CIP Competitiveness and Innovation Framework Programme
CIP-Pilot Actions, 2007-2013
CIP-ICT-PSP-2012-6

Project **CIP-Pilot 325101 / OpenScienceLink**
Deliverable **D7.1.2**
Distribution **Public**

http://opensciencelink.eu

Report on the Operations of the OpenSienceLink Pilot Services and Services/Processes Improvements

Authors: Adomas Bunevicius, Alicja Juskiene, Inesa Birbilaite, George Tsatsaronis, Michael Schroeder, Matthias Zschunke, Julian Mendez, Nikolaos Saklampanakis, Freyja Van den Boom, Griet Verhenneman, Giorgio Iervasi, Todor Tagarev, Iordanis Mourouzis, Constantinos Pantos

Status: Final (Version 1.0)

February 2015

Project

Project ref.no.	CIP-Pilot 325101
Project acronym	OpenScienceLink
Project full title	Open Semantically-enabled, Social-aware Access to Scientific Data
Project site	http://opensciencelink.eu
Project start	February 2013
Project duration	3 years
EC Project Officer	David Guedj

Deliverable

Deliverable type	Report
Distribution level	Public
Deliverable Number	D7.1.2
Deliverable title	OpenScienceLink Evaluation Framework
Contractual date of delivery	M24
Actual date of delivery	February 2015
Relevant Task(s)	WP7/Tasks 7.1, 7.2, 7.3, 7.4, 7.5
Partner Responsible	LUHS
Other contributors	TUD, TI, PROCON, CNR, KU Leuven
Number of pages	29
Authors	Adomas Bunevicius, Alicja Juskiene, Inesa Birbilaite, George Tsatsaronis, Michael Schroeder, Matthias Zschunke, Julian Mendez, Nikolaos Saklampanakis, Freyja Van den Boom, Griet Verhenneman, Giorgio Iervasi, Todor Tagarev, Iordanis Mourouzis, Constantinos Pantos
Reviewers	Vassiliki Andronikou, Efstathios Karanastasis
Status & version	Final (Version 1.0)
Keywords	Operation of Pilot Services, Service Improvements, Validation of Pilot Operations

Executive Summary

This deliverable reflects the outcomes of tasks T7.1-T7.5 in the second year of the project. The deliverable focuses in the pilot operations of the OpenScienceLink project and the services that implement those via the OpenScienceLink platform. The validation of the five main pilot scenarios of the project is presented via real case studies. In addition, the deliverable reflects the outcome of task T7.5 by illustrating the suggested improvements in the various added-value services, and reports on the improvements made within Year 2, as well as the improvements that are planned for the third year of the project. The overall presentation of the reported results is based on five use cases that stem from the requirements specification documents of the OpenScienceLink project.

Contents

List of Figures

List of Tables

1 Introduction

This deliverable focuses on the operation of the OpenScienceLink pilot services. In particular, the following are reported:

- the current operations of all five pilots;
- the validation of the current state of affairs of every pilot;
- improvements that were made to the pilot services within year 2, and suggestions of improvements to be conducted within year 3.

The reporting activities for each pilot correspond directly to the five tasks of the Work Package 7, thus the deliverable will be structured in a task-wise manner. Below is the description of the tasks.

T7.1 Data Journals Lifecycle Management: Validation of the pilot scenarios that are associated with data journals creation and management. As part of the task the five pilot sites engaged researchers in the publishing, classification and linking of experimental datasets. The validation took place via the launched open access Biomedical Data Journal.

T7.2 Research Reviews and Evaluation: Validation of the added-value review process, which enables reviewers and editors to have flexible and instant access to the resources required for the objective, fast and effective completion of the review process. Editors of the Biomedical Data Journal are facilitated by the OpenScienceLink platform in selecting competent reviewers.

T7.3 Management and Visualization of Metrics: Validation of identifying research trends. Different scenarios associated with trends and metrics were piloted in order to assess the efficiency of the services and their ability to adapt to user requirements/needs (regarding trends and metrics). The scenarios that were adopted utilized the focus and the research interest of the scope that the Biomedical Data Journal has.

T7.4 Researchers Collaboration and Linking: Validation of suggesting collaborations between researchers. The operation of this pilot scenario leverages the social networking and semantic search services of the OpenScienceLink platform, via the notion of the co-authorship graphs as thee can be inferred by the literature.

T7.5 Services and Processes Improvements: This task was devoted to the reengineering of processes associated with the pilot scenarios that are validated in the above tasks. The task documents on issues identified as part of the previous tasks and it devises remedial actions at the process level. Process refinements are accordingly integrated to the OpenScienceLink platform, in order to be validated and (re)assessed. Additional suggestions of improvements for year 3 are identified and graded in complexity.

2 Per-Pilot Operation Activities Analysis

In this section we report the progress of every pilot and list the respective activities carried out within year 2.

2.1 Pilot 1 Operations

In the preparation of the pilot Procon Ltd., the publisher of the Biomedical Data Journal (BMDJ), launched the journal website, http://biomed-data.eu (also the domain http://datajournals.eu is reserved for the launching of additional journals in the future), agreed with consortium partners of journal policies and the composition of the BMDJ Editorial Board, as well as to waive the authors' fee for publication in the BMDJ for 2014 and 2015.

Jointly with the Editorial Board, it prepared and launched four calls for papers: three calls for special issues (one - for policy papers on open access to biomedical data, and two calls for data sets - see the table below) and one standing call. Another Call for Papers is being prepared jointly with LUHS. The four calls for papers were disseminated via the journal website, by editors and other consortium partners, and via the Wiki for calls for papers, as shown in the table below.

Editor(s)	Topic	URL at the Wiki CFP	Views by August 20, 2014
Constantinos Pantos	Open Access to Experimental and Clinical Data	http://wikicfp.com/cfp/servlet/event.showcfp?eventid=39857	607
Yixin Zhang	High Throughput Drug Screening	http://wikicfp.com/cfp/servlet/event.showcfp?eventid=39856	706
	Standing Call	http://wikicfp.com/cfp/servlet/event.showcfp?eventid=39858	899
Alessandro Pingitore and Giorgio Iervasi	Heart Failure and Stress Response	http://wikicfp.com/cfp/servlet/event.showcfp?eventid=39859	897

Table 1: Call for papers of the BMDJ journal announced in 2014.

To be able to get indication on the interest, a Google Analytics account for the BMDJ website was set, as well as a site counter. The site of Wiki CFP also counts views (in the last column of the table) and registered users who track the respective call for papers.

During the second year of the project, the following features were introduced to the platform implementing operations related to Pilot 1, namely the submission process:

Operation	Description
Inclusion of categorisation of datasets.	The author can choose between research data and clinical trial. The platform is prepared to add more categories, if this is needed in the future.
Stricter verification of fields belonging to the dataset metadata.	This stricter verification helps the author find incomplete or wrongly formatted fields.
Removal of owner's name of submitted dataset.	The submitter must be also author of the submitted dataset.
Improvement of explanations to fill out the dataset metadata.	The page now includes a small explanation of the fields that are not self-explanatory.
Integration of counter to measure the number of downloads per month.	Provides useful statistics regarding the data traffic on the platform.
Inclusion of dates of submission, review, acceptances, and publication of datasets.	Operations are now time-stamped.

