Towards an open data on how the research data are used: CRIS-CERIF based approach

Sergey Parinov, CEMI RAS
WHAT DOES IT MEAN: “THE OPEN DATA ON HOW THE RESEARCH DATA ARE USED”?

1. An openness of how scientists manipulate with research outputs (RO), e.g. anchoring, selecting artifacts, etc.
2. An openness of researchers motivations to use (or not) RO in producing new scientific knowledge.
3. A guaranteed awareness of researchers on all facts of using their research outputs.
4. An openness of detailed usage statistics (including motivations, etc.)
ON THE CRIS-CERIF SIDE

- Ecosystems:
  - (1) manipulations, (2) motivations,
  - (3) notifications, (4) usagemetrics

CRIS-CERIF response:
- Classification of existed types and techniques
- Use cases for CRIS users
- CERIF support
- Specific type of CRIS interoperability and interactions
- Necessary requirements for CRIS to support the ecosystem
MANIPULATIONS WITH RESEARCH OUTPUTS

• Known techniques and ecosystems
  • Open Annotation
  • Semantic linking (LOD, research relationships, etc.)
  • etc.

• On the CRIS-CERIF side
  • Information about manipulations made within one CRIS should be shared with a metasystem (research e-infrastructure) and should be available for other CRIS systems
  • CERIF data model should support such data exchange
The purpose of the Open Annotation Community Group is to work towards a common, RDF-based, specification for annotating digital resources. The effort will start by working towards a reconciliation of two proposals that have emerged over the past two years: the Annotation Ontology [1] and the Open Annotation Model [2]. Initially, editors of these proposals will closely collaborate to devise a common draft specification that addresses requirements and use cases that were identified in the course of their respective efforts. The goal is to make this draft available for public feedback and experimentation in the second quarter of 2012. The final deliverable of the Open Annotation Community Group will be a specification, published under an appropriate open license, that is informed by the existing proposals, the common draft specification, and the community feedback.

[1] http://code.google.com/p/annotation-ontology/
[2] http://www.openannotation.org/spec/beta/
Annotation Use Cases, W3C Editor's Draft 03 March 2014

• Requirements - Services
  • Discovery of Annotations Relevant to Publication
  • Persistence of Annotations
  • Service for Sharing Annotations, either with others or between platforms
  • Management of Shared Annotations
  • Authentication and Authorization in the Annotation ecosystem

Source: http://www.openannotation.org/usecases.html
ANNOTATION DATA MODEL

Source: http://www.openannotation.org/spec/core/core.html
ANNOTATIONS ON THE WEB

Annotator

The copy-editor of the web

Annotator is an open-source JavaScript library to easily add annotation functionality to any webpage. Annotations can have comments, tags, links, users, and more. Annotator is designed for easy extensibility so it's a cinch to add a new feature or behaviour. Annotator also fosters an active developer community with contributors from four continents, building 3rd party plugins allowing the annotation of PDFs, EPUBs, videos, images, sound, and more.

Download    GitHub
MANIPULATIONS WITH LINKAGES

Source: http://socionet.ru/pub.xml?h=repec:rus:ecoper:parinov_serigney.56054-1&l=en
| №   | Role                                      | Authors                                           | Date         | Notes                           |
|-----|-------------------------------------------|---------------------------------------------------|--------------|---------------------------------|
| 1   | Project admin                            | Liz Allen, Amy Brand, Jo Scott, Micah Altman and  | 2014-04-26   |                                 |
|     |                                           | Marjorie Hlava                                   |              |                                 |
| 2   | Computation                              | Liz Allen, Amy Brand, Jo Scott, Micah Altman and  | 2014-04-26   |                                 |
|     |                                           | Marjorie Hlava                                   |              |                                 |
| 3   | Investigation: performed the experiments  | Liz Allen, Amy Brand, Jo Scott, Micah Altman and  | 2014-04-26   |                                 |
|     |                                           | Marjorie Hlava                                   |              |                                 |
| 4   | Investigation: data/evidence collection   | Liz Allen, Amy Brand, Jo Scott, Micah Altman and  | 2014-04-26   |                                 |
|     |                                           | Marjorie Hlava                                   |              |                                 |
| 5   | Study conception                         | Liz Allen, Amy Brand, Jo Scott, Micah Altman and  | 2014-04-26   |                                 |
|     |                                           | Marjorie Hlava                                   |              |                                 |
| 6   | Data curation                            | Liz Allen, Amy Brand, Jo Scott, Micah Altman and  | 2014-04-26   |                                 |
|     |                                           | Marjorie Hlava                                   |              |                                 |
| 7   | Methodology                               | Liz Allen, Amy Brand, Jo Scott, Micah Altman and  | 2014-04-26   |                                 |
|     |                                           | Marjorie Hlava                                   |              |                                 |
| 8   | Supervision                              | Liz Allen, Amy Brand, Jo Scott, Micah Altman and  | 2014-04-26   |                                 |
|     |                                           | Marjorie Hlava                                   |              |                                 |
| 9   | Funding acquisition                       | Liz Allen, Amy Brand, Jo Scott, Micah Altman and  | 2014-04-26   |                                 |
|     |                                           | Marjorie Hlava                                   |              |                                 |
| 10  | Writing/manuscript preparation:          | Liz Allen, Amy Brand, Jo Scott, Micah Altman and  | 2014-04-26   |                                 |
|     | visualization/data presentation          | Marjorie Hlava                                   |              |                                 |
| 11  | Writing/manuscript preparation:          | Liz Allen, Amy Brand, Jo Scott, Micah Altman and  | 2014-04-26   |                                 |
|     | critical review, commentary or revision  | Marjorie Hlava                                   |              |                                 |
| 12  | Writing/manuscript preparation:          | Liz Allen, Amy Brand, Jo Scott, Micah Altman and  | 2014-04-26   |                                 |
|     | writing the initial draft               | Marjorie Hlava                                   |              |                                 |
| 13  | Resources                                | Liz Allen, Amy Brand, Jo Scott, Micah Altman and  | 2014-04-26   |                                 |
|     |                                           | Marjorie Hlava                                   |              |                                 |

