

# How Openness can Change Scientific Practice

by Robert Viseur and Nicolas Devos

*The term ‘open data’ refers to “information that has been made technically and legally available for reuse”. Open data is currently a hot topic for a number of reasons, namely: the scientific community is moving towards reproducible research and sharing of experimental data; the enthusiasm, especially within the scientific community, for the semantic web and linked data; the publication of datasets in the public sector (e.g., geographical information); and the emergence of online communities (e.g., OpenStreetMap). The open data movement engages the public sector, as well as business and academia. The motivation for opening data, however, varies among interest groups.*

The publication of scientific data is the logical continuation of open access initiatives. Such initiatives, e.g., the Budapest Open Access Initiative, wish to achieve open access to scientific publications - i.e., access without financial, legal or technical barriers. Three criteria must be met for data to be considered open: (i) access to publications should be free; (ii) the data should not be protected by digital rights management (digital handcuffs) or scrambled, (iii) access policies should be clearly communicated and allow the copy, distribution and modification of data.

Murray-Rust notes that the lack of a consistent definition of the terms ‘open’ and ‘open access’ may cause confusion. The term ‘open’ should imply that the scientific data belongs to the scientific community and is made freely available. The accessibility of scientific publications and the related data is of great interest both for the validation of the research and the reuse of the data in new studies that may (or may not) have been foreseen by the original authors. The practice of data sharing is more common in some scientific disciplines, such as biosciences where the data is published and consolidated into databases. Nevertheless, some publishers aggressively defend their copyright and are opposed to new scientific practices that publicly disseminate

the research results with their source code, the data structures, the experimental design, the parameters, the documentation and figures [3].

The journal “Insight” shows how open access, open source and open data could change scientific practice. “Insight” is a peer reviewed online publication that is associated with the software Insight Segmentation and Registration Toolkit (ITK). ITK, which is supported by the company Kitware, is an open source software tool for image analysis. Scientific results are published with each article as per usual, but are also accompanied by the source code and the data in order to enhance the reproducibility of the research (‘reproducible research’) [1]. The technical infrastructure automates the source code compilation and testing. Several authors have developed the idea of ‘executable papers’, based on Linked Data and the Cloud infrastructure [2].

Open access to data can be fostered by governments. Governments can impose open access to research units, as illustrated by the UK Department for International Development, which in 2012, opened access to development research data in order to stimulate innovation.

Several issues will have to be addressed to facilitate the spread of open access to scientific data. Few peer reviewed journals support open data or possess the technical infrastructure needed to power the automation of the code execution, compilation and testing. Moreover, high ranking journals benefit from their advantageous market position and have no incentive to develop new publication methods. As a result, the mobilization of the entire scientific community, its commitment to play the openness game, and its support to high-quality innovative journals, will be fundamental to the success of these new collaborative practices.

## Links:

Insight: <http://www.insight-journal.org>  
Insight Segmentation and Registration Toolkit: <http://www.itk.org>  
Open Knowledge Foundation: <http://www.okfn.org>  
Panton Principles for Open scientific data: <http://www.pantonprinciples.org>

## References:

- [1] J. Jomier, A. Bailly, M Le Gall, et al.: “An open-source digital archiving system for medical and scientific research”, Open Repositories, 2010, vol. 7.
- [2] T. Kauppinen, G.M. de Espindola: “Linked open science-communicating, sharing and evaluating data, methods and results for executable papers”, Procedia Computer Science, 2011, vol. 4, p. 726-731.
- [3] V. Stodden: “The legal framework for reproducible scientific research: Licensing and copyright”, Computing in Science & Engineering, 2009, vol. 11, no 1, p. 35-40.

## Please contact:

Robert Viseur, Nicolas Devos  
CETIC, Belgium  
E-mail: [robert.viseur@cetic.be](mailto:robert.viseur@cetic.be),  
[nicolas.devos@cetic.be](mailto:nicolas.devos@cetic.be)

The screenshot shows a web interface for the Insight journal. It features two main sections: 'Code' and 'Reviews'.  
**Code Section:** Contains two entries for 'Automatic Testing Results' by 'Insight-Journal Dashboard'. The first entry is dated 'Fri Jan 22 22:26:54 2010 for revision #9' and has a 5-star rating. The second entry is dated 'Tue Oct 27 14:44:12 2009 for revision #6' and also has a 5-star rating.  
**Reviews Section:** Contains one entry titled 'Good contribution' by 'Dženan Zukić' dated '2011-05-30 09:14:31 for revision #11' with a 3-star rating. Below the review is a 'Summary:' section with the text: 'A valuable contribution, which spares us from implementing the N3 method. It has basic documentation and a very complex example, but it lacks a simplest example.' Below the summary are three sub-sections: 'Open Science:' with the text 'Source code provided, as well as some images.', 'Reproducibility:' with the text 'It compiles after ITK was configured to use reiew code.'

Figure 1: An extract from Insight illustrating the automatic testing and contribution process.

