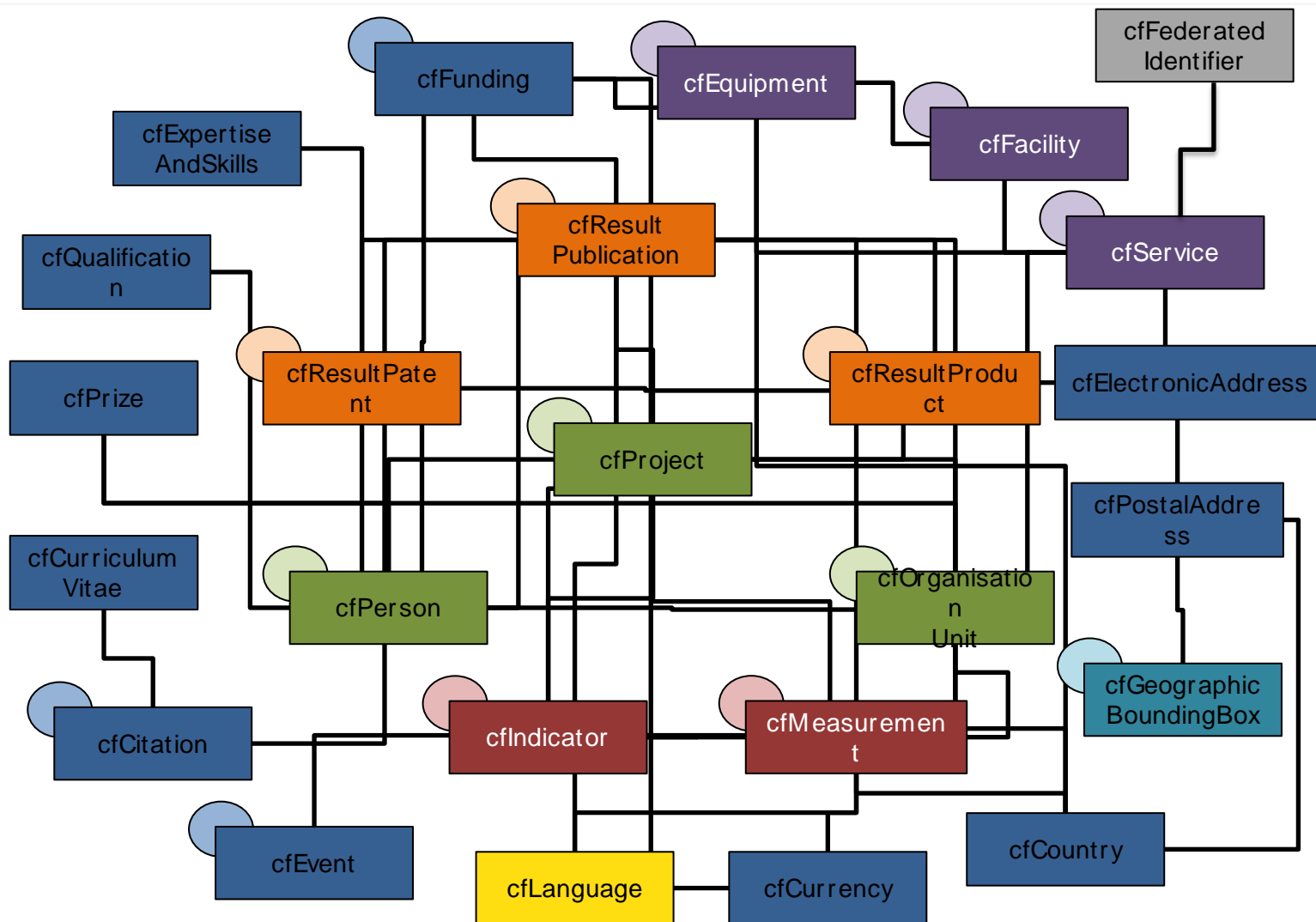
– economic and commercial

- economic



Extract from the MICE List of Indicators

CERIF 1.5

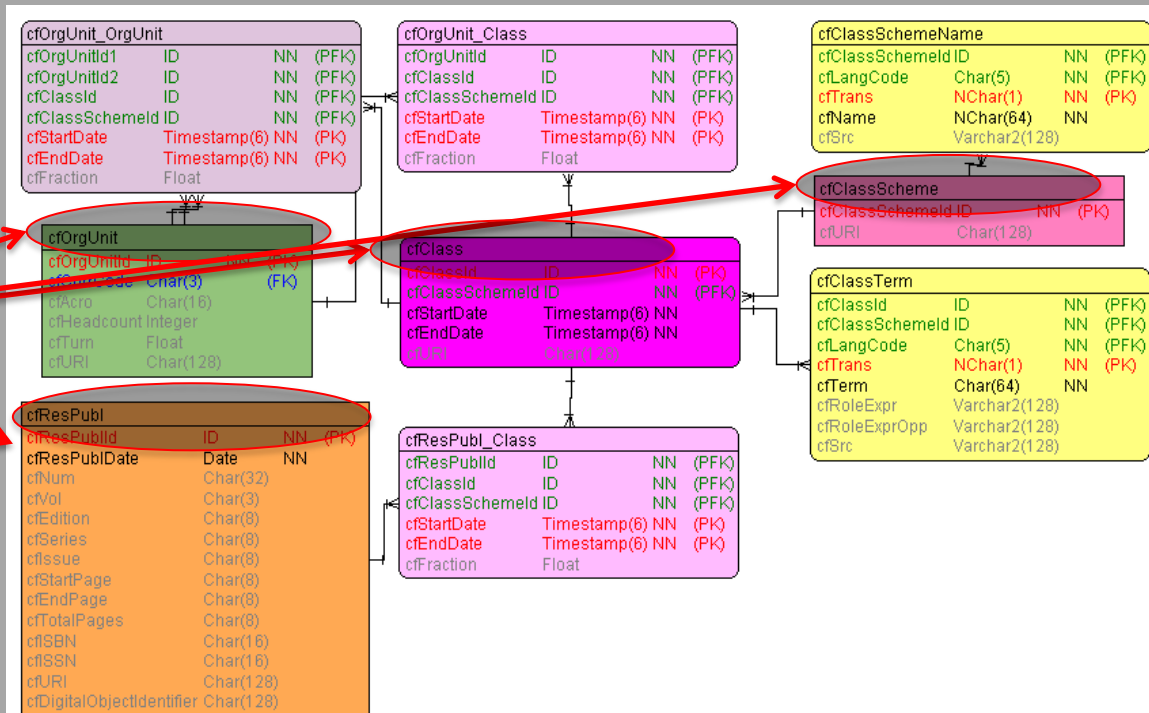


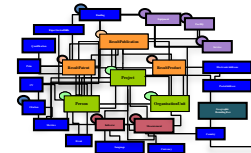
