



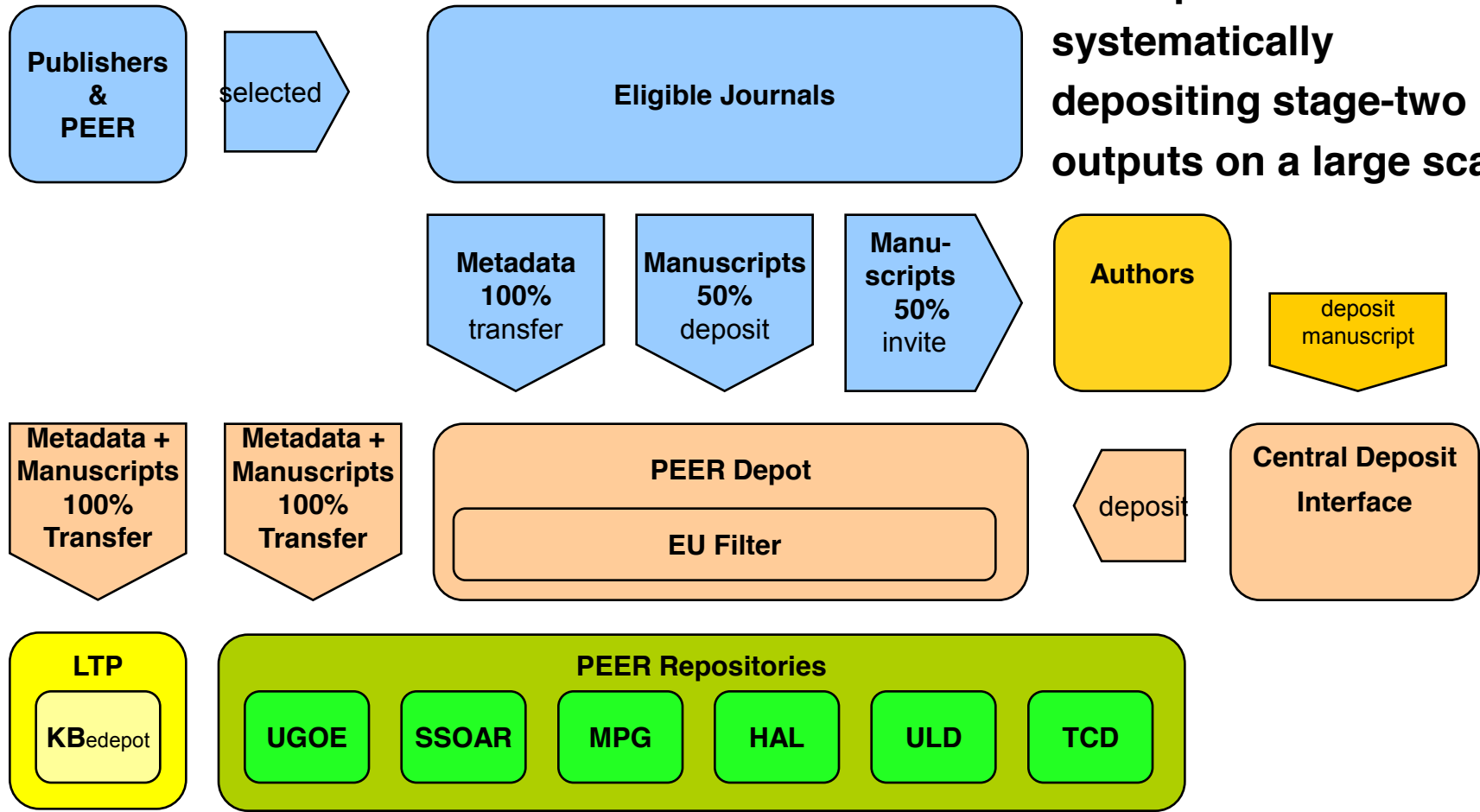
What PEER is teaching us about Green OA

STM Frankfurt conference 2011

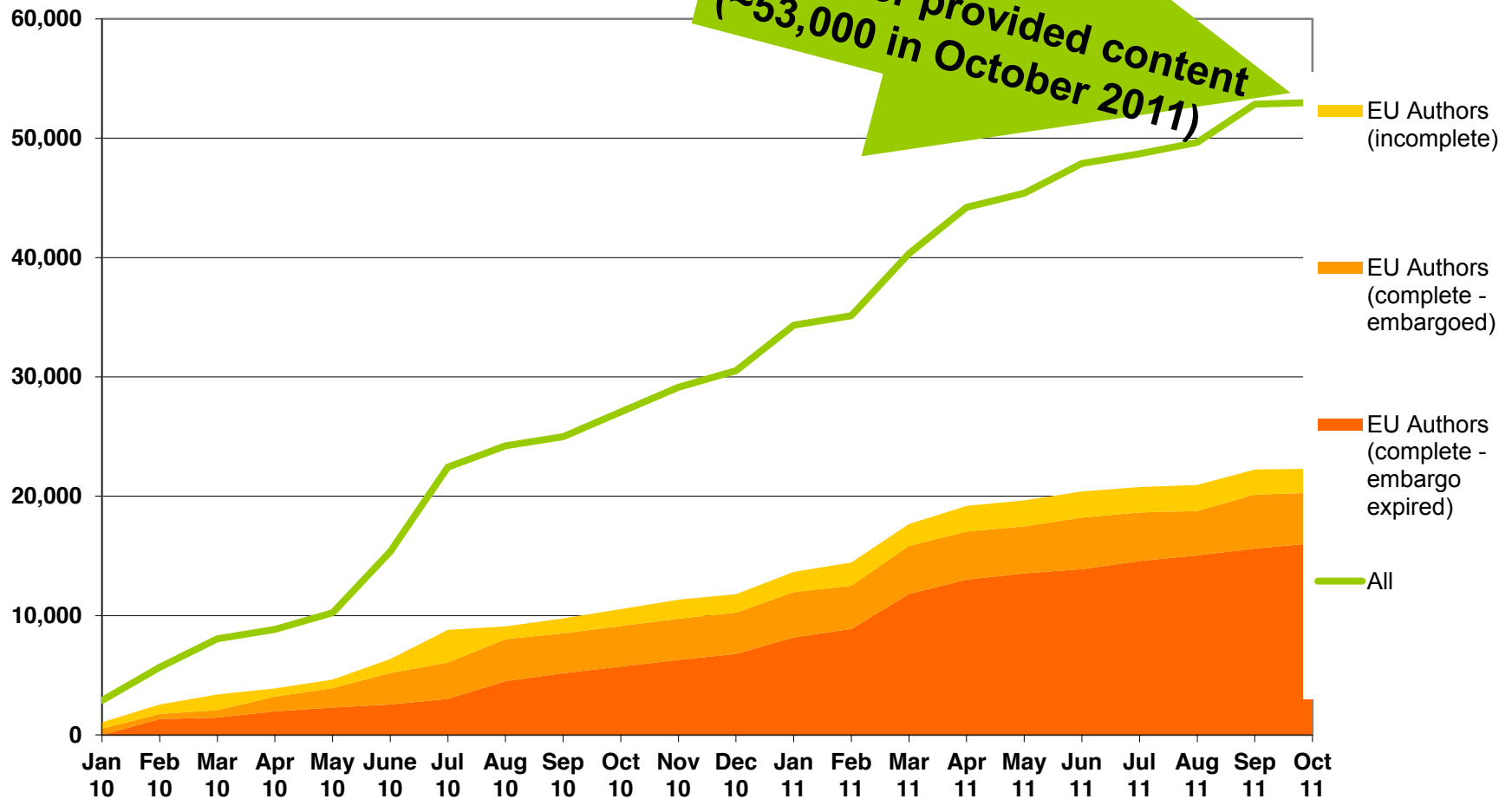
Chris Armbruster, Research Manager

The PEER Observatory

“Observatory”
developed to monitor
the impact of
systematically
depositing stage-two
outputs on a large scale

