

CERIF 2008 - 1.1 Full Data Model (FDM)

Introduction and Specification

Editors:

Brigitte Jörg	DFKI GmbH, Berlin, Germany
Keith Jeffery	Science and Technology Facilities Council, Didcot, UK
Geert van Grootel	Flemish Government, Brussels, Belgium
Anne Asserson	University of Bergen, Bergen, Norway

Abstract:

CERIF (the **C**ommon **E**uropean **R**esearch **I**nformation **F**ormat) is a formal model to setup Research Information Systems and to enable their interoperation. Research Information is information about research entities such as People, Projects, Organisations, Publications, Patents, Products, Funding, or Equipment and the relationships between them. Information Systems are means to structure, store, maintain, access, retrieve, extract, exchange, disseminate or assess the information they contain. We consider CERIF, its entities and their rich relationships as well as the management of their semantics a powerful instrument for setting up scalable and quality research information systems.

This latest CERIF upgrade 2008 – 1.1 adds funding related entities like call or grant to the model, it incorporates new attributes (a fraction attribute allows for %-relationships at link entities, a birthdate attribute enriches the person entity), and extends the ID attribute to char(128). This document captures the whole range and structure of the CERIF 2008–1.1 model and release.

The CERIF model is considered a standard; recommended by the European Union to its Member States. It has been developed with support by the European Commission in two major phases: 1987-1990 and 1997-1999. In 2000 the European Commission handed over care and custody of CERIF to euroCRIS (<http://www.eurocris.org>) a not-for-profit organisation dedicated to the promotion of CRISs (**C**urrent **R**esearch **I**nformation **S**ystems).

Status:

CERIF model improvements are based on discussions among euroCRIS CERIF task group members. This document will be updated alongside major model updates.

Location:

http://www.eurocris.org/fileadmin/cerif-2008/CERIF2008_1.1_FDM.pdf

Table of Contents

1.	Introduction and Concise History	3
1.1	Purpose of this Document	3
1.2	CERIF Components	4
1.3	CERIF Upgrades	5
2.	The CERIF 2008 – 1.1 Model	6
2.1	CERIF Conceptual Structure	6
2.2	CERIF Core Entities	7
2.2.1	CERIF Entity Project	8
2.2.2	CERIF Entity Person	10
2.2.3	CERIF Entity OrganisationUnit	12
2.3	CERIF Result Entities	14
2.3.1	CERIF Entity ResultPublication	15
2.3.2	CERIF Entity ResultPatent	19
2.3.3	CERIF Entity ResultProduct	20
2.4	CERIF 2 nd Level Entities	21
2.5	CERIF Link Entities	23
2.6	CERIF Multiple Language Features	25
2.7	CERIF Semantic Layer [Semantic Features]	26
2.8	Additional Features	27
3.	CERIF-based SQL scripts	28
4.	CERIF XML	29
5.	CERIF Semantics	31
6.	CERIF Extensions	32
7.	Next Steps	33
8.	Appendix	34
8.1	List of CERIF Entities	34
8.1.1	CERIF Core Entities (Logical (PhysicalName))	34
8.1.2	CERIF Result Entities (Logical (PhysicalName))	34
8.1.3	CERIF 2 nd Level Entities (Logical (PhysicalName))	34
8.1.4	CERIF Link Entities (Logical (PhysicalName))	34
8.1.5	CERIF Multiple Language Features (Logical (PhysicalName))	36
8.1.6	Additional Entities (Logical (PhysicalName))	37
8.1.7	CERIF Classification Entities (Logical (PhysicalName))	37
8.1.8	CERIF Attributes	38
8.1.9	Attribute in all Link Tables	38
8.2	Logical / Physical CERIF Entity Names	39
9.	References	43

1. Introduction and Concise History

Most nation-states have publicly-supported research programmes. It is realised that public sponsorship of research and development leads to wealth creation and improvement in the quality of life. Because public funding is involved, it is necessary for there to be appropriate governance, and for the related information to be available to the public. Broadly, each nation state has a similar research process of: strategic planning; programme announcement; call for proposals; proposal evaluation and awarding; project result monitoring; project result exploitation. However, research is international. A research project in country A is likely to be based on previous research in several other countries. Many research projects are now transnational: well-known examples include the human genome and climate change, but there are many others, especially where expensive infrastructure is utilised such as particle physics or space science. Furthermore, knowledge of the research activity in country A may influence the strategy towards research – including priorities and resources provided – in country B. Thus, there is a need to share research information across countries, or even between different funding agencies in the same country. Research Information is used by researchers (to find partners, to track competitors, to form collaborations); research managers (to assess performance and research outputs and to find reviewers for research proposals); research strategists (to decide on priorities and resourcing compared with other countries); publication editors (to find reviewers and potential authors); intermediaries/brokers (to find research products and ideas that can be carried forward with knowledge/technology transfer to wealth creation); the media (to communicate the results of R&D in a socio-economic context) and the general public (for interest). Most European countries collect and store their research information in digital repositories; these may be national, regional, institutional, functional, or thematic in their range, where each system builds upon a particular format or structure to serve for special requests. Research Information is relevant for actors in scientific environments as well as for decision makers to support related organization, management and planning. We consider Research Information as the transmitter between Science and Society and as such as a powerful instrument for governance. Having such an impact, Research Information has to be collected carefully and preserved systematically, in order to most effectively support society and the individuals within [1, 2, 4, 5, 7].

CRIS and CERIF approaches to enable advances into this direction are not new. The first release of CERIF has been published in 1991 with the aim of facilitating data exchange of records on research projects between European Member States, and to serve as a format to allow for the networking of databases. The European Working Group on Research Databases has recommended the CERIF format as a result of a workshop held in 1987. The CERIF 1991 data model which described project records only has been applied in the ERGO project¹ and the needs for an extension were recognised. In 1997 revision work was entrusted to unit D2 DG XIII of the European Commission. The revisions in the model were based on reflections of user requirements and led to a recommendation for CERIF 2000² to Member States and a handover of CERIF to euroCRIS³. The CERIF 2000 release has added person and organisation as entities and many other entities relevant in the research context, such as publication, service, equipment, patent, country, language, event, etc., and classification. Additionally, the entities had types and the relationships assigned roles to capture their semantics. In the CERIF 2006 release these roles and types at entities have been re-organised within the so called Semantic Layer to supply the needed flexibility for capturing different application semantics and views. Alongside the 2006 model, the *CERIF XML* interchange format has been introduced [9, 11]. This CERIF 2008 release extends its predecessors with substantial elaboration on the publication entity, and thus establishes the long requested connectivity to repositories for scholarly publications. Additionally, CERIF 2008 – 1.0 introduces the *CERIF Semantics* [12] for publication entities as a first step towards a formal vocabulary to manage the semantics inherent in the relationships between research entities.

This document will walk through the CERIF model by following a conceptual structure. The physical presentations of database levels and some real life examples will support the understanding of the model in a more applied context.

1.1 Purpose of this Document

This document provides a detailed description of the CERIF model 2008 – 1.1 and demonstrates potential use cases and application scenarios.

¹ ERGO project: <http://cordis.europa.eu/ergo/>

² EC Recommendation: <http://cordis.europa.eu/cerif/>

³ euroCRIS: <http://www.eurocris.org/>

1.2 CERIF Components

The current CERIF 2008 – 1.1 release comprises the following components:

- CERIF 2008 – 1.1 FDM: Model Introduction and Specification
this document
- CERIF 2008 – 1.1 FDM: SQL scripts for most common databases
available from the euroCRIS website for members only
- CERIF 2008 – 1.1 XML: Data Exchange Format Specification
separate document available from the euroCRIS website [11]
- CERIF 2008 – 1.1 XML Examples
available from the euroCRIS website for members only
- CERIF 2008 – 1.1 XML Schema Files
CERIF XML validation files available from the euroCRIS website
- CERIF 2008 – 1.1 Semantics
separate document available from the euroCRIS website [12]

CERIF 2008 – 1.1 related files and more documents and background information about CERIF and CRISs can be downloaded from the public euroCRIS website: <http://www.eurocris.org/cerif/cerif-releases/cerif-2008/>. The physical SQL scripts and XML examples files are available for members only^{4*}.

* The CERIF 2008 – 1.1 release was modeled with Toad Data Modeler⁴ by Quest Software⁴ which allows to draw ERM diagrams, to generate SQL scripts for most common databases (Oracle, Microsoft, IBM, etc.), to reverse engineer from databases, to create screenshots of the model and model parts, and to model at physical and logical level. The resulting CERIF SQL scripts are generated automatically from the physical level.

1.3 CERIF Upgrades

Compared to its preceding version CERIF 2008 1.0, the current upgrade to CERIF 2008 – 1.1 incorporates the following features:

- **Addition of new Entities in the Context of Funding (needs further discussion)**

- cfCall
- cfCallName
- cfCallDescription
- cfCallKeywords
- cfCall_FundingProgramme
- cfGrant
- cfGrantName
- cfGrantDescription
- cfGrantKeywords
- cfGrant_FundingProgramme

- **Addition of new Entities for Model Consistency**

- cfService_FundingProgramme
- cfResultPatent_ResultPatent
- cfResultProduct_ResultProduct

- **Addition of Attributes or Changes in Attributes**

- all identifiers in CERIF now have Char(128) instead of Char (32)
- cfBirthdate (added)
- cfProductInternId (now non-mandatory)
- cfFraction to all Link Entities (added)
- default values removed from
 - cfTranslation (o,m,h)
 - cfSex (u,m,f)
 - cfSkillReading (m,f,w)
 - cfSkillSpeaking (m,f,w)
 - cfSkillWriting (m,f,w)

- **Addition to CERIF Semantics**

- “is part of” added to cfResultPublication-ResultPublication Roles

2. The CERIF 2008 – 1.1 Model

To reduce the complexity of the model towards a better understanding, this introduction and specification document follows a conceptual structure. The conceptual structure allows for different perspectives and views when talking about parts of the model and enables to emphasize particular model features. This conceptual structure is only a virtual structure and as such not inherent in the physical data model, and therefore, also not incorporated in the physical SQL scripts. It is used for organizing this document and considered an instrument to support understanding of the CERIF model.

2.1 CERIF Conceptual Structure

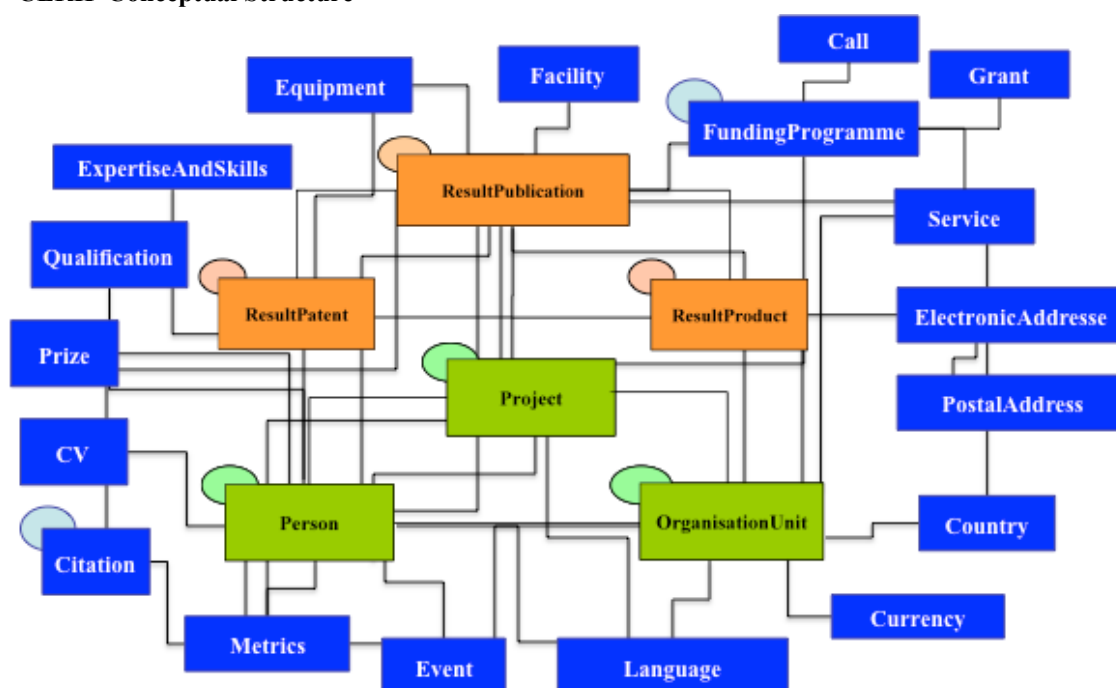


Figure 1: CERIF 2008 – 1.1 entities and some relationships

We conceptually structure the CERIF model into entity types and features. In between the types we distinguish core, result, link and 2nd level entities, as features we consider multilinguality and semantics. This conceptual structure is also supported by colors.

CERIF Entity Types		CERIF Features	
	Core Entities [core]		Multiple Language [lang]
	Result Entities [result]		Semantics [class]
	2 nd Level Entities [2nd]		Additional [add]
	Link Entities [link]		

The conceptual model parts will subsequently be presented in abstract views. For the rather technical details at logical or physical/database level (attributes, datatypes, keys) the corresponding screenshots from Toad Modeler will be incorporated. Whereas the entity names in abstract views are presented in full length, the table names in the screenshots are abbreviated and include a prefix ‘cf’ for CERIF. Because in some databases the length of a table name is restricted to a particular number of characters, we have shortened the table names at physical level to ensure the consistency of SQL scripts by avoiding uncontrolled truncations. The CERIF XML element names map with the physical (short) names of the model. The CERIF XML specification applies the same conceptual structure for a recommended ordering and clustering of the XML files in the XML file names [11].

A complete list of the CERIF entities is attached in the Appendix including their conceptual type or feature; a HTML reference to the model, including the conceptual images, is available from the public euroCRIS website for navigation: <http://www.eurocris.org/>.

2.2 CERIF Core Entities

The core CERIF entities are Person, OrganisationUnit and Project. Figure 2 shows the core entities, as well as their recursive and linking relationships. Each core entity recursively links to itself and maintains relationships with other core entities. The core entities allow for a representation of scientific actors and their different kinds of interactions.

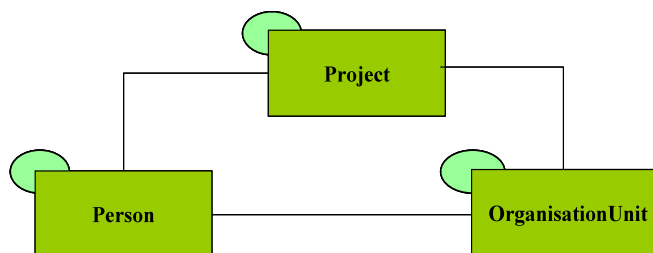


Figure 2: CERIF Core Entities

Figure 3 below shows the core entities (cfProj, cfPers, cfOrgUnit) and some related entities from a physical perspective. The little circles in figure 2 represent their recursiveness; that is, their relationships in between each other; within projects, within persons, and within organisations (cfProj_Proj, cfPers_Pers, cfOrgUnit_OrgUnit). The recursive as well as the interlinking relations presented as cfPers_OrgUnit, cfProj_Pers, and cfProj_OrgUnit in figure 3 are link type entities and will be introduced in section 2.5. The yellow colored entities cfProjTitle, cfProjAbstr, cfProjKeyw, cfOrgUnitName, cfOrgUnitResAct, and cfOrgUnitKeyw support multiple languages features and will be explained in section 2.6.

