



PARSE.Insight

International Workshop on Preservation, Access and Re-use of Scientific Data

21 - 22 September 2009

**at the Georg-Christoph-Lichtenberg-Haus
Darmstadt, Germany**

PROGRAMME

Monday, 21st September 2009

09:30 Coffee and croissants

10:30 **Welcome and introduction**

David Giaretta (STFC)

10:45 **Scientific Data as a basic e-infrastructure for e-Science: an overview of European activities**

Carlos Morais Pires (European Commission)

11:30 **Overview of PARSE.Insight: context, objectives, results so far**

David Giaretta (STFC)

12:00 **Community Insight – approach and overview of results**

Jeffrey van der Hoeven (Koninklijke Bibliotheek)

13:00 Lunch

13:45 **Parallel group discussion sessions: “What has been missed and has been got right”**

14:45 **Case Study: High Energy Physics**

Salvatore Mele (CERN)

15:05 **Case Study: Social Sciences and Humanities**

Sabine Schrimpf (Deutsche Nationalbibliothek)

15:25 **Case Study: Earth Observation**

Vincenzo Beruti (ESA ESRIN)

15:45 Coffee break

16:00 **Parallel group discussion sessions: “Commonalities and differences between disciplines”**

17:00 **Feedback plenary session**

20:00 **Dinner at “Khan Restaurant – Mongolian Grill”** (Goebelstr. 21, Darmstadt)

Bus Shuttle to City Center and Restaurant for the Workshop Dinner

17:30 Conference room → City Center (Luisenplatz) *for recreation, shopping or checking-in*

19:30 City Center (Luisenplatz) → Restaurant Khan *to get there comfortably...*

23:00 Restaurant Khan → City Center (Luisenplatz) *...and return safely.*

PROGRAMME

Tuesday, 22nd September 2009

09:00 **Piloting an E-journals Preservation Registry Service (PEPRS)**

Peter Burnhill (EDINA)

09:30 **Threats and outline of the Roadmap**

David Giaretta (STFC)

10:15 **Parallel group discussion sessions: “Technical solutions”**

11:15 Coffee break

11:30 **Preservation, access and re-use of research data. A publisher's Perspective**

Joep Verheggen (Elsevier)

12:00 **Parallel group discussion sessions: “Non-technical solutions”**

13:00 Lunch

14:00 **Feedback plenary session**

14:30 **Gap Analysis: methodology, tool and first application**

Moritz Gomm (FernUniversität Hagen)

15:15 Coffee break

15:30 **Parallel group discussion sessions: “Gaps and how to fill them”**

16:15 **Final plenary discussion and plans on how to take things forward**

17:00 End of the programme

Workshop Documentation:

All presentations and documents from the workshop will be available after the workshop on the project web site: www.parse-insight.eu

Invited Talk on Monday, 21st at 10:45

Scientific Data as a basic e-infrastructure for e-Science: an overview of European activities

Carlos Morais Pires (European Commission)

Information Society and Media Directorate-General - Unit INFSO F3 (GÉANT and e-Infrastructures)

Scientific Data as a basic e-infrastructure for e-Science

The European Commission, in a coordinated effort with EU Member states, is investing in different domains of e-Infrastructures. The aim is to connect researchers, scholars, educators and students through high speed research networks (GÉANT), provide access to cooperative grid-computing platforms (e-Science grids), develop supercomputing capacity for very demanding applications (PRACE) and help the emergence of a robust platform for access and preservation of scientific information (*Scientific Data e-Infrastructure - SDI*).

In the Communication "*On Scientific Information in the Digital Age: Access, Dissemination and Preservation*" [1], the European Commission raises the attention to the importance of providing wide access and ensuring long term preservation of scientific information for building in Europe a dynamic information society. The document stresses that "*The Internet makes instant access to and dissemination of scientific information possible and new information and communication tools offer innovative ways to add value. They have opened new ways to use masses of data resulting from experiments and observations in the scientific process and to extract meaning from this data stored in repositories in combination with other scientific information resources. This leads to a 'continuum' of scientific information space from raw data to publications across different communities and countries*".

Many international initiatives confirm the growing importance of scientific data and its value as a resource for modern science and have been followed by key initiatives worldwide (US, Australia, etc).

At the European level the *Scientific Data e-Infrastructure* domain was launched in 2007 under the Capacities of FP7 (Seventh Framework Programme for research and technological development). Incorporating e-Science digital repositories and their holdings into an open information ecosystem will deepen and broaden Europe's *Single Information Space*. It will support new scientific methods and paradigms, improving the efficiency of the scientific process and its impact.