CERIF – Generic Entity Structure

Generic

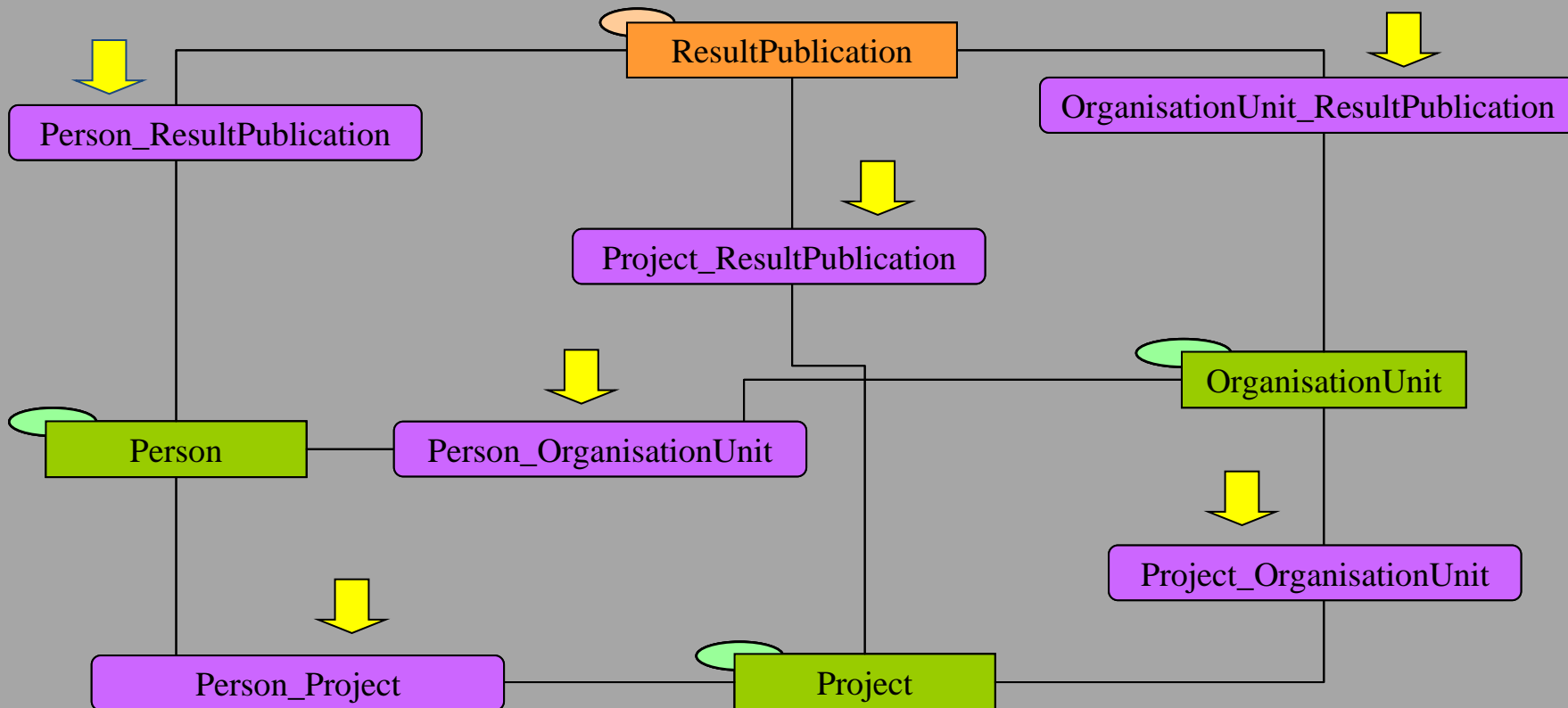
Identifier
URI
Attributes

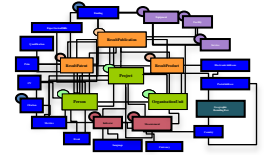
Multilingual Entities
Relationships (Links)