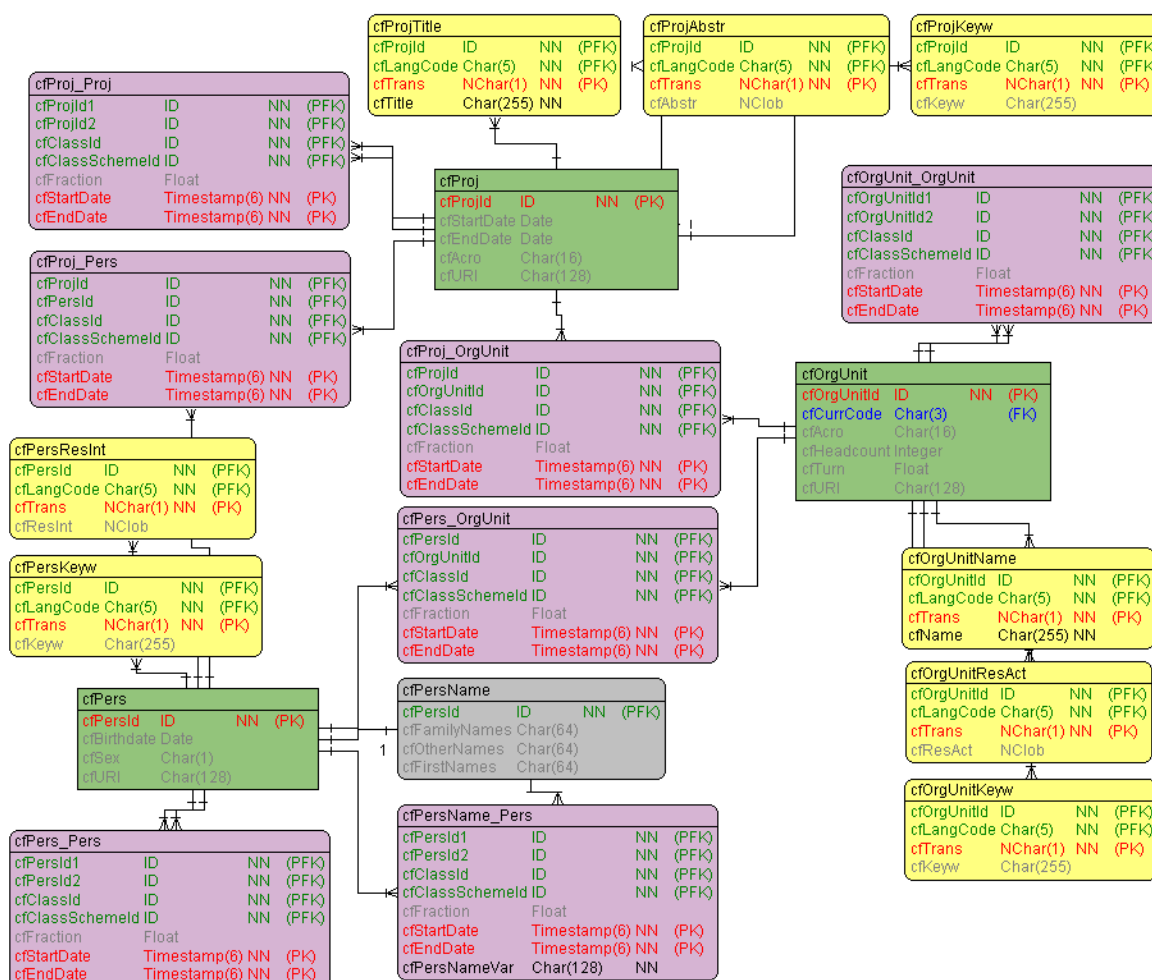


Figure 3: CERIF Core Entities, their Recursion and some Link Entities

Each core entity cfProj, cfPers, cfOrgUnit will subsequently be presented and some examples will be provided to support understanding.

2.2.1 CERIF Entity Project

For an identification of project records, the core entity (cfProj) foresees an id attribute (cfProjId). Besides, the attributes acronym, uri, and start/end date (cfAcro, cfURI, cfStartDate, cfEndDate) are considered as common project attributes. The project entity maintains many relationships with other entities: project, person, organisation, publication, patent, product, funding programme, equipment, facility, service, event, prize and classification (cfProj_Proj, cfProj_Pers, cfProj_OrgUnit, cfProj_ResPubl, cfProj_ResPat, cfProj_ResProd, cfProj_FundProg, cfProj_Equip, cfProj_Facil, cfProj_Srv, cfProj_Prize, cfProj_Class) as shown in figure 4. Each relationship or link entity carries semantics with a time-stamped reference to the CERIF Semantic Layer by cfClassId and cfClassSchemId and a cfFraction attribute to assign fractional values to a classification reference. Additionally, the project entity supports multilingual features for title, abstract, and keywords (cfProjTitle, cfAbstr, cfProjKeyw).

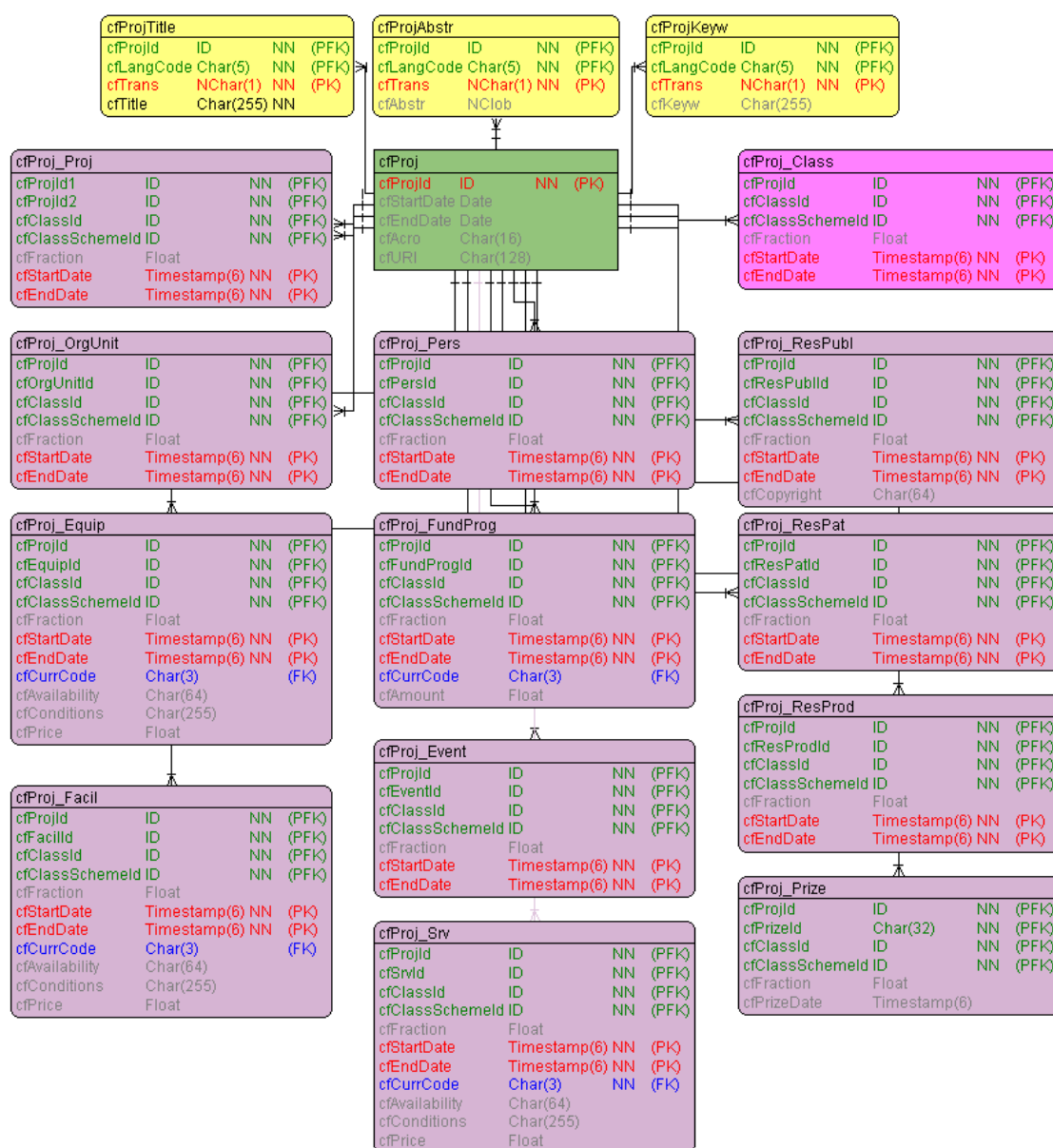


Figure 4: CERIF Core Entity Project

Table 1 shows an example project record from a database perspective where common [core] and multilingual [lang] attributes are stored in the upper rows, and the lower rows show example relationships [link] including their semantics. Links are established by ids (i.e. cfClassId, cfResPublId, cfOrgUnitId, cfFundProgId) as indicated in the Attribute column, the carrying link entities are named in the Table column, the Type column indicates the conceptual type (core, link, lang), the semantic values (i.e. isOriginatorOf, isCoordinatedBy, isFundedBy) and some corresponding fractions are indicated in the Classification column, where each value belongs to a defined scheme (i.e. FP6-IST, PROJ-PUBL, etc).

Table 1: CERIF Project Example Record

CERIF Project example database entry	Attribute	Table	Type	Semantic Layer (CERIF Semantics)	
				Classification (ClassIds)	Classification Scheme
project-ist-world	cfProjId	cfProj	core		
IST World	cfAcro	cfProj	core		
http://www.ist-world.org/	cfURI	cfProj	core		
2005-04-01	cfStartDate	cfProj	core		
2007-11-30	cfEndDate	cfProj	core		
Knowledge Base for RTD Competencies in IST	cfTitle	cfProjTitle	lang[en]		
Wissensbasis für RTD Kompetenzen im Bereich IST	cfTitle	cfProjTitle	lang[de]		
IST, Research Information, NMS, Portal, Information System	cfKeyw	cfProjKeyw	lang[en]		
The objective of the project is to set up and populate an information portal with innovative functionalities that helps to promote RTD competencies in IST in the New Member States (NMS) and Associate Candidate Countries (ACC) in order to facilitate and foster the involvement of different research entities in joint RTD activities. The IST World portal is built on the CERIF standard ...	cfAbstr	cfProjAbstr	lang[en]		
class-2004-ist-3	cfClassId	cfProj_Class	link	2004-IST-3	FP6-IST
publ-analyzing-european-rt-d-in-ist	cfResPublId	cfProj_ResPubl	link	Originator	PROJ-PUBL
publ-cris-is-for-research-activity	cfResPublId	cfProj_ResPubl	link	Originator	PROJ-PUBL
publ-analytic-services-for-the-era	cfResPublId	cfProj_ResPubl	link	Originator	PROJ-PUBL
fundprog-fp6	cfFundProgId	cfProj_FundProg	link	Funder	PROJ-FPROG
orgunit-dfki	cfOrgUnitId	cfProj_OrgUnit	link	Coordinator	PROJ-ORG
orgunit-dfki	cfOrgUnitId	cfProj_OrgUnit	link	2006-[fraction=0.5]	Proj-Budget-Allocation
orgunit-dfki	cfOrgUnitId	cfProj_OrgUnit	link	2007-[fraction=0.2]	Proj-Budget-Allocation

The example record shows some common and multilingual project attributes: id, acronym, uri, start and end date, title, abstract and keywords; the lower rows present some relationship examples. With cfClassId=2004-IST-3, the example record is classified according to a FP6-IST scheme. CERIF entities store their semantics by reference ids with interlinking [link] entities. The given example record is linked with some publications in the role of an originator. In the same way, it is linked with the FP6 funding programme in the role of the funder, and with an organisation in the role of a coordinator. Moreover, it demonstrates a possible budget allocation over a two years period. The example record only gives some relationships; the entire model allows for many more. The linkage mechanism by link entities is consistent across the model and will be explained in detail within section 2.5; for the semantic features we refer to section 2.7.

2.2.2 CERIF Entity Person

For an identification of person records the core entity (cfPers) offers an id attribute (cfPersId). Besides, the attributes birthdate, sex and uri (cfBirthdate, cfSex, cfURI) are considered as common person attributes. CERIF allows for the maintenance of multiple person names or name variants with cfPersName and cfPersName_Pers.

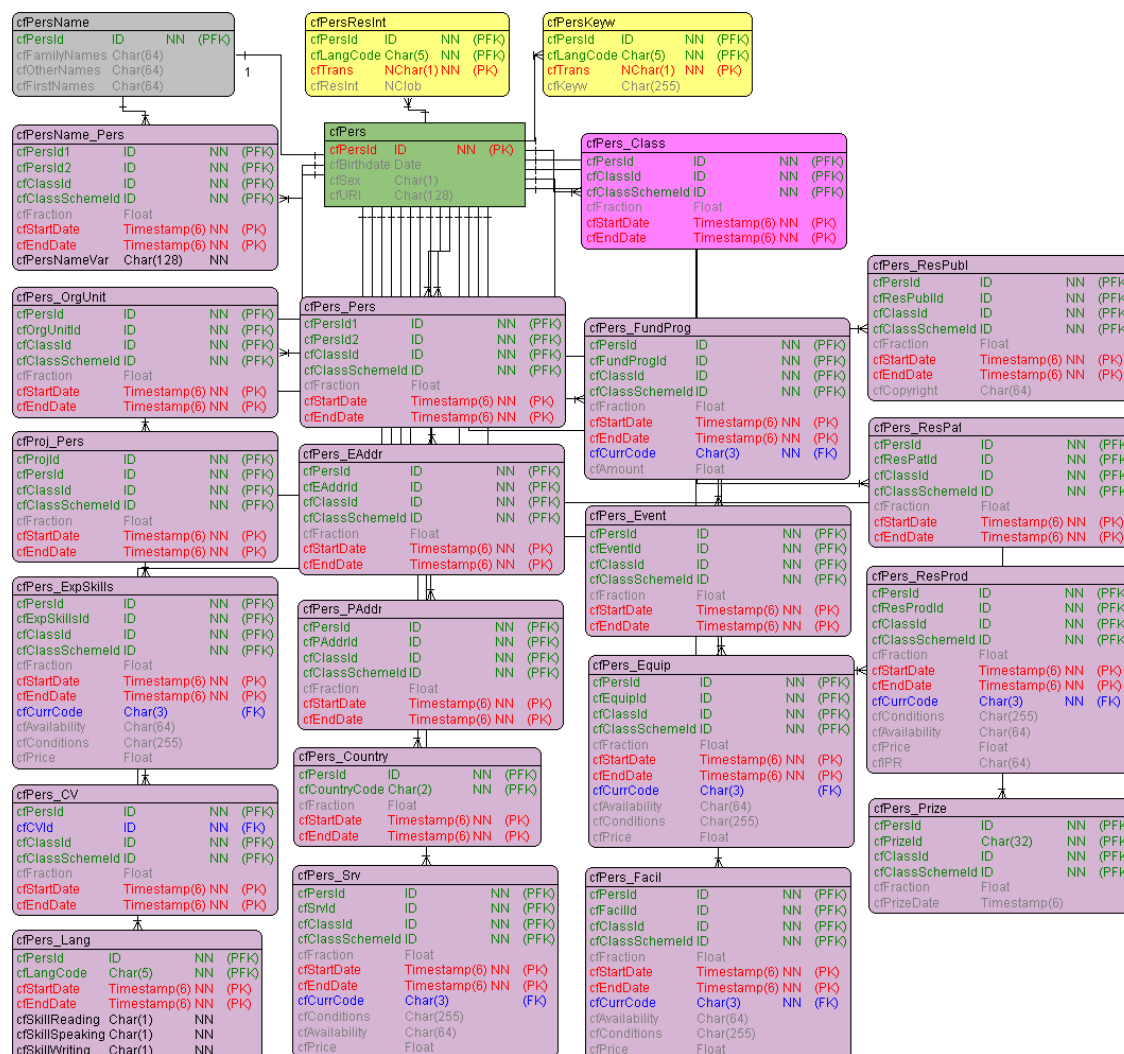


Figure 5: CERIF Core Entity Person

The entity person maintains many relationships: person, project, organisation, publication, patent, product, funding programme, equipment, facility, service, event, prize, electronic address, physical address, expertise and skills, cv, language, country and classification (cfPers_Pers, cfPers_Proj, cfPers_OrgUnit, cfPers_ResPubl, cfPers_ResPat, cfPers_ResProd, cfPers_FundProg, cfPers_Equip, cfPers_Facil, cfPers_Srv, cfPers_Event, cfPers_Prize, cfPers_EAddr, cfPers_PAddr, cfPers_ExpSkills, cfPers_CV, cfPers_Lang, cfPers_Country, cfPers_Class) shown in figure 5. Each relationship or link entity carries semantics with a time-stamped reference to the CERIF Semantic Layer by cfClassId and cfClassSchemId and a cfFraction attribute to assign fractional values to a classification reference. Additionally, the person entity supports multilingual features for research interest descriptions and keywords (cfPersResInt, cfPersKeyw). Table 2 shows one example person record from a database perspective. The common and the multilingual attributes are stored in the upper rows; the lower rows show example relationships including their semantics. Relationships are established by ids (i.e. cfPersId2, cfResPublId, cfOrgUnitId, cfProjId) indicated in the Attribute column, the carrying link entites are named in the Table column, the Type column indicates the conceptual type (core, link, lang), semantic values (spelling Variant, M.A. isAuthorOf, isAffiliatedWith, isMemberOfBoard, isTG-LeaderOf, isCoordinatedBy, hasParticipant) and fractions are indicated in the Classification column, where each value belongs to a particular scheme (PERS_PERSNAME, ACADEMIC-TITLES, PERS_PUBL, etc).

