

# CERIF COURSE

## Session4: DataModel 2

Keith G Jeffery,

Director, IT CLRC

[k.g.jeffery@rl.ac.uk](mailto:k.g.jeffery@rl.ac.uk)

Anne Asserson,

University of Bergen

[anne.asserson@ub.uib.no](mailto:anne.asserson@ub.uib.no)

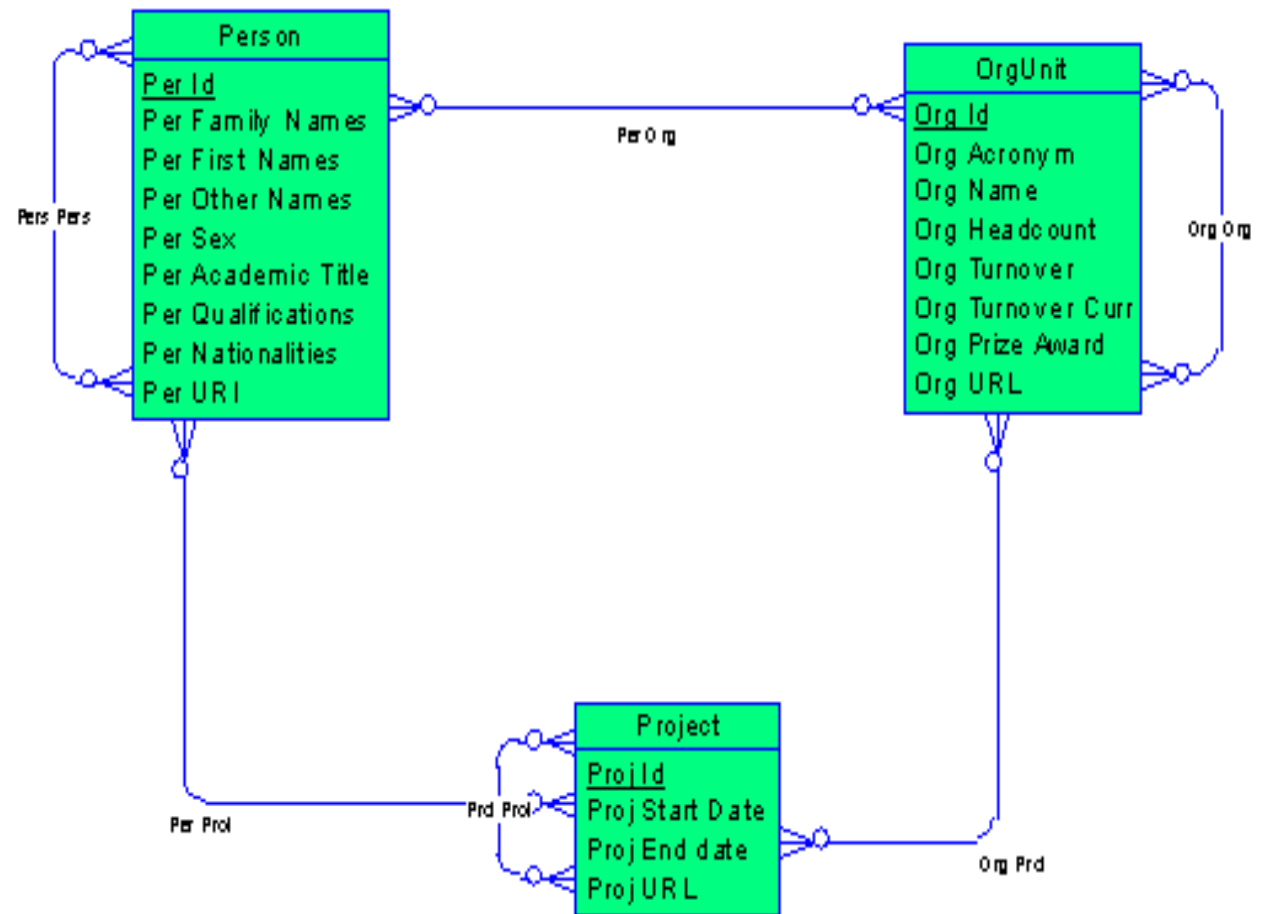
- In this session
  - We shall use extended E-R diagrams generated by the EC staff for CERIF2000 using Oracle Designer and available at [www.cordis.lu/cerif](http://www.cordis.lu/cerif)
    - In some cases errors have been found, these will be indicated
  - We shall use Database Description tables
    - With columns entity, attribute, type, constraints, comment / meaning
    - In some cases errors have been found, these will be indicated
- At all times the latest production version of CERIF is available from the CERIF Task Group pages under [www.eurocris.org](http://www.eurocris.org)

- Full model – primary base entities
- Person – secondary base entities
- Person– Language-base entities
- Person– Lookup tables
- Person– linking relations: at one level, recursion, at different levels
- Exchange Model and Metadata Model

- Full model – primary base entities
- Person – secondary base entities
- Person– Language-base entities
- Person– Lookup tables
- Person– linking relations: at one level, recursion, at different levels
- Exchange Model and Metadata Model

# Full model – primary base entities

## Base Entities



# Full model – primary base entities

## Project

Entity	Attribute	Type	Constraint	Comments / Meaning
<b>BASE TABLES</b>				Primary 3 object types (entities)
				Note because of multilingual, text fields are separated out into language-field-base tables
Project	Project Id	char(32)	m, pk	
Project	Start Date	date	m	
Project	End Date	date	o	
Project	Status	char(8)	m, enumlist	stalled, completed...
Project	URI	char(128)		project homepage

