

WHAT IS BEHIND THE JOINT RESEARCH CENTRE'S PUBLICATIONS REPOSITORY?

Jesus Felix POZUELO MORENO

Document Management Officer, European Commission's Joint Research Centre (JRC)

▪ Le *Répertoire des Publications du CCR* est un service en ligne qui donne accès aux données des publications des recherches réalisées par le Centre Commun de Recherche (CCR) de la Commission Européenne. Ce service est mis en place pour aider la Centrale de stockage numérique à la gestion et à la recherche de nos publications scientifiques. Dans la mesure du possible, l'accès à la version intégrale du texte est disponible pour le grand public.

▪ De *JRC Publications Repository* is een onlinedienstverlening die toegang verschaft tot de wetenschappelijke publicaties van het Joint Research Centre (JRC) van de Europese Commissie. Deze dienst werd in het leven geroepen ter ondersteuning van de centrale opslag, het beheer en het opzoeken in de gepubliceerde JRC-onderzoeksresultaten. De resultaten worden indien mogelijk full text aangeboden.

The mission of the Joint Research Centre (JRC) is to provide customer-driven scientific and technical support for the conception, development, implementation and monitoring of European Union policies. As a service of the European Commission, the JRC functions as a reference centre of science and technology for the Union. Close to the policy-making process, it serves the common interest of the Member States, while being independent of special interests, whether private or national.

The principal customers of the JRC are the policy-making Directorates-General of the European Commission, in particular Directorates-General for Energy, Mobility and Transport, Enterprise and Industry, Climate Action, Environment, Agriculture and Rural Development, Health and Consumers, and Research, Technology and Development.

Into this scenario and during the summer of 2004, the JRC¹ decided to create a system to support the management of publications of all JRC institutes located in different countries (Belgium, the Netherlands, Germany, Spain and Italy). This new repository is aligned with the JRC's mission.

The main objectives of the repository are to improve the visibility and availability of the research conducted by the JRC, to promote knowledge sharing between the JRC staff, and to store and preserve institutional digital assets and internal documents which are otherwise easily lost ("grey" literature, e.g. technical reports, technical notes, oral presentations, and posters).

The *JRC Publications Repository* on the internet was built to share with the general public the wealth of knowledge produced in the JRC.

Corporate publications management system

Publications Management System Architecture

The structure of the JRC repository is based on three pillars. The first pillar is an intranet portal which helps authors with information about the rules, categories of publications, last publication registered in the system, useful links, like JRC Corporate templates, *Social²Sciences Citation Index* and *Science³ Citation Index* expanded, etc. The second pillar is an approval workflow built with Java and Oracle.

The third pillar is the repository built with DSpace⁴, an open source software which provides tools for managing digital collections and commonly used as an institutional repositories solution. This part is accessible to the general public.

Workflow

The **P**ublications Management **S**ystem (PUBSY) is the corporate publications management system of the JRC that gives visibility and availability of the research conducted by the JRC researchers.

The *PUBSY* workflow (fig.1) is composed of four main steps, supported by the *PUBSY* workflow system.

1. **Publication request** (done by the applicant by submitting a request in PUBSY). Submitting a request creates an electronic record card for the publication, in which are included meta-data like: kind of publication, title, abstract, authors, format, language/s, year and each typology of publication has and specific group

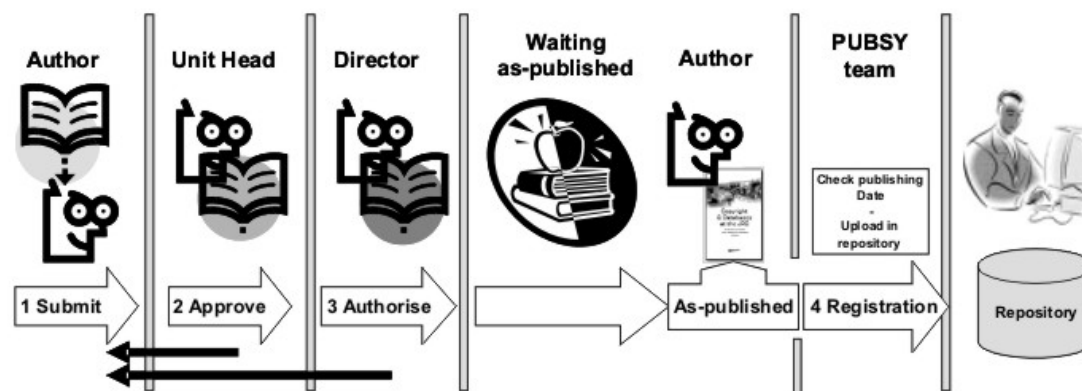


Fig. 1: PUBSY workflow.

of secondary metadata, i.e. a poster or oral presentation has the event name, where and when has been presented the work, etc...The card, in which all further operations shall be reported, is given a unique request number.