Table 2: Features introduced in Year 2 in OSL platform for Pilot 1.

2.2 Pilot 2 Operations

This pilot aims at offering a novel peer-review process methodology, which consists of the following four steps: Automated search for and suggestion of appropriate reviewers, support in carrying out the review process, organization of review form submission and post-review discussion.

The current version of the OSL platform distributes the functionality of pilot 2 between the roles of a reviewer and an editor – and between the respective tabs of the platform. The following services of pilot 3, described in D3.1, section 4.2.1, have been implemented so far:

- Issuing a review call (4.2.1.1)
- Relevant reviewer identification through literature (4.2.1.3)
- Reviewer pool suggestion (4.2.1.5)
- Reviewer selection and invitation (4.2.1.7)
- Reviewer form submission (4.2.1.10)

With regards to the operations of Pilot 2, the following features were introduce to the platform:

Operation	Description
Inclusion of review form.	The review form that must be filled by the reviewer is provided by the platform. The filled-out review is transformed into sentences before being sent to the author.

Implementation of the DOI publication module.	The editor can publish a submitted dataset and article giving them a DOI (Digital Object Identifier). The published material is openly accessible without the need of registration.
Integration of reminders for reviewers.	The editor can send e-mail reminders to reviewers.
Integration of management of issues.	The editor can create, modify, and delete issues.
GUI improvements.	Improvements of the visual interface, especially in the flow and navigability.

Table 3: Features introduced in Year 2 in OSL platform for Pilot 2.

2.3 Pilot 3 Operations

The aim of Pilot 3 is to detect and visualize research trends in biomedical domain. The Platform integrates all pilot services under the same search tab. Hence, as in other pilots, the trend detection functionality can be queried with any textual string, be it a biomedical term, a gene name, a name of a researcher or a research organization, a location etc. After the user enters a query, the platform analyzes PubMed data sources, and the trend search results are retrieved and displayed on the same page, in a user-friendly graphical manner. For example, Figure 1 illustrates the trend analysis results for querying “High Throughput Drug Screening”, which is the main focus of one of the BMDJ special issues, and shows a very promising trend.

The graph presents the number of published articles dealing with the specific scientific topics per year, arranged in chronological order, which also includes an estimated number for the current year. It also shows the relative research interest per year. These values are used in order to create a sweep, which is then smoothed. From the form of the sweep, it is possible can make some presumptions regarding the trendiness of the specific research topic.

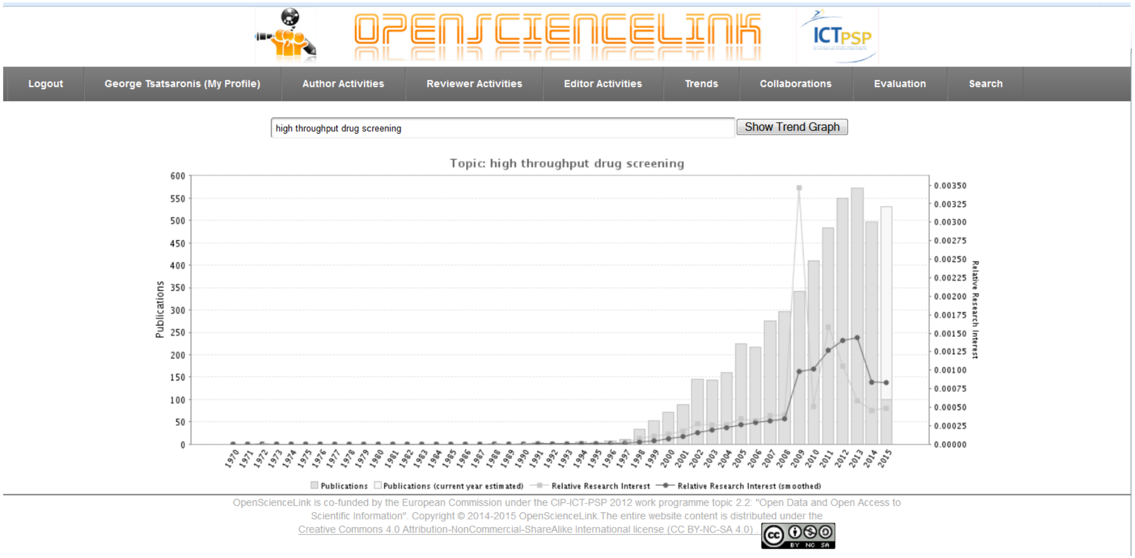


Figure 1. An example of the OSL trend graph for the query "High Throughput Drug Screening".

The following improvements were made in the platform regarding the operations of Pilot 3:

Operation	Description
Improvement of the visual interface.	This is mainly concerning the initial query and the case when no results are found.
Integration of counter to measure the number of trend searches.	A counter that measures the number of trend searches performed by each user.

Table 4: Features introduced in Year 2 in OSL platform for Pilot 3.

2.4 Pilot 4 Operations

This Pilot aims at assisting the networking and collaboration of researchers and scholars working on similar scientific fields. It facilitates the creation and tracking of networks of researchers. However, a key distinguishing characteristic of this Pilot from conventional research information systems is the ability to infer relationships between researchers and research groups, including non-obvious, non-declared relationships.

A researcher logs in the Platform in order to identify researchers working on similar fields or sharing similar interests. Identification of network of researchers based on expertise and/or interest could be useful to establish new collaborations, define partners in grant proposals or organize lectures in congresses and meetings. After specifying his/her fields of interest and expertise, the Platform provides a list of suggested researchers for potential collaboration. Alternatively, without specifying fields of interest, the Platform infers the fields based on the publications of the user and provides the list of suggested researchers.

A user wishes further to identify research groups or laboratories interested in other scientific areas and fields in order to possibly expand his research or organize a workshop. In this case, the researcher specifies his field of interest to be taken into account manually and the Platform provides a list of the top authors in this field. The results are presented along with their relevant data, and the researcher can filter and sort the results based on these data and according to their personal preferences (e.g. relevance to the subfields, excellence in the field, location, affiliation, date). The results are also presented in the form of a network, where connections between different researchers are provided (see Figure 2).

The suggested collaborations among researchers are on the basis of the degree of relevance of the research topics and fields they work on, as indicated across their published work and/or their participation in research communities. The researcher can request for receiving regular notifications about suggestions of scientific collaborations with researchers, research groups and/or communities via e-mail.

The operations referring to pilot 4 that were improved are listed in the following table:

Operation	Description
Improvement of the visual interface.	This is mainly concerning the initial query and the case when no results are found.
Integration of counter to measure the number of collaboration searches.	A counter that measures the number of collaboration searches performed by each user.

Table 5: Features introduced in Year 2 in OSL platform for Pilot 4.