## Full model – primary base entities

# Person

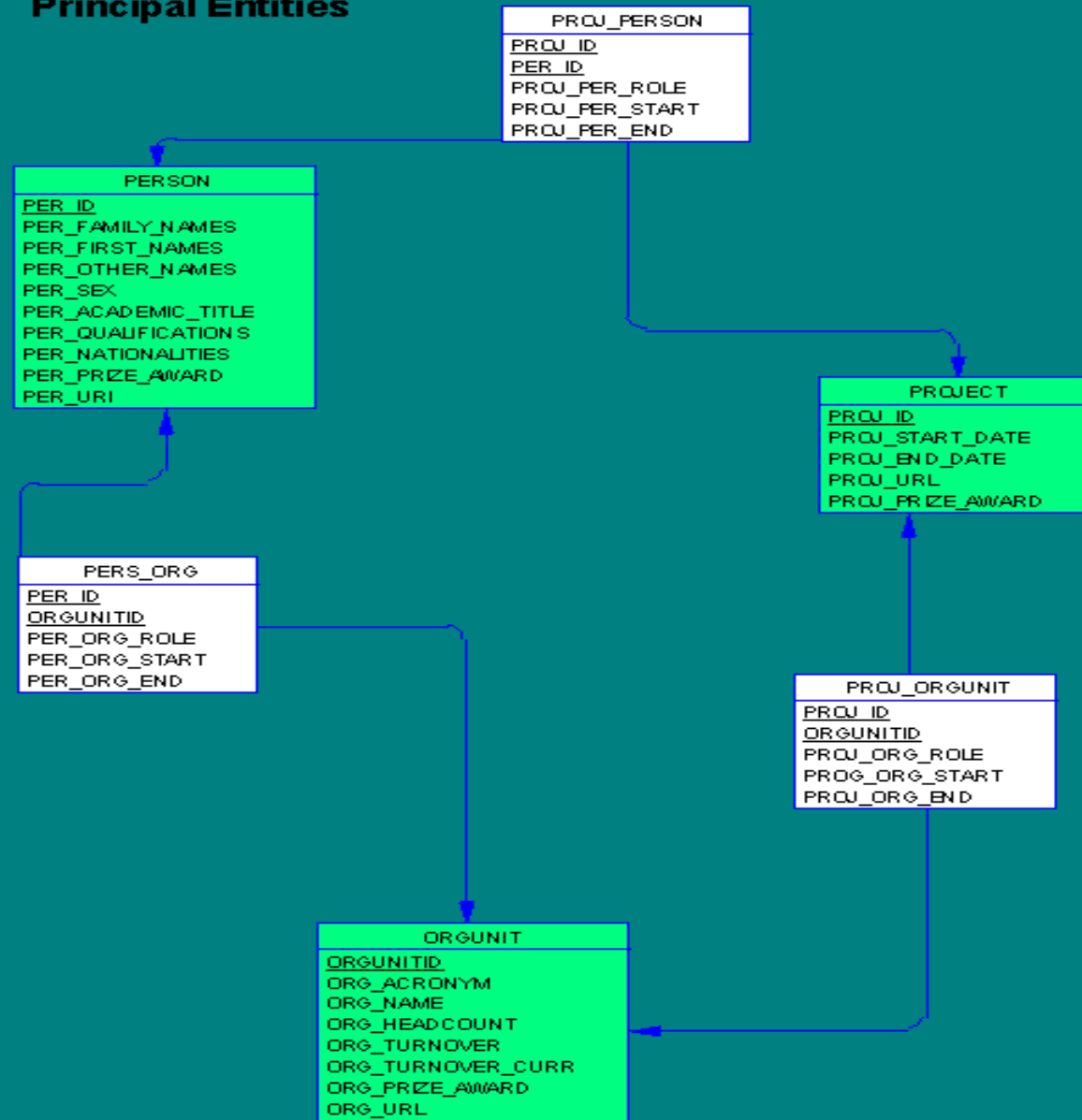
Person	Person Id	char(32)	m, pk	
Person	Family Names	char(32)	m	list, separated
Person	First Names	char(16)	o	list, separated
Person	Other Names	char(32)	o	list, separated
Person	Sex	char(1)	o, enumlist	M F
Person	Honorific Title	char(12)	o, enumlist	enumlist, separated
Person	Academic Title	char(12)	o, enumlist	enumlist, separated
Person	Qualifications	char(24)	o, enumlist	enumlist, separated
Person	Nationalities	char(16)	o, enumlist	enumlist, ISO 3166 standard, two letter codes, separated
Person	URI	char(128)		Person homepage

- Subsequent to CERIF 2000 it was decided that the comma separated list type attributes should be removed to separate tables linked to person to allow a clear 1:n relationship
- i.e. they were fully normalised
- This allows easy retrieval or update for a person with >1 value for these attributes

# OrgUnit

OrgUnit	OrgUnit Id	char(32)	m, pk	
OrgUnit	Acronym	char(16)	o	
OrgUnit	Type	char(8)	m, enumlist	enumlist, separated
OrgUnit	Headcount	char(2)	o	The harmonised package for the European Union Fifth Framework Programme provides definitions for Organization/Company size, which are used here.
OrgUnit	Turnover	float	o	total annual working budget
OrgUnit	Currency	char(3)	m if value else o	3 letter ISO 4217 currency codes to be used.
OrgUnit	URI	char(128)		OrgUnit homepage

### Principal Entities



## Full model – primary base entities

# Project-Person

Project-Person	Project Id	char(32)	m, fk, pk(part)	
Project-Person	Person Id	char(32)	m, fk, pk(part)	
Project-Person	Role	char(16)	o, enumlist	e.g. project leader, investigator, accountant
Project-Person	Start Date	date	o	
Project-Person	End Date	date	o	

## Full model – primary base entities

# Project-OrgUnit

Project-OrgUnit	Project Id	char(32)	m, fk, pk(part)	
Project-OrgUnit	OrgUnit Id	char(32)	m, fk, pk(part)	
Project-OrgUnit	Role	char(16)	o, enumlist	List of values are defined for EU Fifth Framework Programme  CO : Co-ordinator (=scientific, administrative and financial co-ordinator) CF: Only financial co-ordinator (if different from co-ordinator) AC: Associate contractor CR: Contractor (other than the co-ordinator)
Project-OrgUnit	Start Date	date	o	
Project-OrgUnit	End Date	date	o	

## Full model – primary base entities

# Person-OrgUnit

Person-OrgUnit	Person Id	char(32)	m, fk, pk(part)	
Person-OrgUnit	OrgUnit Id	char(32)	m, fk, pk(part)	
Person-OrgUnit	Role	char(16)	o, enumlist	e.g. division head, group leader
Person-OrgUnit	Start Date	date	o	
Person-OrgUnit	End Date	date	o	

## Secondary Base Entities

- There are two secondary base entities with special characteristics because they are not functional sub-entities of any of {Project | Person | OrgUnit}
- These are
  - Funding Programme
  - Contact

# Funding Programme

Secondary Base  
Entity  
Funding  
Programme

It has link  
relations to  
Project and  
OrgUnit

Funding/Programme	Funding/Programme Id	char(32)	m, pk	
Funding/Programme	Start Date	date	o	
Funding/Programme	End Date	date	o	
Funding/Programme	Budget	float	o	
Funding/Programme	Currency	char(3)	m if value else o	
Funding/Programme	URI	char(128)	o	contact URI

Secondary Base  
Table  
Contact

Has link  
relations to  
Person and  
OrgUnit

Contact	Contact Id	char(32)	m, pk	
Contact	Addrline1	char(80)	o	
Contact	Addrline2	char(80)	o	
Contact	Addrline3	char(80)	o	
Contact	Addrline4	char(80)	o	
Contact	Addrline5	char(80)	o	
Contact	City Town	char(32)	m	
Contact	State of Country	char(32)	o	
Contact	Postcode	char(16)	m	
Contact	Country code	char(2)	m, enumlist	
Contact	NUTS Code	char(10)	m, enumlist	Eurostat region locator code, mandatory in EU
Contact	Telephone	char(24)	o	list, separated
Contact	Fax	char(24)	o	list, separated
Contact	Email	char(32)	o	
Contact	URI	char(128)	o	contact URI