Table 2: CERIF Person Example Record

CERIF Person example database entry				Semantic Layer (CERIF Semantics)	
Data	Attribute	Table	Type	Classification (ClassIds)	Classification Scheme
person-brigitte-joerg	cfPersId	cfPers	core		
f	cfSex	cfPers	core		
http://www.dfki.de/~brigitte/	cfURI	cfPers	core		
Joerg	cfFamilyNames	cfPersName	add		
Brigitte	cfFirstNames	cfPersName	add		
Brigitte is interested in Research Information and Research Information Systems.	cfResInt	cfPersResInt	lang		
Information Systems, Research Information, Information Science, Science, Knowledge, Bibliometrics	cfKeyw	cfProjKeyw	lang		
person-brigitte-joerg	cfPersId2	cfPersName_Pers	link	spellingVariant	PERS_PERS NAME
classification-MA	cfClassId	cfPers_Class	link	M.A.	ACADEMIC-TITLES
publication-analyzing- european-research- competencies-in-ist	cfResPublId	cfPers_ResPubl	link	Author	PERS-PUBL
publication-analytic- services-for-era	cfResPublId	cfPers_ResPubl	link	Author	PERS-PUBL
organisation-dfki	cfOrgUnitId	cfPers_OrgUnit	link	Affiliation	PERS_ORGUNIT
organisation-It-lab	cfOrgUnitId	cfPers_OrgUnit	link	Subaffiliation	PERS_ORGUNIT
organisation-euroCRIS	cfOrgUnitId	cfPers_OrgUnit	link	Board-Member	PERS_ORGUNIT
organisation-CERIF-TG	cfOrgUnitId	cfPers_OrgUnit	link	TG-Leader-CERIF	PERS_ORGUNIT
project-ist-world	cfProjId	cfProj_Pers	link	Coordinator[fract=0.7]	PROJ_PERS [2007DISTRIBUTION]
project-It-world	cfProjId	cfProj_Pers	link	Participant[fract=0.3]	PROJ_PERS [2007DISTRIBUTION]

The example record shows some common and multilingual person attributes id, sex, family name, first name, research interest and keywords; the lower rows present some relationship examples. A reference cfPersId2= 'person-brigitte-joerg' in the cfPersName_Pers table allows for the storage of person name spelling variants. CERIF entities store their semantics by reference ids with interlinking (link) entities. The example record shows that the person is author of articles, has co-ordinated and participated in projects, and is active with different organisations. Additionally, the fraction examples indicate a distribution of work across two projects in the year 2007. The example record only gives some relationships; the entire model allows for many more. The linkage mechanism by link entities is consistent across the model and will be explained in detail within section 2.5; for the semantic features we refer to section 2.7.

2.2.3 CERIF Entity OrganisationUnit

For an identification of organisation records, the core entity (cfOrgUnit) offers an id attribute (cfOrgUnitId). Besides, the attributes acronym, currency, headcount, turnover and uri (cfCurrCode, cfAcro, cfHead, cfTurn, cfURI) are considered as common organisation attributes.

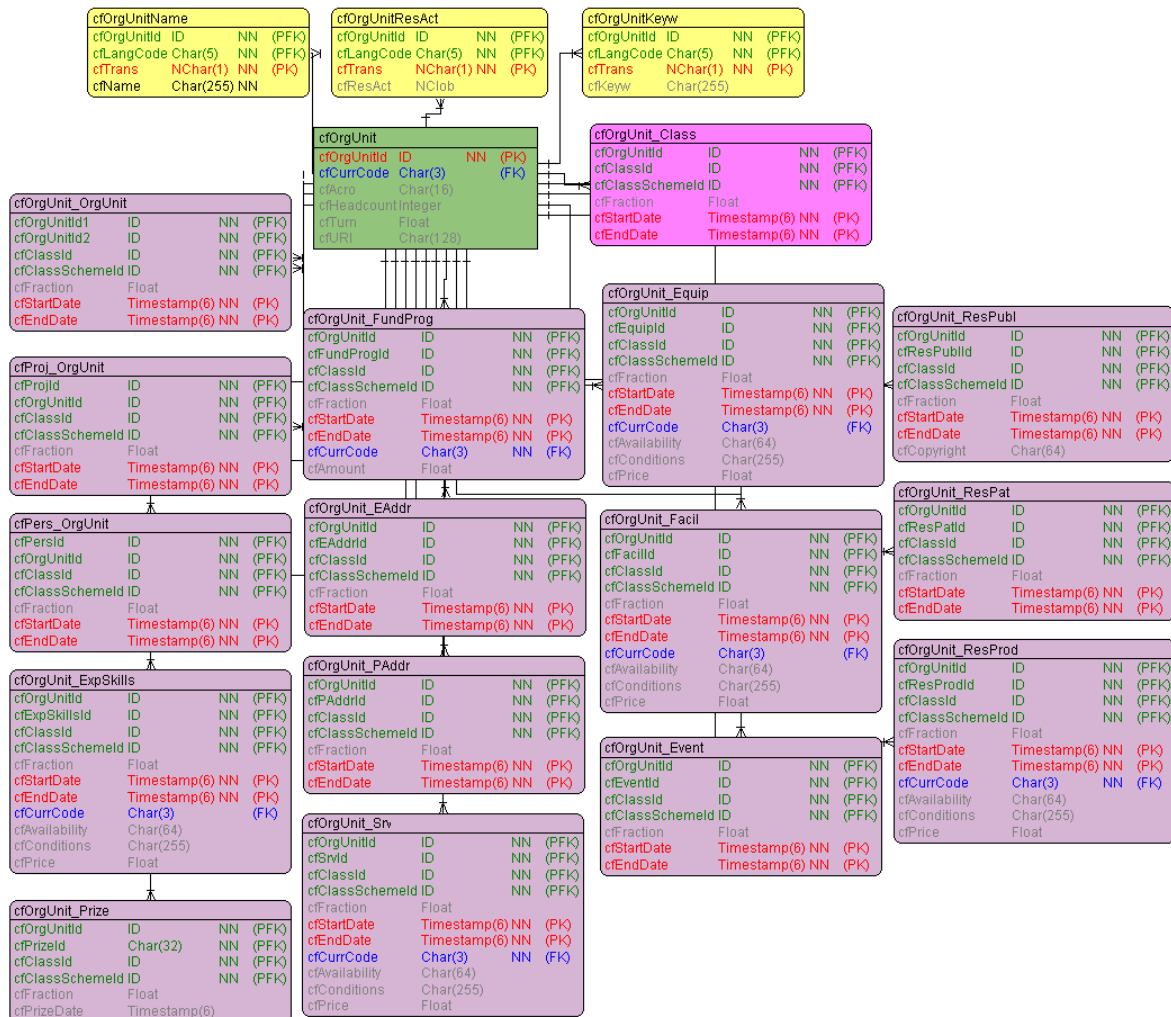


Figure 6: CERIF Core Entity OrganisationUnit

The organisation entity maintains many relationships with other entities: person, project, organisation, publication, patent, product, funding programme, equipment, facility, service, event, prize, electronic address, physical address, expertise and skills, cv, language, country and classification (cfPers_Pers, cfPers_Proj, cfPers_OrgUnit, cfPers_ResPubl, cfPers_ResPat, cfPers_ResProd, cfPers_FundProg, cfPers_Equip, cfPers_Facil, cfPers_Srv, cfPers_Event, cfPers_Prize, cfPers_EAddr, cfPers_PAddr, cfPers_ExpSkills, cfPers_CV, cfPers_Lang, cfPers_Country, cfPers_Class), as shown in figure 6. Each relationship or link entity carries semantics with a time-stamped reference to the CERIF Semantic Layer by cfClassId and cfClassSchemId and a cfFraction attribute to assign fractional values to a classification reference. Additionally, the organisation entity supports multilingual features for name, research activity descriptions and keywords (cfPersResInt, cfPersKeyw). Table 3 shows one example organisation record from a database perspective. The common and multilingual organisation attributes are stored in the upper rows; the lower rows show some example relationships including their semantics. The relationships are established by ids (i.e. cfPersId, cfOrgUnitId, cfProjId) as indicated in the Attribute column, the carrying link entites are named in the Table column, the Type column indicates the conceptual entity type (core, link, lang), the semantic values (notForProfit, isPresident, isSecretary, isTreasurer, isExecutiveStrategy, etc.) are indicated in the Classification column, where each value belongs to a particular scheme (PERS_ORGUNIT, ORGUNIT_ORGUNIT, etc). The organization example does not explicitly include any fraction values; the cfFraction attribute is not mandatory.

Table 3: CERIF OrganisationUnit Example Record

CERIF OrganisationUnit example database entry	Attribute	Table	Type	Semantic Layer (CERIF Semantics)	
				Classification (ClassIds)	Classification Scheme
organisation-eurocris	cfOrgUnitId	cfOrgUnit	core		
EUR	cfCurrCode	cfOrgUnit	core		
http://www.eurocris.org/	cfURI	cfOrgUnit	core		
euroCRIS	cfAcro	cfOrgUnit	core		
European Current Research Information Systems	cfName	cfOrgUnitName	lang		
euroCRIS as the professional association of CRIS experts and custodian of CERIF is dedicated to improvement of research information availability.	cfResAct	cfOrgUnitResAct	lang		
classification-nfp	cfClassId	cfOrgUnit_Class	link	notForProfit	ORGUNIT_CLASS
person-keith-jeffery	cfPersId	cfPers_OrgUnit	link	President	PERS-ORGUNIT
person-harrie-lalieu	cfPersId	cfPers_OrgUnit	link	Secretary	PERS-ORGUNIT
person-geert-van-grootel	cfPersId	cfPers_OrgUnit	link	Treasurer	PERS-ORGUNIT
person-anne-asserson	cfPersId	cfPers_OrgUnit	link	Strategy	PERS-ORGUNIT
person-wolfgang-adamczak	cfPersId	cfPers_OrgUnit	link	Conference	PERS-ORGUNIT
person-maximilian-stempfhuber	cfPersId	cfPers_OrgUnit	link	Workshops	PERS-ORGUNIT
person-nikos-houssos	cfPersId	cfPers_OrgUnit	link	TG-Leader-Projects	PERS-ORGUNIT
person-brigitte-joerg	cfPersId	cfPers_OrgUnit	link	TG-Leader-CERIF	PERS-ORGUNIT
person-sergey-parinov	cfPersId	cfPers_OrgUnit	link	TG-Leader-Best-Practice	PERS-ORGUNIT
person-ed-simons	cfPersId	cfPers_OrgUnit	link	TG-Leader-IR-CERIF	PERS-ORGUNIT
paddr-Voorschoten	cfPAddrId	cfOrgUnit_PAddr	link	PostOfficeBox	ORGUNIT_PADDR
eaddr-eurocris@eurocris.org	cfEAddrId	cfOrgUnit_EAddr	link	Email	ORGUNIT_EADDR
eaddr-eurocris	cfEAddrId	cfOrgUnit_EAddr	link	Skype	ORGUNIT_EADDR

The example record shows common and multilingual organisation attributes like id, currency, uri, acronym, name, research activity; the lower rows present some relationship examples. With a reference cfClassId='classification-nfp' the organisation record is classified as 'not for profit'. CERIF entities store their semantics by reference ids with interlinking [link] entities. The record maintains many person relationships with different roles: president, secretary, treasurer, etc. For person records, CERIF allows for the storage of address types: electronic addresses (email, skype) or postal addresses (post-office-box). The example record only gives some relationship examples; the entire model allows for many more. The linkage mechanism by link entities is consistent across the model and will be explained in detail within section 2.5; for the semantic features we refer to section 2.7.

2.3 CERIF Result Entities

The CERIF result entities are ResultPublication, ResultPatent and ResultProduct. Figure 7 shows the result entities and their linking relationships. The ResultPublication entity like a core entity recursively links to itself. The result entities represent research output.

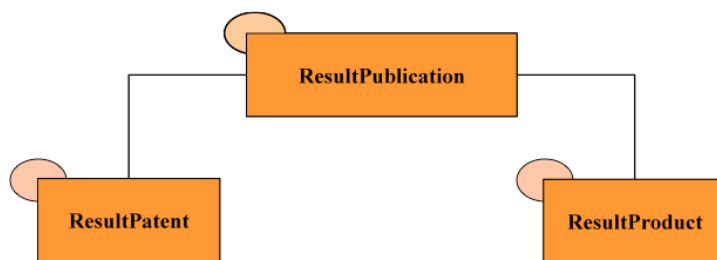


Figure 7: CERIF Result Entities

Figure 8 shows the result entities (cfResPubl, cfResPat, cfResProd) and their related entities from a physical perspective. The circle in figure 7 represents recursiveness; that is, the relationships in between publications (cfResPubl_ResPubl). The recursive and the interlinking relations (cfResPubl_ResProd, cfResPubl_ResPat) in figure 8 are link type entities to be introduced in section 2.5. The yellow colored entities (cfResPublTitle, cfResPublSubtitle, cfResPublAbstr, cfResPatTitle, etc.) support the feature of multiple languages and will be introduced in section 2.6.

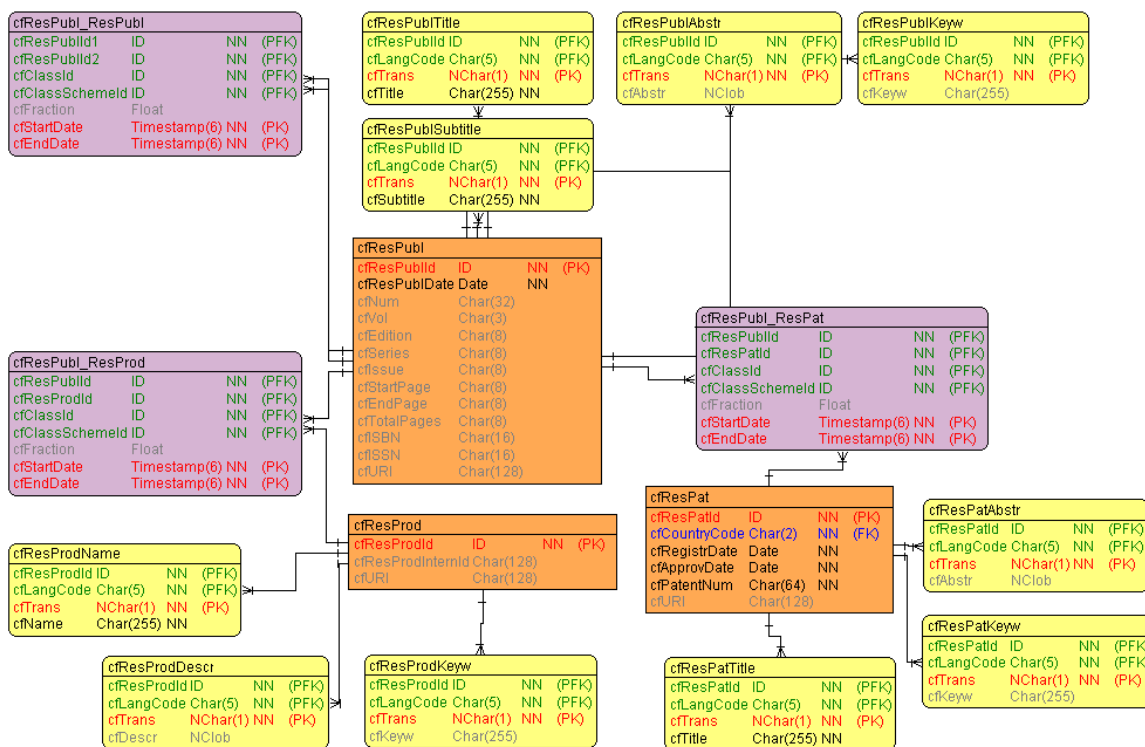


Figure 8: CERIF Result Entities, their Recursion and some Link Entities

Each result entity (cfResPubl, cfResPat, cfResProd) will subsequently be presented and some examples for the publication entity will be provided for a better understanding.