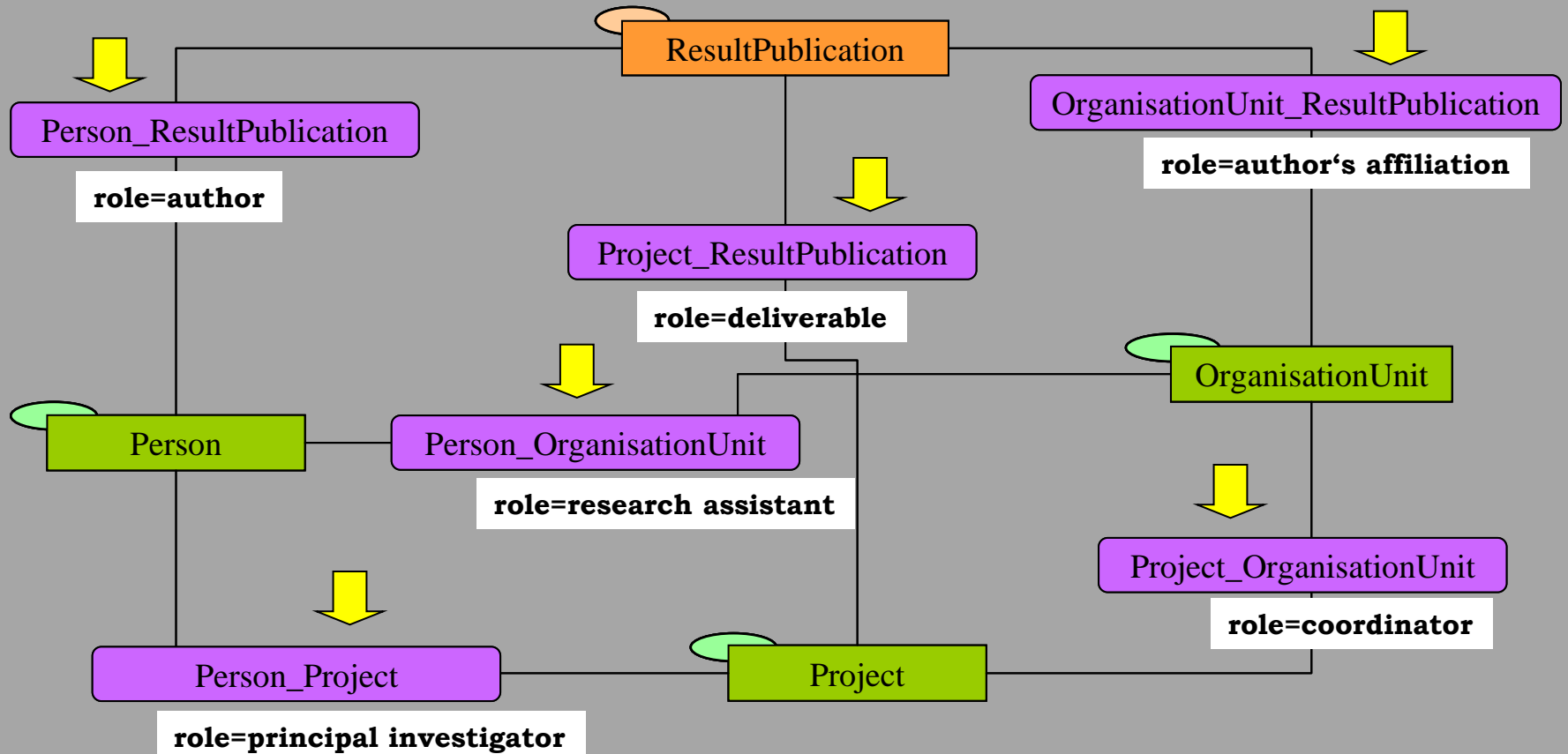


Some CERIF Link Entities





Some CERIF Link Entities



Some CERIF Link Entities

cfPers_Pers			
cfPersId1	ID	NN	(PFK)
cfPersId2	ID	NN	(PFK)
cfClassId	ID	NN	(PFK)
cfClassSchemeld	ID	NN	(PFK)
cfFraction	Float		
cfStartDate	Timestamp(6)	NN	(PK)
cfEndDate	Timestamp(6)	NN	(PK)

cfProj_OrgUnit			
cfProjId	ID	NN	(PFK)
cfOrgUnitId	ID	NN	(PFK)
cfClassId	ID	NN	(PFK)
cfClassSchemeld	ID	NN	(PFK)
cfFraction	Float		
cfStartDate	Timestamp(6)	NN	(PK)
cfEndDate	Timestamp(6)	NN	(PK)

cfOrgUnit_Equip			
cfOrgUnitId	ID	NN	(PFK)
cfEquipId	ID	NN	(PFK)
cfClassId	ID	NN	(PFK)
cfClassSchemeld	ID	NN	(PFK)
cfFraction	Float		
cfStartDate	Timestamp(6)	NN	(PK)
cfEndDate	Timestamp(6)	NN	(PK)
cfCurrCode	Char(3)		(FK)
cfAvailability	Char(64)		
cfConditions	Char(255)		
cfPrice	Float		