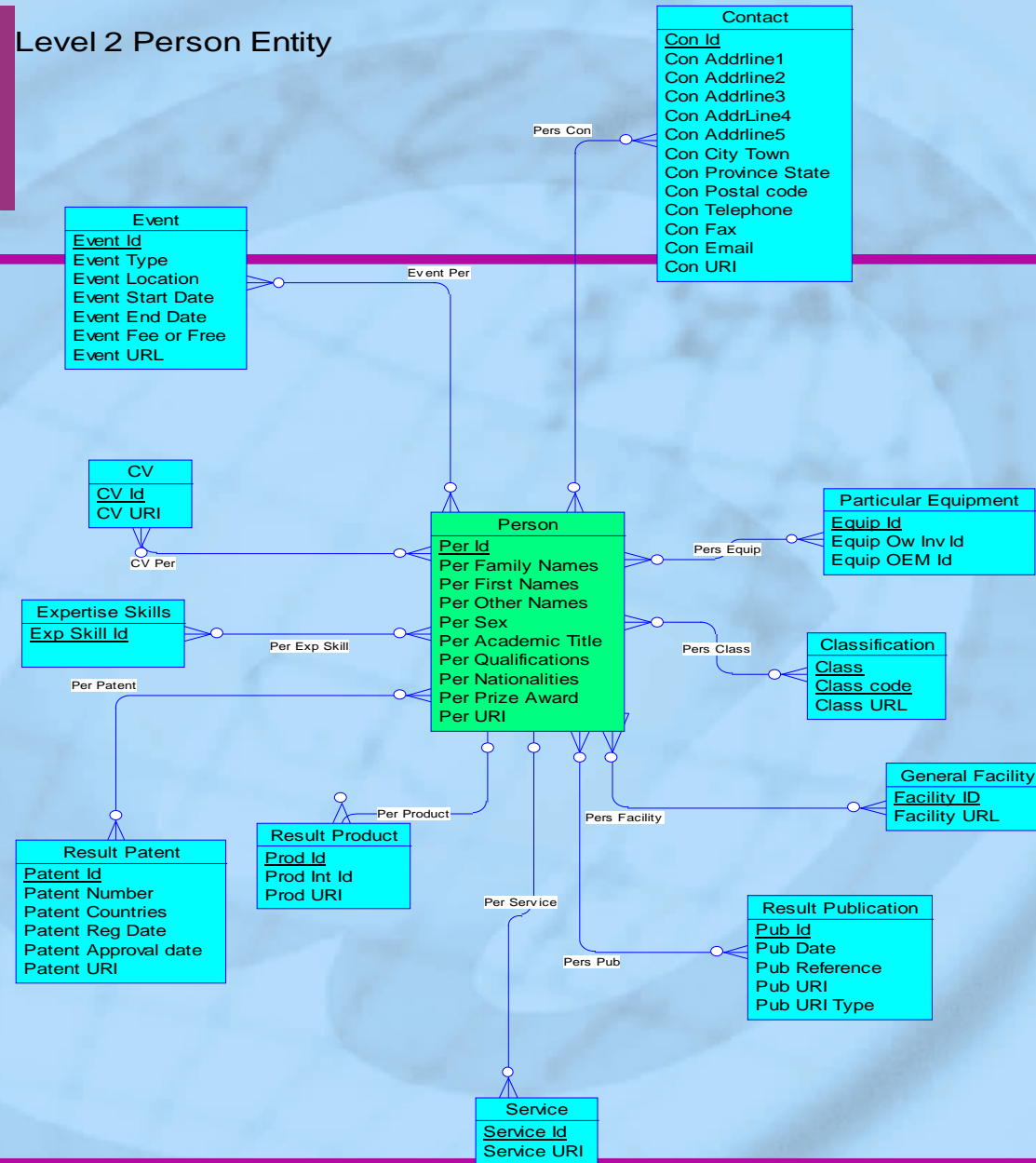
- We could go into depth with the
  - Secondary base tables
  - Language base tables
  - Lookup tables
  - Linking relation tables
- From any of the three primary entities
- For the purpose of this course we shall do this for PERSON

- Take Base Entity Person
  - Secondary base tables
  - Secondary base tables and language base tables
  - Secondary base tables and lookup tables (without links)
  - Lookup tables and links
  - Secondary Base tables and links

- Full model – primary base entities
- **Person** – secondary base entities
- Person– Language-base entities
- Person– Lookup tables
- Person– linking relations: at one level, recursion, at different levels
- Exchange Model and Metadata Model

- Take Base Entity Person
  - Secondary base tables
  - Secondary base tables and language base tables
  - Secondary base tables and lookup tables (without links)
  - Lookup tables and links
  - Secondary Base tables and links