# ERCIM NEWS



www.ercim.eu



## Special theme: Scientific Data sharing and re-use

### Also in this issue:

**Keynote:**  
by Carlos Morais, European  
Commission

**Joint ERCIM Actions:**  
100 issues of ERCIM News  
ERCIM 25 Years Celebration

**Research and Innovation:**  
Simulations Show How  
Lightning Creates Antimatter

*ERCIM News is the magazine of ERCIM. Published quarterly, it reports on joint actions of the ERCIM partners, and aims to reflect the contribution made by ERCIM to the European Community in Information Technology and Applied Mathematics. Through short articles and news items, it provides a forum for the exchange of information between the institutes and also with the wider scientific community. This issue has a circulation of about 6,000 printed copies and is also available online.*

ERCIM News is published by ERCIM EEIG  
BP 93, F-06902 Sophia Antipolis Cedex, France  
Tél: +33 4 9238 5010, E-mail: [contact@ercim.eu](mailto:contact@ercim.eu)  
Director: Jérôme Chailloux  
ISSN 0926-4981

#### Editorial Board:

Central editor:

Peter Kunz, ERCIM office ([peter.kunz@ercim.eu](mailto:peter.kunz@ercim.eu))

Local Editors:

Austria: Erwin Schoitsch, ([erwin.schoitsch@ait.ac.at](mailto:erwin.schoitsch@ait.ac.at))

Belgium: Benoît Michel ([benoit.michel@uclouvain.be](mailto:benoit.michel@uclouvain.be))

Cyprus: Ioannis Krikidis ([krikidis.ioannis@ucy.ac.cy](mailto:krikidis.ioannis@ucy.ac.cy))

Czech Republic: Michal Haindl ([haindl@utia.cas.cz](mailto:haindl@utia.cas.cz))

France: Thierry Priol ([thierry.priol@inria.fr](mailto:thierry.priol@inria.fr))

Germany: Michael Krapp ([michael.krapp@scai.fraunhofer.de](mailto:michael.krapp@scai.fraunhofer.de))

Greece: Eleni Orphanoudakis ([eleni@ics.forth.gr](mailto:eleni@ics.forth.gr)),

Artemios Voyiatzis ([bogart@isi.gr](mailto:bogart@isi.gr))

Hungary: Erzsébet Csuhaaj-Varjú ([csuhaj@inf.elte.hu](mailto:csuhaj@inf.elte.hu))

Italy: Carol Peters ([carol.peters@isti.cnr.it](mailto:carol.peters@isti.cnr.it))

Luxembourg: Thomas Tamsier ([tamsier@lippmann.lu](mailto:tamsier@lippmann.lu))

Norway: Poul Heegaard ([poul.heegaard@item.ntnu.no](mailto:poul.heegaard@item.ntnu.no))

Poland: Hung Son Nguyen ([son@mimuw.edu.pl](mailto:son@mimuw.edu.pl))

Portugal: Joaquim Jorge ([jorgej@ist.utl.pt](mailto:jorgej@ist.utl.pt))

Spain: Sílvia Abrahão ([sabrahao@dsic.upv.es](mailto:sabrahao@dsic.upv.es))

Sweden: Kersti Hedman ([kersti@sics.se](mailto:kersti@sics.se))

Switzerland: Harry Rudin ([hrudin@smile.ch](mailto:hrudin@smile.ch))

The Netherlands: Annette Kik ([Annette.Kik@cwi.nl](mailto:Annette.Kik@cwi.nl))

W3C: Marie-Claire Forgue ([mcf@w3.org](mailto:mcf@w3.org))

#### Contributions

Contributions should be submitted to the local editor of your country

#### Copyright notice

All authors, as identified in each article, retain copyright of their work

#### Advertising

For current advertising rates and conditions, see

<http://ercim-news.ercim.eu/> or contact [peter.kunz@ercim.eu](mailto:peter.kunz@ercim.eu)

#### ERCIM News online edition

The online edition is published at

<http://ercim-news.ercim.eu/>

#### Subscription

Subscribe to ERCIM News by sending an email to

[en-subscriptions@ercim.eu](mailto:en-subscriptions@ercim.eu) or by filling out the form at the ERCIM

News website: <http://ercim-news.ercim.eu/>

#### Next issue

April 2015, Special theme: *The Internet of Things & the Web of Things*

## JOINT ERCIM ACTIONS

### 4 Our First 25 Years: 100 Issues of ERCIM News

by Peter Kunz, ERCIM News editor

### 6 ERCIM 25th Anniversary Celebration

### 8 Juan Reutter Receives the 2014 Cor Baayen Award

#### ERCIM White Papers

### 9 Big Data Analytics: towards a European Research Agenda

Editors: Mirco Nanni, Costantino Thanos, Fosca Giannotti and Andreas Rauber

### 10 Cyber-Security and Privacy

Editors: Javier Lopez, Fabio Martinelli, Pierangela Samarati

## KEYNOTE

### 12 e-Infrastructures Enabling Trust in the Era of Data-Intensive Science

by Carlos Morais, European Commission

## SPECIAL THEME

The special theme section "Scientific Data Sharing and Reuse" has been coordinated by Costantino Thanos, ISTI-CNR and Andreas Rauber, TU Vienna.

