Open academic community in Poland: social aspects of new scholarly communication models as observed during the transformation period
Maciej Ostaszewski, Information Processing Center, OPI, Poland

Abstract
Digital revolution has contributed to fundamental changes in the way research is conducted and its results are propagated. The emergence of Open Science paradigms has reflected the very essence of those processes. It is only the last decade that the ideas of treating outcomes of research work as common goods have gained ever wider acceptance. The recent recommendation position statement of the European Commission sets a new landscape for future presentation of research results and their sharing. Member countries of EU have been requested there to define an agenda of the related implementation process. There are several actions in Poland in the course of implementation, ranging from a set-up of the national IT platform for scientific and technical information that will constitute a core of the national knowledge infrastructure, up to appropriate underlying legislation changes. Apart from ever more numerous activities contributing to the development of specific technical solutions, analyses are lacking in Poland that would show and provide diagnosis of the acceptance for those changes. The same lack is the case for the overall distribution of Open Science models in research work. Those transformation processes assume wide adoption of Open Access models. The latter refers not only to research publications, but also encompass diverse documentation forms for the research outcomes; educational content and looking into further future also research data. It is our aim to present initial results of a study addressing the social context of that transformation. We are conducting a survey on the attitude of the Polish academic community towards introduction of the Open Science concepts, with focus groups representatively selected. A study undertaken focused on:

- Diagnosis of the attitude and awareness level within the Polish academic and research communities against broadly viewed open models (open access, open data, open research, Science 2.0).
- Analysis of the range and level of the open science paradigms and tools implementation to research.

We use quantitative survey to collect data. The e-questionnaire based on LimeSurvey platform was sent to over 48K scientists holding at least PhD degree, registered in the “Polish Science” database operated by the OPI. The resulting conclusions are stratified according to disciplines, institutions and their categories, demographic criteria, in particular. Based on these filters, results will be attributed to several variables. The analytic data resulting from the research should enable diagnosis of potential barriers, bottlenecks and sources of fear within the academic community.

A Challenge of Research Outputs in GL Circuit: From Open Access to Open Use
Sergey Parinov, CEMI, Central Economics and Mathematics Institute, RAS Russian Academy of Sciences; Mikhail Kogalovsky, Market Economy Institute of Russian Academy of Sciences; Victor Lyapunov, Institute of Economics and Industrial Engineering Siberian Branch of Russian Academy of Sciences, Russia

Abstract
Open Access movement and currently formed GL circuit provide the scientific community with unique opportunity to modernize a fundamental part of research life-cycle: processes by which the scientists re-use research outputs when they produce new knowledge and then the community assesses their impact. Cameron Neylon marked in his comments («Nature», 2012, №492, pp. 348–349) an important for GL domain challenge: “Open access must enable open use” that means the “innovators can manipulate the material” and from the technical side it requests a “standardizing the representations of data and knowledge in ways that make them easily transferable”. When scientists mentally manipulate the research outputs, outcomes and other objects of scientific information space they discover relationships between the objects and thereby they re-use it to produce a new scientific knowledge. Some of these relationships become visible in scientists’ articles (e.g. by citations). Most of them are directly not observable and may exist in a mental form only.
CRIS technology, CERIF-based semantic linkage technique and available ontologies allow scientists to express explicitly their knowledge, opinions and hypotheses about scientific relationships between the objects as multilayer networks of semantic linkages over research information systems (RIS) content. Accumulated semantic linkages present a new data source for scientometric studies, which can partly compensate an absence of traditional peer-reviewing for GL instances. Started in 1997 as an open repository for Russian-writing researchers in Social Sciences the Socionet CRIS (socionet.ru) is currently developed to allow scientists an explicit expression of their opinions and/or scientific hypotheses on inference/deduction, impact/usage, logical hierarchy/associations, etc. about relationships between research information objects. The Socionet system uses the CERIF model of standardizing for the representations of data and knowledge and also the CERIF Semantics and SPAR ontologies converted to a form of semantic vocabularies. The CERIF based semantic linkage technique after some upgrades allows scientists to link different pairs of information objects from RIS content. The semantic meanings assigned by the scientists to the created linkages carry information about classes of relationships between research information objects. The scientific relationship classes are defined by taxonomy based on controlled semantic vocabularies produced from available ontologies. By setting semantic linkages between information objects scientists provide some kind of professional tagging and folksonomy over research information space. This data is accumulated and processed by the Socionet Statistics unit. This public service gives the community an opportunity of direct assessment of research usage/impact that complements the traditional mechanism of research assessment for publications and scientific journals.

Information support of research information interactions of PhD students in Slovakia Jela Steinerová and Andrea Hrčková, Comenius University Bratislava, Slovakia

