### The XML/RDF encoding schemes for CERIF. Application profiles for scientific metadata.

Andrei Lopatenko CSTIT RAS. andrei@ccas.ru andrei@derpi.tuwien.ac.at

> CSTIT RAS Andrei S. Lopatenko EuroCRIS May 2001

#### Integration of CRIS

- Reasons
  - Why we need integrate CRISes
- Requirements
  - How we should integrated them
  - What are primary requirements for integration solution
- Solutions
  - What our solution provides new for integration

#### Current situation

- There are a number of CRIS in Europe and national areas (Russian, as example)
- Usually each CRIS stores geographically local (university, district, country) data. Some of them store sector data scientific data (for innovations only, chemical data)
- Each CRIS provides a limited number of services
- The users of CRIS has access to data stored in it only, no CRIS works as a gateway or as a mediator

#### Current situation

- As usual researchers need an access to all information related to their topic, not local only
- Required information is stored in a number of CRIS
- To get information researcher need to know about and to have access to a number of CRIS

#### Current situation

- So to give to researcher ability to get required information we should
- 1. to give him list of all related to his research CRISes(develop CRIS portals)
- 2. to give him access to CRISes
- 3. to teach him to work with CRISes
- Such solution is difficult, expensive, timeconsumable, people-consumable and maybe impropriate in some situations

Russian

 ISIR (Integrated System of Informational Resources) of Russian Academy of Science (<u>http://isir.ras.ru</u>) - Central Datawarehouse for RAS

holds data about

• Persons, Institutes, Departments, Projects, Publications

#### ISIR

- Central Node (Presidium of RAS). Data about all Academy of Science
- Regional Nodes (Kazan). Data of institutes of Kazan Scientific Center
- Sectorial Nodes (Mathematics, joint with MathNET, Telecommunications)

#### Russia

- Innovation projects of Russian Fund of Basic Research (innovation projects, results, patents)
- Oriental Studies (epigraphic sources)

- Data of Innovation Projects system and regional node of ISIR are a great asset to central node of ISIR, because those data describes activities in Academy of Science
- Some data (staff) of the central node of ISIR are a great asset for innovation projects system

### Primary requirements

Data integration technique

- should be platform and language independent
- work in case of independent development of each CRIS
- able to solve problems of semantic and schematic differences between systems, work in heterogeneous environment with independently developing information systems

### Solution

- To exchange data between CRIS
- To use RDF as a language for encoding data
- To use kernel metadata model(CERIF), which provides basic metadata model for interoperation of CRIS
- To use application profiles as a method for expanding metadata schemas to application needs



### RDF for encoding

- Power of RDF in expressing data.
- <u>RDF is semantic.</u>
- There is XML encoding of RDF. XML standard applications for signing, linking, identifying, searching, transforming can be applied
- A lot of developed and emerging tools allow to easily develop solutions on RDF

#### CERIF

- A lot of CRIS store same kinds of data(semantically), which are described in CERIF standard
- CERIF is about ten-years subject of discussion and using by European CRIS project teams
- As author experience shows CERIF is detailed enough to define kernel, satisfying needs of CRIS and not detailed so to be schematic conflicted
- CERIF is multilevel (metadata model, data exchange model, Full model), what makes solution flexible<sub>CSTIT RAS Andrei S. Lopatenko EuroCRIS May 2001</sub>

### Application profiles

- To combine elements from a number of "namespace schemas"
- To override the default definitions of element with a definition that makes sense in a local context.
- To describe schema of application
- To specify dependencies among vocabulary terms
- To define multiple entity classes in a particular data model.
- To provide guidelines, in free text, about the local use of particular terms

CSTIT RAS Andrei S. Lopatenko EuroCRIS May 2001

# Developing application-specific schemas

- Need to specify a set of elements and their properties
- Need to specify restrictions/integrity rules
- Need to specify which vocabulary and where
- Need to specify relation between elements