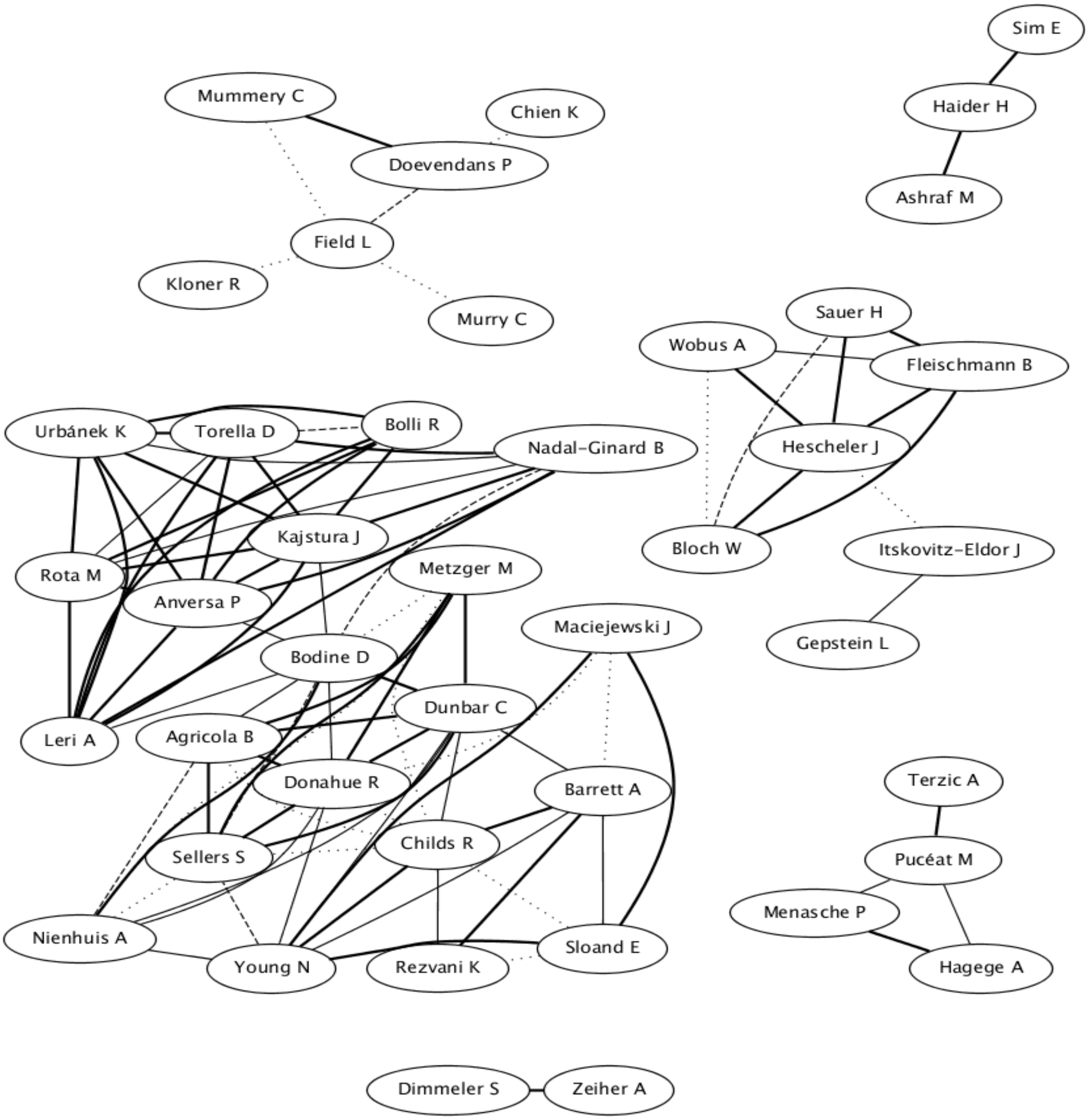


Figure 2: Network of top authors for the field "stem cells and heart".

2.5 Pilot 5 Operations

The aim of Pilot 5 is to introduce, produce and track new objective metrics of research and scientific performance, beyond conventional metrics associated with conventional indices and impact factors.

The current version of the OSL platform provides evaluation of the following **entities**: countries, cities, journals and authors. As with other pilots, the query term that is evaluated can be of different kinds, e.g., a research area, a scientist or an institution. The **evaluation metric** that is currently implemented is OpenScore (deployed in the current version only for authors). This metric takes into account a large set of openly accessible features (e.g., number of papers, journals, collaborators) to provide a final evaluation for each author. In Figure 3 a screen shot produced for the evaluation of all entities using the query "stem cells and heart" is illustrated, showing also the OpenScore values for the authors.

Future operations on Pilot 5 will mainly focus on expanding the OpenScore functionality to the rest entities. In particular, we are going to expand the evaluation for different research units: universities, institutes, centers, laboratories etc. For this purpose, a special partonomy has been created from all the affiliation strings of PubMed, that identifies all individual entities (universities, labs, departments, etc.). The partonomy will be included along with the rest OpenScore measures within Year 3.

1 2 3 ...17

Countries	Publications
United States	94,698
Japan	26,297
China	23,863
Germany	17,980
United Kingdom	15,185
California, United States	13,667
Italy	12,899
France	11,982
Massachusetts, United States	9,730
Canada	9,038
New York, United States	8,715
Maryland, United States	7,791
South Korea	6,456
Texas, United States	5,761
Pennsylvania, United States	5,562
Spain	5,498
Netherlands	5,323
Australia	5,229
Sweden	3,833
Illinois, United States	3,509

1 2 3 ...17

1 2 3 ...122

Cities	Publications
Boston	6,587
Tokyo	5,829
New York City	5,670
London	5,225
Beijing	4,140
Paris	4,045
Los Angeles	3,998
Bethesda	3,948
Seoul	3,610
Houston	3,523
Shanghai	3,418
Philadelphia	3,162
Baltimore	2,703
Seattle	2,609
Toronto	2,606
Chicago	2,437
Roma	2,228
Guangzhou	2,204
Milano	2,056
San Francisco	2,022

1 2 3 ...122

1 2 3 ...299

Journals	Publications
Blood	10,481
J Immunol	5,579
Plos One	5,343
J Biol Chem	4,591
Bone Marrow Transpl	3,593
Exp Hematol	3,523
Stem Cells	3,384
P Natl Acad Sci Usa	3,081
Leukemia	2,778
Cancer Res	2,680
Brit J Haematol	2,491
Nature	2,191
J Exp Med	2,049
Development	2,016
Biochem Bioph Res Co	1,891

1 2 3 ...5473

Authors	Publications	OpenScore
Lee S	8	14
Passweg J	7	60
Socié G	7	77
Humphreys B	6	18
Soleimani M	6	22
Asakura A	5	18
Bornhäuser M	5	73
Ghavamzadeh A	5	12
Redl H	5	73
Zhdanov V	5	18
Baharvand H	4	18
Bianco P	4	40
Clevers H	4	63
Fibbe W	4	54
Fischer A	4	73

Figure 3: Evaluation of entities based on the results of the query "stem cells and heart".

Finally, the following enhancements were introduced to the platform during the second year, we regards to pilot 5 operations:

Operation	Description
Implementation and integration of the OpenScore.	An evaluation rating for the author as calculated by the platform, based on a score produced by a supervised machine learning model that uses a set of 12 publicly available features to predict the h-index of authors.
Integration of counter to measure the number of evaluation searches.	A counter that measures the number of evaluation searches performed by each user.

