

CERIF 1.5 XML

Data Exchange Format Specification



Editors

Jan Dvořák InfoScience Praha, CZ
Faculty of Arts, Charles University in Prague, CZ
Brigitte Jörg UKOLN, University of Bath, UK

Contributors

Thomas Vestdam Atira, DK (acquired by Elsevier)

Abstract

CERIF (the Common European Research Information Format) XML is the data exchange format to enable the transport of CERIF data between CRISs (Current Research Information Systems) or between CRISs and non-CRISs. It builds on standardised XML technologies recommended by the W3C. The CERIF XML format has been consolidated and updated in the CERIF 1.4 release and is thus a major upgrade compared to CERIF 1.3. The CERIF 1.5 XML complies with the CERIF 1.5 Full Data Model (FDM).

CERIF is a formal conceptual model to support the management of Research Information, including the set up of and the interoperation between Research Information Systems. 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 the care and custody of CERIF to euroCRIS (www.eurocris.org), a not-for-profit organisation dedicated to the promotion of Current Research Information Systems (CRISs).

Covered XML Namespace

urn:xmlns:org:eurocris:cerif-1.5-1

Locations:

The CERIF XML 1.5 XML Schema:

http://www.eurocris.org/Uploads/Web%20pages/CERIF-1.5/CERIF_1.5_1.xsd

This document:

http://www.eurocris.org/Uploads/Web%20pages/CERIF-1.5/CERIF1.5_XML.pdf

Table of Contents

1	Introduction	3
1.1	Purpose of CERIF XML	3
2	CERIF XML Structure	4
2.1	The Namespace	4
2.2	The XML Schema location	4
2.3	CERIF markup root: The “CERIF” element	4
2.4	Contents of the “CERIF” element	5
2.4.1	Simple CERIF entities	5
2.4.2	Multilingual CERIF entities	5
2.4.3	CERIF link entities	6
2.5	Relationship between the CERIF ER Model and CERIF XML	6
2.6	Backward compatibility	7
2.7	Special cases	8
2.7.1	CERIF XML data types	8
2.7.2	Valid time intervals	8
2.7.3	Financial amounts	8
2.7.4	Federated Identifiers	8
2.7.5	Person Names	9
3	Examples	10
4	Acknowledgements	15
5	Components of the CERIF 1.5 Release	15
6	References	16

1 Introduction

The CERIF Data Exchange Format supports interchange of research information. It is one of the components of CERIF (Common European Research Information Format) [1]. It is based on XML, the widely used format recommended by the W3C [3].

1.1 *Purpose of CERIF XML*

CERIF XML aims at enabling consistent data interchange across systems in the Research Information domain. It is a structured XML format that follows the CERIF model. It aims at supporting the whole spectrum of data interchange scenarios, ranging from one-object-at-a-time messaging up to dumps and bulk loads of entire databases. It is inspired by the CERIF Entity-Relationship model [1] but maintains its own technological characteristics, e.g. embedding of link entities.

2 CERIF XML Structure

This section describes the CERIF XML features and how it links to the CERIF Entity-Relationship model. CERIF XML uses the following technologies: XML itself [3], XML Namespaces [4], and XML Schema [5].

2.1 The Namespace

The CERIF XML 1.5 uses the following XML namespace for its XML elements:

```
urn:xmlns:org:eurocris:cerif-1.5-1
```

In the sequel we have this namespace URI in mind wherever we don't mention a namespace prefix i.e., as if the namespace was declared using the following declaration:

```
xmlns="urn:xmlns:org:eurocris:cerif-1.5-1"
```

Additionally, we will refer to the `xs` namespace prefix as referring to the XML Schema namespace itself i.e., as if the following XML namespace declaration was in place:

```
xmlns:xs="http://www.w3.org/2001/XMLSchema"
```

2.2 The XML Schema location

The primary location of the XML Schema for the CERIF XML 1.5 namespace is:

```
http://www.eurocris.org/Uploads/Web%20pages/CERIF-1.5/CERIF_1.5_1.xsd
```

Validating XML processors can be pointed to this location by including the following XML attributes in the “CERIF” element (see below) or one of its ancestors:

```
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="urn:xmlns:org:eurocris:cerif-1.5-1
http://www.eurocris.org/Uploads/Web%20pages/CERIF-1.5/CERIF_1.5_1.xsd"
```

CERIF XML processors should not, however, rely on this association being present in all CERIF XML files. The mapping between the namespace URI and the XML Schema URL should be made part of the XML parser configuration.¹

2.3 CERIF markup root: The “CERIF” element

The CERIF data is represented as descendants of a single XML element, “CERIF”. This element encloses the CERIF XML message. This element supports the following attributes:

<i>Attribute name</i>	<i>Data type</i>	<i>Mandatory?</i>	<i>Meaning</i>
release	xs:string	yes	1.5 (fixed value) – the CERIF XML release
date	xs:date	yes	the date of production of the XML message
sourceDatabase	xs:string	yes	an identification of the source of the XML message

¹ Such as the <http://apache.org/xml/properties/schema/external-schemaLocation> settable property for the Apache Xerces parser (<http://xerces.apache.org/xerces2-j/>).