The Communication "*ICT infrastructures for e-Science*" [2] highlights the key role of ICT based infrastructures (*e-Infrastructures*) as a crucial asset underpinning European research and innovation policies. Building in Europe a high-quality and reliable repository infrastructure is a key component of a renewed strategy for *e-Science*, *e-Infrastructures*, and *innovation*. As there are many open challenges it is very important to pursue a consistent European strategy in close cooperation with EU Member States to strengthen e-infrastructures in support of e-Science.

References

[1] *On Scientific Information in the Digital Age: Access, Dissemination and Preservation*: Communication from the Commission to the European Parliament and the Council, COM(2007)56 final

[2] *ICT Infrastructures for e-Science*: Communication from the Commission to the European Parliament and the Council, COM(2009)108 final

Invited Talk on Tuesday, 22nd at 09:00

Piloting an E-Journals Preservation Registry Service (PEPRS)

Peter Burnhill

University of Edinburgh

Who is looking after e-journal content? The shift to e-journals means ease of online access to scholarly articles, but how to ensure that the world's cultural and scientific record will continue to exist and still be readable? PEPRS is a project to meet that challenge, by investigating and testing a registry with key information about preservation actions on e-journal content.

ABOUT THE PROJECT

Building on their long-standing association, the ISSN International Centre (ISSN-IC) and EDINA, the UK national academic data centre based at the University of Edinburgh, decided to work together in a project funded by the UK Joint Information Systems Committee (JISC). Although funded by a national body, the product of this cooperation promises international benefit.

The aim of this two-year project, which began in August 2008, is to enable librarians and policy-makers, initially in the UK but intended worldwide, to find out about archival provision and to help identify 'endangered' e-journals. The plan has been to design, build and to pilot an online facility, based upon self-reporting by the archiving agencies and use of the ISSN Register as the authoritative list of e-journals: 'if its worth preserving it should have an ISSN'.

Five archiving agencies agreed to participate: CLOCKSS & Portico; the British Library & e-Depot at the Koninklijke Bibliotheek(KB)/National Library of the Netherlands; the UK LOCKSS Alliance. These are working together with the project to help standardise statements about the type of archiving scheme used; extent of e-journal content being archived; access/release arrangements.

The outputs from the projects include:

1. project report, including definition of user/stakeholder requirements
2. architecture, with standards and interoperability (m2m) protocols
3. presentations at conferences and meetings across the globe
4. prototype and working demonstrator
5. business plan, with statement of the value proposition
6. project-to-service plan, for roll-out and launch of service

Halfway, the project is at Outputs 3 & 4. A review is scheduled in February 2010 to decide whether & how to support the preservation registry as a service.

Project Website & Publications: <http://edina.ac.uk/projects/peprs/index.html>

Peter Burnhill, Françoise Pelle, Pierre Godefroy, Fred Guy, Morag Macgregor, Christine Rees, Adam Rusbridge. *Piloting an e-journals preservation registry service*. *Serials* 22(1) March 2008.

Burnhill, P. *Tracking e-journal preservation: archiving registry service anyone?* *Against the Grain*. 21(1) February 2009. pp. 32,34,36

Archiving Agencies Participating in PEPRS:

- **CLOCKSS** <http://www.clockss.org/clockss/Home>

- Portico <http://www.portico.org/>
- British Library e-journals digital archiving initiative
- <http://www.bl.uk/reshelp/atyourdesk/docsupply/publisher/blejournal/index.html>
- e-Depot <http://www.kb.nl/dnp/e-depot/e-depot-en.html>
- UK LOCKSS Alliance <http://www.edina.ac.uk/lockss/>

Background Reports

Loughborough University and rightscom. *Scoping study for a registry of electronic journals that indicates where they are archived*, Rightscom Ltd. and Loughborough University, January 2008; available from: www.jisc.ac.uk/media/documents/programmes/preservation/ejournalregstudy.pdf

Jones, M, *Archiving E-Journals Consultancy – Final Report*. Report commissioned by the Joint Information Systems Committee, October 2003; available from: www.jisc.ac.uk/uploaded_documents/ejournalsfinal.pdf

Jones, M, *E-Journal Archiving (Review and Analysis of the CLIR Report E-Journal Archiving Metes and Bounds: A Survey of the Landscape)*, February 2007; available from: www.jisc.ac.uk/whatwedo/programmes/preservation/ejournalarchiving.aspx

Kenney, A R, et al *E-Journal Archiving Metes and Bounds: A Survey of the Landscape*. Washington DC.: Council on Library and Information Resources, September 2006; available from: <http://www.clir.org/pubs/reports/pub138/pub138.pdf>