Table 6: Features introduced in Year 2 in OSL platform for Pilot 5.

3 Case-Study Validation of Pilots

In this section we carry out a validation of pilot activities based on respective case studies. The section acts as a transition between the description of current pilot activities and future improvements, motivating the latter. The validation is done in a case-study manner, focusing on scenarios that involve the launched Biomedical Data Journal (BMDJ). The starting point of every case is similar: an editor of Biomedical Data Journal wants to create a new issue of the journal, and investigate trends and key people. This task is complex and involves multiple subtasks, e.g., determining the relevant issue topics, evaluating trends and impact, and managing the submission and publication of data papers via the platform. Every pilot case will focus on one of such subtasks.

3.1 Pilot 1 Case Study

Situation: The first step towards the creation of a new journal issue is *to open a call* for it. As soon as the topic of the issue is chosen (see Pilot 3 Use Case), the editor creates a description of the issue with further details about its scope, expected format of submissions etc. and publishes it as an official call for papers.

Actions: The OSL platform provides the functionality to create a new call for papers from scratch. The editor logs into the platform and goes to the “Editor Activities” tab, where the management of calls can be conducted. Every call has a title, submission and review deadlines, and the publication date. In order to open a new call, the editor clicks on the “Add Issue” button. The “Add Issue” page has a form with meta-information fields: title, volume, description, keywords, and respective deadlines. The editor fills in the title and description fields (in our example the issue is dedicated to lymph node biopsy procedures, as it is the outcome of Pilot 3 Use Case). Based on the textual data from these two fields, the platform is able to suggest relevant keywords (see Figure 4); alternatively, the editor enters the keywords himself. After all the necessary data is filled in, the editor clicks on “Save Changes” button.

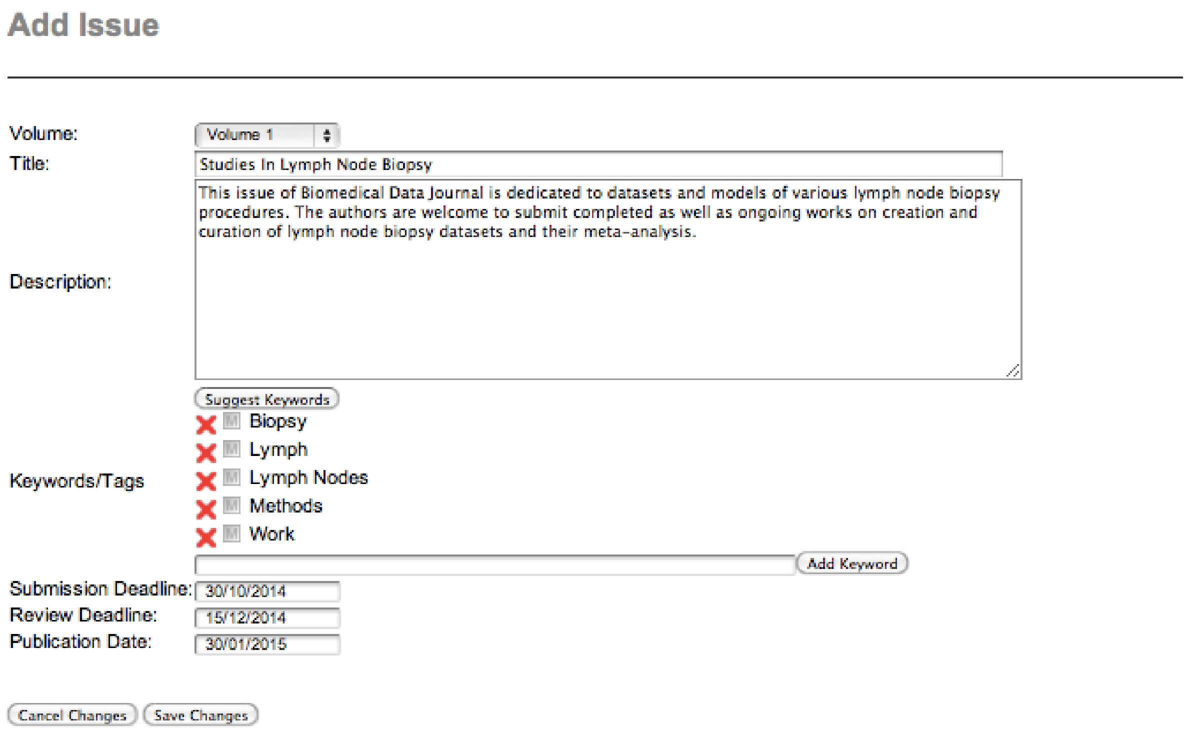


Figure 4: The "Add Issue" page of the OSL platform.

Results: The newly created call is now visible in the “Editor Activities” tab. The editor is able to modify and delete the call, as well as to see the submissions made for this call. Once the call is created, it is considered active and it is visible from authors in the submission process.

Challenges: The logical controls of contradicting deadlines between issues, e.g., two new calls for special issues that have very similar deadlines. This problem can only be addressed by the editors at this level with internal communication of the journal’s editorial board.