#### Problems

• There can be semantic and schematic differences between namespaces

# Our application profiles consist of.

- A set of included namespaces
- Definitions semantic relations between elements from different namespaces
- Overriding definition of elements
- Vocabularies definitions and vocabularies content
- Definitions semantic relations between elements from different vocabularies
- Rules for resolving semantic and schema conflicts

### Application profiles. Advantages

- Allow to reuse schemas
- Proved to be useful (SCHEMAS, UKOLN DESIRE, Z39.50 applications)
- Close to Digital Library projects such as Digital Library Warwick Framework

#### Current state

Have developed

- 1. RDF schema for CERIF.
- 2. importing and exporting tools from CERIF database (published on cordis) into CERIF RDF
- 3. vocabulary tools
- 4. application profiles managing tools

### Representing CERIF in RDF.

#### Possible usage

- Data exchange between CRIS
- Semantic representation of research data in the web pages for data gathering. RDF-DC is not sufficient
- Translating data into RDF for additional functionality, which can not be provided by CRIS. Developing Value-added services



🚈 http://128.130.73.166/RFBR/Dictionaries/voc_item.asp?parent=4945&Action=Edt - M	ficrosoft Internet Explorer 📃 🗗 🔀
<u>F</u> ile <u>E</u> dit ⊻iew F <u>a</u> vorites <u>T</u> ools <u>H</u> elp	
Image: the second state     Image: the second state     Image: the second state     Image: the second state       Back     Forward     Stop     Refresh     Home     Search     Favorites     Histor	ry Mail Print Edit Discuss
Address 🛃 http://128.130.73.166/RFBR/Dictionaries/voc_item.asp?parent=4945&Ac	ction=Edt 💽 🖉 Go 🗍 Links 🎽
Registry Defining app	plication profiles
Name ISIR App	
Description	
<ul> <li>Âêëþ÷àåò</li> <li>iðíñòðàíñòâà</li> <li>èìẩí</li> <li>CERIF:RDF-scher : a for CERIF metadamodel</li> <li>ISIR-vCard:vCard schema</li> <li>Innovations:the namespace intended to describe data r</li> <li>RDFS:W3 RDFS schema</li> </ul>	related to inno∨ations. intellectual rights
Èçìåíèòü	
	▼
🗉 Done	📔 👘 İnternet

🚈 http://128.130.73.166/F	RFBR/Dictionaries/voc_item.asp?parent=2828&Action=Edt - Microsoft Internet Explorer 📃 🗗 🔀
∫ <u>F</u> ile <u>E</u> dit <u>V</u> iew F <u>⊰</u>	avorites <u>T</u> ools <u>H</u> elp
Back Forward	Stop Refresh Home Search Favorites History Mail Print Edit Discuss
Address 🖉 http://128.1	30.73.166/RFBR/Dictionaries/voc_item.asp?parent=2828&Action=Edt 💿 🎓 Go 🗍 Links »
Name	Publications
Description	Property describing list of person's publications
Type	Property namespaces
See also	No
Defined by	No
Constraint resource	No
Constraint property	No
Subclass of	RDFS:Bug Go
Subproperty	Íåò 🔽
rdfs:Domain	No CERIF:PersonPublication CERIF:Project-abstract CERIF:ResearchInterest



![](_page_25_Picture_0.jpeg)

#### Management facilities

# Organizing information. Catalogs. Search services. Example:

#### Directories at http or ftp server

- Projects
  - Computer Science
    - 2001
      - Auris.rdf
    - 2000
  - Chemical
- Persons
  - Vienna University of Technology
  - Computing Center of Russian Academy of Science

![](_page_27_Picture_0.jpeg)

![](_page_28_Figure_0.jpeg)

### Representing in CERIF RDF FoDok data (TUW)

Projects

- Title
- Persons take part in
- Acronym
- Begin and end data
- Description
- URL
- etc