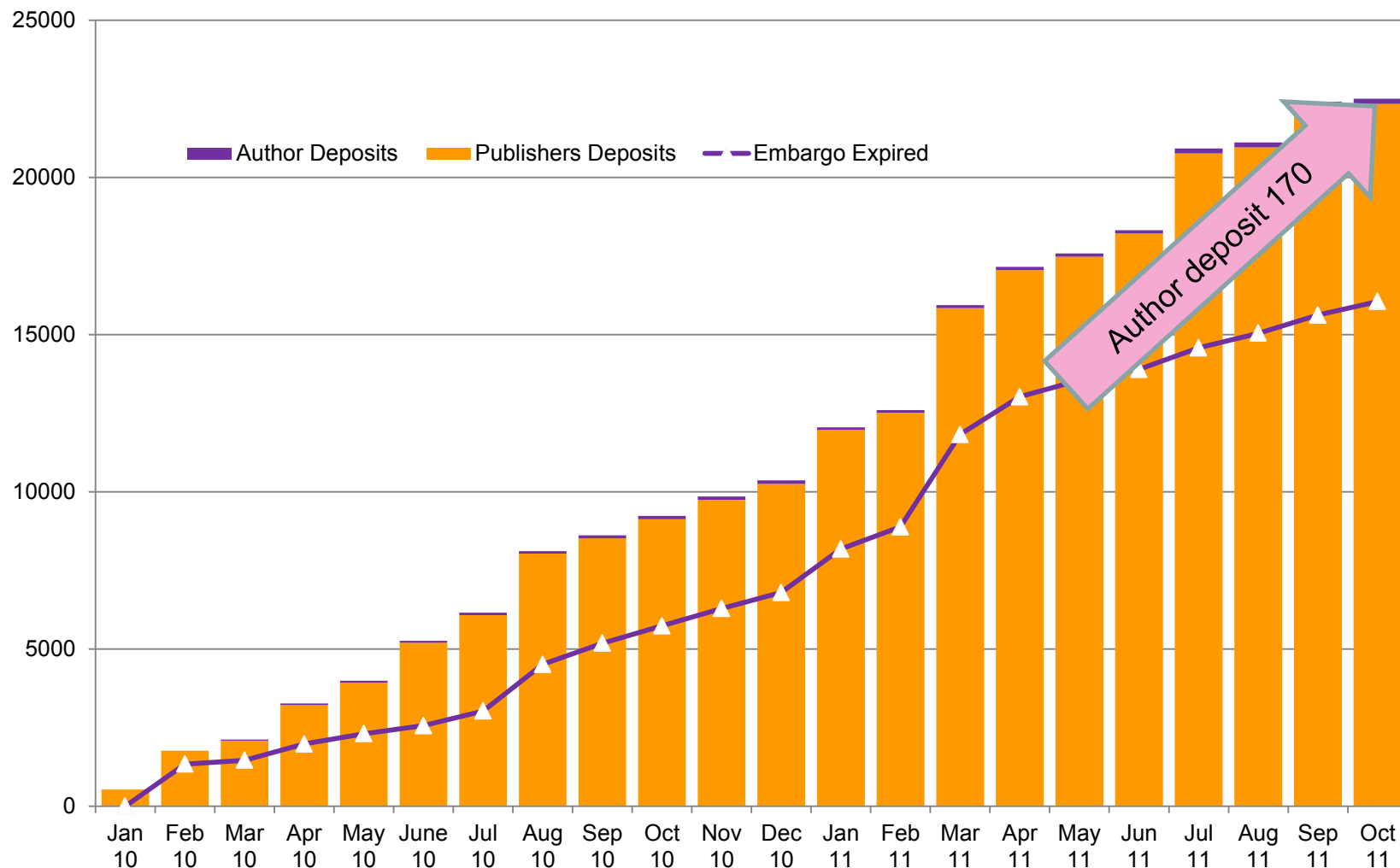


Publisher deposits (cumulated)