Level 2 Person Entity

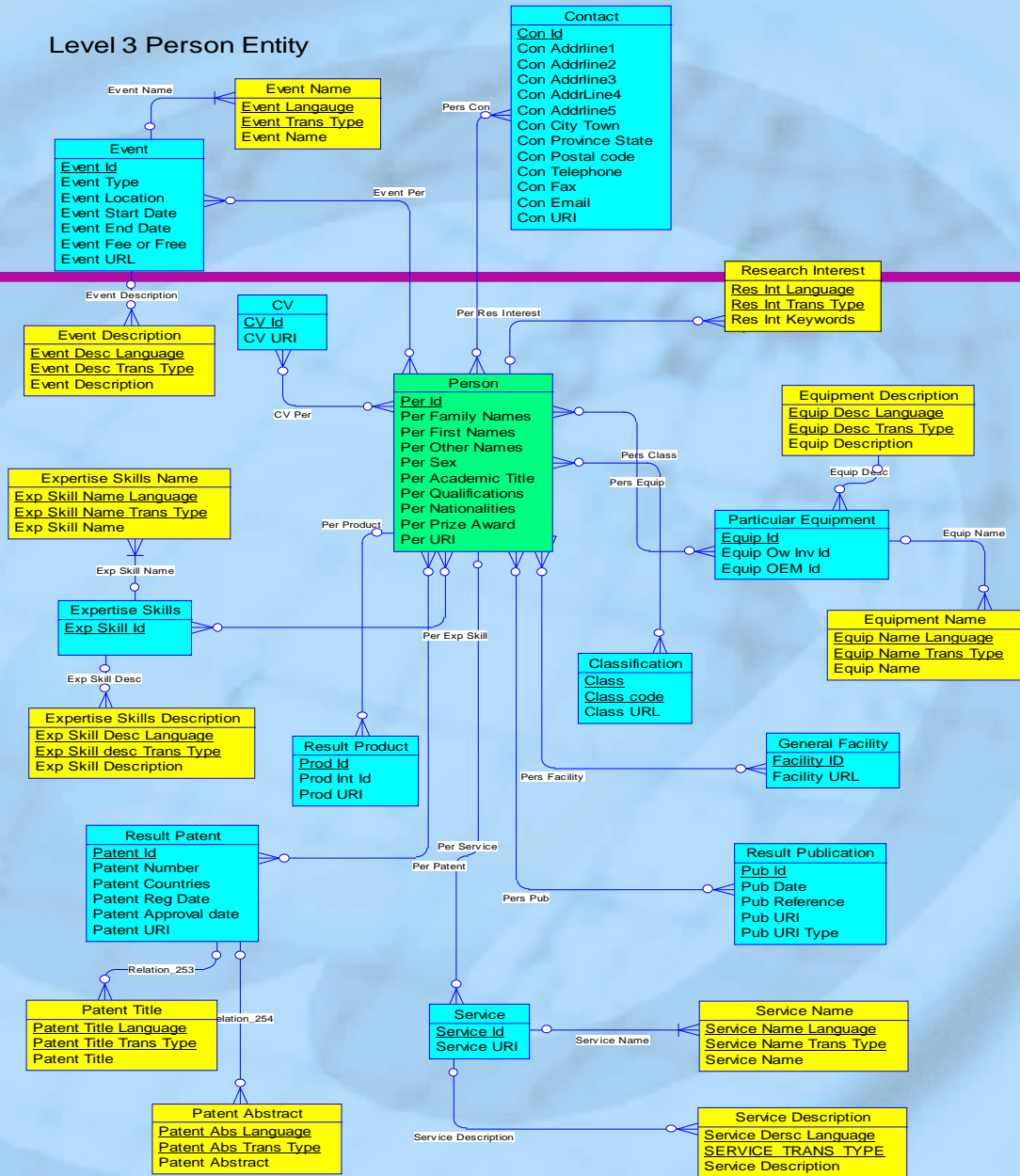


# Secondary Base Example CV (Curriculum Vitae)

CV	CV Id	char(32)	m, pk	
CV	CV	Multi-media	m	to repeat fields in person + profess history
CV	URI	char(128)		CV homepage

- Full model – primary base entities
- Person – secondary base entities
- Person– Language-base entities
- Person– Lookup tables
- Person– linking relations: at one level, recursion, at different levels
- Exchange Model and Metadata Model

Level 3 Person Entity



# Person– Language-base entities

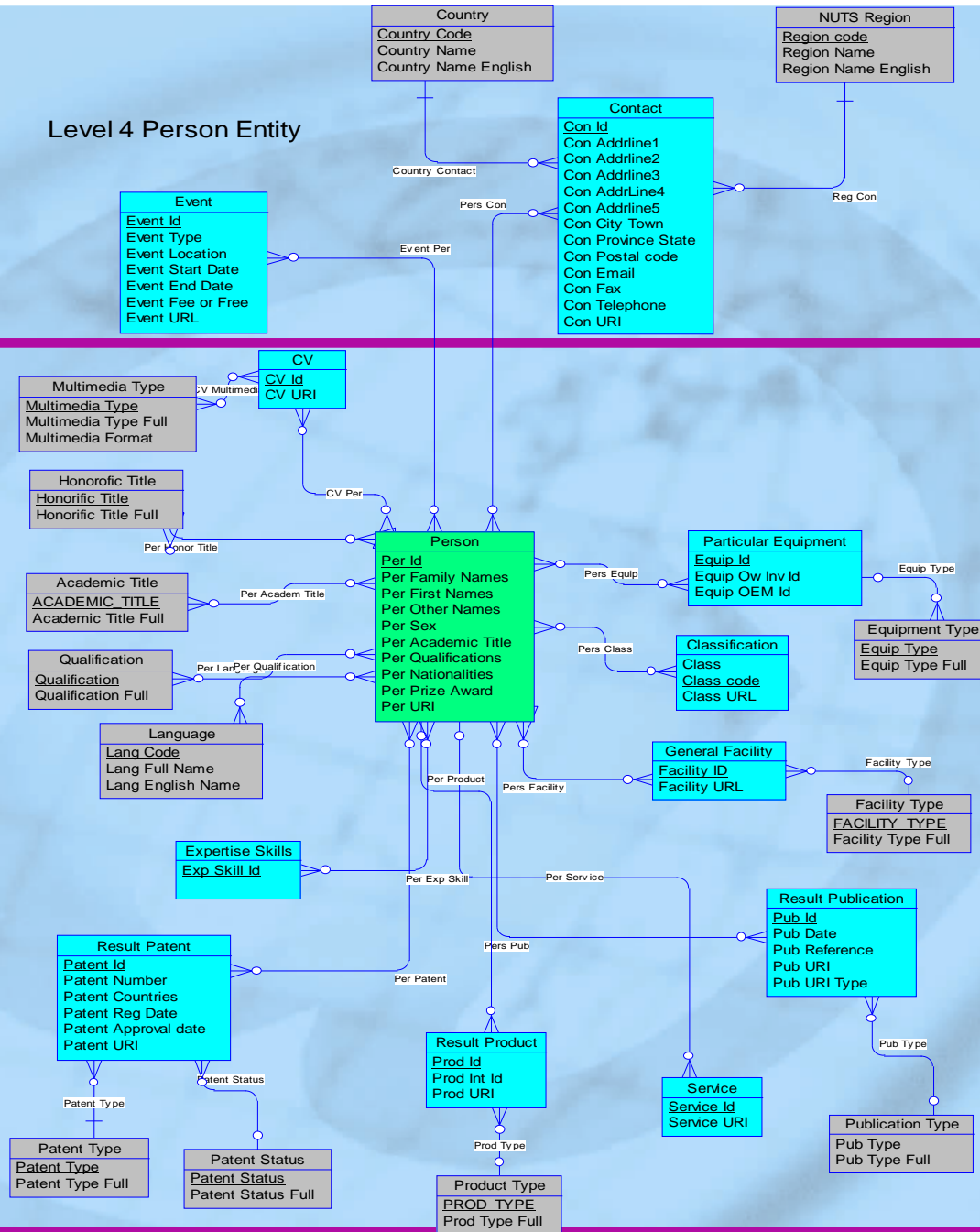
## Language Field Base

### Example: Research Interest

Person-Research interest	Person Id	char(32)	m, pk(part)	
	Language	char(2)	m, pk(part)	
	Translation	char(1)	m, pk(part)	o(rig), h(uman), m(achine)
	Keywords	char(10 24)		