#### 🔕 🖻 🧭 💁 🎒 2 <u></u> $(\mathbf{x})$ Forschungsdokumentation Außeninstitut der Technischen Universität Wien Home Search Update Products Feedback Help About Technische Universität Wien Dienstleistungseinrichtungen und Senatsinstitute Außeninstitut Forschungs-Projekt E015-01-08 Titel: Multimediale Neugestaltung und Erweiterung von AURIS (Austrian Research Information System) zur Steigerung der Attraktivität und Bedienerfreundlichkeit der österreichischen Forschungsdokumentation Titel (englisch): Austrian Research Information System: Multimedial Enhancement Leiter (Kontaktperson): Name: Niedermayer, Walter (Dipl.-Ing.) Spezialgebiet: Forschungsdokumentation; Current research information systems E-Mail-Adresse: walten@derpi.tuwien.ac.at Name: Krieger, Andreas Spezialgebiet: Datenverarbeitung; Programmierung; Datenbank E-Mail-Adresse: kriegen@derpi.tuwien.ac.at Auftraggeber / Forschungsprogramm / Programm-Akronym / Projektnummer: BMVK Beginn-Datum: 2000-02 Ende-Datum: 2001-12

![](_page_31_Figure_0.jpeg)

```
📋 rfbr.RDF - Notepad
```

\_ 8 ×

<u>File Edit Search Help</u> <first-name> Walter </first-name> <familu-name> Niedermauer </familu-name> <uri>email:walter@derpi.tuwien.ac.at </uri> <nationality> Austrian </nationality> <qualifications>Dipl.-Inq.</qualifications> <sex>M</sex> </person> <person id="person:2"> <first-name> Andreas </first-name> <family-name>Krieger </family-name> <uri>email:krieger@derpi.tuwien.ac.at </uri> <nationality>Austrian </nationality> <sex>M</sex> </person> <project id="project:E015-01-08"></project <status>Execution</status> <URL>http://arge.tuwien.ac.at</URL> <participants> <rdf:Bag> <rdf:li> <participant> <person rdf:resource="person:1"/> <role>head</role> </participant> </rdf:li> <rdf:li> <participant> cperson rdf:resource="person:2"/> <role>co-worker</role> </participant> </rdf:li> </rdf:Bag> </participants>

#### **RDF** Semantic

- Allows to automatically include new CRIS in data exchange, specifying semantic definition of its entities withour programming
- Allows to automatically include new types of entities existing CRIS specifying their semantic relations with defined in schema
- Subclassing
- Other semantic descriptions(MetaNet)

#### Russian CRIS

ISIR RAS (Integrated System of Informational Resources of Russian Academy of Science )

(http://isir.ras.ru)

Data about institutions, departments, persons, projects, publications

<u>Distributed</u> (Presidium, Computing Center, Kazan, Mathematical Institute MathNET)

Primary audience: Researchers, Administrators, Public

### Russian CRIS. ISIR RAS.

#### Services

- Searching informational resources (web)
- Navigating in information space (web)
- Publishing information through the web
- Data exporting and importing in RDF for organizing vertical and horizontal information infstructures (application profile based on CERIF)
- Distributed services for data gathering

![](_page_36_Figure_0.jpeg)

![](_page_37_Figure_0.jpeg)

← • → • 🕺 🔂 🛣	1 3 🖪 - 🍠 🛒	× 🗐		) <u>- 문 ×</u>
	Structure: Organiza	ations Search: Orga	nizations Persons	s Projects Publications ? 🔺
Comput	ing Centre of t	he Russian Aca	demy of Scie	nces
General info	Structure	Employees	Projects	Publications
	CC RAS 117967, 40, Vav phone: 1 fax: (0 e-mail address: w www-page:h	ilova st., GSP-1, Mose 350440 , (095) 13524 095) 1356159 vcan@c?as.ru , evt@c ttp://www.ccas.ru/	cow 189 , (095) 135615 cas.ru	Links to list of related resources
Page describing resource INSTITUTE	Director Deputy director Deputy director Deputy director Deputy director Scientific secret	ary	corresponding men phone: 1352489 academician Zhura academician Krası cpms Mikhailov phone: (095)1356 fax: (095)1356159 e-mail address: gn Tyapkin Sergei A phone: 1350152 dpms Shurshalov phone: 1350228	nber Evtushenko Yurii Gavri avlev Yurii Ivanovich aoshchekov Pavel Sergeevich Gurii Mikhailovich 234 9 nickail@ccas.ru Arkad'evich Lev V.
How to reach				_
•				• •