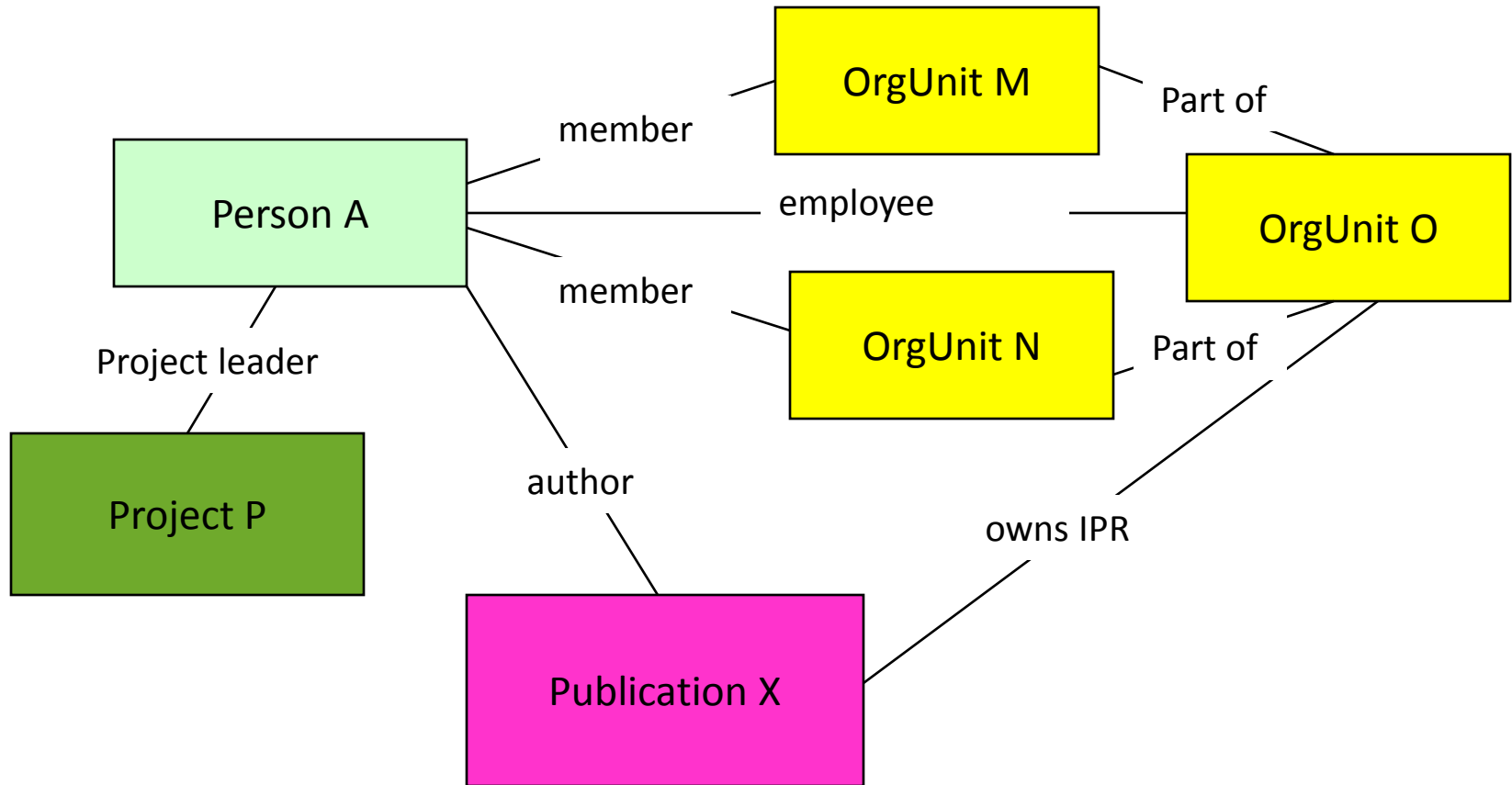
cfOrgUnit_Class			
cfOrgUnitId	ID	NN	(PFK)
cfClassId	ID	NN	(PFK)
cfClassSchemeld	ID	NN	(PFK)
cfFraction	Float		
cfStartDate	Timestamp(6)	NN	(PK)
cfEndDate	Timestamp(6)	NN	(PK)

cfClass_Class			
cfClassId1	ID	NN	(PFK)
cfClassId2	ID	NN	(PFK)
cfClassSchemeld1	ID	NN	(PFK)
cfClassSchemeld2	ID	NN	(PFK)
cfClassId	ID	NN	(PFK)
cfClassSchemeld	ID	NN	(PFK)
cfFraction	Float		
cfStartDate	Timestamp(6)	NN	(PK)
cfEndDate	Timestamp(6)	NN	(PK)

cfProj_Pers			
cfProjId	ID	NN	(PFK)
cfPersId	ID	NN	(PFK)
cfClassId	ID	NN	(PFK)
cfClassSchemeld	ID	NN	(PFK)
cfFraction	Float		
cfStartDate	Timestamp(6)	NN	(PK)
cfEndDate	Timestamp(6)	NN	(PK)