- Full model – primary base entities
- Person – secondary base entities
- Person– Language-base entities
- Person– Lookup tables
- Person– linking relations: at one level, recursion, at different levels
- Exchange Model and Metadata Model

Level 4 Person Entity



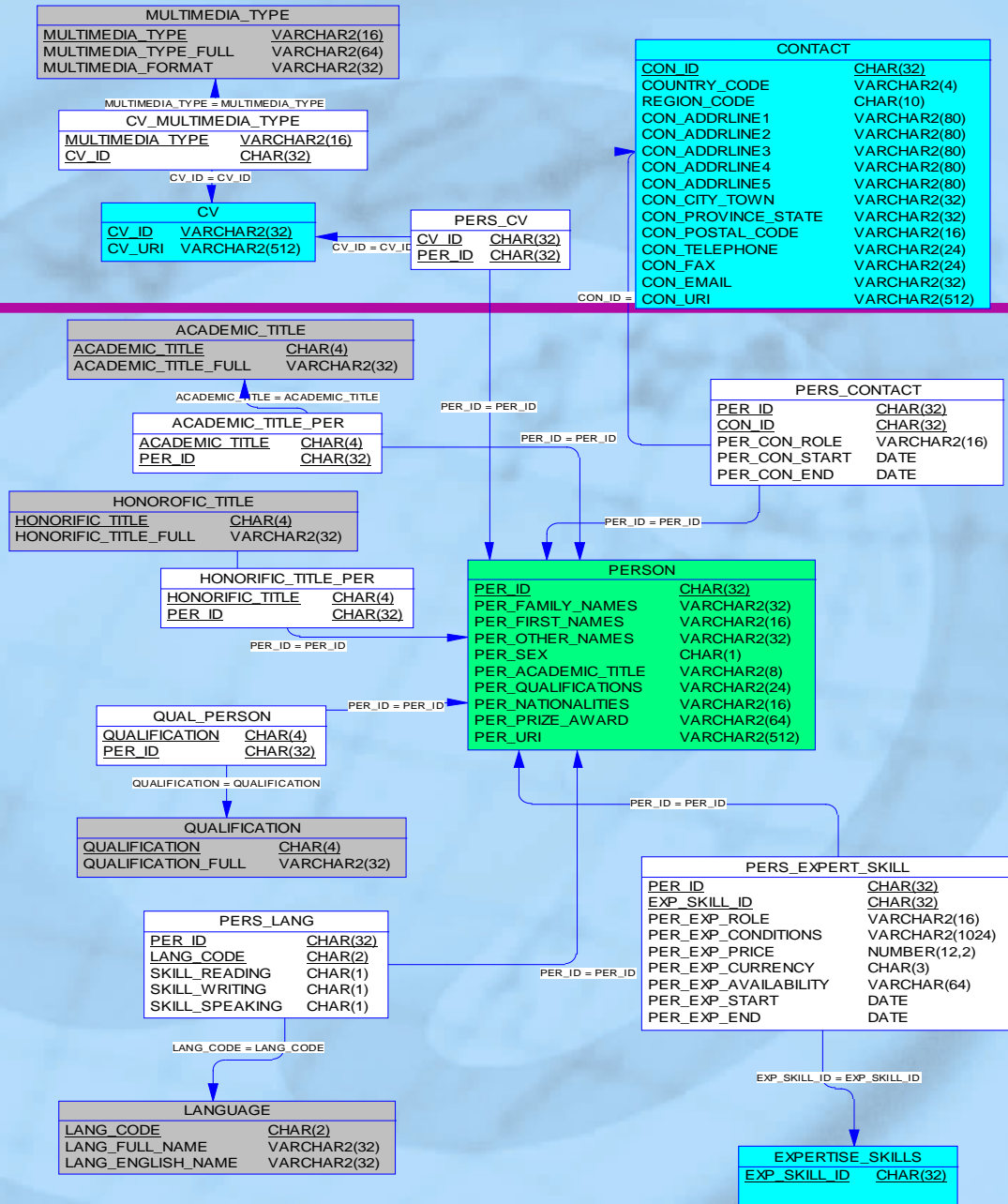
# Person– Lookup-base entities

## Lookup Table

### Examples: Person-related

Person-honorific title	Honorific Title	char (4)	m, pk(part)	Sir, Lady...
Person-honorific title	Title full	char(32)	m, pk(part)	
Person-academic title	Academic Title	char (4)	m, pk(part)	Dr, Prof...
Person-academic title	Title full	char(32)	m, pk(part)	
Person-qualification	Qualification	char(4)	m, pk(part)	
Person-qualification	Qualification full	char(32)	m, pk(part)	ISCED (UNESCO International Classification of Education, level of education)