The applicant cannot modify data once he/she has submitted a publication request, until this has been authorised by the Institute Director, or unless it has been rejected.

The applicant can access the record card again after temporary rejection, to make the necessary changes, or after authorisation to complete the bibliographic information and to attach the as-published version.

2. **Publication approval** (by the applicant's Head of Unit, responsible for the scientific and technical part) Unit Heads can modify the bibliographic data of a submitted request but only before approving it. They cannot modify data once the request is approved.
3. **Publication authorisation** (by the applicant's Director, responsible for policy and corporate image of their institute and by default of the JRC). Directors can modify the bibliographic data of an approved request but only before authorisation. They cannot modify data once the request is authorised.
4. **Publication registration** in a repository (by the JRC corporate publication office, after the work has been published).

In principle, steps 1 to 3 must be completed before the work is sent out for publication (to a journal, a conference, a publisher, etc.) as

they relate to the authorization process.

Step 4 can only be completed after the work has been made public. When the work has been published, the process is concluded.

Data management (Data quality requirements and responsibilities)

PUBSY is a tool for actors (authors, speakers and editors) and managers to track and make queries about their publications. The data quality is important to ensure good queries and it is directly reflected in the *JRC Publications Repository* available to external users.

It is the responsibility of the applicant to ensure data quality and completeness in the first place before submitting the request for approval. Instructions on how to introduce data in *PUBSY* are provided in each field of the request with a small definition. As sample and returning to the oral presentation or poster in the field "Event venue",

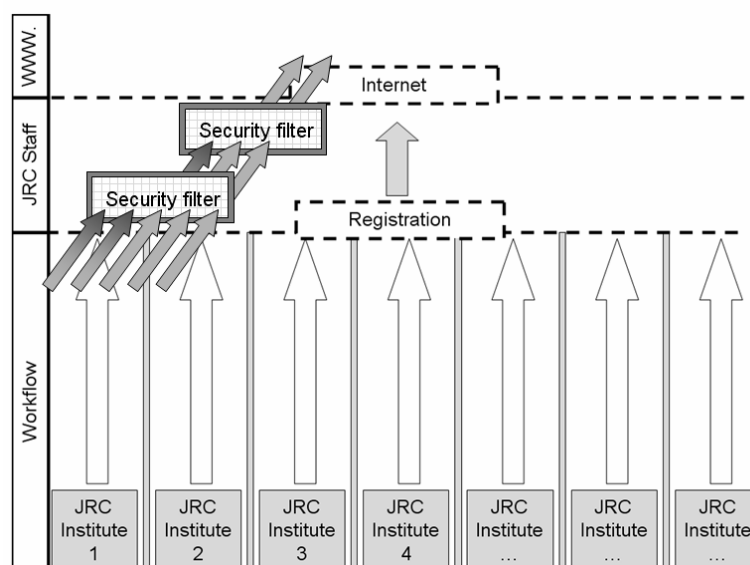


Fig. 2: Security filters and visibility levels.

we will obtain a pop up when the mouse pointer is over field with the following text "The venue refers to the place where the conference is held. The format is: City (Country)". These instructions need to be rigorously adhered to in order to ensure data quality.

The purpose of collecting data is to store all metadata related to publications and to keep a track of the life cycle of JRC scientific publications, from conception to publication. Thus preserving both a full-text digital copy and bibliographic information including references to the R&D programmes and JRC Institutes' activities as well as the approval and authorization from the hierarchy.

The database is a key instrument to generate statistical, non-nominative, reports on the publication activity in the JRC.

Security and preservation

Authentication of users of the system and access rights

All users must be authenticated to use the system, and the user's roles and privileges are derived from this authentication and information about his/her position in the organisation. During the approval process a given publication, record is only visible to the hierarchy involved. After successful registration, the final publication, unless marked for limited distribution, becomes visible to all staff in the JRC.

If intellectual property rights (IPR) provisions permits it subsequently gets published also on the JRC internet site otherwise the repository can not offer the final publication only the bibliographic information to the publisher. Security filters are put in place at different stages in order to ensure that internal documents and publications with a security marking or specific protection are not disclosed inadvertently (Fig. 2).

Preservation of data?

The long-term preservation of data requires tackling complex technology issues. For instance to avoid technical obsolescence of hardware and software regular updates and data migration operations must be enacted, on top of the ordinary data backups.

This, however, is regarded as necessary, not only in order to preserve the scientific knowledge contained

in the registered publications, but also for a matter of liability.

The impact of research is often seen over a very long term: a scientific publication may be quoted or questioned, either on the content or in intellectual propriety rights terms, and/or for liability reasons for the duration of ten years, or more.