The trivial XML message reads:

```
<?xml version="1.0" encoding="UTF-8"?>
<CERIF xmlns="urn:xmlns:org:eurocris:cerif-1.5-1"
  xsi:schemaLocation="urn:xmlns:org:eurocris:cerif-1.5-1
    http://www.eurocris.org/Uploads/Web%20pages/CERIF-1.5/CERIF_1.5_1.xsd"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  release="1.5" date="2012-07-11" sourceDatabase="empty">
</CERIF>
```

2.4 Contents of the “CERIF” element

The top level of CERIF markup can contain data from any CERIF entity, in any number and any order.

```
<CERIF xmlns="urn:xmlns:org:eurocris:cerif-1.5-1"
  release="1.5" date="2012-07-11" sourceDatabase="example">  <!-- ... -->
  <cfOrgUnit/>  <!-- ... -->
  <cfPers/>  <!-- ... -->
  <cfProj/>  <!-- ... -->
  <cfResProd/>  <!-- ... -->
</CERIF>
```

2.4.1 Simple CERIF entities

An instance of a CERIF entity is represented by an enclosing element whose name is the physical name of the CERIF entity in the CERIF Full Data Model. E.g. the data of a CERIF `cfService` instance is enclosed in the `<cfSrv> ... </cfSrv>` pair of tags, the data of a CERIF `cfOrgUnit` instance is enclosed in the `<cfOrgUnit> ... </cfOrgUnit>` pair of tags.

Such an instance XML element can contain the attributes of the entity, again mapped as XML elements named after the physical names of the corresponding attributes. E.g.

```
<cfOrgUnit>
  <cfOrgUnitId>organisation-1-id</cfOrgUnitId>
  <cfAcro>euroCRIS</cfAcro>
</cfOrgUnit>
```

The values of the ERM attribute shall be given in the order they are listed in the CERIF Full Data Model. The (ERM’s) key attributes are mandatory, with a few exceptions (see later). The (ERM’s) non-key attributes are optional, even if the attribute is denoted as NOT NULL in the data model. If an ERM attribute has the value of NULL, the corresponding XML element is not present in the CERIF XML at all.

In addition and inline with the basic elements derived from ERM attributes, one can embed XML markup representing multilingual and linking entities, as described in the rest of this subsection.

2.4.2 Multilingual CERIF entities

CERIF multilingual entities are transformed to XML using a standardized construct: an XML element with multilingual attributes containing the text value itself; i.e. the XML element has additional qualifying XML attributes:

- (1) `cfLangCode` (the code of language and possibly also the country variant of the value),
- (2) `cfTrans` (the translation mode: o – the original language, h – human translation, m – machine translation).

E.g.

```

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

```

2.4.3 CERIF link entities

The XML markup representing a CERIF link entity can be embedded under either end of the relationship they represent. The identifier of the enclosing entity instance is omitted as it is found one level higher in the XML. E.g.

```

<cfOrgUnit>
  <cfOrgUnitId>system-internal-organization-identifier</cfOrgUnitId>
  <cfName cfLang="en" cfTrans="o">organization name</cfTerm>
  <cfProj_OrgUnit>
    <cfProjId>system-internal-project-identifier</cfProjId>
    <cfClassId>classification-uuid</cfClassId>
    <cfClassSchemeld>classification-scheme-uuid</cfClassSchemeld>
    <cfStartDate>date-time when the relationship starts</cfStartDate>
    <cfEndDate>date-time when the relationship ends</cfEndDate>
  </cfProj_OrgUnit>
</cfOrgUnit>

```

Embedding under the other end of the relationship is equally possible, e.g.:

```

<cfProj>
  <cfProjId>system-internal-project-identifier</cfProjId>
  <cfProj_OrgUnit>
    <cfOrgUnitId>system-internal-organization-identifier</cfOrgUnitId>
    <cfClassId>classification-uuid</cfClassId>
    <cfClassSchemeld>classification-scheme-uuid</cfClassSchemeld>
    <cfStartDate>date-time when the relationship starts</cfStartDate>
    <cfEndDate>date-time when the relationship ends</cfEndDate>
  </cfProj_OrgUnit>
</cfProj>

```

Unary link entities (such as `cfProj_Class`) can also be embedded under the base object, e.g.:

```

<cfProj>
  <cfProjId>system-internal-project-identifier</cfProjId>
  <cfProj_Class>
    <cfClassId>classification-uuid</cfClassId>
    <cfClassSchemeld>classification-scheme-uuid</cfClassSchemeld>
    <cfStartDate>date-time when the relationship starts</cfStartDate>
    <cfEndDate>date-time when the relationship ends</cfEndDate>
  </cfProj_Class>
</cfProj>

```

The other end of the relationship as well as the classification² referenced in a link entity shall either be present in the same XML message, or be reachable by some other means for the recipients of the XML message.

2.5 Relationship between the CERIF ER Model and CERIF XML

The XML Schema for the updated CERIF XML is generated from the ER Model representation. This is possible due to its highly regular structure, namely the three basic kinds of CERIF entities: Research (Base, 2nd Level, other), Multilingual, and Link entities. An overview of the relationship for the different entities is given in Figure 1.

² euroCRIS started to develop standard vocabularies and formed strategic partnerships with e.g. CASRAI and VIVO for the re-use of already defined vocabularies. euroCRIS recommends the re-use of existing vocabularies – and hence the identifiers that come with the terms [2].