2.3.1 CERIF Entity ResultPublication

For an identification of records the result publication entity (cfResPubl) foresees an id attribute (cfResPublId). Besides, the attributes publication date, number, volume, edition, series, issue, startpage, endpage, total pages, isbn, issn, and uri (cfResPublDate, cfNum, cfVolume, cfEdition, cfSeries, cfIssue, cfStartPage, cfEndpage, cfTotalPages, cfISBN, cfISSN, cfURI) are considered as common publication attributes. The result publication entity maintains many relationships with other entities: publication, patent, product, organisation, project, person, funding programme, equipment, facility, event, classification (cfResPubl_ResPubl, cfResPubl_ResPat, cfResPubl_ResProd, cfOrgUnit_ResPubl, cfProj_ResPubl, cfPers_ResPubl, cfResPubl_Equip, cfResPubl_Facil, cfResPubl_FundProg, cfResPubl_Class) as shown in figure 9. Each relationship or link entity carries semantics with a time-stamped reference to the CERIF Semantic Layer by cfClassId and cfClassSchemId and a cfFraction attribute to assign fractional values to a classification reference. Additionally, the publication entity supports multilingual features for title, subtitle, abstract, note, abbreviation and keywords (cfResPublTitle, cfResPublSubtitle, cfResPublAbstr, cfResPublBiblNote, cfResPublKeyw) as shown in figure 9.

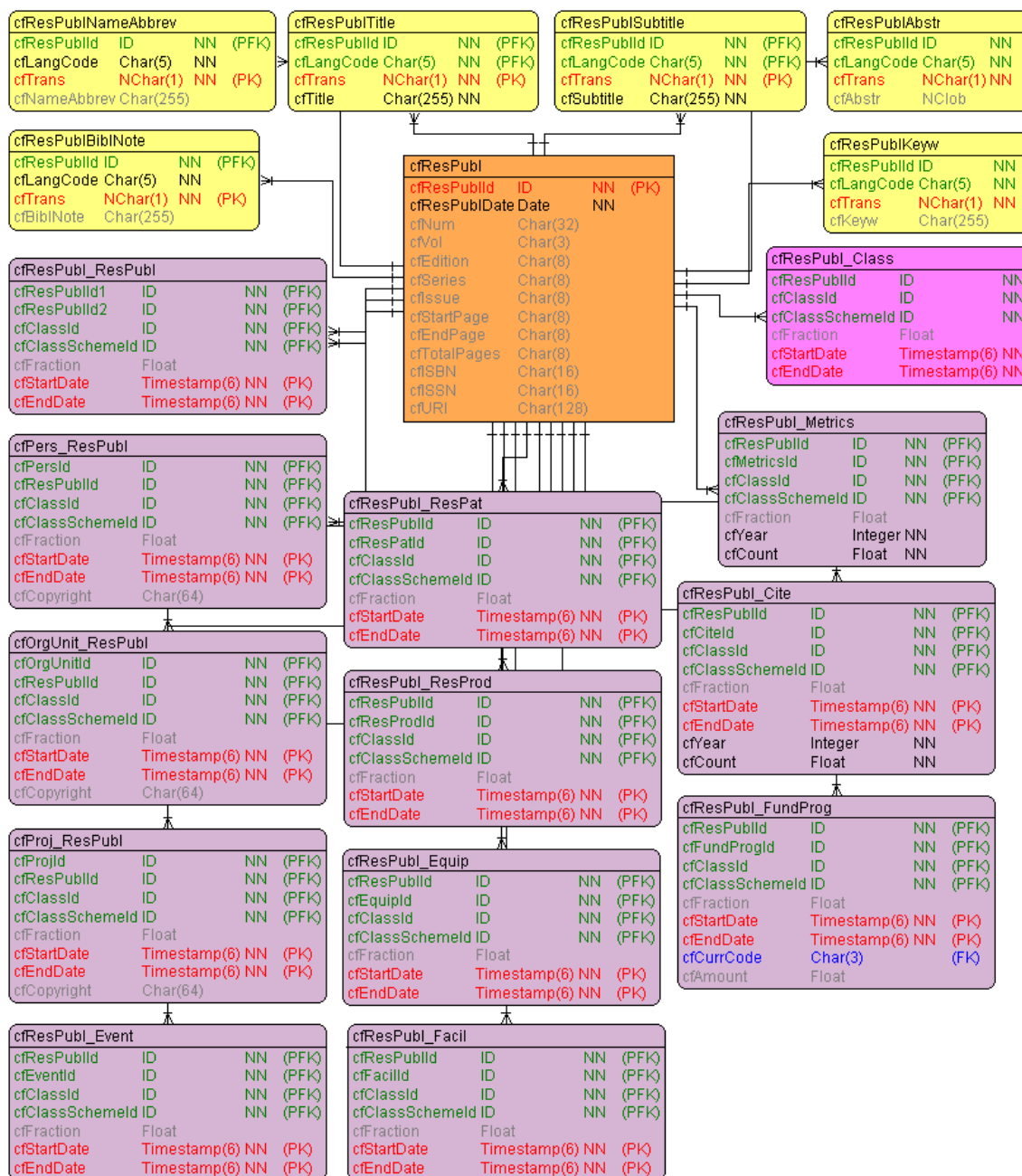


Figure 9: CERIF Result Entity ResultPublication

Table 4 shows one example publication record from a database perspective. The common and multilingual publication attributes are stored in the upper rows; the lower rows show some example relationships including their semantics. The relationships are established by ids (i.e. cfPersId, cfOrgUnitId, cfProjId, cfEventId) as indicated in the Attribute column, the carrying link entities are named in the Table column, the Type column indicates the entity type (core, link, lang), the semantic values (Conference Proceedings Article, isPartOf, isAuthorOf, isOriginatorOf, PresentedAt etc.) and fractions are indicated in the Classification column where each value belongs to a scheme (CERF2008-PUBL-TYPE, RESPUBL_STRUCTURE, etc.).

Table 4: CERIF ResultPublication Example Record

CERIF ResultPublication example database entry				Semantic Layer (CERIF Semantics)	
Data	Attribute	Table	Type	Classification (ClassIds)	Classification Scheme
publication-joerg-et-al	cfResPublId	cfResPubl	result		
2008-01-01*	cfResPublDate	cfResPubl	result		
107	cfStartPage	cfResPubl	result		
123	cfEndPage	cfResPubl	result		
978-961-6133-38-8	cfISBN	cfResPubl	result		
http://www.eurocris.org/fileadmin/Upload/Events/Conferences/CRIS2008/Papers/cris2008_Joerg.pdf	cfURI	cfResPubl	result		
Analyzing European Research Competencies	cfTitle	cfResPublTitle	lang		
Results from a European SSA Project	cfSubtitle	cfResPublSubtitle	lang		
With this paper we will present the approach of analyzing research competencies across EU countries	cfAbstr	cfResPublAbstr	lang		
IST, ERA, CRIS, CERIF, Research Competencies, NMS, Analysis	cfKeyw	cfResPublKeyw	lang		
classification-conf-proc-article	cfClassId	cfResPubl_Class	link	Conference Proceedings Article	CERIF2008-PUBL-TYPE
publication-get-the-good-cris	cfResPublId2	cfResPubl_ResPubl	link	isPartOf	RESPUBL-RESPUBL
person-brigitte-joerg	cfPersId	cfPers_ResPubl	link	FirstAuthor [fract=0.25]	PERS-RESPUBL-%ALLOCATION
person-hans-uszkoreit	cfPersId	cfPers_ResPubl	link	Author	PERS-RESPUBL
person-jure-ferlez	cfPersId	cfPers_ResPubl	link	Author	PERS-RESPUBL
person-mitja-jermol	cfPersId	cfPers_ResPubl	link	Author	PERS-RESPUBL
project-ist-world	cfProjId	cfProj_ResPubl	link	Originator	PERS-RESPUBL
event-cris-2008	cfPersId	cfResPubl_Event	link	Presenter	RESPUBL-EVENT

The example record shows the common and multilingual publication attributes id, date, startpage, endpage, isbn, number, title, abstract and keywords. The lower rows present some relationship examples. With a reference cfClassId='classification-conf-proc-article', the publication record is classified as a Conference Proceedings Article. A recursive relationship cfResPublId2='publication-get-the-good-cris' refers to the entire proceedings. The example shows some person relationships with different roles. The fraction example shows a %-allocation in the person-publication relationship link with the role of first author. A reference to project cfProj='project-ist-world' reveals the project as originator of the publication, an event link indicates that the paper was presented at cfEventId=event-cris-2008. The record only gives some relationship examples; the entire model allows for many more. The linkage mechanism by link entities is consistent across the model and will be explained in detail within section 2.5; for the semantic features we refer to section 2.7.

* The cfPublicationDate attribute is specified as a "date" datatype. It requires a date format: YYYY-MM-DD For a year 2007 or 2008 – we recommend the following values: 2007-01-01; 2008-01-01.

Another example record in table 5 below again shows the common and multilingual result publication attributes id, date, no, volume, startpage, endpage, isbn and issn number, title, abstract and keywords; the lower rows present some relationship examples. The example publication record is classified as a ‘Journal Article’ and a recursive relationship via cfResPubId2=’publication-vldb-journal’ indicates the linkage to the journal of which the article is part. The example record is classified by the Springer subject scheme into ‘Computer Science’. A person link carries the author role, and the link to the organisation record ‘organisation-springer’ indicates the publisher of the article.

Table 5: CERIF ResultPublication Example Record of a Journal Article

CERIF ResultPublication example database entry	Attribute	Table	Type	Semantic Layer (CERIF Semantics)	
				Classification (ClassIds)	Classification Scheme
publication-veda-c-storey	cfResPubId	cfResPubl	result		
1993	cfResPublDate	cfResPubl	result		
4	cfNum	cfResPubl	result		
2	cfVol	cfResPubl	result		
455	cfStartPage	cfResPubl	result		
488	cfEndPage	cfResPubl	result		
1066-8888	cfISSN	cfResPubl	result		
http://www.springerlink.com/content/j23263j02m850617/	cfURI	cfResPubl	result		
Understanding Semantic Relationships	cfTitle	cfResPublTitle	lang		
To develop sophisticated database management systems, there is a need to incorporate more understanding of the real world in the information that is stored in a database. Semantic data models have been developed to try to capture some of the meaning, as well as the structure, of data using abstractions such as inclusion, aggregation, and association. Besides these well-known relationships, a number of additional semantic relationships have been identified by researchers in other disciplines ...	cfAbstr	cfResPublAbstr	lang		
Database design, entity relationship model, relational model, semantic relationships, database design systems	cfKeyw	cfResPublKeyw	lang		
classification-journal-article	cfClassId	cfResPubl_Class	link	Journal Article	CERIF2008-PUBL-TYPES
classification-computer-science	cfClassId	cfResPubl_Class	link	Computer Science	SPRINGER-SUBJECTS
publication-vldb-journal	cfResPubId2	cfResPubl_ResPubl	link	isPartOf	RESPUBL-RESPUBL
person-veda-c-storey	cfPersId	cfPers_ResPubl	link	Author	PERS-RESPUBL
organisation-springer	cfOrgUnitId	cfOrgUnit_ResPubl	link	Publisher	PERS-RESPUBL

Table 6: CERIF ResultPublication Example Record of a Journal

CERIF ResultPublication example database entry	Attribute	Table	Type	Semantic Layer (CERIF Semantics)	
				Classification (ClassIds)	Classification Scheme
publication-vldb-journal	cfResPubId	cfResPubl	result		
1992-07-01	cfResPubDate	cfResPubl	result		
http://www.vldb.org/dblp/db/journals/vldb/	cfURI	cfResPubl	result		
The VLDB Journal	cfTitle	cfResPublTitle	lang		
Published on behalf of the VLDB Endowment, this journal contains scholarly contributions that examine information system architectures, the impact of technological ...	cfAbstr	cfResPublAbstr	lang		
Persistent Object Systems, Multimedia Databases, Databases and the Web, E-Services, XML Data Management, Data Stream Processing, Data Management, Analysis and Mining for the Life Sciences, Information Retrieval	cfKeyw	cfResPublKeyw	lang		
classification-journal-article	cfClassId	cfResPubl_Class	link	Journal	CERIF2008-PUBL-TYPES
publication-veda-c-storey	cfResPubId2	cfResPubl_ResPubl	link	hasPart	RESPUBL-RESPUBL
person-kyu-young-whang	cfPersId	cfPers_ResPubl	link	ChiefEditor	PERS-RESPUBL
person-philip-a-bernstein	cfPersId	cfPers_ResPubl	link	ChiefEditor	PERS-RESPUBL
person-christian-s-jensen	cfPersId	cfPers_ResPubl	link	ChiefEditor	PERS-RESPUBL
organisation-springer	cfOrgUnitId	cfOrgUnit_ResPubl	link	Publisher	PERS-RESPUBL

The link entities as semantic carriers are a major strength of the CERIF model. In the example record only some relationships have been presented where the entire model allows for many more, according to system context and needs. The linkage mechanism by link entities is consistent across the model and will be explained in detail within section 2.5; for the semantic features we refer to section 2.7. With the current release, a semantic scheme for publication types and related roles has been introduced: CERIF Semantics [12].

The CERIF result publication entity allows for the generation of complete publication reference records like BibTeX, as shown in table 7.

Table 7: BibTeX example records generated from CERIF publication examples

BibTeX example record generated from table 4	BibTeX example record generated from table 5
<pre>@article{ , author = {Joerg Brigitte, Uszkoreit Hans, Ferlez Jure, Jermol Mitja}, title = {Analyzing European Research Competencies in IST: Results from a European SSA Project}, year = {2008}, isbn = { 978-961-6133-38-8}, pages = {107--123}, publisher = {IZUM, Institut of Information Science}, address = {Maribor, Slovenia}, }</pre>	<pre>@article{ , author = {Veda C. Storey}, title = {Understanding semantic relationships}, journal = {The VLDB Journal}, volume = {2}, number = {4}, year = {1993}, issn = {1066-8888}, pages = {455--488}, publisher = {Springer-Verlag New York, Inc.}, address = {Secaucus, NJ, USA}, }</pre>

2.3.2 CERIF Entity ResultPatent

For an identification of records the result patent entity (cfResPat) foresees an id attribute (cfResPatId). Besides, the attributes country code, registration date, approval date, patent number and uri (cfCountryCode, cfRegistrDate, cfApprovDate, cfPatentNum, cfURI) are considered common patent attributes. The result patent entity maintains many relationships with other entities: patent, publication, organisation, project, person, funding programme (cfResPat_ResPat, cfResPat_Class, cfResPubl_ResPat, cfOrgUnit_ResPat, cfProj_ResPat, cfResPat_FundProg, cfPers_ResPat) as shown in figure 10. Each relationship or link entity carries semantics with a time-stamped reference to the CERIF Semantic Layer by cfClassId and cfClassSchemId and a cfFraction attribute to assign fractional values to a classification reference. Additionally, the result patent entity supports multilingual features for title, abstract, and keywords (cfResPatTitle, cfResPatAbstr, cfResPatKeyw).