![](_page_39_Figure_0.jpeg)

#### (+ • + • ⊗ 🕑 🟠 🍳 🖻 🧐 💁 🚽 🗐 • 🗐

#### General info

#### N. N. Olenev, N. V. Belotelov AN ECOLOGICAL DEMOGRAPHIC ECONOMIC MODEL (paper at conference) information

Abstracts of the 3rd Moscow International Conference on Operations Research (ORM2001) (Moscow, April 4-6, 2001), 2001, p.91, english **resource** 

#### Abstract

#### **PUBLICATION**

A non-linear model of ecological, demographic, and economic processes is presented to give a useful tool for global analysis of international labour mobility and trade. The model considers a set of points (countries) on the plane, each determined by dynamics of their own variables, and interacted with each other. The basis for this model is determined on the hypotheses that:

- 1. population can migrate from one point to another to get the best level of living;
- 2. population consist of two employment groups with skilled and unskilled labour;
- 3. production function converts skilled and unskilled labour units into effective labour units.

The level of living at the specified point is determined by economic activity and quality of environment.

#### Keywords

non-linear model, ecological, demographic, economic, international labour mobility, trade, skilled and unskilled labour, economic activity, environment quality

#### Full text

🔊 Search for pe	rson - Microsoft Inte	ernet Explore	r									- 8 ×
<u> </u>	<u>V</u> iew F <u>a</u> vorites	<u>T</u> ools <u>H</u>	elp									
Back F	r⇒ - ⊗ orward Stop	🕼 Refresh	Home	Q Search	Favorites	🧭 History	Mail	int 🕒	<b>⊡∕</b> Edit	• 🛄 Discuss		
🛛 A <u>d</u> dress 🖉 ht	tp://isir.ras.ru/en/o	lb/search_p	er.asp?P=.oi	i-50. vi fi In	n-en						💌 🤗 Go	Links »
Language $\bigtriangledown$	👖 🖥 ላ	🕴 🏟 St	ructure:	Organiz	ations	Search: (	Organiza	ations	Persons	Projects	Publications	? 🔺
				Sea	arch fo	r pers	on					
Query Search result												
Last name	Lopatenko					Rank				•		
First name	_				Do	ctorate						-
rn st name	1				d	egrees						
Middle name						ph/fax						
				e	Search	over the	Academ	v of scie	ences			
				0	Search	in organi	zation	,				
					BRAI	NCHES	OF RU	SSIAN	ACADI	EMY OF S	SCIENCES	
S	Search <b>p</b>	age										
f.	or Dorge	nc		_ 5	Search	Clea	ır					
		)115										
												~
e]											nternet	

![](_page_42_Figure_0.jpeg)

### Innovation projects

Holds data about

- Innovation projects
- Innovation results
- Patents

To support process of development innovations projects, for technology dissemination Audience: investors, researchers, funds, innovation relay centers

#### Innovation projects

- Data can be entered by
  - Internet forms (secure)
  - RDF description
- Data can be viewed by
  - Web
  - RDF files

#### The Creation of the Logical Model of the Database for the Thematic Epigraphic Set (Referring to the North Indian Early Mediaeval Land Grants)

Presentation provided by prof. Stolyarov A. A. Institute for Oriental Studies Russia EuroCRIS May 2001

The method of investigating thematic subsets of epigraphic sources having the homogenous inner structure permits the scholar to treat them as certain metatext

![](_page_47_Figure_0.jpeg)

CSTIT RAS Andrei S. Lopatenko EuroCRIS May 2001 Using this model the scholar can according to his interests concentrate his attention not upon all layers of his DB but upon one or two of them. The others could be "switched in" according to the necessity.

