



PIRUS 2

organizational, economic and political issues

Peter Shepherd

COUNTER

The PIRUS project has been funded by JISC. the UK Joint Information Systems Committee



PIRUS: the challenge

- An article may be available from:-
 - The main journal web site
 - Ovid
 - ProQuest
 - PubMed Central
 - Authors' local Institutional Repositories
- If we want to assess article impact by counting usage, how can we maximise the actual usage that we capture?



PIRUS: benefits

- Reliable usage data will be available for journal articles, wherever they are held
- Repositories will have access to new functionality from open source software that will allow them to produce standardised usage reports from their data
- Publishers will be able to provide their authors with more reliable usage statistics
- Digital repository systems will be more integral to research and closely aligned to research workflows and environments
- The authoritative status of PIRUS2 usage statistics will enhance the status of article-level usage reports
- The standard can be extended to cover other categories of content stored by repositories



PIRUS1 (Sept 2008-Jan 2009): -outcome

The PIRUS1 project demonstrated that it is *technically* feasible to create, record and consolidate usage statistics for individual articles using data from repositories and publishers, despite the diversity of organizational and technical environments in which they operate. If this is to be translated into a new, implementable COUNTER standard and protocol, further research and development will be required, especially into the technical, organizational and economic issues.



PIRUS2: issues to be addressed

- **Technical**: further tests, with a wider range of repositories and a larger volume of data, will be required to ensure that the proposed protocols and tracker codes are scalable/extensible and work in the major repository environments.
- **Organizational**: the nature and mission of the central clearing house/houses proposed by PIRUS1 has to be developed, and candidate organizations identified and tested
- **Economic**: assess the costs for repositories and publishers of generating the required usage reports, as well as the costs of any central clearing house/houses; investigate how these costs could be allocated between stakeholders
- **Political**: the broad support of all the major stakeholder groups (repositories, publishers, authors, etc) will be required.



PIRUS2 (Oct 2009-Feb 2011): - aims and objectives

The aim of PIRUS2 is to address these issues and by doing so specify standards, protocols, an infrastructure and an economic model for the recording, reporting and consolidation of online usage of individual articles hosted by repositories, publishers and other entities.

Specific objectives:

- Develop a suite of free, open access programmes to support the generation and sharing of COUNTER-compliant usage data and statistics that can be extended to cover any and all individual items in repositories
- Develop a prototype article-level publisher/repository usage statistics service the Central Clearing House (CCH)
- Define a core set of standard useful statistical reports that repositories should produce for internal and external consumption

PIRUS2: organizational issues

- Specifications for the Governance of PIRUS, going forward
- define the nature and mission of the Central Clearing House(s) (CCH) in more detail, in discussion with publishers and repositories
- Develop a specification for the technical, organizational and business models for the CCH
- identify candidate organizations for involvement in the CCH



PIRUS2: governance going forward

Principles

- Independent, not-for-profit organization
- International
- Representation of the main stakeholder groups
 - Repositories
 - Publishers
 - Research Institutions