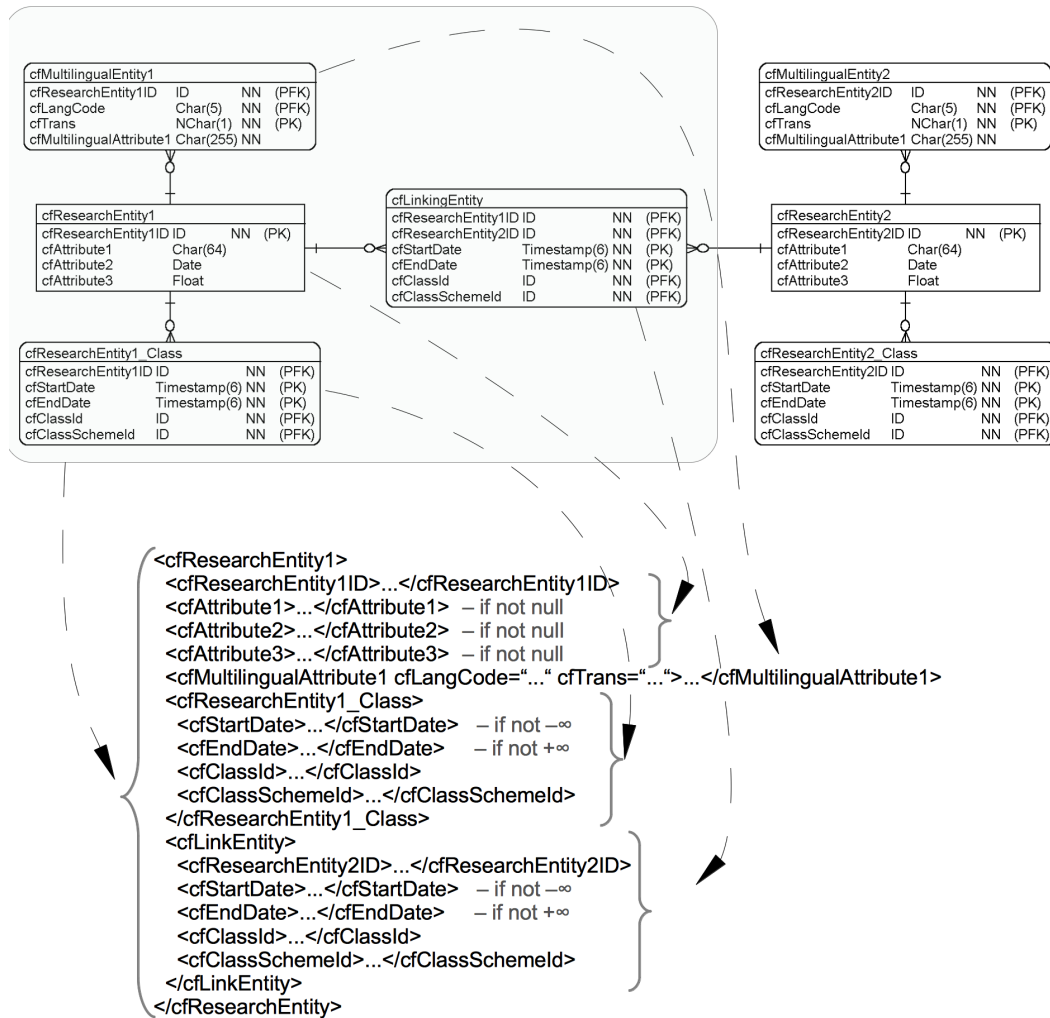


Figure 1: The relational CERIF entities and embedded CERIF XML

2.6 Backward compatibility

Pre-existing CERIF XML is valid with respect to the CERIF 1.5 XML Schema, if the XML namespace is changed to the one of this release.

The embedding style, although encouraged, is not the only one supported. Some data interchange scenarios are indeed optimally served using the original style of one XML file per CERIF entity [6].

All CERIF entities can be represented on the top level (under the “CERIF” element) using the constructs already described. E.g. a multilingual CERIF entity:

```

<cfProjTitle>
  <cfProjId> ... </cfProjId>
  <cfTitle cfLangCode="..." cfTrans="..."> ... </cfTitle>
</cfProjTitle>
  
```

A unary classification link entity:

```

<cfProj_Class>
  <cfProjId> ... </cfProjId>
  <cfClassId> ... </cfClassId>
  <cfClassSchemeld> ... </cfClassSchemeld>
</cfProj_Class>
  
```

A binary link entity:

```
<cfProj_OrgUnit>
  <cfProjId> ... </cfProjId>
  <cfOrgUnitId> ... </cfOrgUnitId>
  <cfClassId> ... </cfClassId>
  <cfClassSchemeld> ... </cfClassSchemeld>
  <cfStartDate> ... </cfStartDate>
</cfProj_OrgUnit>
```

2.7 Special cases

2.7.1 CERIF XML data types

The database types of the CERIF attributes are mapped to XML Schema built-in types. As a rule, the XML Schema does not restrict the maximum lengths of character strings or size and precision of numbers. The only exception is the `cfid_Type` type (corresponds to the ID dictionary type in the ER model): it limits the lengths of the identifiers of CERIF research entities to 128 characters.