![](_page_49_Figure_0.jpeg)

![](_page_50_Figure_0.jpeg)

1. inVocatio - God devotion,

2. inTitulatio - donor,

3. insCriptio - donee,

4. saLutatio - greetings,

5. Arenga - preamble,

6. Promulgatio - public announcement,

7. Narratio - account of circumstances,

8. dIspositio - decree,

9. Sanctio - prohibition of infringement,

10. corroBoratio - signs of certification,

11. Datum - place and time,

12. appRecatio - conclusion = wishing of good.

By Caps in Latin terms those characters are marked, which are expected to be used in the process of preparing the inner form description for short designation of that or other segment of inner structure of the document under research. CSTIT RAS Andrei S. Lopatenko EuroCRIS May 2001

	Microsoft Excel - CP-Bibl												
	) <u>Ф</u> айл <u>П</u> равка <u>В</u> ид Вст <u>а</u> вка Фор <u>м</u> ат С <u>е</u> рвис <u>Д</u> анные <u>О</u> кно <u>С</u> г	травка				_ B ×							
	) C≓ E A A A A A V X B B A √ > Q Σ		<b>43</b> 80% ▼ <b>?</b>	-									
	A211 = Itaunja (Lucknow museum) CP Inscription of Govindachandradeva, V.S. 1186												
1	GO Name Name	State 🚽	District 🗨	Centur 🚽	Year 🚽	Dynas							
2	Taxila CP Inscription of Patika, y.78	West Punjab	Rawalpindi	1	~21	Kusha							
3	Kalawan CP Inscription, y. 134	West Punjab	Rawalpindi	1	~77	Kusha							
4	Taxila Silver Scroll Inscription of a Kushana King, y. 136	West Punjab	Rawalpindi	1	~79	Kusha							
5	Sui Vihar CP Inscription of Kanishka 1, y. 11	West Pakistan	Bahawalpur	1	~89	Kusha							
6	Nalanda Spurious CP Inscription of Samudragupta, y. 5	Bihar	Nalanda	67		Gupt							
7	Gaya Spurious CP Inscription of Samudragupta, y. 9	Bihar	Gaya	67		Gupta							
8	Damodarpur CP Inscription of the time of Budhagupta, Gupta y. 163	Bangladesh	Dinajpur	5	~482	Gupt							
9	Damodarpur CP Inscription of the time of Kumaragupta 1, Gupta y. 124	Bangladesh	Dinajpur	5	~444	Gupt							
10	Damodarpur CP Inscription of the time of Budhagupta	Bangladesh	Dinajpur	5	~476-94	Gupt							
11	Nalanda Seal of Budhagupta	Bihar	Patna	5	~476-94	Gupt							
12	Damodarpur CP Inscription of the time of Kumaragupta 1, Gupta y. 128	Bangladesh	Dinajpur	5	~447	Gupt							
13	Dhanaidaha CP Inscription [of Kumaragupta 1], Gupta y.113	Bangladesh	Rajshahi	5	~432	Gupt							
14	Kalaikuri-Sultanpur CP Inscription of the Gupta y. 120	Banglaringsh	Rajshahi	5	~439	Gupt							
15	Paharpur CP Inscription of the Gupta y. 159	Bangladesh	Rajshahi	5	~479	Gupt							
16	Gunaighar CP Inscription of Vainyagupta, Gupta y. 188	Bangladesh	Tippera (Comilla)	6	~507	Gupta							
17	Nandapur CP Inscription of the Gupta y. 169	Bihar	Monghir	5	~488	Gupt							
18	Baigram CP Inscription of the Gupta y. 128	Bangladesh	Bogra	5	~448	Gupt							
19	Indor CP Inscription of Skandagupta & Sarvanaga, Gupta y. 146	Uttar Pradesh	Bulandshahr	5	~466	Gupt							
20	Bhitari Seal of Kumaragupta 2	Uttar Pradesh	Ghazipur	56		Gupt							
21	Nalanda Seal of Vishnugupta	Bihar	Patna	56		Gupt							
22	Nalanda Fragmentary Seal Inscription of Suvratavarman	Bihar	Patna	7	605-10	Maukharis c							
23	Damodarpur CP Inscription of the Gupta y. 224	Bangladesh	Dinajpur	6	~543	Gupt;							
24	Faridpur CP Inscription of the time of Dharmaditya, R.y. 3	Bangladesh	Faridpur	6	540-60	A-1							
25	Faridpur CP Inscription of the time of Dharmaditya	Bangladesh	Faridpur	6	540-60	Y'N							
26	Faridpur CP Inscription of the time of Gopachandra, R.y. 18	Bangladesh	Faridpur	6	(	Jul 1							
27	Mallasarul CP Inscription [of Vijayasena] of the time of Gopachandra, R.y. 33	West Bengal	Burdwan	6		Tent							
28	Khoh CP Inscription of Sarvanatha, Gupta (or K-C) y. 193	Madhya Pradesh	Satna	6	~512	ccaka							
29	Khoh CP Inscription of Samkshobha, Gupta y. 209	Madhya Pradesh	Satna	6	~529	amo na lvra							
30	Asirgadh Copper Seal Inscription of Sarvavarman	Madhya Pradesh	East Nimar	6	575-85	Maukharis c							
31	Bhamodra-Mohota CP Inscription of Dronasimha, G-V v. 183	Guiarat	Bhavnadar I	6	~502	Maitrakas o							
	()   PI \List / Bibliography /		I										
Го	ТОВО												