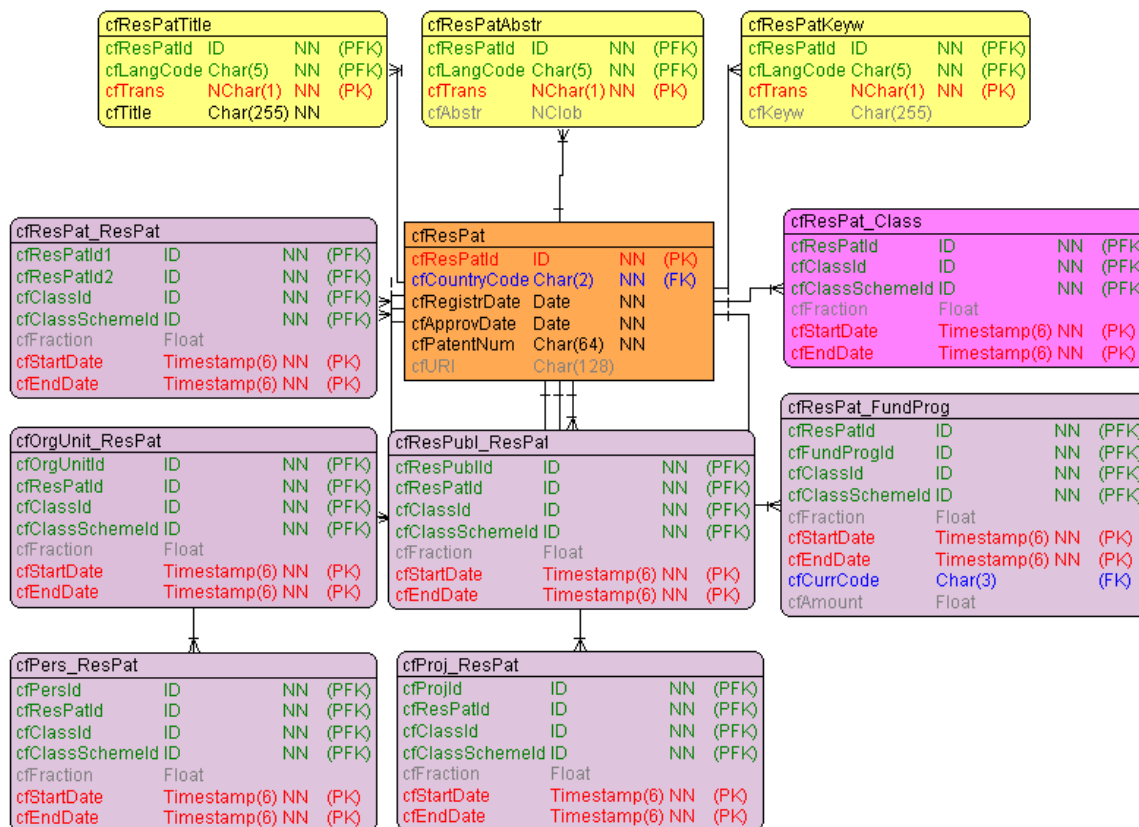


Figure 10: CERIF Result Entity ResultPatent

2.3.3 CERIF Entity ResultProduct

For an identification of records the result product entity (cfResProd) foresees an id attribute (cfResProdId). Besides, the attributes internal identifier and uri (cfResProdInternId, cfURI) are considered as common product attributes. The result product entity maintains many relationships with entities: publication, organisation, project, person, funding programme (cfResProd_Class, cfResPubl_ResProd, cfProj_ResProd, cfPers_ResProd, cfOrgUnit_ResProd, cfResProd_FundProg) as shown in figure 11. Each relationship or link entity carries semantics with a time-stamped reference to the CERIF Semantic Layer by cfClassId and cfClassSchemId and a cfFraction attribute to assign fractional values to a classification reference. Additionally, the result product entity supports multilingual features for the name, for description, and keywords (cfResProdName, cfResProdDescr, cfResProdKeyw).

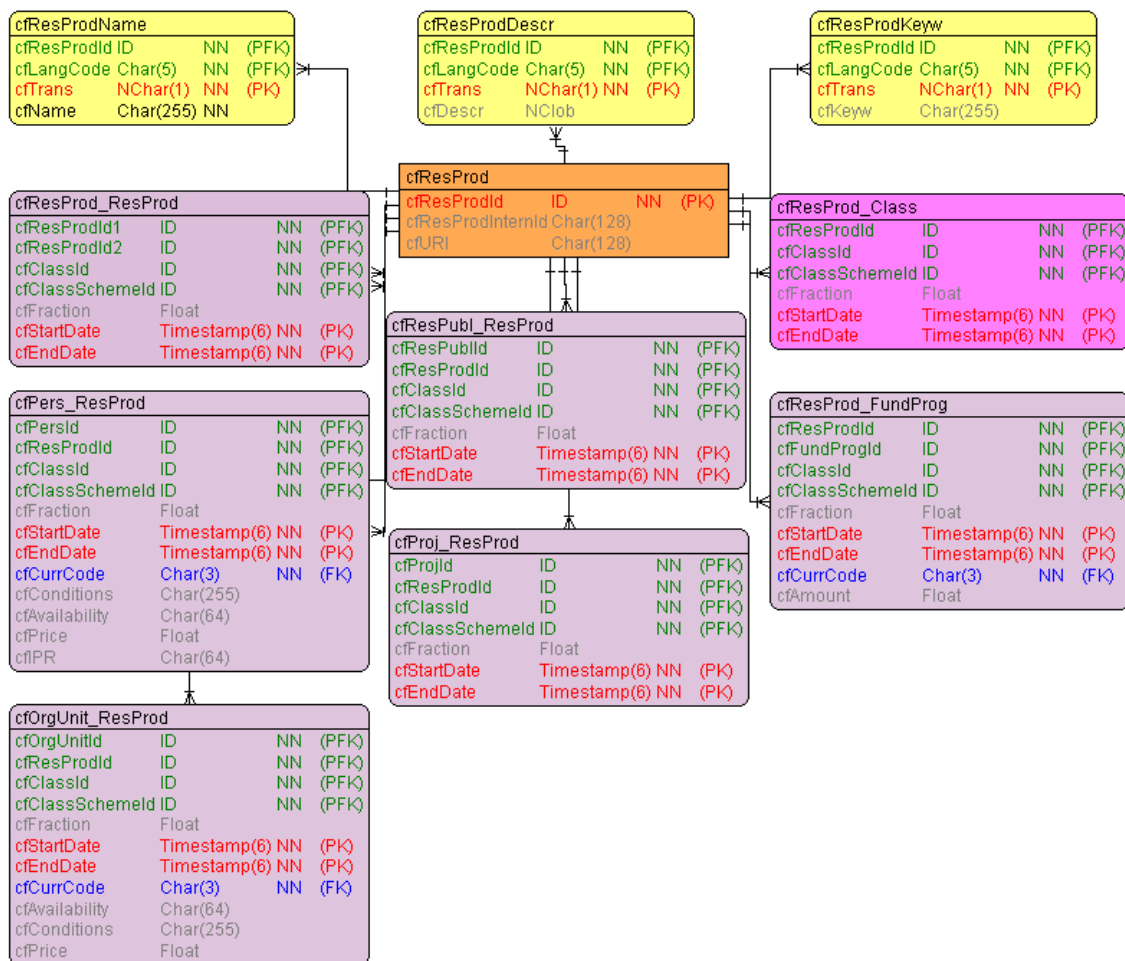


Figure 11: CERIF Result Entity ResultProduct

2.4 CERIF 2nd Level Entities

Beyond the core and result entities, CERIF employs many so called 2nd level entities. In figure 12 the 2nd level entities are presented as a circle surrounding the core and result entities in blue color.

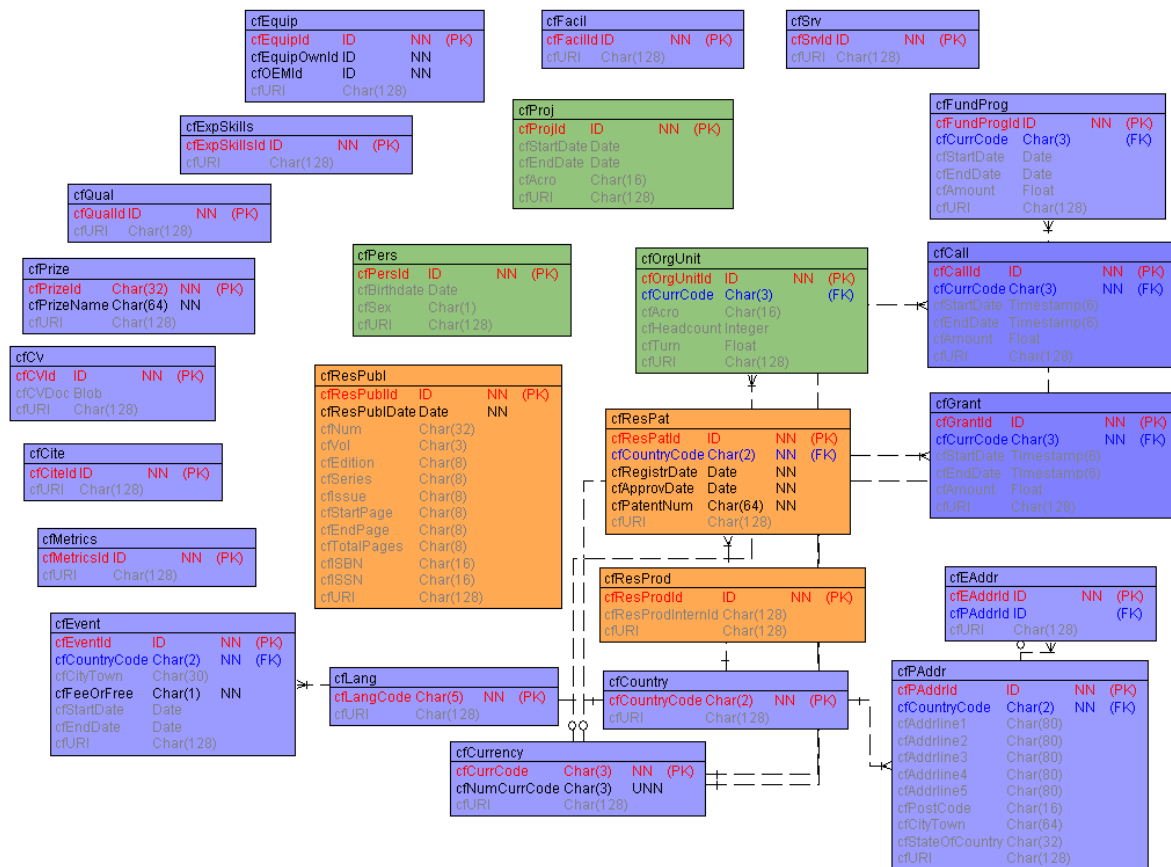


Figure 12: CERIF 2nd Level Entities organised as a circle around core and result entities

The 2nd level entities allow for the representation of the research context by linking to them from core and result entities. Each 2nd level entity supplies some basic attributes; at least an id and an uri attribute. The linkage mechanism and the multilingual features of 2nd level entities – not shown in figure 12 – are equal to the mechanism and features presented with core and result entities. For more details about the link entities and their function as semantic carriers we refer to the following sections.

2.5 CERIF Link Entities

The relationships or links between CERIF entities are called Link Entities. Link entities are considered a major strength of the CERIF model. A link entity always connects two entities, either core, result, or second 2nd level entities. Figure 14 shows an abstract view of some link entities (Person_ResultPublication, Person_Project, Person_OrganisationUnit, Project_ResultPublication, OrganisationUnit_ResultPublication, Project_OrganisationUnit) connecting the core entities and the result publication entity.

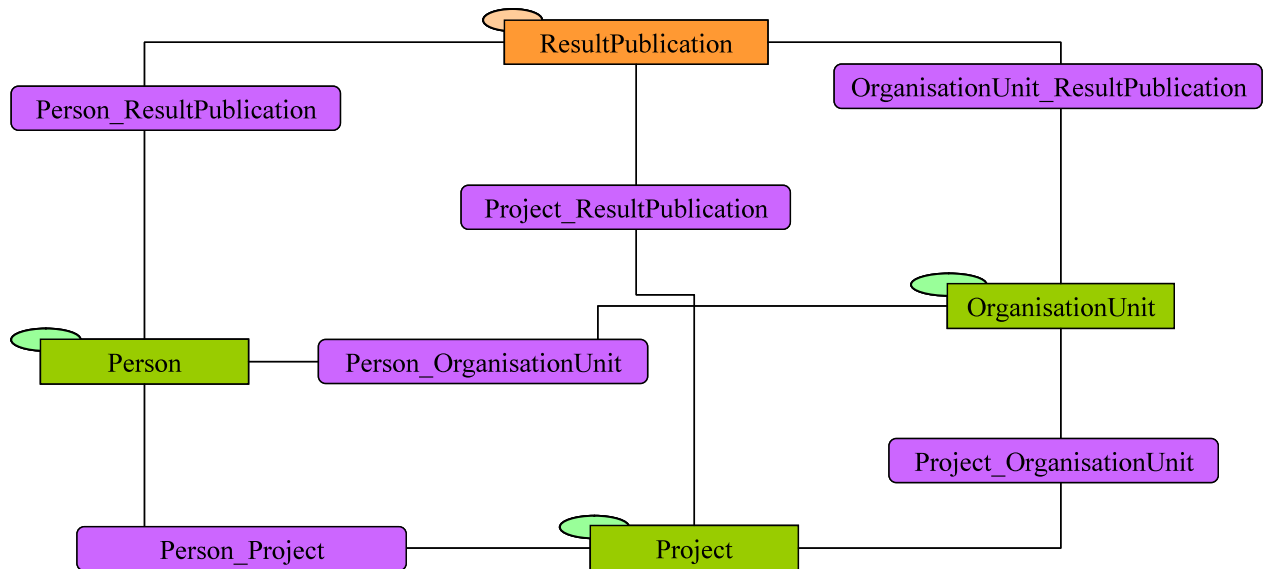


Figure 14: CERIF Link Entities in the context of the core entities and a result entity

The CERIF link entities have been mentioned in the context of the presented core, result and 2nd level entities; their structure and functionality at physical level is consistent all over the model as demonstrated with some example link entities in figure 15.

<table border="1"> <tr><th colspan="4">cfPers_Pers</th></tr> <tr><td>cfPersId1</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfPersId2</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfClassId</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfClassSchemId</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfFraction</td><td>Float</td><td></td><td></td></tr> <tr><td>cfStartDate</td><td>Timestamp(6)</td><td>NN</td><td>(PK)</td></tr> <tr><td>cfEndDate</td><td>Timestamp(6)</td><td>NN</td><td>(PK)</td></tr> </table>	cfPers_Pers				cfPersId1	ID	NN	(PFK)	cfPersId2	ID	NN	(PFK)	cfClassId	ID	NN	(PFK)	cfClassSchemId	ID	NN	(PFK)	cfFraction	Float			cfStartDate	Timestamp(6)	NN	(PK)	cfEndDate	Timestamp(6)	NN	(PK)	<table border="1"> <tr><th colspan="4">cfProj_OrgUnit</th></tr> <tr><td>cfProjId</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfOrgUnitId</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfClassId</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfClassSchemId</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfFraction</td><td>Float</td><td></td><td></td></tr> <tr><td>cfStartDate</td><td>Timestamp(6)</td><td>NN</td><td>(PK)</td></tr> <tr><td>cfEndDate</td><td>Timestamp(6)</td><td>NN</td><td>(PK)</td></tr> </table>	cfProj_OrgUnit				cfProjId	ID	NN	(PFK)	cfOrgUnitId	ID	NN	(PFK)	cfClassId	ID	NN	(PFK)	cfClassSchemId	ID	NN	(PFK)	cfFraction	Float			cfStartDate	Timestamp(6)	NN	(PK)	cfEndDate	Timestamp(6)	NN	(PK)												
cfPers_Pers																																																																													
cfPersId1	ID	NN	(PFK)																																																																										
cfPersId2	ID	NN	(PFK)																																																																										
cfClassId	ID	NN	(PFK)																																																																										
cfClassSchemId	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)																																																																										
cfClassSchemId	ID	NN	(PFK)																																																																										
cfFraction	Float																																																																												
cfStartDate	Timestamp(6)	NN	(PK)																																																																										
cfEndDate	Timestamp(6)	NN	(PK)																																																																										
<table border="1"> <tr><th colspan="4">cfOrgUnit_Equip</th></tr> <tr><td>cfOrgUnitId</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfEquipId</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfClassId</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfClassSchemId</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfFraction</td><td>Float</td><td></td><td></td></tr> <tr><td>cfStartDate</td><td>Timestamp(6)</td><td>NN</td><td>(PK)</td></tr> <tr><td>cfEndDate</td><td>Timestamp(6)</td><td>NN</td><td>(PK)</td></tr> <tr><td>cfCurrCode</td><td>Char(3)</td><td></td><td>(FK)</td></tr> <tr><td>cfAvailability</td><td>Char(64)</td><td></td><td></td></tr> <tr><td>cfConditions</td><td>Char(255)</td><td></td><td></td></tr> <tr><td>cfPrice</td><td>Float</td><td></td><td></td></tr> </table>	cfOrgUnit_Equip				cfOrgUnitId	ID	NN	(PFK)	cfEquipId	ID	NN	(PFK)	cfClassId	ID	NN	(PFK)	cfClassSchemId	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			<table border="1"> <tr><th colspan="4">cfOrgUnit_Class</th></tr> <tr><td>cfOrgUnitId</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfClassId</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfClassSchemId</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfFraction</td><td>Float</td><td></td><td></td></tr> <tr><td>cfStartDate</td><td>Timestamp(6)</td><td>NN</td><td>(PK)</td></tr> <tr><td>cfEndDate</td><td>Timestamp(6)</td><td>NN</td><td>(PK)</td></tr> </table>	cfOrgUnit_Class				cfOrgUnitId	ID	NN	(PFK)	cfClassId	ID	NN	(PFK)	cfClassSchemId	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)																																																																										
cfClassSchemId	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)																																																																										
cfClassSchemId	ID	NN	(PFK)																																																																										
cfFraction	Float																																																																												
cfStartDate	Timestamp(6)	NN	(PK)																																																																										
cfEndDate	Timestamp(6)	NN	(PK)																																																																										
<table border="1"> <tr><th colspan="4">cfClass_Class</th></tr> <tr><td>cfClassId1</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfClassId2</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfClassSchemId1</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfClassSchemId2</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfClassId</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfClassSchemId</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfFraction</td><td>Float</td><td></td><td></td></tr> <tr><td>cfStartDate</td><td>Timestamp(6)</td><td>NN</td><td>(PK)</td></tr> <tr><td>cfEndDate</td><td>Timestamp(6)</td><td>NN</td><td>(PK)</td></tr> </table>	cfClass_Class				cfClassId1	ID	NN	(PFK)	cfClassId2	ID	NN	(PFK)	cfClassSchemId1	ID	NN	(PFK)	cfClassSchemId2	ID	NN	(PFK)	cfClassId	ID	NN	(PFK)	cfClassSchemId	ID	NN	(PFK)	cfFraction	Float			cfStartDate	Timestamp(6)	NN	(PK)	cfEndDate	Timestamp(6)	NN	(PK)	<table border="1"> <tr><th colspan="4">cfProj_Pers</th></tr> <tr><td>cfProjId</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfPersId</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfClassId</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfClassSchemId</td><td>ID</td><td>NN</td><td>(PFK)</td></tr> <tr><td>cfFraction</td><td>Float</td><td></td><td></td></tr> <tr><td>cfStartDate</td><td>Timestamp(6)</td><td>NN</td><td>(PK)</td></tr> <tr><td>cfEndDate</td><td>Timestamp(6)</td><td>NN</td><td>(PK)</td></tr> </table>	cfProj_Pers				cfProjId	ID	NN	(PFK)	cfPersId	ID	NN	(PFK)	cfClassId	ID	NN	(PFK)	cfClassSchemId	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)																																																																										
cfClassSchemId1	ID	NN	(PFK)																																																																										
cfClassSchemId2	ID	NN	(PFK)																																																																										
cfClassId	ID	NN	(PFK)																																																																										
cfClassSchemId	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)																																																																										
cfClassSchemId	ID	NN	(PFK)																																																																										
cfFraction	Float																																																																												
cfStartDate	Timestamp(6)	NN	(PK)																																																																										
cfEndDate	Timestamp(6)	NN	(PK)																																																																										