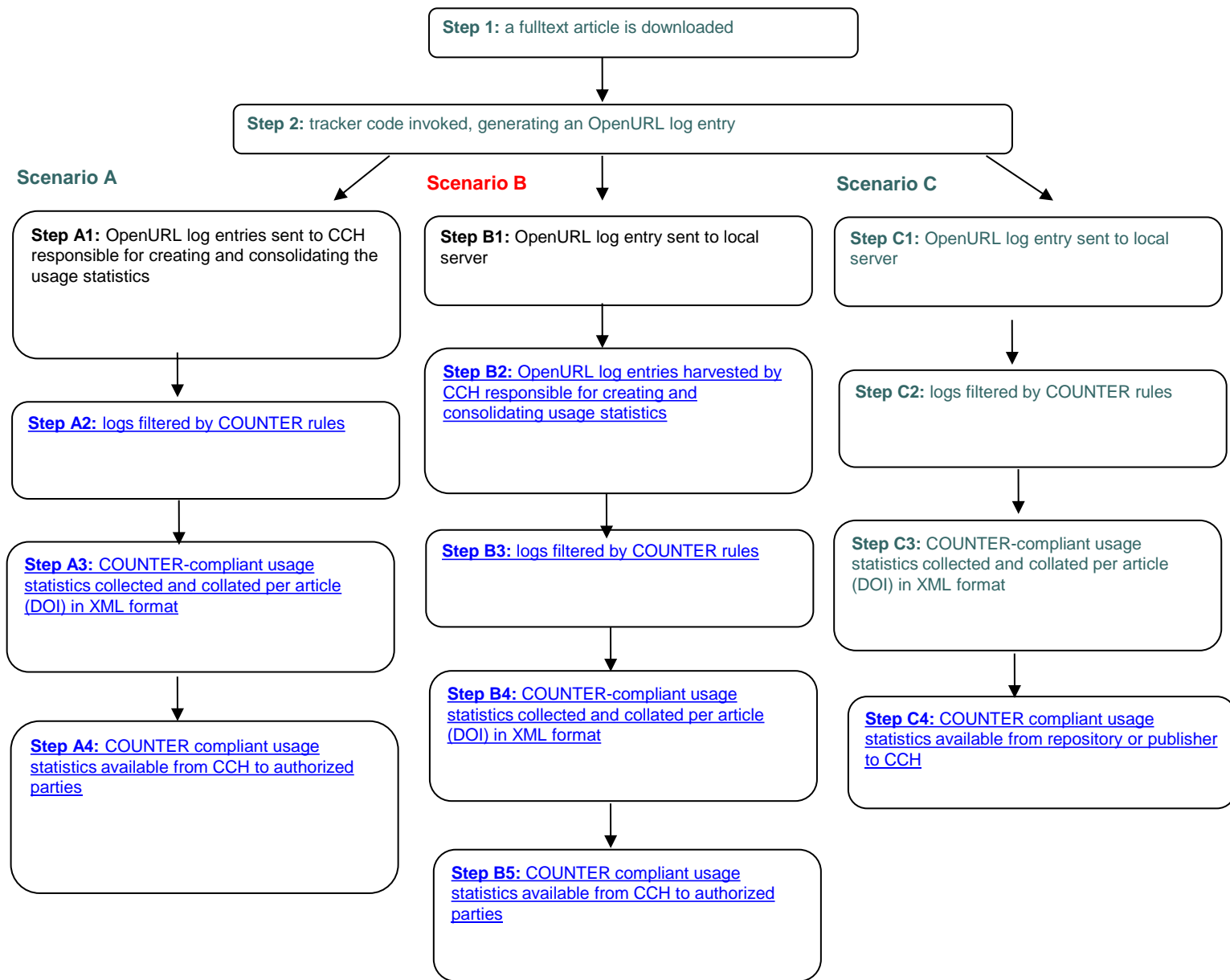
Role

- Define and implement mission
- Strategic oversight
- Set and monitor standards
- Set fees and manage finances
- Select and monitor suppliers



PIRUS2: nature and mission of the Central Clearing House(s)

- **One global CCH**
 - Cost-effective
 - Industry is global, with global standards
 - Easier to set and modify standards
 - Simpler interface with publishers and repositories
 - Can be outsourced
 - Existing organizations exist with the required capabilities
- **Scenarios to be supported**
 - See next slide.....





PIRUS2: CCH operating principles

- The “bucket” of usage data should be controlled by the participants - they can decide whether to compile the usage reports themselves or to delegate that role to the CCH
- Access to the CCH should be limited to authorised parties
- Usage reports must state the sources from which they have been compiled to ensure transparency



PIRUS2: role of CCH

Role 1: consolidated processing (applies to most repositories and to some publishers)

- Relies on all journal article downloads invoking a tracker code that sends data to a single big bucket
 - Consolidated usage reports can be generated by the CCH
- Single data **standard**, not necessarily data tool
 - Requirements can be met by various counting or analytics packages
 - Compliance with the standard can be checked by the “data gathering” audit
- All data in one place allows mining - deeper insights into data and easy integration of other projects, e.g. JUF
- Publishers who use this option could lose control of own data and report compilation
 - Terms and Conditions could handle some aspects of this
- All steps are auditable:
 - Data gathering
 - Process of sending data packet to bucket
 - Profile of data packet – does it meet standards?
 - Counting
 - Correct interpretation of data packets received
 - Compilation of usage reports
 - Correctness, completeness
- Audit overhead lower due to standard system



PIRUS2: role of the CCH

Role 2: distributed processing (applies to most publishers and some repositories)

- Relies on repositories and publishers gathering data in own buckets
 - Publishers
 - count and produce own usage reports according to the specifications of Article Report 1.
 - Repositories
 - count and produce own usage reports and send reports to CCH OR
 - send data to CCH who count and produce usage reports (and return to repositories?)
 - CCH sends repository reports to publishers
- All steps are auditable:
 - Data gathering
 - Process of sending data packet to bucket
 - Profile of data packet – does it meet standards?
 - Counting
 - Correct interpretation of data packets received
 - Compilation of usage reports
 - Correctness, completeness
- Many possible risk areas due to multiple supply points



PIRUS2: outputs from the CCH

- Usage reports for publishers
- Usage reports for repositories
- Usage reports for research institutions

Key requirements:

- Set of core reports
- Flexibility in outputs

File Edit View Insert Format Tools Data Window Help

Type a question for help

Calibri 10.5 B I U

Reply with Changes... End Review...