### Level 5 Person Characteristics

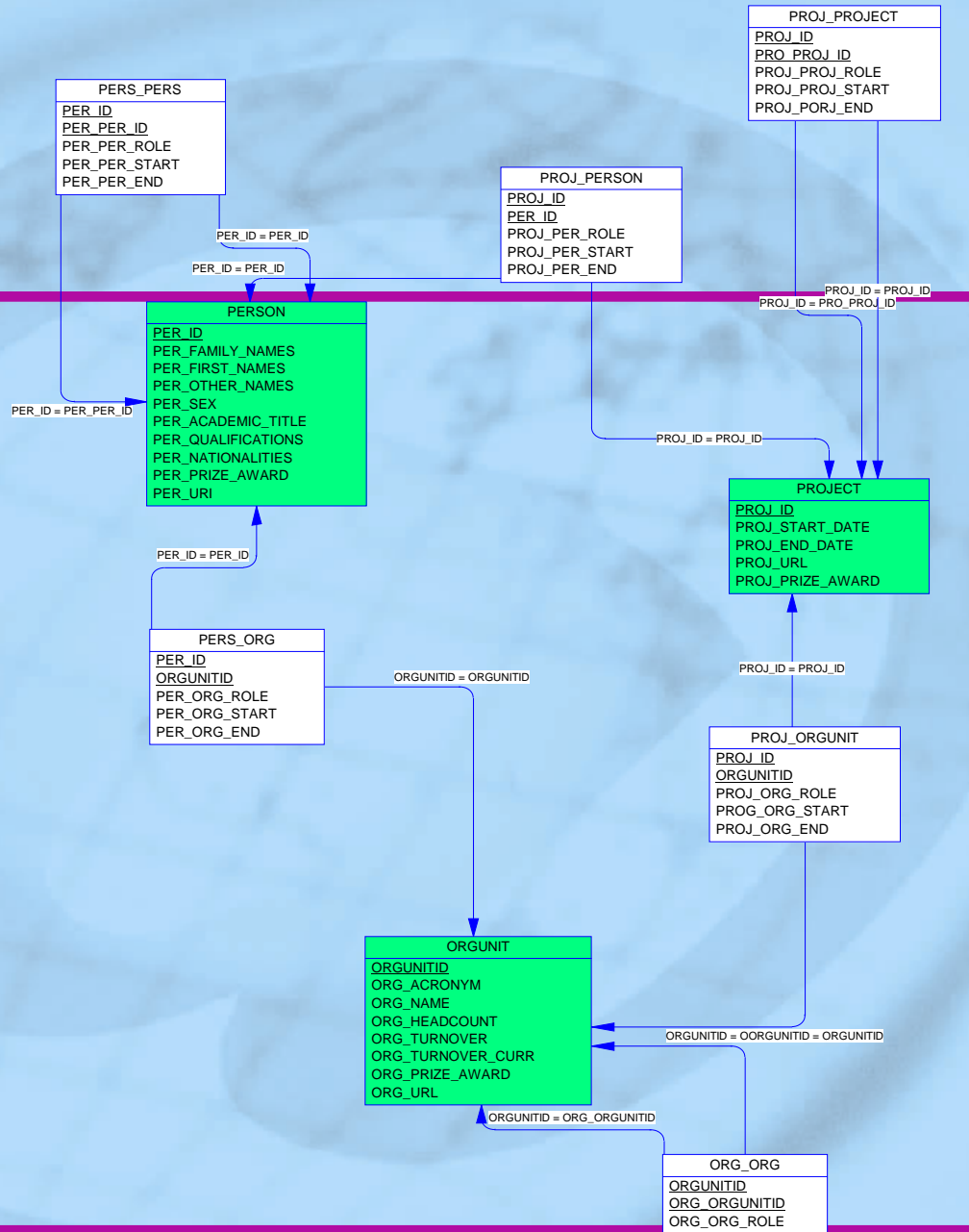


## △ Links to Lookup Tables Example: Person-Related

---

- Note the diagram is CERIF2000
- A mistake was later noticed – Person contains attributes academic title and qualifications
- These are separated out and linked to Person
- Therefore the attributes within Person are redundant

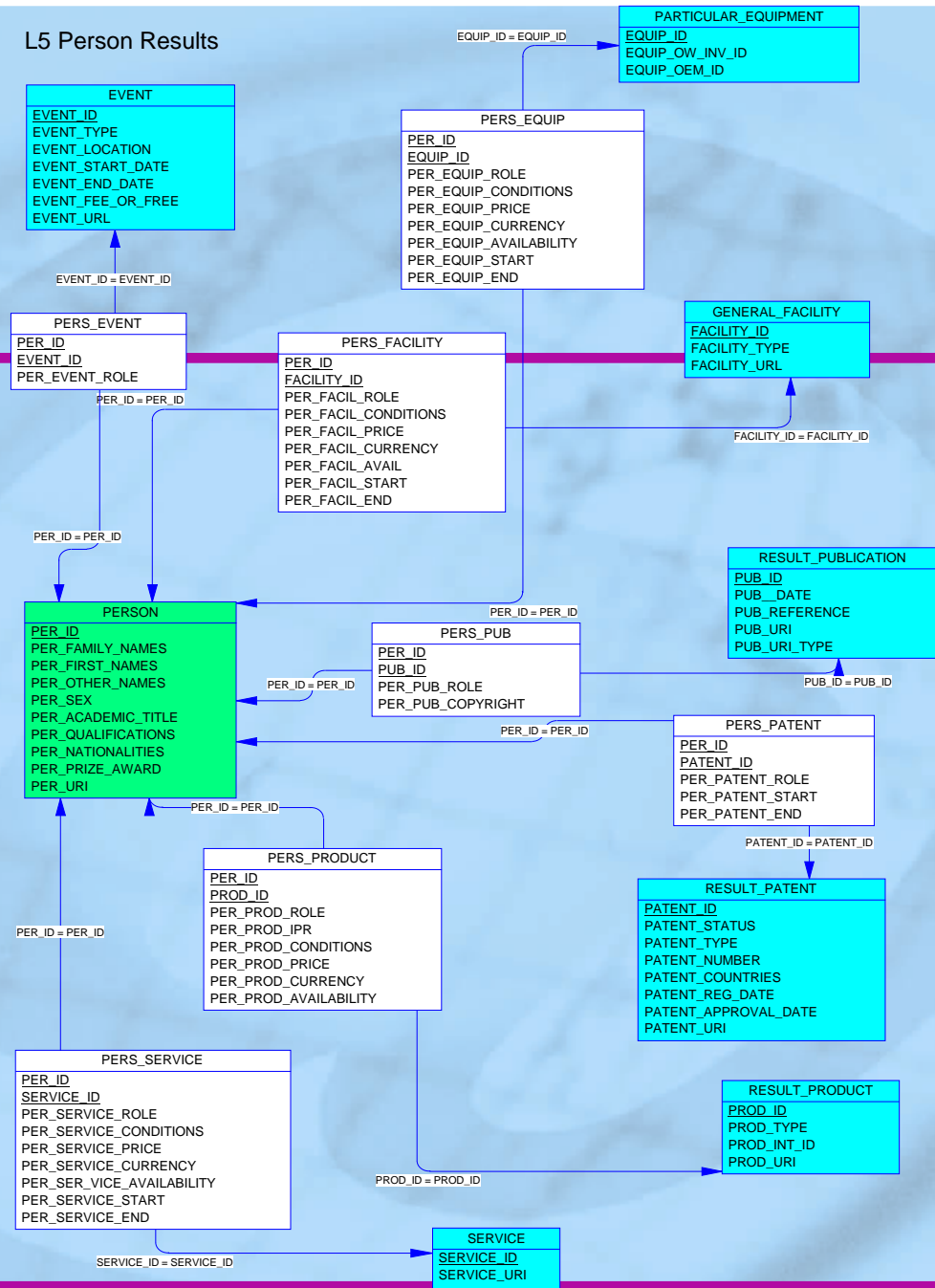
- Full model – primary base entities
- Person – secondary base entities
- Person– Language-base entities
- Person– Lookup tables
- Person– linking relations: at one level, recursion, at different levels
- Exchange Model and Metadata Model



# Linking Relations Person-Person

Person-Person	Person Id	char(32)	m, fk, pk(part)	
Person-Person	Person Id	char(32)	m, fk, pk(part)	
Person-Person	Role	char(16)	o, enumlist	e.g. co-worker, manager
Person-Person	Start Date	date	o	
Person-Person	End Date	date	o	