2.7.2 Valid time intervals

Valid time intervals of the relationships represented by the CERIF linking entities are expressed using the `cfStartDate` and `cfEndDate` XML elements. Although they are key attributes in the Full Data Model, they need not be present in CERIF XML. An omitted value of `cfStartDate` is to be interpreted as “since time immemorial”. An omitted value of `cfEndDate` means “until things change”.

2.7.3 Financial amounts

A financial amount has two components: the currency unit and the number of these units. The ER attributes `cfAmount`, `cfPrice` or `cfTurn` are represented as XML elements with the accompanying `cfCurrCode` forming an XML attribute on it. The number of the currency units is expressed as a decimal number in the contents of the XML element.

For instance, the price of 30,000 £ is expressed as:

```
<cfPrice cfCurrCode="GBP">30000</cfPrice>
```

2.7.4 Federated Identifiers

Federated identifiers are a new feature in the CERIF 1.5 Full Data Model. They record identifiers under which a CERIF base object is known in other contexts. We refer for more general information on this feature to the CERIF 1.5 FDM specification [1]; here we concentrate on the XML representation.

As with all CERIF entities, federated identifiers can be either recorded stand-alone (as direct children of the ‘CERIF’ element, cf. Section 2.6), or embedded. In the embedded case it is immediately clear to which object the federated identifier is related – it is the enclosing object. E.g.:

```
<cfResPubl>
  <cfResPublId>42</cfResPublId>
  <cfTitle cfLangCode="en" cfTrans="o">KEGGconverter: a tool for the in-silico
    modelling of metabolic networks of the KEGG Pathways database</cfTitle>
  <cfFedId>
    <cfFedIdId>14123451</cfFedIdId>
    <cfFedId>10.1186/1471-2105-10-324</cfFedId>
    <cfStartDate>2009-10-31T24:00:00</cfStartDate>
  <cfFedId_Class>
    <cfClassId>31d222b4-11e0-434b-b5ae-088119c51189</cfClassId><!-- DOI -->
```



```

    <cfClassSchemeld>bccb3266-689d-4740-a039-c96594b4d916</cfClassSchemeld><!-- Identifier Types -->
  </cfFedId_Class>
  <cfFedId_Srv>
    <cfSrvld>5123452</cfSrvld><!-- CrossRef -->
    <cfClassId>eda2b2e2-34c5-11e1-b86c-0800200c9a66</cfClassId><!-- Issuer -->
    <cfClassSchemeld>5a270628-f593-4ff4-a44a-95660c76e182</cfClassSchemeld><!-- Ident Service Roles -->
  </cfFedId_Srv>
</cfFedId>
</cfResPubl>

```

With the stand-alone approach, the identifier has to also contain explicit information about the object the identifier refers to: the CERIF entity and the ID. The above example rewritten as stand-alone (the added XML in bold):

```

<cfResPubl>
  <cfResPublId>42</cfResPublId>
  <cfTitle cfLangCode="en" cfTrans="o">KEGGconverter: a tool for the in-silico
    modelling of metabolic networks of the KEGG Pathways database</cfTitle>
</cfResPubl>
<cfFedId>
  <cfFedIdId>14123451</cfFedIdId>
  <cfInstId>42</cfInstId>
  <cfFedId>10.1186/1471-2105-10-324</cfFedId>
  <cfStartDate>2009-10-31T24:00:00</cfStartDate>
  <cfFedId_Class>
    <cfClassId>cf7799e5-3477-11e1-b86c-0800200c9a66</cfClassId><!-- cfResultPublication -->
    <cfClassSchemeld>6e0d9af0-1cd6-11e1-8bc2-0800200c9a66</cfClassSchemeld><!-- CERIF Entities -->
  </cfFedId_Class>
  <cfFedId_Class>
    <cfClassId>31d222b4-11e0-434b-b5ae-088119c51189</cfClassId><!-- DOI -->
    <cfClassSchemeld>bccb3266-689d-4740-a039-c96594b4d916</cfClassSchemeld><!-- Identifier Types -->
  </cfFedId_Class>
  <cfFedId_Srv>
    <cfSrvld>5123452</cfSrvld><!-- CrossRef -->
    <cfClassId>eda2b2e2-34c5-11e1-b86c-0800200c9a66</cfClassId><!-- Issuer -->
    <cfClassSchemeld>5a270628-f593-4ff4-a44a-95660c76e182</cfClassSchemeld><!-- Ident Service Roles -->
  </cfFedId_Srv>
</cfFedId>

```

2.7.5 Person Names

`cfPersName` has become a regular base entity with CERIF 1.5 FDM release enabling multiple structured person names. However, given its most frequent use together with the `cfPers` entity, a certain shortcut was introduced. Aside from all the modes described above, the intimate relationship between the person's name and the person has led to allowing the `cfPersName` attributes to appear within the `cfPersName_Pers` XML element, like this:

```

<cfPers>
  <cfPersId>internal-pers-jan-dvorak-identifier1</cfPersId>
  <cfGender>m</cfGender>
  <cfPersName_Pers>
    <cfPersNameId>internal-persname-jan-dvorak-identifier1</cfPersNameId>
    <cfClassId>1ce717bf-436b-4f02-9854-de9aea97b10a</cfClassId><!-- Known-as name -->
    <cfClassSchemeld>f91bbe72-fd36-4349-8e85-89176a244274</cfClassSchemeld><!-- Person Name Types-->
    <cfFamilyNames>Dvořák</cfFamilyNames>
    <cfFirstNames>Jan</cfFirstNames>
  </cfPersName_Pers>
</cfPers>