Morrow, T, Beagrie, N, Jones, M and Chruszcz, J, *A Comparative Study of e-Journal Archiving Solutions*; a JISC funded investigation. Final Report May 2008; available from: http://www.jiscollections.ac.uk/media/documents/jisc_collections/reports/e_journals_archiving_%20solution_s_report_final_080518.pdf

Morrow, T, *Electronic Journals access and preservation: roles responsibilities and emerging solutions*. Presentation made at the United Kingdom Serials Group (UKSG) 2009 conference. Available at: http://www.uksg.org/event/conference09/breakout_sessions/

Research Information Network. *E-journals: their use, value and impact*. A Research Information network report. April 2009. Available at: http://www.rin.ac.uk/files/E-journals_use_value_impact_Report_April2009.pdf

Project Publications

Burnhill, P. *Piloting an e-journals preservation registry service*. Serials 22(1) March 2008. Available at: <http://uksg.metapress.com/link.asp?id=350487p5670h0v61>

Burnhill, P. *Tracking e-journal preservation: archiving registry service anyone? Against the Grain*. 21(1) February 2009. pp. 32,34,36

PEPRS Partner Organisations

EDINA www.edina.ac.uk and ISSN INTERNATIONAL CENTRE (ISSN IC) www.issn.org/

See also:

<http://www.issn.org/2-22638-ISSN-and-electronic-publications.php>
<http://www.issn.org/2-22637-What-is-an-ISSN-L.php>

Invited Talk on Tuesday, 22nd at 10:45

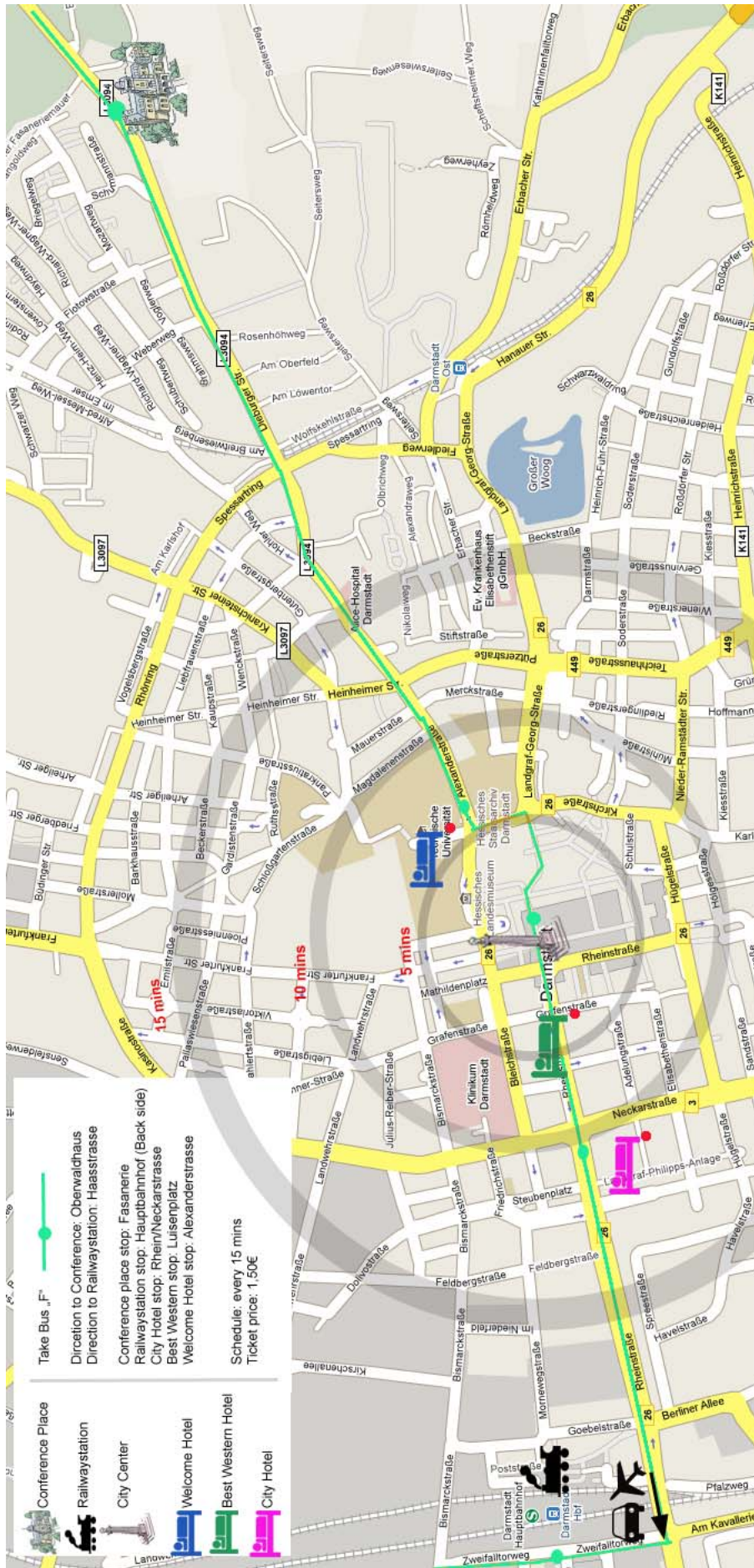
Preservation, access and re-use of research data. A publisher's Perspective.