L5 Person Results



# Secondary Base Links

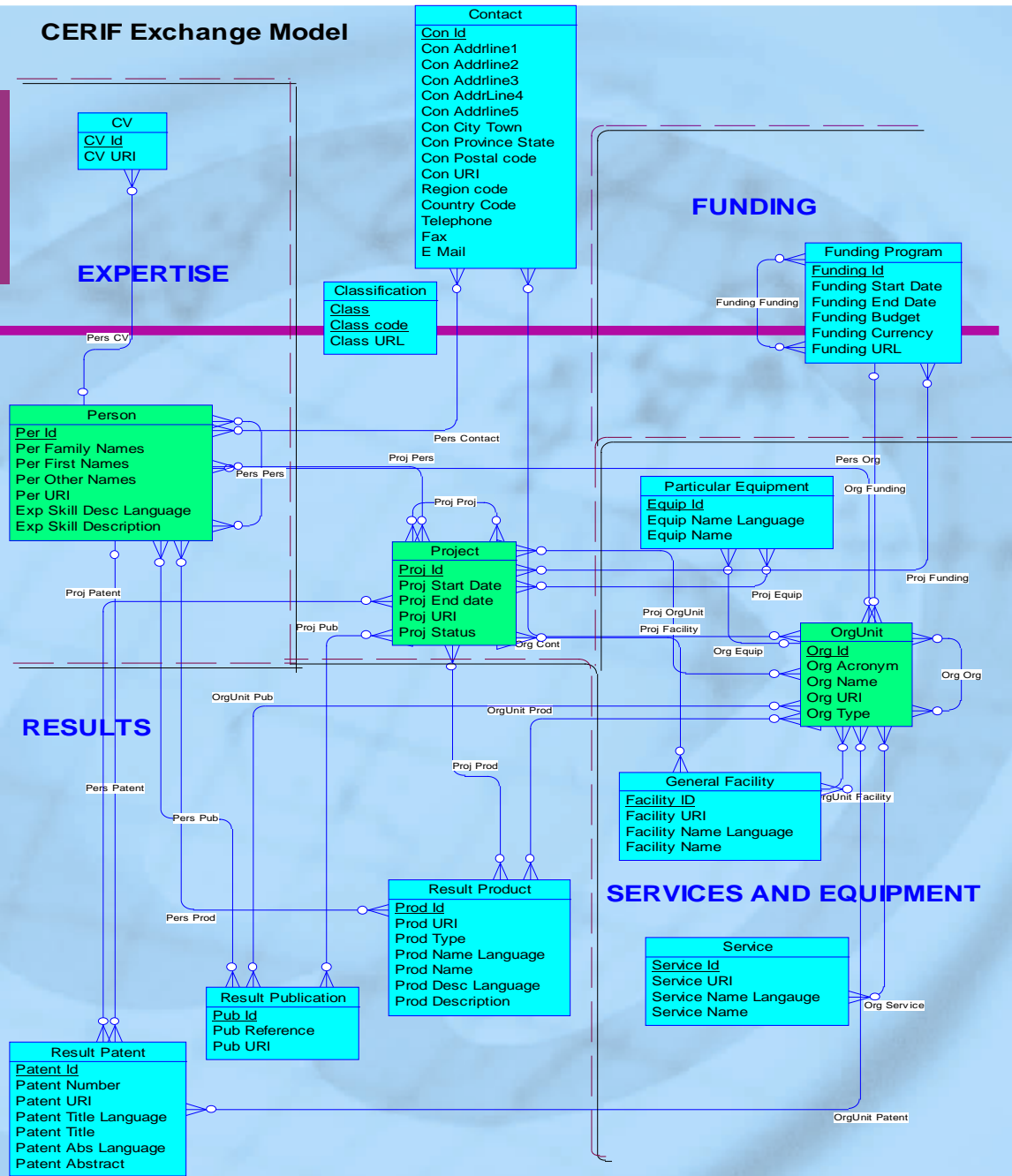
## Example: Person-Result\_Publication

Person-Result_Publication	Person Id	char(32)	m, fk, pk(part)	
Person-Result_Publication	Result_Publication Id	char(32)	m, fk, pk(part)	
Person-Result_Publication	Role	char(16)	o, enumlist	UNIMARC Manual – Bibliographic Format 1994 (IFLA-UBCIM)
Person-Result_Publication	Copyright	char(32)	m	

- Full model – primary base entities
- Person – secondary base entities
- Person– Language-base entities
- Person– Lookup tables
- Person– linking relations: at one level, recursion, at different levels
- Exchange Model and Metadata Model

- A set of subsets (each for a different purpose) of the full CERIF model
- To provide a template for exchanges between CRISs
- Each CRIS converts to / from the CERIF Exchange Model agreed between the CRISs for the exchange