```

3 Examples

The following example expresses the history of euroCRIS presidents.

```
<?xml version="1.0" encoding="UTF-8"?>
<CERIF xmlns="urn:xmlns:org:eurocris:cerif-1.5-1"
  xsi:schemaLocation="urn:xmlns:org:eurocris:cerif-1.5-1
    http://www.eurocris.org/Uploads/Web%20pages/CERIF-1.5/CERIF_1.5_1.xsd"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  release="1.5" date="2012-07-25" sourceDatabase="euroCRIS db">

  <cfClass>
    <cfClassId>6ba88422-214c-4080-a070-9ac15e21d636</cfClassId>
    <cfClassSchemeld>994069a0-1cd6-11e1-8bc2-0800200c9a66</cfClassSchemeld><!-- Person Org Roles -->
    <cfTerm cfLangCode="en" cfTrans="o">President</cfTerm>
  </cfClass>

  <cfOrgUnit>
    <cfOrgUnitId>1023904</cfOrgUnitId>
    <cfAcro>euroCRIS</cfAcro>
    <cfName cfLangCode="nl" cfTrans="o">Vereniging euroCRIS</cfName>
    <cfOrgUnit_Class>
      <cfClassId>eda2b2f5-34c5-11e1-b86c-0800200c9a66</cfClassId><!-- Private non-profit -->
      <cfClassSchemeld>759af939-34ae-11e1-b86c-0800200c9a66</cfClassSchemeld><!-- Org Types -->
    </cfOrgUnit_Class>

    <cfPers_OrgUnit>
      <cfPersId>1238948</cfPersId>
      <cfClassId>6ba88422-214c-4080-a070-9ac15e21d636</cfClassId><!-- President -->
      <cfClassSchemeld>994069a0-1cd6-11e1-8bc2-0800200c9a66</cfClassSchemeld><!-- Person Org Roles -->
      <cfEndDate>2013-01-01T01:00:00</cfEndDate>
    </cfPers_OrgUnit>

    <cfPers_OrgUnit>
      <cfPersId>21390482</cfPersId>
      <cfClassId>6ba88422-214c-4080-a070-9ac15e21d636</cfClassId><!-- President -->
      <cfClassSchemeld>994069a0-1cd6-11e1-8bc2-0800200c9a66</cfClassSchemeld><!-- Person Org Roles -->
      <cfStartDate>2013-01-01T01:00:00</cfStartDate>
    </cfPers_OrgUnit>

  </cfOrgUnit>

  <cfPers>
    <cfPersId>1238948</cfPersId>
    <cfPersName_Pers>
      <cfPersNameId>423934950485</cfPersNameId>
      <cfClassId>55f90543-d631-42eb-8d47-d8d9266cbb26</cfClassId><!-- Presented Name -->
      <cfClassSchemeld>7375609d-cfa6-45ce-a803-75de69abe21f</cfClassSchemeld><!-- Person Names -->
      <cfFamilyNames>Keith</cfFamilyNames>
      <cfFirstNames>Jeffery</cfFirstNames>
    </cfPersName_Pers>
  </cfPers>

  <cfPers>
    <cfPersId>21390482</cfPersId>
    <cfPersName_Pers>
      <cfPersNameId>43995599405</cfPersNameId>
      <cfClassId>55f90543-d631-42eb-8d47-d8d9266cbb26</cfClassId><!-- Presented Name -->
      <cfClassSchemeld>7375609d-cfa6-45ce-a803-75de69abe21f</cfClassSchemeld><!-- Person Names -->
      <cfFamilyNames>Ed</cfFamilyNames>
      <cfFirstNames>Simons</cfFirstNames>
    </cfPersName_Pers>
  </cfPers>

</CERIF>
```