Result_Publication Instance Diagram

(slide by Keith Jeffery)



CERIF Semantic Layer

Allows to capture any Schema or Structure

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

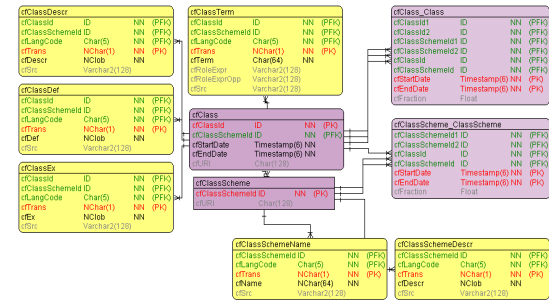
Open / Extensible in all directions

- New Schemas
- New Concepts / Terms
- New Relationships

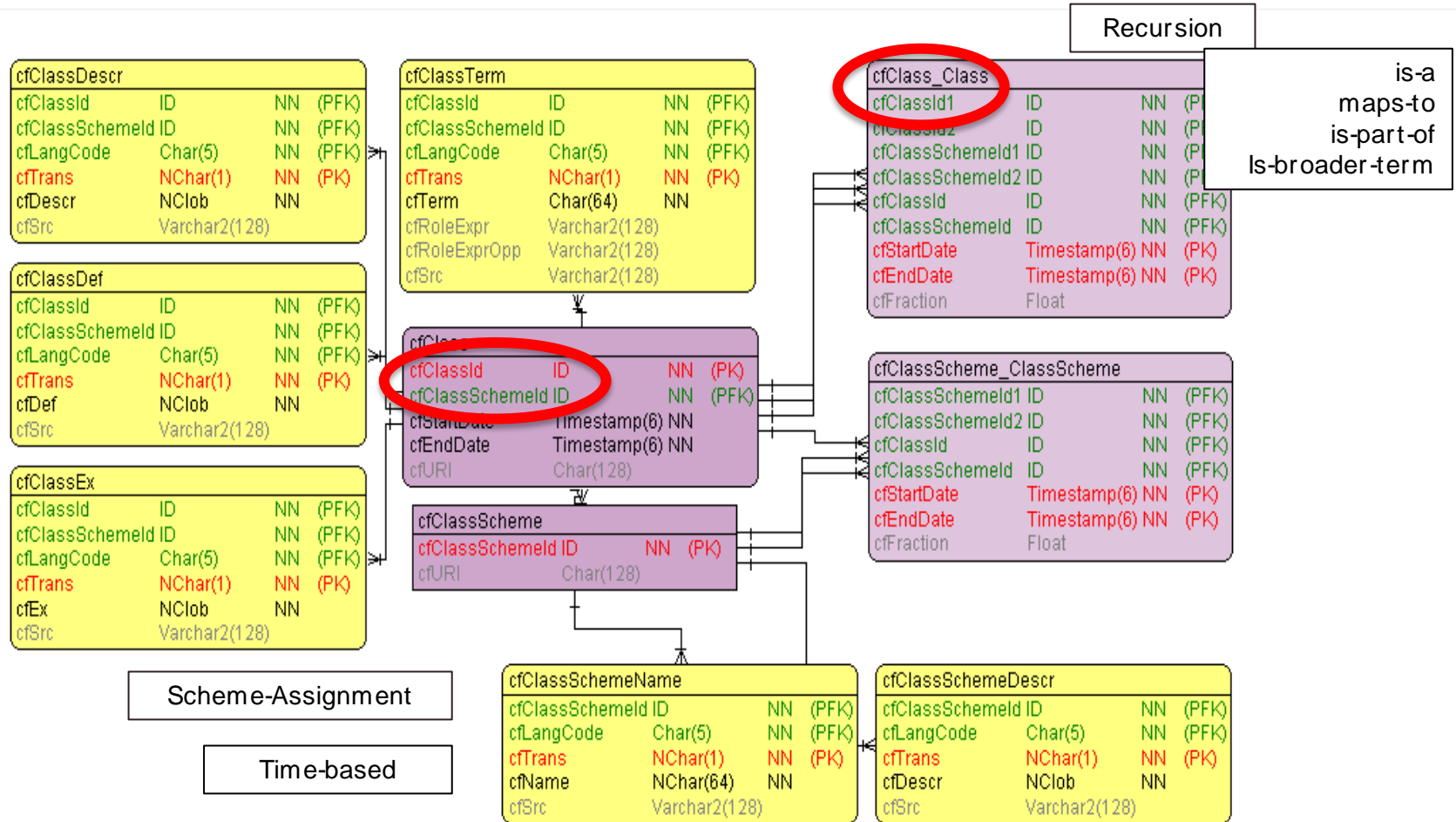
Enables to manage

- Roles / Types Semantics
- Subject Headings
- Archiving (Time component)