Figure 15: Some CERIF Link Entities to demonstrate the consistency in their structure

Where figure 15 shows examples of some link entities at physical level, figure 16 introduces their structure and functionality rather from a meta perspective.

cfEntity1Name_Entity2Name	
cfInheritedEntity1Identifier ID	(PFK)
cfInheritedEntity2Identifier ID	(PFK)
cfInheritedClassificationIdentifier ID	(PFK)
cfInheritedClassificationSchemeIdentifier ID	(PFK)
cfFraction Float	
cfStartDate Timestamp	(PK)
cfEndDate Timestamp	(PK)

Figure 16: Meta perspective towards CERIF Link Entities

The physical name of link entities is composed of the names of the two involved entities, including the CERIF prefix as follows: cfEntity1Name_Entity2Name. The order of the linking entity names implies the order of the both identifier attributes, where the first (cfInheritedEntity1Identifier) is inherited from entity cfEntity1Name, and the second (cfInheritedEntity2Identifier) is inherited from the entity cfEntity2Name. All the identifiers at the meta perspective are labelled as inherited because they do not origin in the link entities themselves but rather are inherited from those entities (cfEntity1, cfEntity2, cfClassification, cfClassificationScheme) where they are maintained. All link entities establish linkage between two entities by id references cfInheritedEntity1Identifier and cfInheritedEntity2Identifier. Additionally, each link entity carries semantics by reference to the so-called CERIF Semantic Layer via the cfInheritedClassificationIdentifier and cfInheritedClassificationSchemeIdentifier (see section 2.7) and a cfFraction attribute to assign fractional values to a classification reference. Whereas the classification and classification scheme references are mandatory, the fraction attribute is not. Besides, each linking record requires a startdate and enddate*. Some link entities allow for attributes like currency or copyright as indicated in figure 14 above. Together, all inherited identifiers and the date attributes build the primary key of link entities.

Real data examples for link entities have been presented in the context of core and result entities with the tables 1-5. Some general linkage examples are provided in table 8. As cfFraction attribute is not mandatory it is not included in the examples of table 8, but has been introduced in previous example tables with core the person and project entities.

Table 8: CERIF Link Entity Examples

Link Table (Link Entity)	Inherited Entity1 Identifier	Inherited Entity2 Identifier	Inherited Classificati on Identifier	Inherited Classification Scheme Identifier	Start Date	End Date
cfOrgUnit1_OrgUnit2	orga-id1	orga-id2	hasPart	Organisation Structure	2001-01-01T12:00:00-05:00	2001-12-31T12:00:00-05:00
cfOrgUnit1_OrgUnit2	orga-id2	orga-id3	isPartOf	Organisation Structure	2009-01-13T12:00:00-05:00	2099-01-13T12:00:00-05:00
cfPers_OrgUnit	person-id1	orga-id1	Head	Organisational Person Roles	2009-01-13T12:00:00-05:00	2099-01-13T12:00:00-05:00
cfPers1_Pers2	person-id1	person-id2	Supervisor	Academic Person Roles	2009-01-13T12:00:00-05:00	2099-01-13T12:00:00-05:00
cfPers_Proj	person-id2	project-id1	Participant	Project-related Person Roles	2009-01-13T12:00:00-05:00	2099-01-13T12:00:00-05:00
cfPers_ResPubl	person-id1	publ-id1	Author	Publication- Person Roles	2009-01-13T12:00:00-05:00	2099-01-13T12:00:00-05:00

* We recommend to add 1901-01-01T00:0000-01:00 as a startdate, in case of unknown, and we recommend to add 2099-12-31T23:59:59-01:00 as an enddate, in case of unknown.

Each record in a link table carries the semantics of the linkage by reference to the Semantic Layer. In table 8, the example records show that there may exist classification schemes for ‘Organisation Structure’, ‘Organisational Person Roles’, ‘Academic Person Roles’, ‘Project-related Person Roles’, ‘Publication-Person Roles’. Each semantic value (classification identifier) has to be assigned to one particular classification scheme. In table 8, the ‘hasPart’ and ‘isPartOf’ classifiers belong to the ‘Organisation Structure’ scheme; the classifier ‘Supervisor’ belongs to the ‘Academic Person Roles’ scheme. Whereas the link entities only carry the semantics because they solely store ids, the real values and classifiers including their scheme assignments are maintained and stored within the CERIF Semantic Layer and will be explained in section 2.7.

2.6 CERIF Multiple Language Features

Much information in research environments needs representation in more than one language. The support of multilingual features is very important in countries where several official languages are spoken and maintained. As indicated in figure 17, CERIF supports multiple language features for names, titles, descriptions, keywords, abstracts, and even for the semantics.

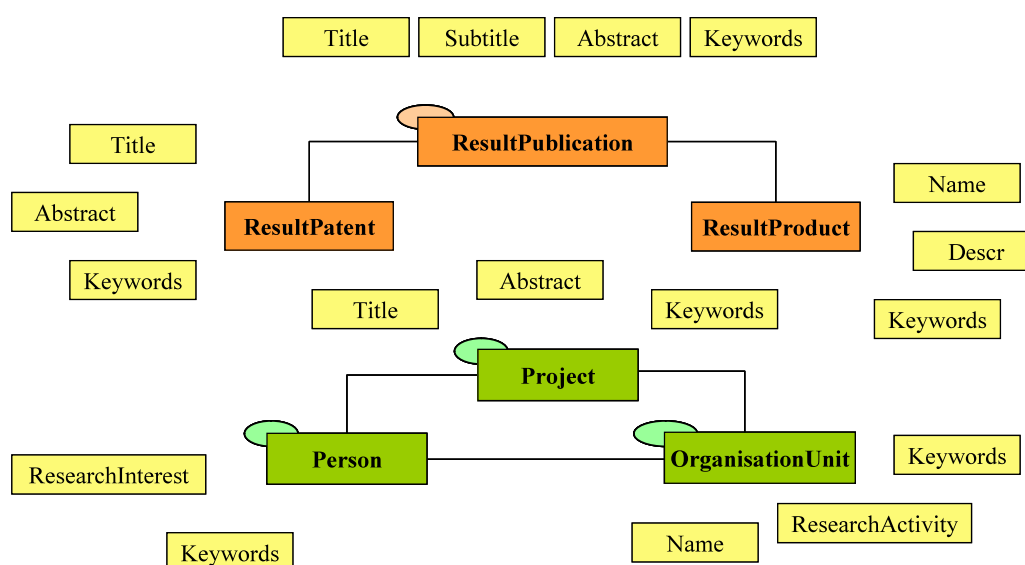


Figure 17: Some CERIF Entities with Multiple Language Features

Figure 18 below shows multilingual features for some selected entities. Their identifiers indicate the assignment towards their originating entities (cfProjId, cfOrgUnitId, cfResPubId). The encoded language is stored with the cfLangCode attribute that allows for five character values (i.e. en, de, fr, si, en-uk, en-us, fr-fr, fr-be, fr-nl). A translation attribute allows for information about the translation type: o=original, h=human, or m=machine. The title, abstract, keyword or research activity attributes (cfTitle, cfAbstract, cfKeyw, cfResAct) store the texts in a particular language.

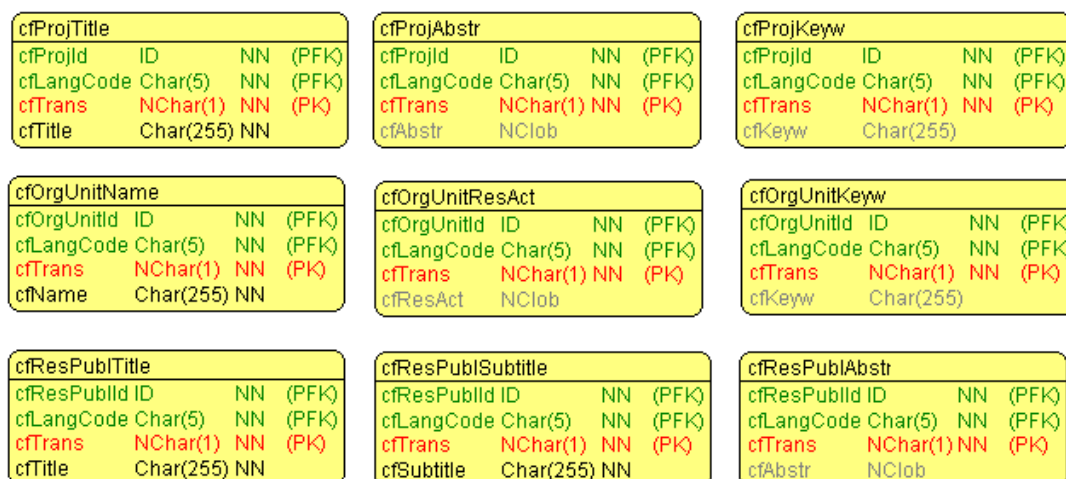


Figure 18: Come CERIF entities with Multiple Language Features

Besides the core, result and 2nd level entities, also the classification entities in the CERIF Semantic Layer allow for multiple language records. It is thus possible to maintain classification schemes in different languages. Even language names and country names can be maintained in several languages: België (cfLangCode=du), Belgien (cfLangCode=de), Belgique (cfLangCode=fr), Belgium (cfLangCode=en).

2.7 CERIF Semantic Layer [Semantic Features]

The so-called CERIF Semantic Layer is a simple but powerful instrument that allows for the representation of relationship kinds [6, 8], application views, subject classifications, any other classification schemes [13, 14, 15], or mappings between schemes. The CERIF Semantic Layer supplies the means for maintaining the CERIF Semantics: any types, roles, terminology, subject classifiers, or mappings. It stores the semantic values that are carried by or referred to from the link entities via the cfClassSchemeId attribute references, and it assigns each semantic value to a particular classification scheme. The CERIF Semantic Layer is constructed by the entities shown in figure 19.

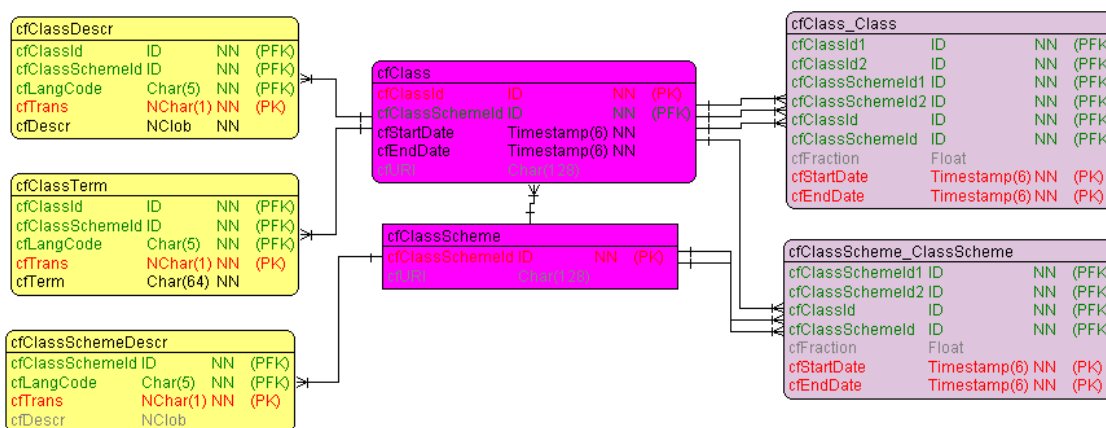


Figure 19: CERIF Semantic Layer

The CERIF Semantic Layer consists of the two class type entities classification (cfClass), and classification scheme (cfClassScheme). Additionally, it allows for a representation of multilingual terms (cfClassTerm) and class descriptions (cfClassDescr). The both class type entities (cfClass, cfClassScheme) are interconnected by two recursive entities (cfClass_Class, cfClassScheme_ClassScheme) to allow for the representation of structures and for the mappings between classifications or classification schemes. The recursive entities of the CERIF Semantic Layer consistently support fractional values for classification references.

The following records in table 9 show examples, including some of the new CERIF publication types and roles [12].

Table 9: CERIF Semantic Layer examples

	cfClassId	cfTerm [cfLangCode=en]	cfClassDescr	Link Entity	cfClass SchemeId
type	class	class	class	link	class
link	class-1	Book	A book is a ...	cfResPubl_Class	publ-types
link	class-2	Book Review	A book review is a ...	cfResPubl_Class	publ-types
link	class-3	Book Chapter Abstract	A book chapter is a ...	cfResPubl_Class	publ-types
link	class-4	Book Chapter Review	A book chapter review ...	cfResPubl_Class	publ-types
link	class-5	Inbook		cfResPubl_Class	publ-types
link	class-6	Anthology		cfResPubl_Class	publ-types
	...				
link	class-10	is author of		cfPers_ResPubl	pers-publ-roles
link	class-11	is author (numbered) of		cfPers_ResPubl	pers-publ-roles
link	class-12	is author (percentage) of		cfPers_ResPubl	pers-publ-roles
link	class-13	is editor (numbered) of		cfPers_ResPubl	pers-publ-roles
link	class-14	is editor of		cfPers_ResPubl	pers-publ-roles
link	class-15	is reviewer of		cfPers_ResPubl	pers-publ-roles
	...				
link	class-20	isSynonymOf	synonym reference for thesauri	cfClass_Class	thesaurus-structure
link	class-21	isBroaderTerm		cfClass_Class	thesaurus-structure
link	class-22	isNarrowerTerm		cfClass_Class	thesaurus-structure

2.8 Additional Features

The current CERIF release contains Dublin Core and Formalised Dublin Core entities and their attributes. With future releases we aim at a mapping from CERIF to Dublin Core, rather than keeping the Dublin Core elements within the physical model. The PersonName entity is currently also categorized as an additional feature, as it does not exactly fit into the conceptual structure otherwise.