Abstract
The support of research process on part of information behavior and information practices has long attracted attention of information science, especially in terms of digital scholarship and science 2.0 (Borgman 2007, Shneiderman 2008, Sonnenwald 2013). However, there is a need to manage large volumes of digital data in science and culture and new strategies and leadership for information environment of young researchers should be designed. The goal of this paper is to report on research of information needs of doctoral students in contexts of information support of the large volumes of digital research data and processes. Special attention is devoted to information interactions in the electronic environment. Based on the concept of information interactions (Fidel 2012) we concentrate on information problems and practices of doctoral students. The study was designed on the basis of previous research projects on relevance judgments of doctoral students and information ecology of the academic environment. The concept of the study conceptualized the following information interactions: research behavior, information use, information seeking, organization of information, information production, social media. We applied qualitative method of semi-structured interviews with doctoral students in different disciplines and visualization of information horizons. Results of data analyses confirm differences in information needs and information strategies in research behavior of doctoral students. Main information problems of doctoral students were identified, namely finding focus, expert support, networking and collaborative information behavior in real and virtual groups. As examples we mention several information horizons of doctoral students. In conclusion we propose special services and tools for the improvement of information interactions in digital environments. The final model of information interactions in different disciplines is described and recommendations for information portal design and information literacy improvement are proposed. Recommendations for modeling information practices in virtual environments with regard to large volumes of data and information behavior are also articulated.
The SK CRIS system as a source of unique information about scientific activities and their outcomes

Danica Zendulkova and Juraj Noge, Slovak Centre of Scientific and Technical Information, CVTI SR, Slovakia

Abstract

Majority of R&D outcomes consist of data, information and documents integrated into a category of Grey Literature. From portfolio of these outcomes, especially publications, patents, products and innovations are interesting for scientific community and wide range of R&D stakeholders indoors and abroad. Current Research Information Systems (CRIS) are one of key software tools for data collection and access, dissemination of information about scientific activities and their outcomes. The EU standards for these systems, primarily the CERIF data format, are covered by the international association euroCRIS. The aim of providing the Information System about Science, Research and Innovation SK CRIS is to integrate all available research information at the national level. The system was designed on a data structure compatible with the CERIF data format and it became a member of the CRIS systems family. The SK CRIS contains data about research projects funded by public resources, a registry of researchers, a registry of research organisations and research and development results. Data acquisition and refilling is possible by using integration interface importing data from external systems. It concerns mainly project data imported from systems of grant agencies. Direct data entry made by research subjects by online forms is the second way how to receive data. The SK CRIS integrates also other available external data sources: registries of researchers from universities and Slovak Academy of Sciences and also publications from universities registered in the Central Registry of Publication Activity. The SK CRIS makes contextual data about science and research accessible during their life cycle. It means that a user can see whether some document was created as a result of the project. The information about its authors and originators, persons and organisations is also available, as well as usage of a document and its citations. The SK CRIS is not limited to work only with metadata. It contains also the functionality allowing the integration of collected data with full text documents. The benefit of the SK CRIS is mostly its ability to offer aggregated information from entered data and to present their relationship. Main characteristic of the SK CRIS is the CERIF based concept of objects (entities) with attributes, with a data model guaranteeing interoperability and full language variability. The relations and semantics allow to record the objects with relationship and to characterise properties and time attributes by roles and time details. However, the system will be used mainly by scientific community but also by research management, decision makers and public. Enlargement of the SK CRIS functionality and integration of the other external data sources is our vision for near future. In the same time we consider the interconnection with other European CRIS systems. The SK CRIS information system was developed within the Activity No 4 of the NISPEZ national project implemented by the Slovak Centre of Scientific and Technical Information: Enlargement of Central Information Portal for Research, Development and Innovation (CIP RDI) with new functionalities complying with EU standards.

Auditing Grey in a CRIS Environment

Keith G. Jeffery, Consultant, United Kingdom and Anne Asserson, University of Bergen, Norway

Abstract

We define grey as information that is not peer reviewed scholarly publications. In the CRIS (Current Research Information System) domain Grey includes not only non-peer reviewed publications (typically technical papers) but also performance art, art artifacts, design documents, models (e.g. for engineering or architecture) and – increasingly importantly – research datasets. Recently national governments - inspired by the work of W3C on LOD (linked open data in the context of the semantic web) – have made government information available for citizen rights (the information should be available to those who paid for its collection) and business leverage – the information is used by businesses for strategic planning and a growing ICT business sector provides applications using LOD for business advantage. Government portals to publicly-funded data collections have been implemented (usually as data.gov.nn sites where nn = country) using ‘flat’ metadata standards such as DC (Dublin Core) and CKAN (Comprehensive Knowledge Archive Network). These metadata standards provide
limited information and – worse – it is very difficult to assure integrity of the associated dataset because there
is no integrity checking mechanism. Much of the presented material is of a summary nature and based on
more detailed research activities; rich metadata is available, both contextual (project, funding, persons,
organisations, related white publications etc.) and detailed (schema level for software to interact with the
dataset. This rich metadata improves dataset integrity in discovery and utilisation. The ENGAGE project
(http://www.engage-project.eu. portal at www.engagedata.eu) aims to enhance the metadata associated with
the usual data.gov portals with the rich metadata available around the supporting research datasets in order to
improve discovery and to ensure appropriate use of the datasets in context. A 3-layer model has been
proposed with discovery metadata (DC, CKAN and several others) generated from the contextual layer (CERIF:
Common European Research Information Format – an EU recommendation to Member States) which in turn
points to the detailed metadata associated with each domain or even individual dataset. The data mappings
from CERIF to/from each metadata format are done once and manually, thereafter conversion is automatic.
This architecture combines the easy browsing / discovery in the semantic web/LOD world with the formal
rigour of underlying rich metadata stored in information system with full integrity constraints. This architecture
allows confidence that any audit of the quality of the semantic web/LOD environment will be favourable
because it is generated from the underlying integrity-rich environment and not just generated by manual input
and linkage.