The following example gives a high-level overview of the CRIS 2012 conference.

```
<?xml version="1.0" encoding="UTF-8"?>
<CERIF xmlns="urn:xmllns:org:eurocris:cerif-1.5-1"
  xsi:schemaLocation="urn:xmllns:org:eurocris:cerif-1.5-1
    http://www.eurocris.org/Uploads/Web%20pages/CERIF-1.5/CERIF_1.5_1.xsd"
  xmllns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  release="1.5" date="2012-12-28" sourceDatabase="CRIS 2012 data">

  <cfEvent>
    <cfEventId>7834895900234</cfEventId>
    <cfCountryCode>CZ</cfCountryCode>
    <cfCityTown>Prague</cfCityTown>
    <cfStartDate>2012-06-06</cfStartDate>
    <cfEndDate>2012-06-09</cfEndDate>
    <cfName cfLangCode="en" cfTrans="o">CRIS 2012</cfName>
    <cfDescr cfLangCode="en" cfTrans="o">The 11th International Conference on Current Research
  Information Systems</cfDescr>

    <cfOrgUnit_Event>
      <cfOrgUnitId>1023904</cfOrgUnitId><!-- euroCRIS -->
      <cfClassId>7b35d959-58eb-4bc1-b13d-9e19aaf8b875</cfClassId><!-- Promoter -->
      <cfClassSchemeld>39cdd720-a0e4-414c-b33f-8bf1c1fd45c4</cfClassSchemeld><!-- Organisation Event
  Involvements -->
    </cfOrgUnit_Event>

    <cfOrgUnit_Event>
      <cfOrgUnitId>9045909068384</cfOrgUnitId><!-- InfoScience Praha -->
      <cfClassId>b4ba809e-60d7-411c-af62-792100c45341</cfClassId><!-- Organizer -->
      <cfClassSchemeld>39cdd720-a0e4-414c-b33f-8bf1c1fd45c4</cfClassSchemeld><!-- Organisation Event
  Involvements -->
    </cfOrgUnit_Event>
    <cfOrgUnit_Event>
      <cfOrgUnitId>8458954893474743</cfOrgUnitId><!-- Action M Agency -->
      <cfClassId>b4ba809e-60d7-411c-af62-792100c45341</cfClassId><!-- Organiser -->
      <cfClassSchemeld>39cdd720-a0e4-414c-b33f-8bf1c1fd45c4</cfClassSchemeld><!-- Organisation Event
  Involvements -->
    </cfOrgUnit_Event>
  </cfEvent>

  <cfEvent>
    <cfEventId>6675747858489328</cfEventId>
    <cfName cfLangCode="en" cfTrans="o">The CRIS conference series</cfName>

    <cfOrgUnit_Event>
      <cfOrgUnitId>1023904</cfOrgUnitId><!-- euroCRIS -->
      <cfClassId>b4ba809e-60d7-411c-af62-792100c45341</cfClassId><!-- Organiser -->
      <cfClassSchemeld>39cdd720-a0e4-414c-b33f-8bf1c1fd45c4</cfClassSchemeld><!-- Organisation Event
  Involvements -->
    </cfOrgUnit_Event>
  </cfEvent>

  <cfEvent_Event>
    <cfEventId1>7834895900234</cfEventId1>
    <cfEventId2>6675747858489328</cfEventId2>
    <cfClassId>eda28bc1-34c5-11e1-b86c-0800200c9a66</cfClassId><!-- Part -->
    <cfClassSchemeld>ac90d633-8ad6-4c23-b45a-cc4f91309843</cfClassSchemeld><!-- Inter-Event Relations-->
  </cfEvent_Event>

  <cfOrgUnit>
    <cfOrgUnitId>1023904</cfOrgUnitId>
    <cfAcro>euroCRIS</cfAcro>
  </cfOrgUnit>
```

```

<cfOrgUnit>
  <cfOrgUnitId>9045909068384</cfOrgUnitId>
  <cfName cfLangCode="cs" cfTrans="o">InfoScience Praha s.r.o.</cfName>
  <cfName cfLangCode="en" cfTrans="h">InfoScience Praha LLC</cfName>
</cfOrgUnit>

<cfOrgUnit>
  <cfOrgUnitId>8458954893474743</cfOrgUnitId>
  <cfName cfLangCode="en" cfTrans="o">Action M Agency</cfName>
  <cfName cfLangCode="cs" cfTrans="h">Agentura Action M</cfName>
</cfOrgUnit>

<cfPers>
  <cfPersId>1238948</cfPersId>
  <cfPersName_Pers>
    <cfPersNameId>423934950485</cfPersNameId>
    <cfClassId>55f90543-d631-42eb-8d47-d8d9266cbb26</cfClassId><!-- Presented Name -->
    <cfClassSchemeld>7375609d-cfa6-45ce-a803-75de69abe21f</cfClassSchemeld><!-- Person Names -->
    <cfFamilyNames>Keith</cfFamilyNames>
    <cfFirstNames>Jeffery</cfFirstNames>
  </cfPersName_Pers>
</cfPers>

<cfPers>
  <cfPersId>984894838</cfPersId>
  <cfPersName_Pers>
    <cfPersNameId>349230994404</cfPersNameId>
    <cfClassId>55f90543-d631-42eb-8d47-d8d9266cbb26</cfClassId><!-- Presented Name -->
    <cfClassSchemeld>7375609d-cfa6-45ce-a803-75de69abe21f</cfClassSchemeld><!-- Person Names -->
    <cfFamilyNames>Jan</cfFamilyNames>
    <cfFirstNames>Dvořák</cfFirstNames>
  </cfPersName_Pers>
</cfPers>

<cfResPubl>
  <cfResPublId>90342909085</cfResPublId>
  <cfResPublDate>2012-05-18</cfResPublDate>
  <cfISBN>978-80-86742-33-5</cfISBN>

  <cfTitle cfLangCode="en" cfTrans="o">E-Infrastructures for Research and Innovation: Linking Information
Systems to Improve Scientific Knowledge Production</cfTitle>
  <cfSubtitle cfLangCode="en" cfTrans="o">Proceedings of the 11th International Conference on Current
Research Information Systems (June 6-9, 2012, Prague, Czech Republic)</cfSubtitle>

  <cfResPubl_Event>
    <cfEventId>7834895900234</cfEventId>
    <cfClassId>eda2d9ec-34c5-11e1-b86c-0800200c9a66</cfClassId><!-- Conference Proceedings -->
    <cfClassSchemeld>759af938-34ae-11e1-b86c-0800200c9a66</cfClassSchemeld><!-- Output Types -->
  </cfResPubl_Event>

  <cfPers_ResPubl>
    <cfPersId>1238948</cfPersId>
    <cfClassId>708b3df0-1cfe-11e1-8bc2-0800200c9a66</cfClassId><!-- Editor -->
    <cfClassSchemeld>b7135ad0-1d00-11e1-8bc2-0800200c9a66</cfClassSchemeld><!-- Person Output
Contributions -->
  </cfPers_ResPubl>
  <cfPers_ResPubl>
    <cfPersId>984894838</cfPersId>
    <cfClassId>708b3df0-1cfe-11e1-8bc2-0800200c9a66</cfClassId><!-- Editor -->
    <cfClassSchemeld>b7135ad0-1d00-11e1-8bc2-0800200c9a66</cfClassSchemeld><!-- Person Output
Contributions -->
  </cfPers_ResPubl>

  <cfOrgUnit_ResPubl>
    <cfOrgUnitId>8458954893474743</cfOrgUnitId>
    <cfClassId>7ef398b2-1cfe-11e1-8bc2-0800200c9a66</cfClassId><!-- Publisher -->