Joep Verheggen,

Academic & Government Markets, Elsevier

The presentation provides a concise overview of the role of the Publisher with regard to the “publication process” of Research data, and more specifically how Publishers currently support submission/depositing and retrieval processes for research data and supplementary information. The presentation reviews how Publishers can contribute to a set of “support functions” for a sustainable and scalable environment for data preservation, access and re-use, while taking discipline specific characteristics into account, and meeting the needs of both “creators and users”. The presenter will demonstrate how involvement of Publishers helps to ensure that research data will become available for preservation and re-use, and can be better located, properly cited and understood in the right context especially in those scientific domains where the science communities are more loosely organised, and/or are more multidisciplinary in nature.

Map of Darmstadt



Restaurant for the Workshop Dinner

Monday, 21st at 20:00

“KHAN – Mongolian Grill”



Adresse:

Goebelstr. 21 in Darmstadt
Tel: +49(0)6151 - 800 30 33

(in front of Cinemaxx close to the main
railway station)



LIST OF PARTICIPANTS

First name	Last name	Organisation
Paul	Ayris	University College London Library Services
Vincenzo	Beruti	European Space Agency
Bob	Boelhouwer	Inst. voor Ned. Lexicologie
Jan	Brase	German National Library of Science and Technology
Holger	Brocks	Fernuniversität Hagen
Paolo	Budroni	University of Vienna
Peter	Burnhill	University of Edinburgh
Jonathan	Clark	Elsevier
Suenje	Dallmeier-Tiessen	Alfred-Wegener-Institut
Matthew	Day	Nature Publishing Group
David	Giaretta	STFC
Helen	Glaves	British Geological Survey
Moritz	Gomm	Fernuniversität Hagen
Katrin	Große	GSI, Helmholtz Association
Veronica	Guidetti	European Space Agency
Elif Aytek	Gürses	TUBITAK-ULAKBIM
Matthias	Hemmje	Fernuniversität Hagen
Benno	Hinnekindt	FWO
Markus	Höckner	University of Vienna
Simon	Hodson	Joint Information Systems Committee
Andras	Holl	Hungarian Scientific Research Fund (OTKA)
Andre	Holzner	CERN
Peter	Igo-Kemenes	CERN
Philipp	Ischer	Infoclio
William	Kilbride	Digital Preservation Coalition
Heike	Klingebiel	Springer
Kimmo	Koivunen	Data Services for Science and Culture
Simon	Lambert	STFC
Martin	Lhotak	Academy of Sciences Library

Salvatore	Mele	CERN
Patricia	Methven	Kings College London
Cliff	Morgan	Wiley Blackwell
Sandra	Oster	Deutsche Nationalbibliothek
Robert	Parker	Royal Society of Chemistry
Louise	Perbal	Ministry of Education, Culture and Science
Carlos Morais	Pires	European Commission
Agnes	Ponsati	Spanish National Research Council (CSIC)
Tomas	Prachar	Moravian Library
Keith	Rochford	Dublin Institute for Advanced Studies
Wouter	Schallier	LIBER, Association of European Research Libraries
Sabine	Schrimpf	Deutsche Nationalbibliothek
Barbara	Sierman	Koninklijke Bibliotheek
Eefke	Smit	International Association of STM Publishers
Stefan	Strathmann	UGOE
Heiko	Tjalsma	Data Archiving and Networked Services
Yasar	Tonta	Hacettepe University
Matthias	Töwe	ETH Zürich, ETH-Bibliothek
Jeffrey	van der Hoeven	Koninklijke Bibliotheek
Joep	Verheggen	Elsevier
Wolfgang	Voges	Max-Planck-Institut für extraterrestrische Physik
Bedrich	Vychodil	National Library
Björn	Werkmann	Fernuniversität Hagen
Stefan	Winkler-Nees	Deutsche Forschungsgemeinschaft
Peter	Wittenburg	Max Planck Institute for Psycholinguistics