3. CERIF-based SQL scripts

From the ERM model in Toad Data Modeler, SQL scripts are generated automatically for most common databases. Some examples extracts are shown in the extracts 19, 20, 21, 22.

```

Create table [cfPersName] (
    [cfPersId] Nchar(128) NOT NULL,
    [cfFamilyNames] Nchar(64) NULL,
    [cfFirstNames] Nchar(64) NULL,
    [cfOtherNames] Nchar(64) NULL,
    Primary Key ([cfPersId])
)

```

Extract 19: SQL Extract for MS SQL7 database

```

Create table "cfPersName" (
    "cfPersId" NChar(128) NOT NULL ,
    "cfFamilyNames" NChar(64),
    "cfFirstNames" NChar(64),
    "cfOtherNames" NChar(64),
    primary key ("cfPersId")
)

```

Extract 20: SQL Extract for Oracle9i database

```

Create table "cfPersName" (
    "cfPersId" Char(128) NOT NULL,
    "cfFamilyNames" Char(64),
    "cfFirstNames" Char(64),
    "cfOtherNames" Char(64),
)

```

Extract 21: SQL Extract for DB2 UDB v.8

```

Create table `cfPers` (
    `cfPersId` Char(128) NOT NULL,
    `Birthdate` Date,
    `cfSex` Char(1),
    `cfURI` Char(128),
    UNIQUE (`cfURI`),
    Primary Key (`cfPersId`)) ENGINE = MyISAM/

```

Extract 22: SQL Extract for mySQL

4. CERIF XML

The CERIF 2008 1.1 – XML: Specification document [11] specifies the interchange of CERIF data in CERIF XML format. The specification document as well as the XML schema [10] files for the validation of CERIF XML files are available for download from the public euroCRIS website: <http://www.euroCRIS.org/>. The XML specification maps to the physical level of the CERIF 2008-1.1 FDM model and is being updated according to CERIF model updates.

The following examples show some CERIFXML representations of some link entity records including semantic references.

```
<cfPers_ResPubl>
  <cfPersId>person-brigitte-joerg</cfPersId>
  <cfResPublId>publication-analytic-information-service-era</cfResPublId>
  <cfClassId>person-isAuthorOf</cfClassId>
  <cfClassSchemeld>pers-publ-roles</cfClassSchemeld>
  <cfStartDate>2009-01-13T00:00:00-00:00</cfStartDate>
  <cfEndDate>2099-01-13T00:00:00-00:00</cfEndDate>
</cfPers_ResPubl>
```

Example 1: CERIF XML Person - Publication Relationship (add cfFraction examples)

```
<cfPers_OrgUnit>
  <cfPersId>person-brigitte-joerg</cfPersId>
  <cfOrgUnitId>organisation-dfki</cfOrgUnitId>
  <cfClassId>class-subaffiliatedWith</cfClassId>
  <cfClassSchemeld>pers-orgunit-roles</cfClassSchemeld>
  <cfStartDate>2009-01-13T00:00:00-00:00</cfStartDate>
  <cfEndDate>2099-01-13T00:00:00-00:00</cfEndDate>
</cfPers_OrgUnit>
```

Example 2: CERIF XML Person -Organisation Relationship

```
<cfClass>
  <cfClassId>class-is-a</cfClassId>
  <cfClassSchemeld>class-scheme-tax-structure</cfClassSchemeld>
  <cfStartDate>2007-09-28T00:00:00-00:00</cfStartDate>
  <cfEndDate>2099-12-31T00:00:00-00:00</cfEndDate>
</cfClass>
<cfClass_Class>
  <cfClassId1>class-information-science</cfClassId1>
  <cfClassId2>class-science</cfClassId2>
  <cfClassSchemeld1>class-scheme-science-tax</cfClassSchemeld1>
  <cfClassSchemeld2>class-scheme-science-tax</cfClassSchemeld2>
  <cfClassId>class-is-a</cfClassId>
  <cfClassSchemeld>class-scheme-tax-structure</cfClassSchemeld>
  <cfStartDate>2007-09-28T00:00:00-00:00</cfStartDate>
  <cfEndDate>2099-12-31T00:00:00-00:00</cfEndDate>
</cfClass_Class>
```

Example 3: CERIF XML Classification Relationship

With CERIF, multiple classification terms and structures can be maintained in parallel and easily identified as semantically different due to their classification scheme assignments. Furthermore, it is possible to map terms across classification schemes like in example 4.

```
<cfClass>
  <cfClassId>class-mappes-to</cfClassId>
  <cfClassSchemeId>class-scheme-CERIF-DC-mapping</cfClassSchemeId>
  <cfStartDate>2007-09-28T00:00:00-00:00</cfStartDate>
  <cfEndDate>2099-12-31T00:00:00-00:00</cfEndDate>
</cfClass>

<cfClass_Class>
  <cfClassId1>class-isAuthorOf</cfClassId1>
  <cfClassId2>class-Creator</cfClassId2>
  <cfClassSchemeId1>class-scheme-CERIF2008</cfClassSchemeId1>
  <cfClassSchemeId2>class-scheme-DC</cfClassSchemeId2>
  <cfClassId>class-mappes-to</cfClassId>
  <cfClassSchemeId>class-scheme-CERIF-DC-Mapping</cfClassSchemeId>
  <cfStartDate>2007-09-28T00:00:00-00:00</cfStartDate>
  <cfEndDate>2099-12-31T00:00:00-00:00</cfEndDate>
</cfClass_Class>
```

Example 4: CERIF XML Classification Mapping

5. CERIF Semantics

The structure and strength of the Semantic Layer as part of the CERIF model has been presented. In close cooperation with the CERIF Best Practice task group some classification schemes will be prepared. With the current CERIF 2008 – 1.1 release, the CERIF publication types and roles have been introduced as CERIF Semantics [12].

6. CERIF Extensions

Contributions, thoughts, error reports or bug reports are very welcome. Incoming feedback will first be discussed within the CERIF task group and subsequently presented to members. A decision towards extension will finally be taken and the CERIF model will be updated accordingly.

7. Next Steps

For the next upcoming release CERIF 2008 – 1.2, we focus on the definition of a CERIF Core. Further upgrades will focus on the context of research funding. More work on proper namespaces is being considered for the CERIF XML specifications in the longer term. The development of a CERIF ontology is foreseen in parallel with the definition of a CERIF Core. A CERIF Core will allow for a better description and handling of CERIF extensions.

8. Appendix

8.1 List of CERIF Entities

Following is a full list of the CERIF entities in alphabetic order, grouped by entity type, giving the Logical and Physical Name of entities in parentheses.

8.1.1 CERIF Core Entities (Logical (PhysicalName))

cfProject (cfProj)
cfPerson (cfPers)
cfOrgUnit (cfOrgUnit)

8.1.2 CERIF Result Entities (Logical (PhysicalName))

cfResultPublication (cfResPubl)
cfResultPatent (cfResPat)
cfResultProduct (cfResProd)

8.1.3 CERIF 2nd Level Entities (Logical (PhysicalName))

cfCall (cfCall)
cfCitation (cfCite)
cfCountry (cfCountry)
cfCurrency (cfCurrency)
cfCurriculumVitae (cfCV)
cfElectronicAddress (cfEAddr)
cfEquipment (cfEquip)
cfEvent (cfEvent)
cfExpertiseAndSkills (cfExpSkills)
cfFacility (cfFacil)
cfFunding (cfFund)
cfGrant (cfGrant)
cfLanguage (cfLanguage)
cfMetrics (cfMetrics)
cfPostalAddress (cfPAddr)
cfPrizeAward (cfPrize)
cfPublicationReference (cfPublRef)
cfQualification (cfQQual)
cfService (cfSrv)

8.1.4 CERIF Link Entities (Logical (PhysicalName))

cfCall_FundingProgramme (cfCall_FundProg)
cfCitation_Classification (cfCite_Class)
cfClassification_Classification (cfClass_Class)
cfClassScheme_ClassScheme (cfClassScheme_ClassScheme)
cfCountry_Classification (cfCountry_Class)
cfCurrency_Classification (cfCurrency_Class)
cfCV_Classification (cfCV_Class)
cfElectronicAddress_Classification (cfEAddr_Class)
cfEquipment_Classification (cfEquip_Class)
cfEquipment_FundingProgramme (cfEquip_FundProg)
cfEvent_Event
cfEvent_Classification (cfEvent_Class)
cfEvent_FundingProgramme (cfEvent_FundProg)
cfEvent_ResultPublication (cfEvent_ResPubl)

cfExpertiseAndSkills_Classification (cfExpSkills_Class)
cfFacility_Classification (cfFacil_Class)
cfFacility_FundingProgramme (cfFacil_FundProg)
cfFundingProgramme_Classification (cfFundProg_Class)
cfFundingProgramme_FundingProgramme (cfFundProg_FundProg)
cfGrant_FundingProgramme (cfGrant_FundProg)
cfLanguage_Classification (cfLanguage_Class)
cfMetrics_Classification (cfMetrics_Class)
cfOrganisationUnit_Classification (cfOrgUnit_Class)
cfOrganisationUnit_DublinCore (cfOrgUnit_DC)
cfOrganisationUnit_ElectronicAddress (cfOrgUnit_EAddr)
cfOrganisationUnit_Equipment (cfOrgUnit_Equip)
cfOrganisationUnit_Event (cfOrgUnit_Event)
cfOrganisationUnit_ExpertiseAndSkills (cfOrgUnit_ExpSkills)
cfOrganisationUnit_Facility (cfOrgUnit_Facil)
cfOrganisationUnit_FundingProgramme (cfOrgUnit_FundProg)
cfOrganisationUnit_OrgUnit (cfOrgUnit_OrgUnit)
cfOrganisationUnit_PostalAddress (cfOrgUnit_PAddr)
cfOrganisationUnit_PrizeAward (cfOrgUnit_Prize)
cfOrganisationUnit_ResultPatent (cfOrgUnit_ResPat)
cfOrganisationUnit_ResultProduct (cfOrgUnit_ResProd)
cfOrganisationUnit_ResultPublication (cfOrgUnit_ResPubl)
cfOrganisationUnit_Service (cfOrgUnit_Srv)
cfPerson_Classification (cfPers_Class)
cfPerson_CV (cfPers_CV)
cfPerson_DublinCore (cfPers_DC)
cfPerson_ElectronicAddress (cfPers_EAddr)
cfPerson_Equipment (cfPers_Equip)
cfPerson_Event (cfPers_Event)
cfPerson_ExpertiseAndSkills (cfPers_ExpSkills)
cfPerson_Facility (cfPers_Facil)
cfPerson_FundingProgramme (cfPers_FundProg)
cfPerson_Language (cfPers_Language)
cfPerson_Country (cfPers_Country)
cfPerson_OrganisationUnit (cfPers_OrgUnit)
cfPerson_Person (cfPers_Pers)
cfPerson_PostAddress (cfPers_PAddr)
cfPerson_PrizeAward (cfPers_Prize)
cfPerson_Qualification (cfPers_Qual)
cfPerson_ResultPatent (cfPers_ResPat)
cfPerson_ResultProduct (cfPers_ResProd)
cfPerson_ResultPublication (cfPers_ResPubl)
cfPerson_Service (cfPers_Srv)
cfPersonName_Person (cfPersName_Pers)
cfPostAddress_Classification (cfPAddr_Class)
cfProject_Classification (cfProj_Class)
cfProject_DublinCore (cfProj_DC)
cfProject_Equipment (cfProj_Equip)
cfProject_Event (cfProj_Event)
cfProject_Facility (cfProj_Facil)
cfProject_FundingProgramme (cfProj_FundProg)
cfProject_OrganisationUnit (cfProj_Orgunit)
cfProject_Person (cfProj_Pers)
cfProject_PrizeAward (cfProj_Prize)
cfProject_Project (cfProj_Proj)
cfProject_Service (cfProj_Srv)
cfProject_ResultPatent (cfProj_ResPat)
cfProject_ResultProduct (cfProj_ResProd)

cfProject_ResultPublication (cfProj_ResPubl)
cfResultPatent_Classification (cfResPat_Class)
cfResultPatent_FundingProgramme (cfResPat_FundProg)
cfResultPatent_ResultPatent
cfResultProduct_Classification (cfResProd_Class)
cfResultProduct_Funding (cfResProd_FundProg)
cfResultProduct_ResultProduct
cfResultPublication_Citation (cfResPubl_Cite)
cfResultPublication_Classification (cfResPubl_Class)
cfResultPublication_DublinCore (cfResPubl_DC)
cfResultPublication_Event (cfResPubl_Event)
cfResultPublication_Equipment (cfResPubl_Equip)
cfResultPublication_Facility (cfResPubl_Facil)
cfResultPublication_FundingProgramme (cfResPubl_FundProg)
cfResultPublication_Metrics (cfResPubl_Metrics)
cfResultPublication_ResultPatent (cfResPubl_ResPat)
cfResultPublication_ResultProduct (cfResPubl_ResProd)
cfResultPublication_ResultPublication (cfResPubl_ResPubl)
cfService_Classification (cfSrv_Class)
cfService_FundingProgramme (cfSrv_FundProg)

8.1.5 CERIF Multiple Language Features (Logical (PhysicalName))

cfCallName (cfCallName)
cfCallDescription (cfCallDescr)
cfCallKeywords (cfCallKeyw)
cfCitationDescription (cfCiteDescr)
cfCitationTitle (cfCiteTitle)
cfClassificationDescription (cfClassDescr)
cfClassificationTerm (cfClassTerm)
cfClassificationSchemeDescription (cfClassSchemeDescr)
cfCountryName (cfCountryName)
cfCurrencyEntityName (cfCurrencyEntityName)
cfCurrencyName (cfCurrencyName)
cfEquipmentDescription (cfEquipPDescr)
cfEquipmentKeywords (cfEquipKeyw)
cfEquipmentName (cfEquipName)
cfEventDescription (cfEventDescr)
cfEventKeywords (cfEventKeyw)
cfEventName (cfEventName)
cfExpertiseAndSkillsDescription (cfExpSkillsDescr)
cfExpertiseAndSkillsKeywords (cfExpSillsKeyw)
cfExpertiseAndSkillsName (cfExpSkillsName)
cfFacilityDescription (cfFacilDescr)
cfFacilityKeywords (cfFacilKeyw)
cfFacilityName (cfFacilName)
cfFundingProgrammeDescription (cfFundProgDescr)
cfFundingProgrammeKeywords (cfFundProgKeyw)
cfFundingProgrammeName (cfFundProgName)
cfGrantName (cfGrantName)
cfGrantDescription (cfGrantDescr)
cfGrantKeywords (cfGrantKeyw)
cfLanguageName (cfLanguageName)
cfMetricsDescription (cfMetricsDescr)
cfMetricsName (cfMetricsName)
cfOrganisationUnitKeywords (cfOrgUnitKeyw)
cfOrganisationUnitName (cfOrgUnitName)
cfOrganisationUnitResearchActivity (cfOrgUnitResAct)

cfPersonResearchInterest (cfPersResInt)
cfPersonKeywords (cfPersKeyw)
cfProjectAbstract (cfProjAbstr)
cfProjectKeywords (cfProjKeyw)
cfProjectTitle (cfProjTitle)
cfResultPatentAbstract (cfResPatAbstr)
cfResultPatentKeywords (cfResPatKeyw)
cfResultPatentTitle (cfResPatTitle)
cfResultProductDescription (cfResProdDescr)
cfResultProductKeywords (cfResProdKeyw)
cfResultProductName (cfResProdName)
cfResultPublicationAbstract (cfResPublAbst)
cfResultPublicationBibliographicNote (cfResPublBiblNote)
cfResultPublicationKeywords (cfResPublKeyw)
cfResultPublicationNameAbbreviation (cfResPublNameAbbrev)
cfResultPublicationSubtitle (cfResPublSubtitle)
cfResultPublicationTitle (cfResPublTitle)
cfServiceDescription (cfSrvDescr)
cfServiceKeywords (cfSrvKeyw)
cfServiceName (cfSrvName)

8.1.6 *Additional Entities (Logical (PhysicalName))*

cfPersonName (cfPersName)
cfDublinCore (cfDC)
cfDCAudience (cfDCAudience)
cfDCContributor (cfDCContributor)
cfDCCoverage (cfDCCoverage)
cfDCCoverageSpatial (cfDCCoverageSpatial)
cfDCCoverateTemporal (cfDCCoverageTemporal)
cfDCCreator (cfDCCreator)
cfDCDate (cfDCDate)
cfDCDescription (cfDCDescription)
cfDCFormat (cfDCFormat)
cfDCLanguage (cfDCLanguage)
cfDCProvenance (cfDCProvenance)
cfDCPublisher (cfDCPublisher)
cfDCRelation (cfDCRelation)
cfDCResourceIdentifier (cfDCResourceIdentifier)
cfDCResourceType (cfDCResourceType)
cfDCRightsHolder (cfDCRightsHolder)
cfDCRightsManagement (cfDCRightsMM)
cfDCRightsManagementAccessRights (cfDCRightsMMAccessRight)
cfDCRightsManagementLicense (cfDCRightsMMLicence)
cfDCSource (cfDCSource)
cfDCSubject (cfDCSubject)
cfDCTitle (cfDCTitle)
cfFormalisedDublinCoreRightsManagementPricing (FDCRightsMMPricing)
cfFormalisedDublinCoreRightsManagementPrivacy (FDCRightsMMPrivacy)
cfFormalisedDublinCoreRightsManagementRights (FDCRightsMM)
cfFormalisedDublinCoreRightsManagementSecurity (FDCRightsMMSecurity)

8.1.7 *CERIF Classification Entities (Logical (PhysicalName))*

cfClassification (cfClass)
cfClassificationScheme (cfClassScheme)

8.1.8 CERIF Attributes

8.1.9 Attribute in all Link Tables

cfFraction (cfFraction)

8.1.9.1 Language-dependent attributes including cflangCode and cfTrans

cfAbstract (cfAbstr)

cfDescription (cfDescr)

cfKeywords (cfKeyw)

cfName (cfName)

cfResearchActivity (cfResAct)

cfResearchInterest (cfResInt)

cfTerm (cfTerm)

cfTitle (cfTitle)

8.1.9.2 Currency-dependent attributes

cfBudget (cfBudget)

cfAmount (cfAmount)

cfPrice (cfPrice)

cfTurnover (cfTurn)

8.2 Logical / Physical CERIF Entity Names

The following table 1 gives an overview of all CERIF 2008 – 1.0 entities, their corresponding attributes with logical and physical names (including cf prefixes).