Source: Credit where credit is due. NATURE | VOL 508 | 17 APRIL 2014, p. 312
### Semantic Linkage between Research Objects

| (1) Source Object | (2) Linked Object |
|-------------------|-------------------|
| **Source Object:** Паринов Сергей Иванович | **Linked Object:** *A Challenge of Research Outputs in GI Circuit: From Open Access to Open Use* |
| **Data Type:** person | **Class of Relationship:** Digital taxonomy to help researchers to identify their contributions to collaborative projects |
| **Semantic Meaning:** Methodology | **Data Type:** paper |
| **Semantic Code:** spz:parinov:adwzse:methodology | **Semantic Meaning:** Writing/manuscript preparation: writing the initial draft |
| **Semantic Code:** | **Semantic Code:** spz:parinov:adwzse:manuscript_writing |

### Author's Information

| **Author of the Linkage** | Parinov Сергей |
| ------------------------- | --------------|
| **Author's e-mail** | sparinov@gmail.com |
| **Author's Workplace** | ЦЭМИ РАН |
| **Author's profile** | Паринов Сергей Иванович |
| **Organization's profile** | Федеральное государственное бюджетное учреждение науки Центральный экономико-математический институт Российской академии наук |
| **Creation Date** | 2014-5-10 |
| **Revision Date** | 2014-5-10 |
SEMANTIC UPGRADE OF A LINKAGE BETWEEN THE AUTHOR AND THE PAPER

Source: http://socionet.ru/publication.xml?h=repec:rus:mqijxk:33&l=en
SHARING DATA ABOUT MOTIVATIONS

- Popular approach -
  - Doing their research scientists wish to share with the community the motivations to use some research outputs, how it was used, and how it impacted on their research

- Innovative approach -
  - Scientist wish to share with the community details on their tries and fails to use some research outputs in their research (sharing data on motivation do not use it)

- On the CRIS-CERIF side
  - Data about using some research output from the content of one CRIS should be shared with a metasystem (research e-infrastructure) and should be available for other CRIS systems
  - CERIF data model should support such data exchange
NOTIFICATION ECOSYSTEM

- A scientist wishes to receive immediate notifications: (1) if other scientists manipulated or used their research outputs; (2) if other scientists changed materials which were annotated or used by the scientist
- A scientist wishes to have an ability to react on cases of annotating (using) their research outputs, e.g. to help/assist with proper using of the outputs, or to protest against wrong using, etc
- On the CRIS-CERIF side
  - Data about authors of used research output from the content of one CRIS should be shared with a metasystem (research e-infrastructure) and should be available for other CRIS systems
  - CERIF data model has supported such data exchange
### USAGEMETRICS ECOSYSTEM

**Source:** [http://socionet.ru/stat-link.xml?h=repec:rus:ecoorg:cemi-ras_admin.45009-org1&l=en](http://socionet.ru/stat-link.xml?h=repec:rus:ecoorg:cemi-ras_admin.45009-org1&l=en)

**Total linkages:**
- By staff: 13,437
- By collections: 3,959

**Date:** 2014-05-12

**Object type:** institution-object

#### Outgoing linkages, total
- By staff: 6,214 (46%)
- By collections: 487 (12%)

#### Ingoing linkages, total
- By staff: 7,223 (54%)
- By collections: 3,472 (88%)

| Outgoing scientific relationships | Outgoing share |
|----------------------------------|---------------|
| research usage (4)               | 0%            |
| research material components (3) | 0%            |
| to organization from the person (1) | 0%        |
| author - research output relationship (3) | 0%     |
| to created collections (44 + 17) | 1%            |
| to relative materials (2)        | 0%            |
| personal opinion (11)            | 0%            |
| to staff from organization (103) | 2%            |
| to materials from the scientist (4963) | 80%       |
| hierarchy and association (4)    | 0%            |
| to cited materials (881)         | 14%           |
| to out-of-date versions (178)    | 3%            |

| Ingoing scientific relationships | Ingoing share |
|----------------------------------|---------------|
| research usage (10)              | 0%            |
| from collections (21)            | 0%            |
| research material components (3) | 0%            |
| from person to the organization (141) | 2%      |
| author - research output relationship (2650) | 37%    |
| from collections (9)             | 0%            |
| from relative object (2)         | 0%            |
| personal opinion (2)             | 0%            |
| from organizations to the person (104) | 1%      |
| hierarchy and association (3)    | 0%            |
| from citing materials (1131)     | 16%           |
| from materials to the organization (3144) | 44%     |
| from new versions (3)            | 0%            |
CHALLENGES OF CRIS-CERIF DEVELOPMENT

- A CRIS system should support necessary data exchange for global manipulations, sharing motivations, notifications and collecting usagemetrics
- We should incorporate them into CRIS architecture and CERIF data model

- A message: euroCRIS agenda on CRIS-CERIF development has to include new areas
  - making manipulations with RO, sharing motivations, sending notifications and processing usagemetrics