It follows that we may need to produce evidence of a scientific statement made by a JRC author that was fully backed by the organisation, and was not merely the author's point of view on the matter. For this reason, all data in PUBSY is preserved without a time limit.

Now available on the Internet

The *JRC Publications Repository*⁶ is now available on the Internet. Bibliographic data of almost 10 000 articles are available to the public and more than 1700 reports of theoretical or experimental studies, on a specific subject, carried out at the JRC or under contract (EUR series) are freely available for download.

Publications can be found by free-text search, or an advanced search based on bibliographic data. The publications can be browsed by institute, publication year, author, or title and cross-links are supplied to find all the works of a specific author.

All publications are available on the internet 24 hours after registration in PUBSY internal database (fig.3).

The JRC Publications Repository has been jointly developed by the PUBSY team in the Knowledge Management and Evaluation Unit and the Informatics Unit.

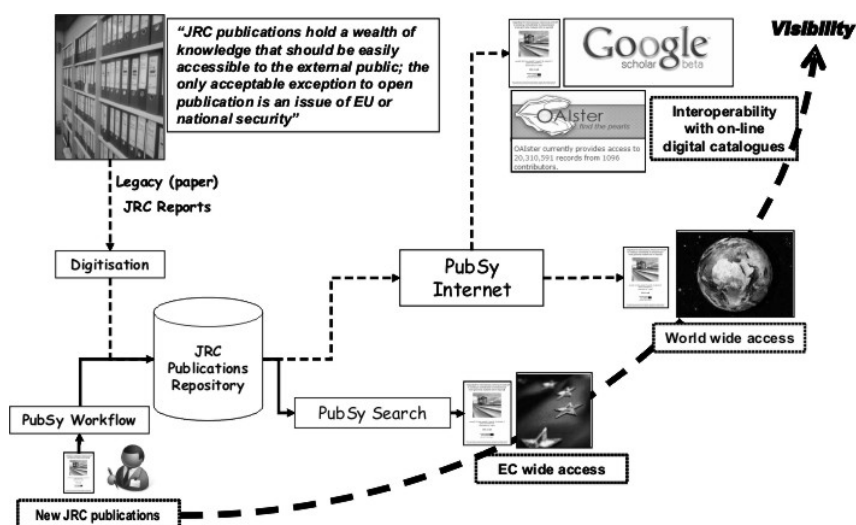


Fig. 3: Accessibility to JRC research output.

The project has been implemented in cooperation with the PUBSY steering committee, the Intellectual Property and Technology Transfer Unit, the Internal and External Communication Unit and the Programme Support Network.

PUBSY now plays a consolidated role for approval, registration, archiving and retrieval of scientific publications in the JRC and is used for monitoring and reporting of the publishing activity.

The JRC Publications Repository will be extended over the next couple of years to provide access to

additional scientific resources and data available in the JRC.

Jesus Felix Pozuelo Moreno

Joint Research Centre

Via Enrico Fermi 2749

21027 Ispra (VA)

Italy

Jesus.pozuelo-moreno@ec.europa.eu

<http://ec.europa.eu/dgs/jrc/>

[http://publications.jrc.ec.europa.eu/
repository/](http://publications.jrc.ec.europa.eu/repository/)

September 2010

References

The Plone Foundation [on line] <<http://plone.org>> (consulted on 10 September 2010).

The web site for the Zope Community [on line] <<http://www.zope.org>> (consulted on 10 September 2010).

Python. *Website of Python programming language* [on line] <<http://www.python.org>> (consulted on 10 September 2010).

Oracle. *Java Platform, Enterprise Edition* [on line] <<http://download-llnw.oracle.com/javaee/>> (consulted on 10 September 2010).

Oracle [on line] <<http://www.oracle.com>> (consulted on 10 September 2010).

The Open Archive Initiative [on line] <<http://www.openarchives.org>> (consulted on 10 September 2010).

Notes

¹ *Joint Research Centre of the European Commission* [on line] <<http://ec.europa.eu/dgs/jrc/index.cfm>> (consulted on 10 September 2010).

² Thomson Reuters. *Social sciences citation index* [on line] <<http://www.thomsonscientific.com/cgi-bin/jrnlst/jloptions.cgi?PC=J>> (consulted on 10 September 2010).

³ Thomson Reuters. *Science citation index* [on line] <<http://www.thomsonscientific.com/cgi-bin/jrnlst/jloptions.cgi?PC=D>> (consulted on 10 September 2010).

⁴ *DSpace software* [on line] <<http://www.dspace.org>> (consulted on 10 September 2010).

⁵ *Joint Research Centre Publications Repository* [on line] <<http://publications.jrc.ec.europa.eu/repository/>> (consulted on 10 September 2010).