CSTIT RAS Andrei S. Lopatenko EuroCRIS May 2001

🖾 Microsoft Excel - CP-Bibl 📃 🖅 🗵														
	<u>Ф</u> айл <u>П</u>	равка	<u>В</u> ид Вс	т <u>а</u> вка с	₽ор <u>м</u> ат	С <u>е</u> рвис	÷ <u>∤</u>	]анные <u>О</u>	кно <u>С</u> пра	вка				_ 8 ×
	i 🛱	84	3 Q :	њ X	e C	<b>S</b>	0	* Ci * (	🍓 Σ f	R A R	l 🛍	🚯 80% 🔹 📿 🗸		
Aria	l Cyr		<b>-</b> 10	- Ж	<u>к ч</u>		≣	≣ 🖬	<b>9 %</b> 0	00 <sup>+</sup> ,00 -0	,8 🗊	律 🛛 • ð • A	· · .	
	A1537 = HGD													
		Α			в	С		D	E	F		G		H 🔺
1	GO Name	Editio	n Nam 🖵	Author's	Name .	, Year	•	Volume, 🚽	Number 🖵	Page 🚽	Name (h	hyper-link)	-	_
523	BLINI			Bhanda	rkar D.R.				213	34	Isaprai	tishhana CP Inscription	of Govindachandradev	a, V.S. 1182
524	JASB			Hall				27		242	Isaprai	tishhana CP Inscription	of Govindachandrad <u>ev</u>	a, V.S. 1182
1536	HGD			Niyogi R		195	9		App.B 24	250	Isaprai	tishhana CP Inscription	of Govindachandradev	a, V.S. 1182
1538														
1539														
1540														
1542														
1543														
1544														
1545														
1546							_							
1547														
1548						_								
1549														
1550														
1551														
1553														
1554														
1555													0.4	
1556													TO S	0
1557													( )	
1558													7	
1559												ር የ	(1)1	
1560												-	dath	in-
1561														
1562														
1564														
		.ist $\lambda$ Bil	bliograp	hy/					•	•				▶
Най,	дено заг	писей:	3 из 15	36										
	CSTIT RAS Andrei S. Lopatenko 54											54		
EuroCRIS May 2001														