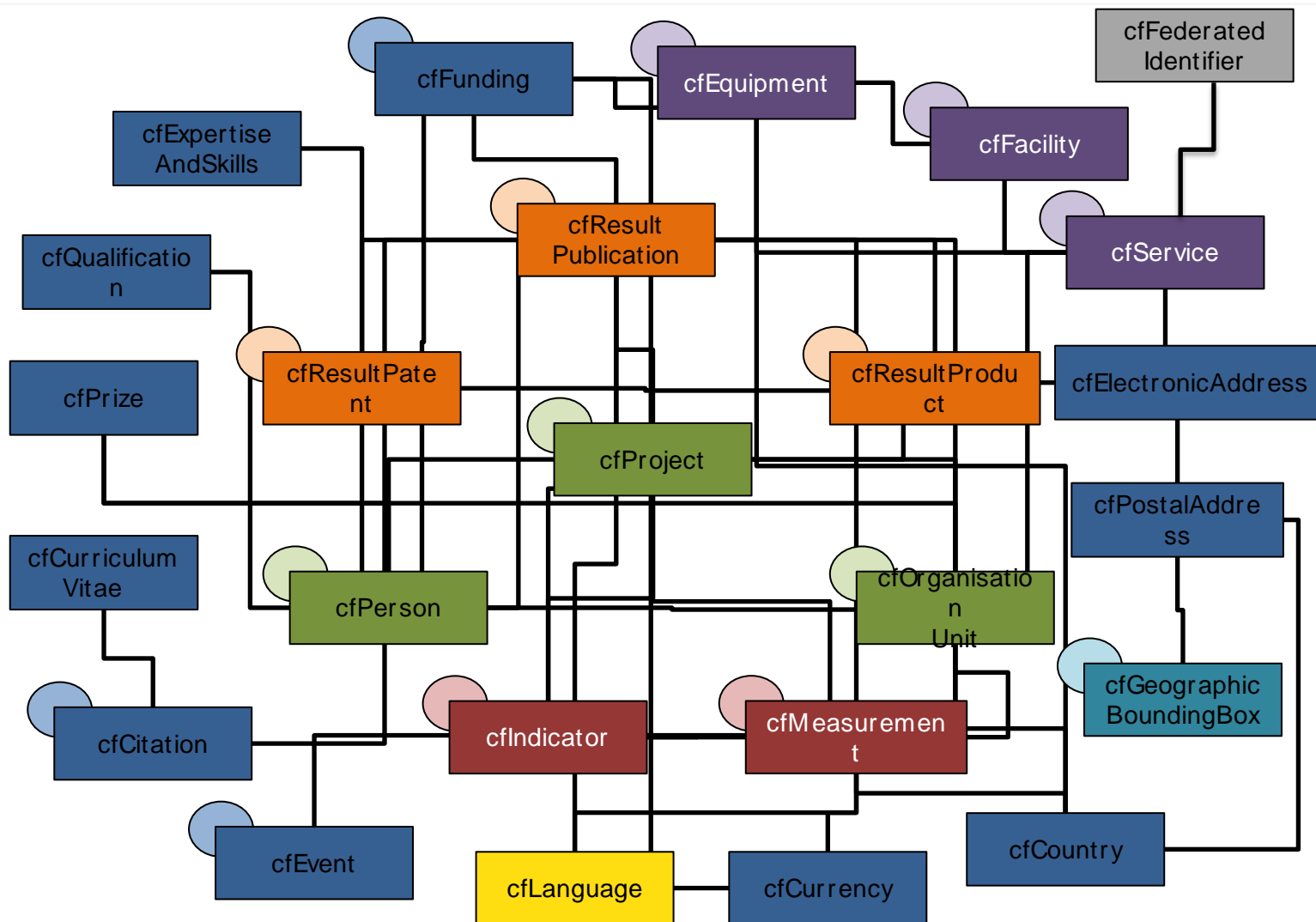
Allows for simple Mappings between Schemes



CERIF Semantic Layer (Declared Semantics)



CERIF 1.5



CERIF Federated Identifiers

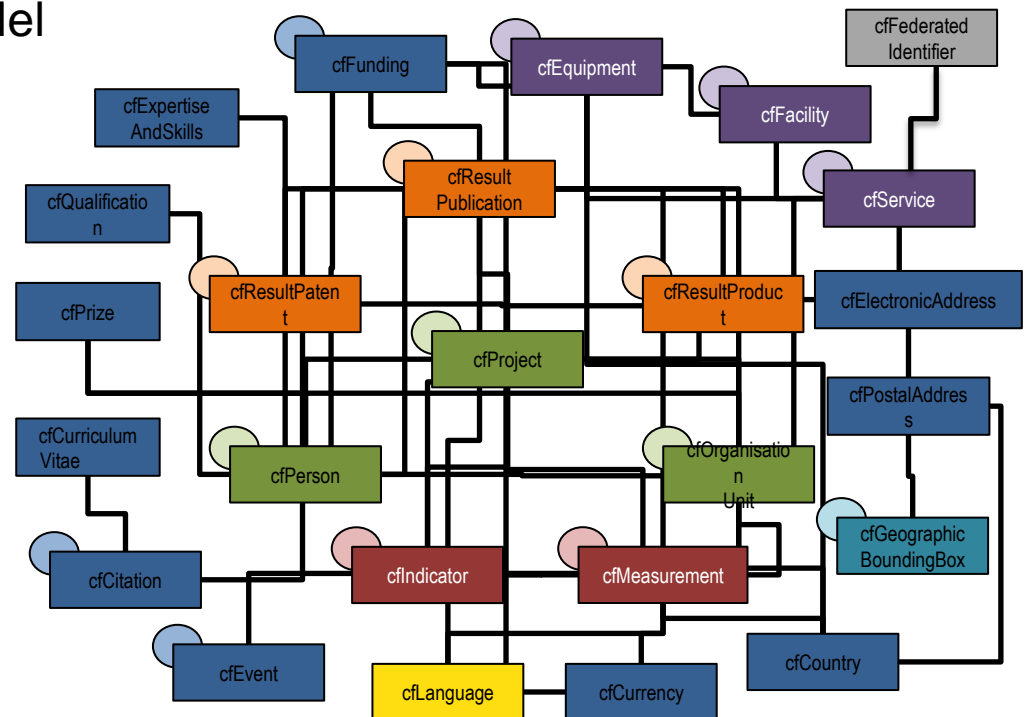
- Result Publication
 - DOI
 - WoS Accession Number
- Person
 - Social Security Number
 - Staff Id in HR system
 - Author identifier
 - ORCID
 - ScopusID
- Project/ Grant
 - Funder's reference number
- Organisation
 - VAT Identification Number
 - Internal Code
 - FundId
- Classification
 - External Code