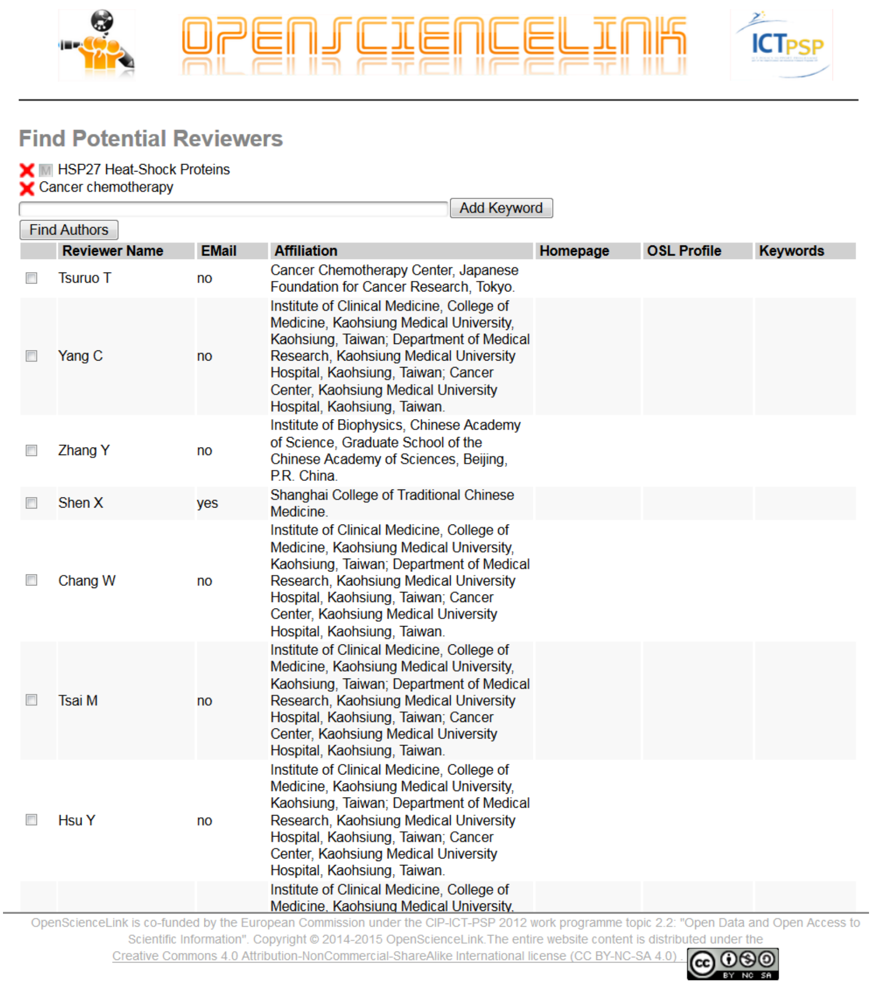
3.2 Pilot 2 Case Study

Situation: After the call for papers for the new journal special issue is published, the platform is ready to receive submissions for this call. Every submission is processed by the editor, who launches the peer-review assessment and makes an accept-reject decision based on the reviews. The complete *review process* is enabled by the OSL platform.

Actions: The user, logged in as editor, goes to the “Editor Activities” tab. For every journal issue that is then visible in the “Available Issues” list it is possible to see all the submitted datasets by clicking on the leftmost icon in the operations column. For every submission it is possible to inspect its metadata and to invite reviewers. Reviewers’ suggestion is made in an automated manner by the platform. By clicking on the right icon in the operations column of a certain submission, the editor is able to launch the review process for this submission (see Figure 5).

Figure 5: The "Invite Reviewers" page of a given submission. The submission meta-data and a list of the already invited reviewers is shown.

The editor clicks on the “Create Invitation” button, and starts the process of finding people with relevant background that can potentially be the reviewers for a given submission by clicking on “Select Reviewers” button. The OSL platform returns a result list of potential reviewers with their affiliations and email (see Figure 6). The editor is able to modify, shrink and expand the original list of keywords provided by the authors of the submission, in order to improve the reviewers search. This is a type of query reformulation process in information retrieval. After several iterations the editor selects several candidate reviewers and sends them an invitation via the platform by clicking on the “Invite Reviewers” button and filling in personalized emails (see Figure 7).



OPENSCIENCELINK

ICTPSP

Find Potential Reviewers

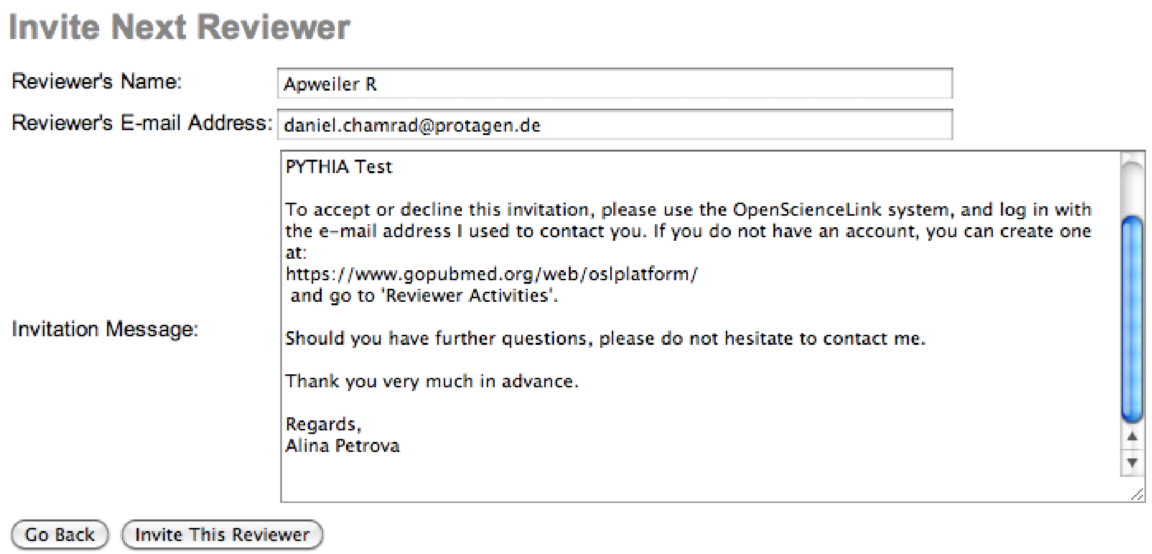
HSP27 Heat-Shock Proteins
Cancer chemotherapy

Add Keyword

Find Authors

	Reviewer Name	EMail	Affiliation	Homepage	OSL Profile	Keywords
	Tsuruo T	no	Cancer Chemotherapy Center, Japanese Foundation for Cancer Research, Tokyo.			
	Yang C	no	Institute of Clinical Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan; Department of Medical Research, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan; Cancer Center, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan.			
	Zhang Y	no	Institute of Biophysics, Chinese Academy of Science, Graduate School of the Chinese Academy of Sciences, Beijing, P.R. China.			
	Shen X	yes	Shanghai College of Traditional Chinese Medicine.			
	Chang W	no	Institute of Clinical Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan; Department of Medical Research, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan; Cancer Center, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan.			
	Tsai M	no	Institute of Clinical Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan; Department of Medical Research, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan; Cancer Center, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan.			
	Hsu Y	no	Institute of Clinical Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan; Department of Medical Research, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan; Cancer Center, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan.			
			Institute of Clinical Medicine, College of Medicine, Kaohsiung Medical University,			

Figure 6: "Find Potential Reviewers" page with a result list of researchers for a given submission.



Invite Next Reviewer

Reviewer's Name: Apweiler R

Reviewer's E-mail Address: daniel.chamrad@protagen.de

Invitation Message:

PYTHIA Test

To accept or decline this invitation, please use the OpenScienceLink system, and log in with the e-mail address I used to contact you. If you do not have an account, you can create one at:
https://www.gopubmed.org/web/oslplatform/
and go to 'Reviewer Activities'.

Should you have further questions, please do not hesitate to contact me.

Thank you very much in advance.

Regards,
Alina Petrova

Go Back | Invite This Reviewer

Figure 7: Sending personalized emails to potential reviewers via the OSL platform.

Results: All submissions that are considered for evaluation are processed by the editor, who assigned reviewers to each submission and is waiting for their replies: whether they accept or reject the review invitation, and in case they accept, which reviewer decision they make.

Challenges: Efficient monitoring and tracking of reviewer activity from the editor perspective

3.3 Pilot 3 Case Study

Situation: An editor of BMDJ wants to *identify promising topics* for the new issue of the journal, i.e., topics that have good dynamics and have gained an increased interest in the research community. The topic should either be popular and trendy at the moment, or it should belong to an area that is not yet very popular, but has been gaining popularity in the recent years. To analyse the research activity in different areas and to select an interesting topic, the editor uses Pilot 3 functionality.