Table 1: List of Entities with Logical (alphabetical order) and Physical Names

Logical CERIF2008 - 1.1 Entities	Physical CERIF2008-1.1 Entities
cfCall	cfCall
cfCall_FundingProgramme	cfCall_FundProg
cfCallDescription	cfCallDescr
cfCallKeywords	cfCallKeyw
cfCallName	cfCallName
cfCitation	cfCite
cfCitation_Classification	cfCite_Class
cfCitationDescription	cfCiteDescr
cfCitationTitle	cfCiteTitle
cfClassification	cfClass
cfClassification_Classification	cfClass_Class
cfClassificationDescription	cfClassDescr
cfClassificationScheme	cfClassScheme
cfClassificationScheme_ClassificationScheme	cfClassScheme_ClassScheme
cfClassificationSchemeDescription	cfClassSchemeDescr
cfClassificationTerm	cfClassTerm
cfCountry	cfCountry
cfCountry_Classification	cfCountry_Class
cfCountryName	cfCountryName
cfCurrency	cfCurrency
cfCurrency_Classification	cfCurrency_Class
cfCurrencyEntityName	cfCurrencyEntName
cfCurrencyName	cfCurrencyName
cfCurriculumVitae	cfCV
cfCurriculumVitae_Classification	cfCV_Class
cfDublinCore	cfDC
cfDublinCoreAudience	cfDCAudience
cfDublinCoreContributor	cfDCContributor
cfDublinCoreCoverage	cfDCCoverage
cfDublinCoreCoverageSpatial	cfDCCoverageSpatial
cfDublinCoreCoverageTemporal	cfDCCoverageTemporal
cfDublinCoreCreator	cfDCCreator
cfDublinCoreDate	cfDCDate
cfDublinCoreDescription	cfDCDescription
cfDublinCoreFormat	cfDCFormat
cfDublinCoreLanguage	cfDCLanguage
cfDublinCoreProvenance	cfDCProvenance
cfDublinCorePublisher	cfDCPublisher
cfDublinCoreRelation	cfDCRelation
cfDublinCoreResourceIdentifier	cfDCResourceIdentifier
cfDublinCoreResourceType	cfDCResourceType
cfDublinCoreRightsHolder	cfDCRightsHolder
cfDublinCoreRightsManagement	cfDCRightsMM
cfDublinCoreRightsManagementAccessRights	cfDCRightsMMAccessRights
cfDublinCoreRightsManagementLicense	cfDCRightsMMLicense
cfDublinCoreSource	cfDCSource
cfDublinCoreSubject	cfDCSubject
cfDublinCoreTitle	cfDCTitle
cfElectronicAddress	cfEAddr

cfElectronicAddress_Classification	cfEAddr_Class
cfEquipment	cfEquip
cfEquipment_Classification	cfEquip_Class
cfEquipment_Funding	cfEquip_Fund
cfEquipmentDescription	cfEquipDescr
cfEquipmentKeywords	cfEquipKeyw
cfEquipmentName	cfEquipName
cfEvent	cfEvent
cfEvent_Classification	cfEvent_Class
cfEvent_Event	cfEvent_Event
cfEvent_Funding	cfEvent_Fund
cfEvent_ResultPublication	cfEvent_ResPubl
cfEventDescription	cfEventDescr
cfEventKeywords	cfEventKeyw
cfEventName	cfEventName
cfExpertiseAndSkills	cfExpSkills
cfExpertiseAndSkills_Classification	cfExpSkills_Class
cfExpertiseAndSkillsDescription	cfExpSkillsDescr
cfExpertiseAndSkillsKeywords	cfExpSkillsKeyw
cfExpertiseAndSkillsName	cfExpSkillsName
cfFacility	cfFacil
cfFacility_Classification	cfFacil_Class
cfFacility_Funding	cfFacil_Fund
cfFacilityDescription	cfFacilDescr
cfFacilityKewords	cfFacilKeyw
cfFacilityName	cfFacilName
cfFormalisedDublinCoreRightsManagementPricing	cfFDCRightsMMPricing
cfFormalisedDublinCoreRightsManagementPrivacy	cfFDCRightsMMPrivacy
cfFormalisedDublinCoreRightsManagementRights	cfFDCRightsMMRights
cfFormalisedDublinCoreRightsManagementSecurity	cfFDCRightsMMSecurity
cfFundingProgramme	cfFundProg
cfFundingProgramme_Classification	cfFundProg_Class
cfFundingProgramme_FundingProgramme	cfFundProg_FundProg
cfFundingProgrammeDescription	cfFundProgDescr
cfFundingProgrammeKeywords	cfFundProgKeyw
cfFundingProgrammeName	cfFundProgName
cfGrant	cfGrant
cfGrant_FundingProgramme	cfGrant_FundProg
cfGrantDescription	cfGrantDescr
cfGrantKeywords	cfGrantKeyw
cfGrantName	cfGrantName
cfLanguage	cfLang
cfLanguage_Classification	cfLang_Class
cfLanguageName	cfLangName
cfMetrics	cfMetrics
cfMetrics_Classification	cfMetrics_Class
cfMetricsDescription	cfMetricsDescr
cfMetricsName	cfMetricsName
cfOrganisationUnit	cfOrgUnit
cfOrganisationUnit_Classification	cfOrgUnit_Class
cfOrganisationUnit_DublinCore	cfOrgUnit_DC
cfOrganisationUnit_ElectronicAddress	cfOrgUnit_EAddr
cfOrganisationUnit_Equipment	cfOrgUnit_Equip
cfOrganisationUnit_Event	cfOrgUnit_Event
cfOrganisationUnit_ExpertiseAndSkills	cfOrgUnit_ExpSkills
cfOrganisationUnit_Facility	cfOrgUnit_Facil
cfOrganisationUnit_FundingProgramme	cfOrgUnit_FundProg
cfOrganisationUnit_OrganisationUnit	cfOrgUnit_OrgUnit

cfOrganisationUnit_PostAddress	cfOrgUnit_PAddr
cfOrganisationUnit_PrizeAward	cfOrgUnit_Prize
cfOrganisationUnit_ResultPatent	cfOrgUnit_ResPat
cfOrganisationUnit_ResultProduct	cfOrgUnit_ResProd
cfOrganisationUnit_ResultPublication	cfOrgUnit_ResPubl
cfOrganisationUnit_Service	cfOrgUnit_Srv
cfOrganisationUnitKeywords	cfOrgUnitKeyw
cfOrganisationUnitName	cfOrgUnitName
cfOrganisationUnitResearchActivity	cfOrgUnitResAct
cfPerson	cfPers
cfPerson_Classification	cfPers_Class
cfPerson_Country	cfPers_Country
cfPerson_CurriculumVitae	cfPers_CV
cfPerson_DublinCore	cfPers_DC
cfPerson_ElectronicAddress	cfPers_EAddr
cfPerson_Equipment	cfPers_Equip
cfPerson_Event	cfPers_Event
cfPerson_ExpertiseAndSkills	cfPers_ExpSkills
cfPerson_Facility	cfPers_Facil
cfPerson_FundingProgramme	cfPers_FundProg
cfPerson_Language	cfPers_Language
cfPerson_OrganisationUnit	cfPers_OrgUnit
cfPerson_Person	cfPers_Pers
cfPerson_PostAddress	cfPers_PAddr
cfPerson_PrizeAward	cfPers_Prize
cfPerson_Qualification	cfPers_Qual
cfPerson_ResultPatent	cfPers_ResPat
cfPerson_ResultProduct	cfPers_ResProd
cfPerson_ResultPublication	cfPers_ResPubl
cfPerson_Service	cfPers_Serv
cfPersonKeywords	cfPersKeyw
cfPersonName	cfPersName
cfPersonName_Person	cfPersName_Pers
cfPersonResearchInterest	cfPersResInt
cfPostAddress	cfPAddr
cfPostAddress_Classification	cfPAddr_Class
cfPrizeAward	cfPrize
cfPrizeAward_Classification	cfPrize_Class
cfProject	cfProj
cfProject_Classification	cfProj_Class
cfProject_DublinCore	cfProj_DC
cfProject_Equipment	cfProj_Equip
cfProject_Event	cfProj_Event
cfProject_Facility	cfProj_Facil
cfProject_FundingProgramme	cfProj_FundProg
cfProject_OrganisationUnit	cfProj_OrgUnit
cfProject_Person	cfProj_Pers
cfProject_PrizeAward	cfProj_Prize
cfProject_Project	cfProj_Proj
cfProject_ResultPatent	cfProj_ResPat
cfProject_ResultProduct	cfProj_ResProd
cfProject_ResultPublication	cfProj_ResPubl
cfProject_Service	cfProj_Srv
cfProjectAbstract	cfProjAbstr
cfProjectKeywords	cfProjKeyw
cfProjectTitle	cfProjTitle
cfPublicationReference	cfPublRef
cfQualification	cfQual

cfQualification_Classification	cfQual_Class
cfQualificationDescription	cfQualDescr
cfQualificationKeywords	cfQualKeyw
cfResultPatent	cfResPat
cfResultPatent_Classification	cfResPat_Class
cfResultPatent_FundingProgramme	cfResPat_FundProg
cfResultPatent_ResultPatent	cfResPat_ResPat
cfResultPatentAbstract	cfResPatAbstr
cfResultPatentKeywords	cfResPatKeyw
cfResultPatentTitle	cfResPatTitle
cfResultProduct	cfResProd
cfResultProduct_Classification	cfResProd_Class
cfResultProduct_FundingProgramme	cfResProd_FundProg
cfResultProduct_ResultProduct	cfResProd_ResProd
cfResultProductDescription	cfResProdDescr
cfResultProductKeywords	cfResProdKeyw
cfResultProductName	cfResProdName
cfResultPublication	cfResPubl
cfResultPublication_Citation	cfResPubl_Cite
cfResultPublication_Classification	cfResPubl_Class
cfResultPublication_DublinCore	cfResPubl_DC
cfResultPublication_FundingProgramme	cfResPubl_FundProg
cfResultPublication_Equipment	cfResPubl_Equip
cfResultPublication_Event	cfResPubl_Event
cfResPubl_Facility	cfResPubl_Facil
cfResPubl_FundingProgramme	cfResPubl_FundProg
cfResPubl_Metrics	cfResPubl_Metrics
cfResPubl_ResultPatent	cfResPubl_ResPat
cfResPubl_ResultProduct	cfResPubl_ResProd
cfResultPublication_ResultPublication	cfResPubl_ResPubl
cfResultPublicationAbstract	cfResPublAbstr
cfResultPublicationBibliographicNote	cfResPublBiblNote
cfResultPublicationKeywords	cfResPublKeyw
cfResultPublicationNameAbbreviation	cfResPublNameAbbrev
cfResultPublicationSubtitle	cfResPublSubtitle
cfResultPublicationTitle	cfResPublTitle
cfService	cfSrv
cfService_Classification	cfSrv_Class
cfService_FundingProgramme	cfSrv_FundProg
cfServiceDescription	cfSrvDescr
cfServiceKeywords	cfSrvKeyw
cfServiceName	cfSrvName

9. References

- [1] Asserson, A.; Jeffery, K.G.; Lopatenko, A. (2002): CERIF: Past, Present and Future: An Overview. In Proceedings: Gaining Insight from Research Information. 6th International Conference on Current Research Information Systems, Kassel, Germany.
- [2] EuroHORCS (2008): Window to Science: Information Systems of European Research Organisations. Report of the EUROHORCS – ESF Working Group on a Joint Research Information System. Strasbourg: European Science Foundation. Report Editing: Alexis-Michel Mugabushaka. ISBN: 2-912049-86-5, October 2008
- [3] Gartner, R. (2008): Metadata for the digital libraries: state of the art and future directions (1.0). Peer reviewed report from the JISC Technology and Standards Watch. April 2008, Bristol, UK
- [4] Jeffery, K.G.; Asserson, A. (2006): CRIS: Central Relating Information System. In: Asserson, A. and Simons, Ed. (eds). Enabling Interaction and Quality: Beyond the Hanseatic League. 8th International Conference on Current Research Information Systems, May 2006, Bergen Norway, Leuven: Leuven University Press, 109-119.
- [5] Jeffery, K.G.; Asserson, A.; Lopatenko, A. S. (2002): Comparative Study of Metadata for Scientific Information: The place of CERIF in CRISs and Scientific Repositories. Gaining Insight from Research Information. 6th International Conference on Current Research Information Systems, Kassel, Germany.
- [6] Storey, V.C. (1993): Understanding Semantic Relationships. The International Journal on Very Large Databases (VLDB). Volume 2, Number 4, October 1993, pages 458-488, Springer Berlin/Heidelberg.
- [7] UNISIST (1971): Study Report on the Feasibility of a World Science Information System. 171 pages, UNIPUB Inc., P.O. Box 433, New York, N.Y.
- [8] Wang, R.Y.; Storey, V.C.; Weber, R.; (1999): An ontological analysis of the relationship construct in conceptual modeling. ACM Transactions on Database Systems (TODS) Journal, Vol. 24, Is-sue 4, December 1999, pages 494-528. New York USA.
- [9] W3C Recommendation: Extensible Markup Language (XML) 1.0, Fourth Edition, 16 August 2006, edited in place, 29 September 2006. <http://www.w3.org/TR/2006/Rec-xml-2006-08-16/>
- [10] W3C XML Schema: <http://www.w3.org/XML/Schema>
- [11] Jörg, B.; van Grootel, G.; Jeffery, K.G.; (2010): CERIF 2008 – 1.1 XML: Data Exchange Format Specification. euroCRIS, March 2010.
- [12] Jörg, B.; Jeffery, K.G.; Asserson, A.; van Grootel, G.; Rasmussen, H.; Price, A.; Vestam, T.; Elbæk, M.K.; Houssos, N.; Voigt, R.; Simons, E.J.; Dvorak, J.; (2010): CERIF 2008 – 1.1 Semantics. euroCRIS, March 2010.
- [13] ISO Country Codes: http://www.iso.org/iso/country_codes.htm
- [14] ISO Language Codes: http://www.iso.org/iso/language_codes.htm
- [15] ISO Currency Codes: http://www.iso.org/iso/currency_codes.htm