```

```

    <cfClassSchemeld>877161b4-00d2-42c8-a368-aaa35262f3a8</cfClassSchemeld><!-- Organisation Output
Roles -->
  </cfOrgUnit_ResPubl>
</cfResPubl>

```

```

<cfClassScheme>
  <cfClassSchemeld>39cdd720-a0e4-414c-b33f-8bf1c1fd45c4</cfClassSchemeld>
  <cfName cfLangCode="en" cfTrans="o">Organisation Event Involvements</cfName>
  <cfClass>
    <cfClassId>b4ba809e-60d7-411c-af62-792100c45341</cfClassId>
    <cfTerm cfLangCode="en" cfTrans="o">Organizer</cfTerm>
    <cfRoleExpr cfLangCode="en" cfTrans="o">organizes</cfRoleExpr>
    <cfRoleExprOpp cfLangCode="en" cfTrans="o">is organized by</cfRoleExprOpp>
  </cfClass>
  <cfClass>
    <cfClassId>7b35d959-58eb-4bc1-b13d-9e19aaf8b875</cfClassId>
    <cfTerm cfLangCode="en" cfTrans="o">Promoter</cfTerm>
    <cfRoleExpr cfLangCode="en" cfTrans="o">promotes</cfRoleExpr>
    <cfRoleExprOpp cfLangCode="en" cfTrans="o">is promoted by</cfRoleExprOpp>
  </cfClass>
</cfClassScheme>

```

```
</CERIF>
```

The following example represents the information about the Value Added Tax Identification Number of InfoScience Praha, an SME. The VATIN had been «005-27131157» until the Czech Republic joined the EU³; it has been «CZ27131157» since.

```

<?xml version="1.0" encoding="UTF-8"?>
<CERIF xmlns="urn:xmlns:org:eurocris:cerif-1.5-1"
  xsi:schemaLocation="urn:xmlns:org:eurocris:cerif-1.5-1
  http://www.eurocris.org/Uploads/Web%20pages/CERIF-1.5/CERIF_1.5_1.xsd"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  release="1.5" date="2012-12-28" sourceDatabase="CRIS 2012 data">

  <cfOrgUnit>
    <cfOrgUnitId>9045909068384</cfOrgUnitId>
    <cfName cfLangCode="cs" cfTrans="o">InfoScience Praha</cfName>
    <cfName cfLangCode="en" cfTrans="h">InfoScience Praha</cfName>
    <cfResAct cfLangCode="en" cfTrans="o">The company provides information systems and services in the
domain of research information processing.</cfResAct>

    <cfOrgUnit_Class><!-- an embedded classification -->
      <cfClassId>eda2b2f2-34c5-11e1-b86c-0800200c9a66</cfClassId><!-- SME -->
      <cfClassSchemeld>759af939-34ae-11e1-b86c-0800200c9a66</cfClassSchemeld><!-- Organisation Types-->
    </cfOrgUnit_Class>

    <cfFedId><!-- the current value of the VAT identification number as a Federated Identifier -->
      <cfFedIdId>23893289493</cfFedIdId><!-- internal identifier of the federated identifier itself -->
      <cfFedId>CZ27131157</cfFedId><!-- this is the current InfoScience VAT Id. No. -->
      <cfStartDate>2004-05-01T02:00:00</cfStartDate><!--midnight local time (CEST) when CZ entered the EU-->
      <cfFedId_Class>
        <cfClassId>eda2b2e2-34c5-11e1-b86c-0800200c9a66</cfClassId><!-- VAT identification number -->
        <cfClassSchemeld>bccb3266-689d-4740-a039-c96594b4d916</cfClassSchemeld><!-- Identifier Types -->
      </cfFedId_Class>
      <cfFedId_Srv>
        <cfSrvId>82394823</cfSrvId><!-- the VATIN issuing service (below) -->
        <cfClassId>eda2b2e2-34c5-11e1-b86c-0800200c9a66</cfClassId><!-- Issuer -->
        <cfClassSchemeld>5a270628-f593-4ff4-a44a-95660c76e182</cfClassSchemeld><!-- Identifier Service
Roles -->
      </cfFedId_Srv>
    </cfFedId>