K17

A B C D E F G H I J K

Article Report 2 ,Number of Successful Full-Text Article Requests by Author, Month and DOI

<Journal>

<Publisher>

<Publisher Platform>

<Author name>

<ORCID Identifier>

<Institutional Identifier>

<Article title>

<DOI>

Date run: 01/04/2010

Source of usage	Jan-09	Feb-09	Mar-09	Total
Publisher	152	226	143	521
Host 1	23	31	29	83
Host 2	15	20	18	53
Host 2	10	15	12	37
Total	200	292	202	694

NOTES

- Article title data is highly recommended but optional
- Usage data should:
 - include: successful full text requests (HTML plus PDF)
 - include: Accepted Manuscript, Proof, Version of Record versions
 - exclude: Author's Original Manuscript and Submitted Manuscript Under Review versions
 - exclude: internal use by publisher and host, downloads from LOCKSS caches, and by robots

2007



PIRUS2: Implementation of the CCH

In view of the technical challenges that the CCH faces, its strong dependency on other initiatives, such as ORCID and institutional identifier and the requirements for publishers to re-engineer some of their processes, it may be prudent to implement the CCH in two Stages:

- Stage 1: gather and consolidate usage data only from repositories and provide the usage statistics generated by the CCH to publishers and other authorised bodies
- Stage 2: and collect usage data from publishers that wish to use the CCH service for this purpose



PIRUS2: organizations that could play a role in the CCH

- Setting the standards for usage reporting and specifications for the usage reports – COUNTER
- data gathering – existing vendor (eg ScholarlyIQ, MPS Technologies, etc)
- DOI information- CrossRef
- counting and reporting- existing vendor (ScholarlyIQ, MPS Technologies, etc)
- final report compilation – existing vendor (ScholarlyIQ, MPS Technologies, etc)
- auditing of the CCH – ABCe
- management of the CCH – existing vendor (ScholarlyIQ, MPS Technologies, etc.) supervised by a PIRUS management board
- customer service/account management – existing vendor (ScholarlyIQ. MPS Technologies, etc)



PIRUS2: economic issues

- estimate the costs of running statistical aggregator services
- assess the costs to repositories and publishers for generating the required usage reports;
- propose a model for the allocation of costs to stakeholders



PIRUS2: model for recovering costs

Possible sources of revenues to support the CCH:

- membership fees that give members the right to use the services of the CCH
- transaction-based fees:
 - From repositories, who provide raw data to the CCH and obtain usage statistics from the CCH
 - From repositories, who provide raw data to the CCH and obtain usage statistics from the CCH
 - from publishers, who obtain usage statistics from the CCH for consolidation into their own usage reports
 - from organizations, such as Thomson ISI or Elsevier (SciVal), who could use the data from the CCH to enhance the citation and usage based performance reports that they provide to institutions.
 - from research institutions, who want global usage reports for content produced by their researchers and departments



PIRUS2: CCH cost allocation model to publishers and repositories

	Large Publisher	Med Publisher	Small Publisher	Repository
Allocation of annual costs				
Annual membership fee	\$15,700	\$9,700	\$1,700	\$750
Reporting services				
Y1	\$808	\$808	\$808	\$323
Y2+	\$515	\$515	\$515	\$206
Transaction-based costs	\$630,000.00	\$264,000	\$4,800	\$288
Total Y1	\$646,508.00	\$274,508.00	\$9,008.00	\$1,361.00
Total Y2+	\$646,215.00	\$274,215.00	\$8,715.00	\$1,244.00



PIRUS2: CCH cost allocation model

Questions:

1. Will the cost allocation model be acceptable?
2. Will the level of costs be acceptable to different stakeholder groups



PIRUS2: political issues

- Support of stakeholder groups
 - Authors
 - Publishers
 - Repositories
 - Research institutions
 - Research funding agencies
- Principle of reporting article-level usage
- Organizational model
- Economic model
- Reports output by the CCH
- Intellectual property and privacy issues



PIRUS2: support for the principle of reporting article-level usage

- Authors: very positive. Evidence from PLoS experience and growing numbers of requests to other publishers
- Publishers: mixed. Willing to provide as a service to authors; concerns about repository statistics
- Repositories: mixed. Some question the need for a global standard
- Research institutions: positive. Interested in measures of the value and impact of research outputs
- Research funding agencies: mixed. Not all would implement usage measures of performance



PIRUS2: support for the organizational model

- Authors: insufficient evidence, but likely to be indifferent
- Publishers: mixed. Some are supportive; others are concerned about repository involvement
- Repositories: mixed. Some see the advantage of involvement; others want to keep a distance from publishers
- Research institutions: too early to say
- Research funding agencies: too early to say



PIRUS2: support for the economic model

- Authors: insufficient evidence, but likely to be indifferent, so long as they do not have to pay
- Publishers: feedback now being obtained
- Repositories: concerns expressed, even about relatively low level of annual fees
- Research institutions: too early to say
- Research funding agencies: too early to say



PIRUS2: support for the reports that could be generated by the CCH

- Authors: positive, provided publishers all adhere to the same standard. Prefer on-demand reports to a specified frequency
- Publishers: mixed. Want flexibility in the types of report they can provide and also the ability to combine usage data with other types of data. Want authors to access reports via publishers. Think that 10 years+ archive of historical usage data is desirable
- Repositories: too early to say
- Research institutions: too early to say
- Research funding agencies: too early to say



PIRUS2: Further work on organizational/political issues

More feedback from stakeholder groups on:

- acceptability of the proposed organizational model
 - Governance
 - Structure
 - Participating organizations
- acceptability of the proposed economic model
 - Level of costs
 - Cost allocation model
- the proposed usage reports to be generated by the CCH
 - Different reports for different stakeholders