Actions: The editor navigates to the "Trends" tab and queries the OSL platform with various biomedical terms, e.g., names of broad areas of research or specific proteins and compounds, with institutions and centers or even single scientists, that conduct research on particular topics. The editor analyses the trend graphs for every query, in particular the "Relevant Research Interest" plots and selects the topic for the issue.

Suppose the editor is interested in studies on lymphatic system. Per se, the area attracts consistent amount of publications every year (see Figure 8), but the research interest in studying the lymphatic system in general is decreasing, which means there are specific problems related to lymphatic system that are actively studied nowadays, and the task of the editor is to single them out. For that, the editor modifies the query in the search tab. He is then provided with context suggestions of MeSH terms that are related to the original query. One of such suggestions is "Sentinel Lymph Node Biopsy". Upon choosing this query, the editor sees the graph that clearly illustrated the research focus on this problem in the past 15 years (see Figure 9). By generalizing the query to "Lymph Node Biopsy" the editor finds the topic, which is popular and has been gaining more and more research attention in the past years (see Figure 10). "Lymph Node Biopsy" is then chosen as the topic for the new issue.

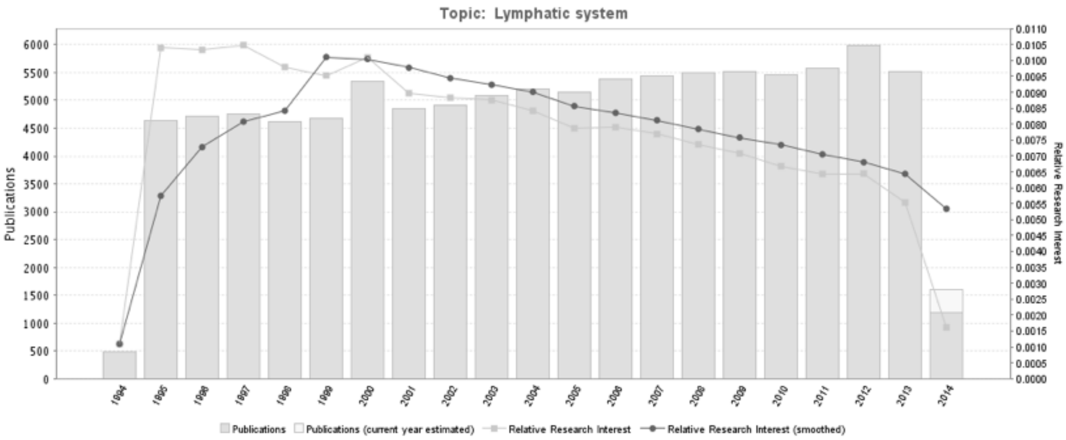


Figure 8: The trend graph for the query "Lymphatic system"

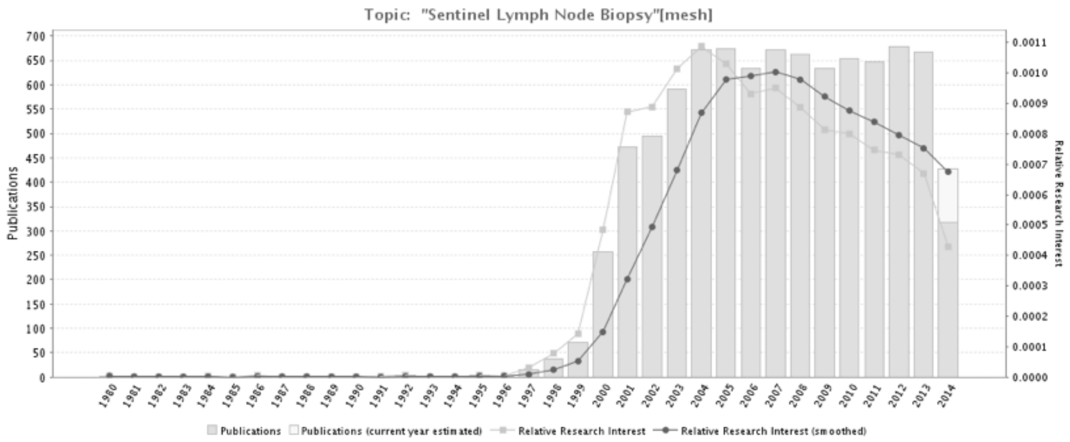


Figure 9: The trend graph for the query "Sentinel Lymph Node Biopsy".

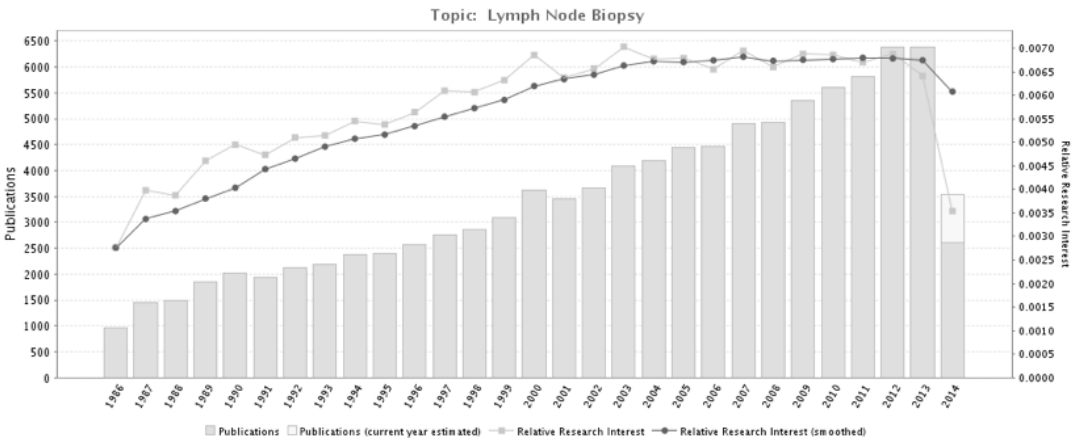


Figure 10: The trend graph for the query "Lymph Node Biopsy".

Results: The new issue is dedicated to a topic that is trendy, hence, it has the potential to attract more submissions and, consequently, to publish more high-impact articles.

Challenges: The main challenge of this case is that, while Pilot 3 provides assistance to the editor, it is still up to the editor: (1) to come up with queries, (2) to analyse and compare the queries, and (3) to make the final decision. Hence, in section 4.3 we have suggested how the functionality of the pilot should be expanded in order to facilitate the decision-making process of the user.

3.4 Pilot 4 Case Study

Situation: In order to create a new issue of the Biomedical Data Journal, the editor needs to collect all submissions and to assess them using the peer-review scheme. As it was mentioned in Pilot 1, the platform provides assistance in finding and inviting the reviewers that are relevant to the topic of the submitted article. However, to ensure that the reviewing is unbiased, the editor who assigns reviewers to articles must verify that there is no conflict of interests involved. In particular, one should not make a positive decision about the paper only because the author and the reviewer have previously worked together. At the final stage of the platform conflicts of interests should be detected and resolved automatically; meanwhile, the editor can use the Pilot 4 functionality to check the co-authorship graph of the reviewer and to look for any of the article authors in it.