```

³ On 1-May-2004

```

<cfFedId><!-- a historical record for the VAT identification number as a Federated Identifier -->
  <cfFedIdId>295553732893</cfFedIdId><!-- internal identifier of the federated identifier itself -->
  <cfFedId>005-27131157</cfFedId><!-- this is the former InfoScience VAT Reg. No. -->
  <cfStartDate>2004-04-07T02:00:00</cfStartDate><!-- midnight local time (CEST) when the company
registered for VAT -->
  <cfEndDate>2004-05-01T02:00:00</cfEndDate><!-- midnight local time (CEST) when CZ entered the EU -->
  <cfFedId_Class>
    <cfClassId>eda2b2e2-34c5-11e1-b86c-0800200c9a66</cfClassId><!-- VAT identification number -->
    <cfClassSchemeld>bccb3266-689d-4740-a039-c96594b4d916</cfClassSchemeld><!-- Identifier Types -->
  </cfFedId_Class>
  <cfFedId_Srv>
    <cfSrvId>82394823</cfSrvId><!-- the VATIN issuing service (below) -->
    <cfClassId>eda2b2e2-34c5-11e1-b86c-0800200c9a66</cfClassId><!-- Issuer -->
    <cfClassSchemeld>5a270628-f593-4ff4-a44a-95660c76e182</cfClassSchemeld><!-- Identifier Service
Roles -->
  </cfFedId_Srv>
</cfFedId>
</cfOrgUnit>

<cfSrv>
  <cfSrvId>82394823</cfSrvId>
  <cfName cfLangCode="en" cfTrans="o">Issuing VATINs</cfName>
  <cfDescr cfLangCode="en" cfTrans="o">The service of issuing Value Added Tax Identification
Numbers.</cfDescr>
  <cfKeyw cfLangCode="en" cfTrans="o">Value Added Tax; VAT Identification Number</cfKeyw>
</cfSrv>

</CERIF>

```

4 Acknowledgements

The CERIF task group wishes to thank all actively involved contributors for highly valuable discussions and input that materialized in the current CERIF XML exchange format. Within this update (CERIF XML 1.5), we want to especially acknowledge the experience and feedback from the CRISPOOL project, and refer to [7] for related experience in the FRIS and the IST World portals.

5 Components of the CERIF 1.5 Release

CERIF is being approached variously, where the developments evolved throughout history and have continuously been influenced by ongoing technological developments and by the practical requirements with implementations. The current CERIF 1.5 release comprises the following components:

- CERIF 1.5 FDM: Model Introduction and Specification [1]
- CERIF 1.5 FDM: SQL scripts for most common databases
(available to euroCRIS members only)
- CERIF 1.5 XML: Data Exchange Format Specification
(this document)
- CERIF 1.5 XML Examples
(available from the euroCRIS website)
- CERIF 1.5 XML Schema File
(available from the euroCRIS website)
- CERIF 1.5 Semantics Documentation [2]
- CERIF 1.5 Vocabulary
(available from the euroCRIS website in Excel and validated CERIF XML)

Additional CERIF related information, files and more documents or background information about CERIF and CRISs are available from the euroCRIS website: <http://www.eurocris.org/>.

6 References

- [1] Jörg, B.; Dvorak, J.; Houssos, N. ; Jeffery, K.G.; Brasse, V.; Höllrigl, T.; Vestdam, T.; Sicilia, M.-A.; Zendulkova, D.; (2012): CERIF 1.5 Full Data Model (FDM): Introduction and Specification. to appear, February 2013.
- [2] Jörg, B.; Kerridge, S.; Cranner, P. Trowel, S., Ginty, K., Jeffery, K.G.; Houssos, N.; Brasse, V.; Höllrigl, T.; Vestam, T.; Baker, D.; Foster S., Dvořák, J.; Asserson, A.; Rasmussen, H.; Zendulkova, D.; Price, A.; Sicilia, M.A.; Ruiz-Rube, I.; van Grootel, G.; Evans, K.; Strijbosch, L.; Cox, M.; Simons, E.J. (2012): CERIF 1.5 Semantics: Research Vocabulary. euroCRIS, February 2013.
- [3] 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/>
- [4] W3C Recommendation: XML Schema: <http://www.w3.org/XML/Schema>
- [5] W3C Recommendation: Namespaces in XML 1.0: <http://www.w3.org/TR/REC-xml-names/>
- [6] Jörg, B.; Dvořák, J.; Vestdam, T.; van Grootel, G.; Jeffery, K.; Clements, A. (2012): CERIF 1.3 XML Data Exchange Format Specification. euroCRIS, January 2012. http://www.eurocris.org/Uploads/Web%20pages/CERIF-1.3/Specifications/CERIF1.3_XML.pdf
- [7] Jörg, B.; Dvořák, J.; Vestdam, T.: Streamlining the CERIF XML Data Exchange Format Towards CERIF 2.0. In: Jeffery, Keith G; Dvořák, Jan (eds.): *E-Infrastructures for Research and Innovation: Linking Information Systems to Improve Scientific Knowledge Production: Proceedings of the 11th International Conference on Current Research Information Systems* (June 6-9, 2012, Prague, Czech Republic). Pp. 221-230. ISBN 978-80-86742-33-5. Available from: http://www.eurocris.org/Uploads/Web%20pages/CRIS%202012%20-%20Prague/CRIS2012_24_full_paper.pdf