CERIF Federated Identifiers

- Records the “tag” by which an object is known elsewhere
- For any Base, Result, Infrastructure, or 2nd Level entity
- Connected to a Service representing the context:
 - The issuer of the identifier
 - Usually an information system

CERIF XML 1.5 Interchange Format

- For point-to-point interchange
- XML namespace
- XML Schema
 - Based on the ER model



CERIF 1.5 XML Interchange Format

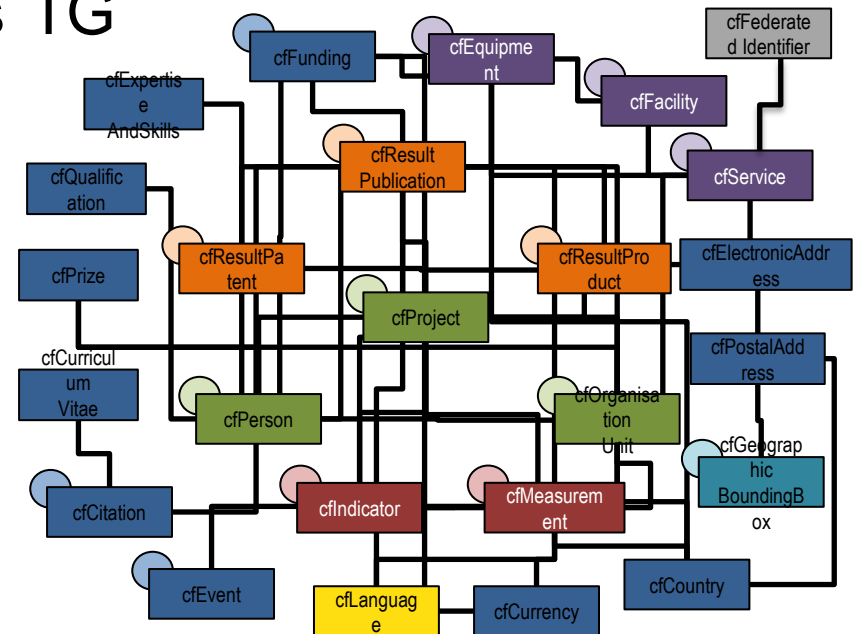
```
<CERIF xmlns="urn:xmlns:org:eurocris:cerif-1.5-1">
  <cfProj>
    <cfProjId>internal-project-identifier</ cfProjId>
    <cfAcro>ACRO</ cfAcro>
    <cfURI>http:// www.project-url.ac.uk/ acro.html</ cfURI>
    <cfTitle cfLangCode="en" cfTrans="o">The title of the project</ cfTitle>
    <cfAbstr cfLangCode="en" cfTrans="o">The goals of the project</ cfAbstr>
    <cfProj_Class>
      <cfClassId>infrastructure-project-uuid</ cfClassId>
      <cfClassSchemeld>-project-types-scheme-uuid</ cfClassSchemeld>
    </ cfProj_Class>
    <cfProj_OrgUnit>
      <cfOrgUnitId>orgunit-1-identifier</ cfOrgUnitId>
      <cfClassId>coordinator-uuid</ cfClassId>
      <cfClassSchemeld>orgunit-project-roles-scheme-uuid</ cfClassSchemeld>
      <cfStartDate>from-datetime</ cfStartDate>
      <cfEndDate>to-datetime</ cfEndDate>
    </ cfProj_OrgUnit>
  </ cfProj>
</ CERIF>
```

CERIF 1.5 Release

- CERIF Model Introduction and Specification **coming**
- CERIF XML Data Exchange Format Specification ✓
- CERIF Formal Semantics (Vocabulary) ✓
- CERIF SQL Scripts ✓ (euroCRIS members only)
- CERIF XML Schemas ✓
- CERIF XML Examples ✓
- CERIF Semantics (Excel) ✓

Ongoing Activities: CERIF 1.6

- Model Cleaning
- Research Data
- Cross-TG Activities
 - Linked Open Data TG
 - Institutional Repositories TG
 - Architectures TG
 - Indicators TG
 - Best Practice TG
- Cooperation with
 - CASRAI
 - VIVO
 - RDA
 - ORCID

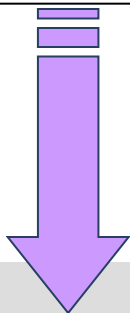


What makes CERIF shine

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

What is a CRIS?