Actions: The editor goes to "Collaborations" tab and queries the platform with the reviewer's name. The platform retrieves the co-authorship graph of the reviewer. Figure 11 illustrates such a graph for Prof. Michael Schroeder of TU Dresden. Note that since the OSL platform is using PubMed as the main source of co-authorship information, the graph below represents only the papers that belong to biomedical domain; all other papers written by Michael Schroeder and belonging to data mining, text mining etc. do not contribute to this graphs.

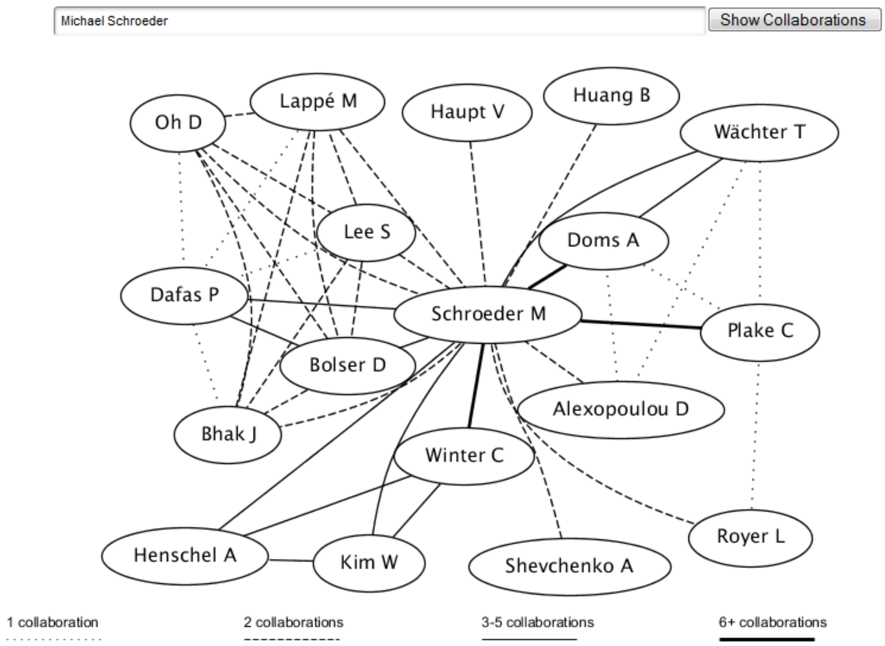


Figure 11: The co-authorship graph for Michael Schroeder.

Results: The editor checks that the graph does not contain any of the potential reviewers that he has chosen, therefore the editor proceeds with sending them official review invitations via the platform (see Pilot 2 Use Case).

Challenges: For now the conflict of interests resolution is not automated; the manual checking of co-authorship graph can help detect certain types of bias that exists between a reviewer and an author, but not all of them. For example, the two researchers may not have written a paper together, but they are still working in the same group, or project and are willing to help each other. These cases will be detected automatically once the affiliation partonomy is fully integrated into the platform (within Year 3).

3.5 Pilot 5 Case Study

Situation: Suppose the editor would like to include into the new issue an article from an invited author. The author could be a well-known specialist in the area to which the new issue is dedicated, and the article could be an overview of the area, or the discussion of the latest advances in the area.

However, the editor is not necessarily an expert in the topic of the issue, thus, he may need assistance in finding a candidate invited author. The Pilot 5 functionality may help him in identifying one.

Actions: The editor goes to the "Evaluation" tab and queries the platform with the main topic of the issue, in our case with "lymph node biopsy" (see Figure 12). The platform returns lists of top countries, cities, journals and authors that are connected to this topic. The editor is mainly interested in the authors list.

Figure 12: The journals and authors with most publications on lymph node biopsy.

For every retrieved author the editor can query the platform for his profile with email address and affiliation, as well as to check whether the author has been occupied with the given topic lately, or (s)he did this research quite some time ago. In our use case the editor goes to the "Trends" tab and queries the platform with the surname of the second author from the list combined with the topic of the issue: "Ross lymph node biopsy". The resulting graph is given on Figure 13. As one can see from the trend curve, for the past 10 years this scientist has been actively working on the topic the editor has chosen, which makes him a perfect candidate for writing an invited article.

Results: The editor is able to identify researchers that are able to write a high-impact and insightful invited article for the new issue of BMDJ, without the editor necessarily being the expert in the topic of the issue.

Challenges: The main challenge pertains with the disambiguation of the entities, a process that is addressed by the clustering-based disambiguation of the integrated GoPubMed platform.