#### Introduction to the Special Theme

### 13 Scientific Data Sharing and Reuse

by Costantino Thanos and Andreas Rauber, guest editors for the special theme section

### 14 Creating the Culture and Technology for a Global Data Infrastructure

by Mark A. Parsons

### 15 If Data Sharing is the Answer, What is the Question?

by Christine L. Borgman

### 16 Enhancing the Value of Research Data in Australia

by Andrew Treloar, Ross Wilkinson, and the ANDS team

### 18 Beyond Data: Process Sharing and Reuse

by Tomasz Miksa and Andreas Rauber

### 19 Open Data – What do Research Communities Really Think about it?

by Marie Sandberg, Rob Baxter, Damien Lecarpentier and Paweł Kamocki

### 20 Providing Research Infrastructures with Data Publishing

by Massimiliano Assante, Leonardo Candela, Paolo Manghi, Pasquale Pagano, and Donatella Castelli

### 22 Sailing Towards Open Marine Data: the RITMARE Data Policy

by Anna Basoni, Stefano Menegon and Alessandro Sarretta

- 23 RDA: The Importance of Metadata**  
by Keith G. Jeffery and Rebecca Koskela
- 25 RDA: Brokering with Metadata**  
by Stefano Nativi, Keith G. Jeffery and Rebecca Koskela
- 26 Asking the Right Questions - Query-Based Data Citation to Precisely Identify Subsets of Data**  
by Stefan Pröll and Andreas Rauber
- 28 Capturing the Experimental Context via Research Objects**  
by Catherine Jones, Brian Matthews and Antony Wilson
- 29 Engineering the Lifecycle of Data Sharing Agreements**  
Mirko Manea and Marinella Petrocchi
- 30 Cross-disciplinary Data Sharing and Reuse via gCube**  
by Leonardo Candela and Pasquale Pagano
- 32 Toward Automatic Data Curation for Open Data**  
by Thilo Stadelmann, Mark Cieliebak and Kurt Stockinger
- 33 An Interactive Tool for Transparent Data Preprocessing**  
by Olivier Parisot and Thomas Tamisier
- 35 e-Infrastructure across Photon and Neutron Sources**  
by Juan Bicarregui and Brian Matthews
- 36 Understanding Open Data CSV File Structures for Reuse**  
by Paulo Carvalho, Patrik Hitzelberger and Gilles Venturini
- 37 How Openness can Change Scientific Practice**  
by Robert Viseur and Nicolas Devos

## RESEARCH AND INNOVATION

This section features news about research activities and innovative developments from European research institutes

- 38 Bridging the Gap between Testing and Formal Verification in Ada Development**  
by Claude Marché and Johannes Kanig
- 40 Simulations Show how Lightning Creates Antimatter**  
by Christoph Köhn and Ute Ebert
- 41 GROBID - Information Extraction from Scientific Publications**  
by Patrice Lopez and Laurent Romary
- 43 European Cyber-Security Research and Innovation**  
by Federico Maggi, Stefano Zanero, and Evangelos Markatos
- 44 A Single Password for Everything?**  
by Jan Camenisch, Anja Lehmann, Anna Lysyanskaya and Gregory Neven

- 45 SIREN – A Network Infrastructure for Emergencies**  
by Ioannis Askoxylakis, Paschalis Papagrigoriou, Diomedes Kastanis, Panos Karampelas, and George Spanoudakis
- 46 MYVISITPLANNER: Cloud-based Recommender for Personalized Tourism**  
by Ioannis Refanidis and Christos Emmanouilidis
- 48 Combinatorial Problem Solving for Fair Play**  
by Mats Carlsson
- 49 Optimizing Text Quantifiers for Multivariate Loss Functions**  
by Andrea Esuli and Fabrizio Sebastiani
- 50 Dynacargo: An Urban Solid Waste Collection Information System**  
by Artemios G. Voyiatzis and John Gialelis

## EVENTS, IN BRIEF

### Reports

- 51 Joint Collaborative Workshops of the ERCIM Dependable Embedded Software-intensive Systems Working Group**  
by Erwin Schoitsch
- 52 MUSCLE Working Group International Workshop on Computational Intelligence for Multimedia Understanding**  
by Maria Trocan, Emanuele Salerno and Enis Cetin
- 53 Networking: IFIP TC6 2014 Dagstuhl Meeting**  
by Harry Rudin

### Announcements

- 54 ASQT 2015 - 13th User Symposium on Software Quality, Test and Innovation**
- 54 Teaching, Education and Training for Dependable Embedded Systems at SEAA/DSD 2015**
- 54 ERCIM/EWICS/ARTEMIS Workshop at SAFECOMP 2015**

### In Brief

- 55 The 2014 Nobel Prize in Physiology or Medicine for two professors at NTNU**
- 55 Antoine Petit new Chairman and CEO of Inria**
- 55 Manfred Hauswirth new Executive Director of Fraunhofer FOKUS**
- 55 The Fondazione Cassa di Risparmi di Lucca is Co-funding two Projects Proposed by two Institutes of the National Research Council in Pisa**