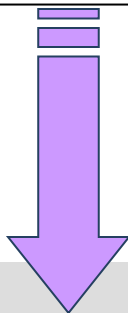
Current



... that means

- of current interest
- not necessarily ongoing

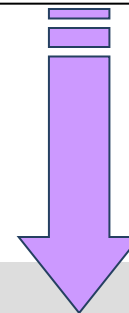
Research Information



... information about

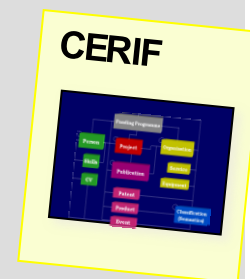
- Researchers
- Organisations (Research-performing, Funding)
- Funding Programmes, Calls
- Projects
- ...

System



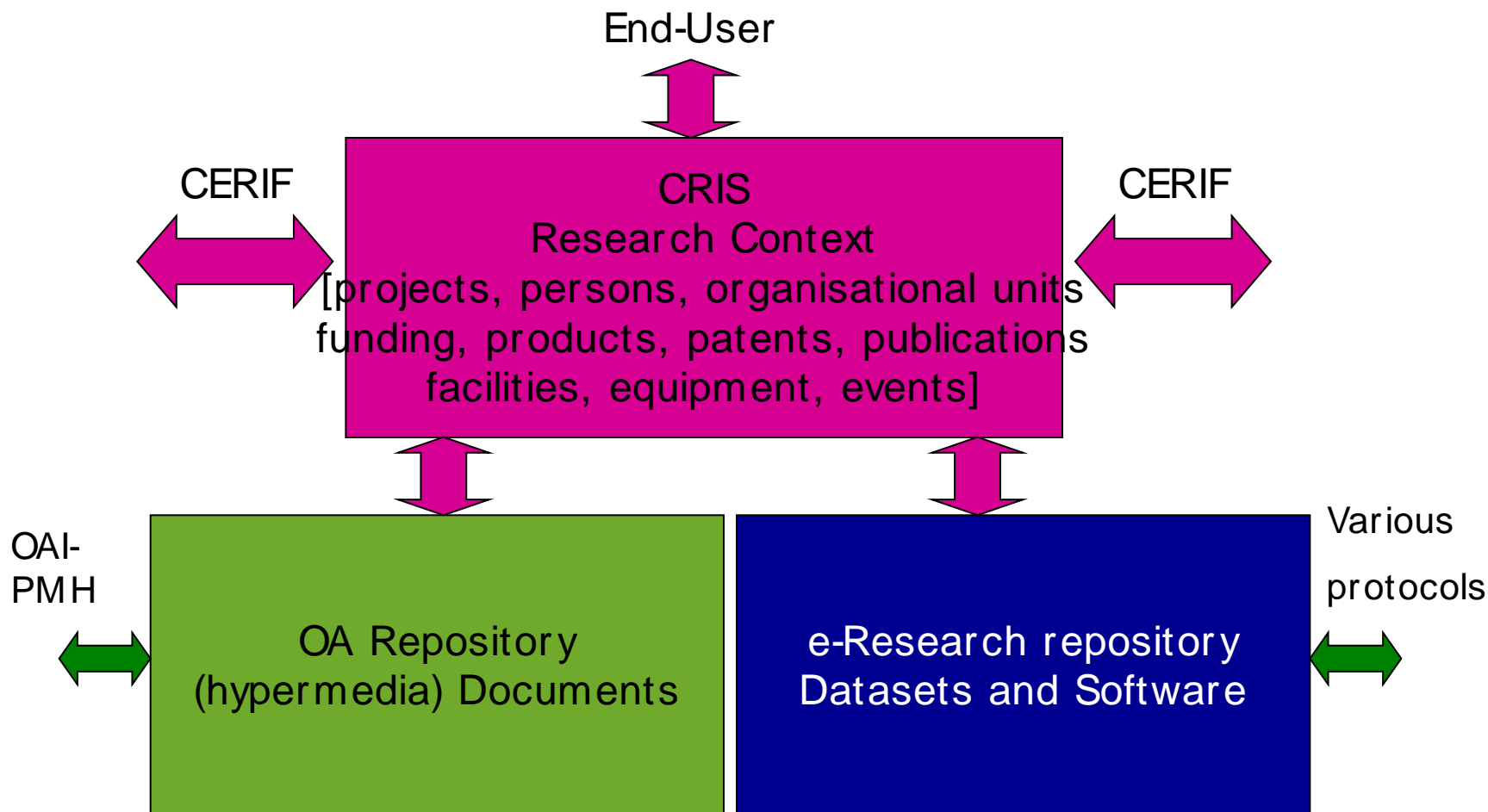
... driven by

- Concepts
- Model
- Implementation (Information System)



an integrated approach towards managing research information

CRIS and Repositories at an institution (slide by Keith Jeffery)



ORCID



*International Council for Science;
Commission on Data Access*



EARMA

*European Association of Research
Managers and Administrators*



www.eurocris.org



All European Academies

← Alliance for Permanent Access →