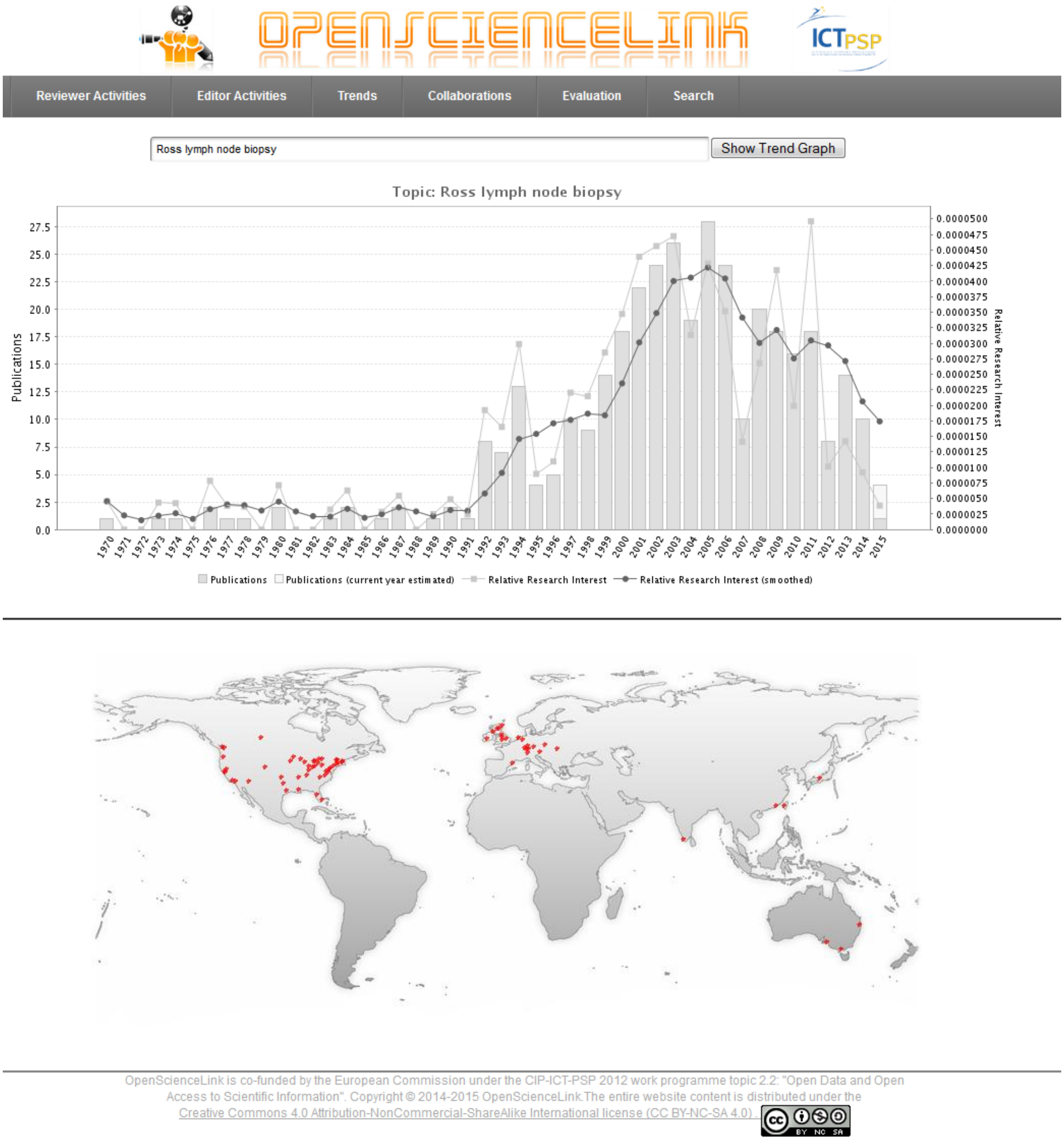


Figure 13: The trend graph for a topic "lymph node biopsy" and a researcher named Ross.

4 Suggested Pilot Improvements

In this section we list the OSL platform improvements that should be integrated in later releases. The improvements are grouped per pilot and are linked to the respective functionalities of the platform that stem from initial user requirements (see D3.1). The improvements that were suggested in year 1 of the project have been analyzed with regards to their implementation complexity. They have been rated in a scale from 1 to 9, with those rated with 1, being considered as easily implementable, and those with 9 being very hard to realize. Final changes will be reported in the updated D2.1 D3.1 D3.2.

4.1 Pilot 1 Improvements

Improvement ID	Platform Improvement Description	Related Functionality ID	Related Component	Status	Complexity
I1.1	The Comment/Discussion functionality for published reviews of research works or datasets should be included in future releases of the platform.	F2.8, F2.9	Datasets Management	Analyzed / implementation pending	5
I1.2	Relevant datasets detection should be included in future releases of the platform.	F1.1	Datasets Management	Analyzed / implementation pending	1
I1.3	Datasets selection and publication should be included in future releases of the platform.	F1.3, F1.4	Datasets Management	implemented	
I1.4	The functionality for specifying the papers in which the dataset to be uploaded has been used, could be improved so that the user selects the datasets rather than having to type manually their identification number.	F1.2	Datasets Management	Analyzed / implementation pending	3

Table 7: Pilot 1 improvements.

4.2 Pilot 2 Improvements

Improvement ID	Platform Improvement Description	Related Functionality ID	Related Component	Status	Complexity
I2.1	Conflict of interest detection should be included in future releases of the platform.	F2.3	Groups management (affiliations), Users Management, Reviews Management	Analyzed / implementation pending	6
I2.2	Use of author networks for reviewer identification should be included in future releases of the platform.	F2.1	Authors Management, Reviews Management	Analyzed / needs to be revised	7
I2.3	Retrieval of related information for the uploaded article should be included in future releases of the platform.	F0.3, F2.6	Reviews Management, Articles Management	Analyzed / implementation pending	1
I2.4	Automated suggestion of missing references should be included in future releases of the platform.	F2.7	Reviews Management	Analyzed / needs to be revised	7
I2.5	Automated retrieval of cited bibliography should be included in future releases of the platform.	F2.6	Reviews Management, Articles Management	Analyzed / implementation pending	5
I2.6	Automated retrieval of related datasets to suggest comparisons should be included in future releases of the platform.	F0.4	Reviews Management, Articles Management	Analyzed / implementation pending	1
I2.7	Reviews of other evaluators should be made available in future releases of the platform.	F2.9	Reviews Management	Analyzed / implementation pending	3
I2.8	Open discussion of reviews and review results should be	F2.9, F2.10, F2.11	Reviews Management	Analyzed / implemen	5

	made available in future releases of the platform.			tation pending	
I2.9	Reviewer's rating. Track the performance of reviewers and present reviewers scores to editors	F1.8	Reviews Management	Analyzed / implementation pending	5

Table 8: Pilot 2 improvements.

4.3 Pilot 3 Improvements

Improvement ID	Platform Improvement Description	Related Functionality ID	Related Component	Status	Complexity
I3.1	Sorting of topics by trendness should be made available in future releases of the platform.	F3.2	Data mining Processes, Articles Management	Analyzed / implementation pending	3
I3.2	A suggestion of topics by trendness (increasing or declining fields) with regard to a given set of topics should be made available in future releases of the platform.	F3.2	Data mining Processes, Articles Management	Analyzed / implementation pending	4
I3.3	A justification of the level of trendness (along with the relevant data) should be presented to the users in future releases of the platform.	F3.2	Web Interface	Analyzed / implementation pending	5

Table 9: Pilot 3 improvements.

4.4 Pilot 4 Improvements

Improvement ID	Platform Improvement Description	Related Functionality ID	Related Component	Status	Complexity
I4.1	Researchers' connections identification through	F4.1	Authors Management	Analyzed / needs to be revised	9

	social networks should be provided in future releases of the platform.				
I4.2	Researchers' connections identification through literature could possibly be improved by presenting detailed connections information to the end users, and enabling them to fine tune the connections identification parameters.	F4.2	Articles Management Authors Management	Analyzed / needs to be revised	9
I4.3	Researchers' collaboration filtering and proposition should be implemented in future releases of the platform.	F4.3	Authors Management	Analyzed / implementation pending	2
I4.4	Research groups identification through both social networks and literature should be supported by future releases of the platform.	F4.4	Groups Management	Analyzed / implementation pending	6

Table 10: Pilot 4 improvements.

4.5 Pilot 5 Improvements

Improvement ID	Platform Improvement Description	Related Functionality ID	Related Component	Status	Complexity
I5.1, I5.2, I5.4, I5.5	Open Score for articles (I5.1), journals (I5.2), research groups (I5.4), and institutions (I5.5).	F5.1, F5.2, F5.3, F5.4, F5.5, F5.6, F5.8, F5.9, F5.10, F5.11, F5.12	Articles Management	pending	5
I5.3	OpenScore for researchers.	F5.7	Authors Management	done	

Table 11: Pilot 5 improvements.

5 Conclusions

In this deliverable we presented the current outcomes of tasks T7.1-T7.4 in terms of the pilot operations implemented by the respective OpenScienceLink platform services. The report detailed the validation of the five main pilot scenarios of the project. In addition, it reflected the outcome of task T7.5 by illustrating the suggested improvements for the five pilot scenarios of the project, coupled with the respective rationale that led to these suggestions. As a result, the current document, along with D5.2.1, will constitute the basis for the improvements of the five pilot operations of the project and the respective OpenScienceLink platform services.

6 References

OpenScienceLink Consortium. (2013). *OpenScienceLink: OpenSemantically-enabled, Social-aware Access to Scientific Data.* EC.

OpenScienceLink D3.1, Revision 1. "Detailed Specification of OpenScienceLink Pilot Scenarios and Services".