**CERIF Exchange Model**

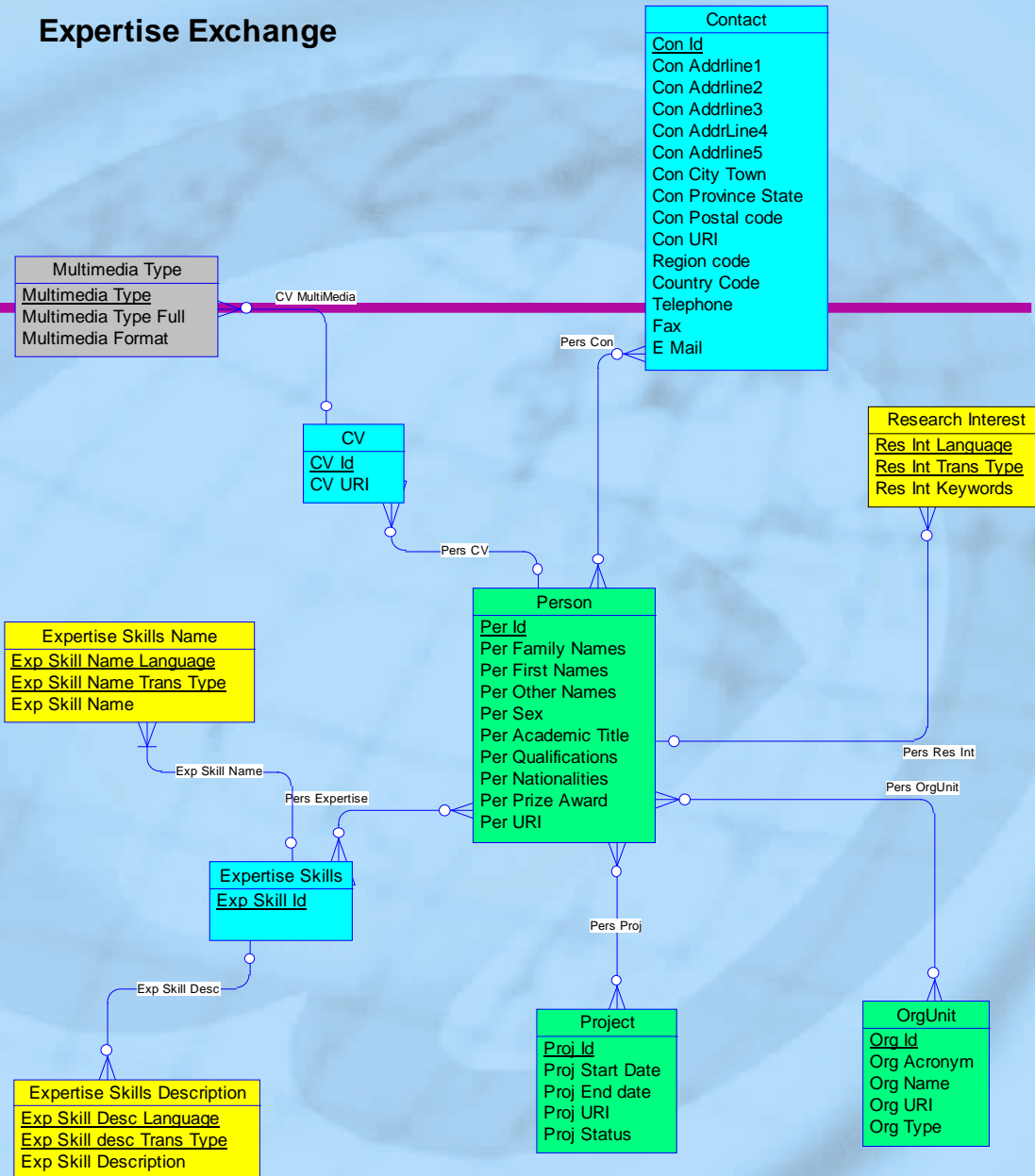


## Exchange Model

---

- Project and Contact Entities always present
  - Can specialise thereafter as
    - Expertise
    - Results
    - Funding
    - Services and Equipment
  - Depending on kind of CRIS and need  
for exchange
-

## Expertise Exchange



## Metadata Model

---

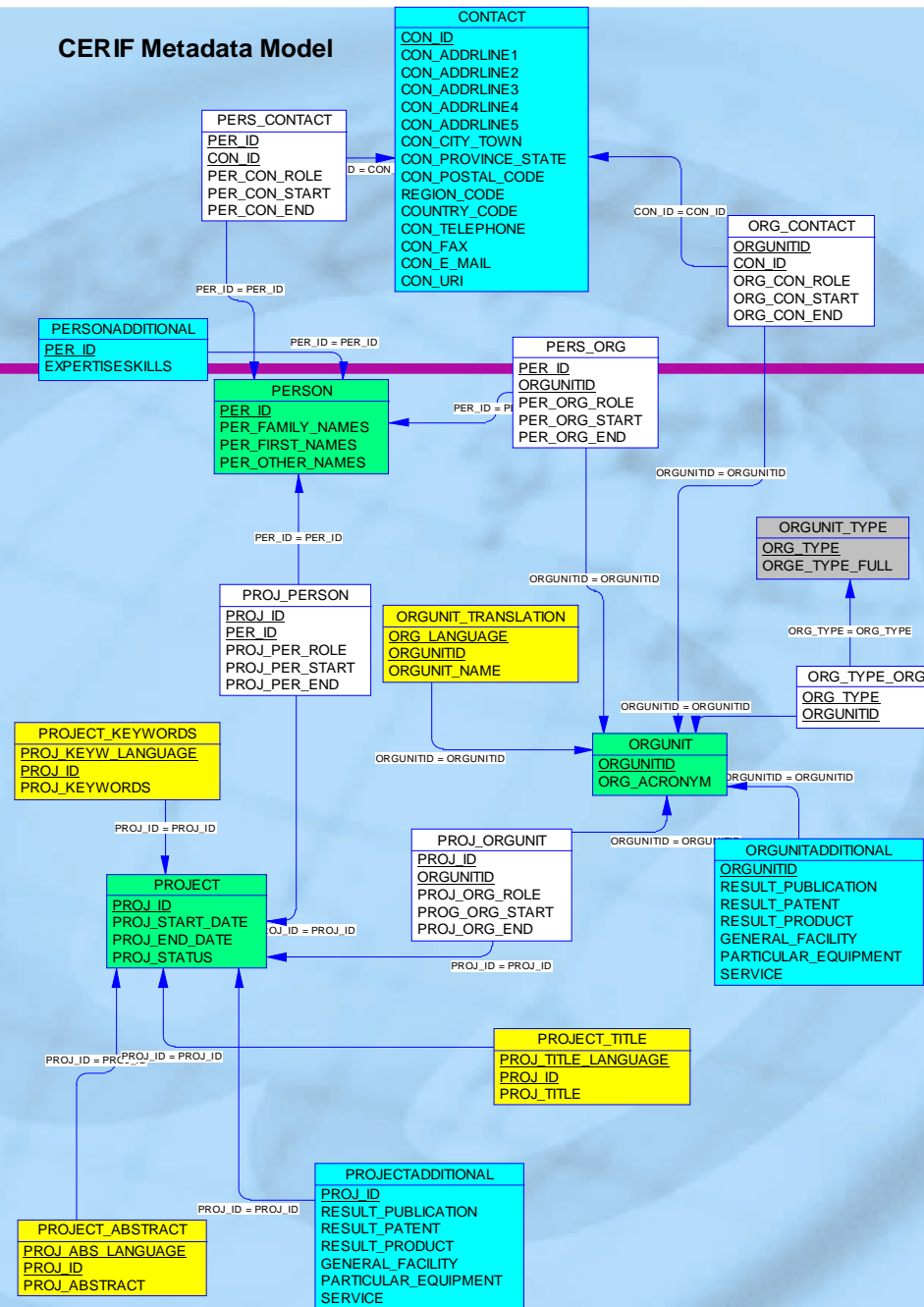
- A model which should be a subset of any CRIS
- Used to provide an entry representing an instance within the CRIS to a catalog system
- Used to provide to harvesting engines (WWW) a description of an instance within the CRIS

## Metadata Model

---

- A trade-off between
- Enough information for the purposes
  - Data exchange
  - Catalog providing homogeneous access over heterogeneous information
- Minimal information
  - Commercial exploitation of the full CRIS represented by the metadata
  - Data storage and communications costs

CERIF Metadata Model



## Metadata Model

---

- Fewer attributes in base entities
  - Person, Project, OrgUnit
- Full attribute set in Contact
- Minimum use of language base tables
- Minimum use of lookup tables
- Use of 'X' Additional (where 'X' is base table) to indicate which entities are present in full CRIS

- CERIF is a data model which
  - represents accurately the world of interest
    - Entities and their attributes
    - With language variants and lookup tables
    - And the structural relationships through linking relations
- CERIF can be used
  - as a full data model to implement a CRIS
  - for data exchange to exchange data with other CRISs
  - as a metadata model providing access to a CRIS