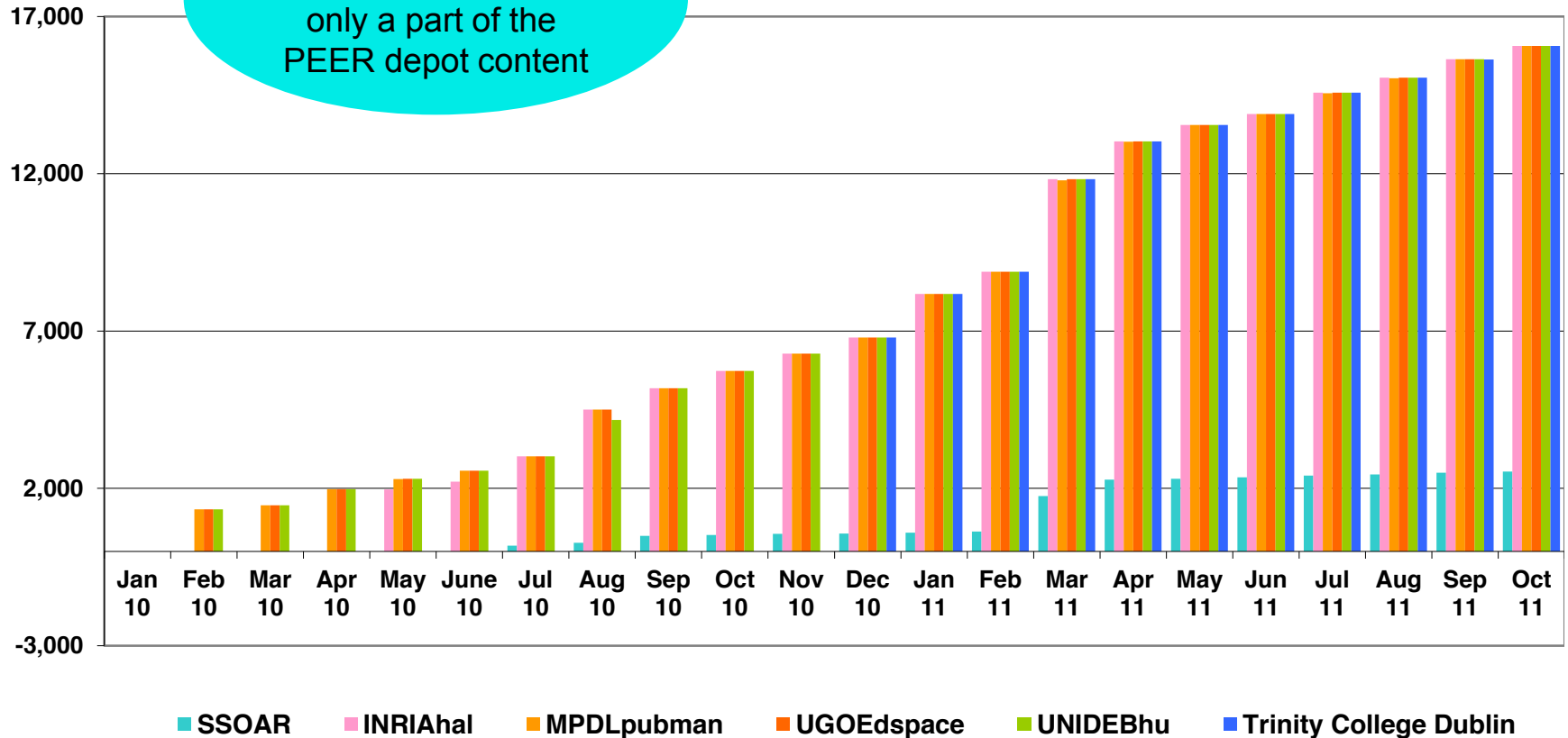
Total amount of publisher provided content (~53,000 in October 2011)

EU-Deposits processed (cumulated)



PEER repository population

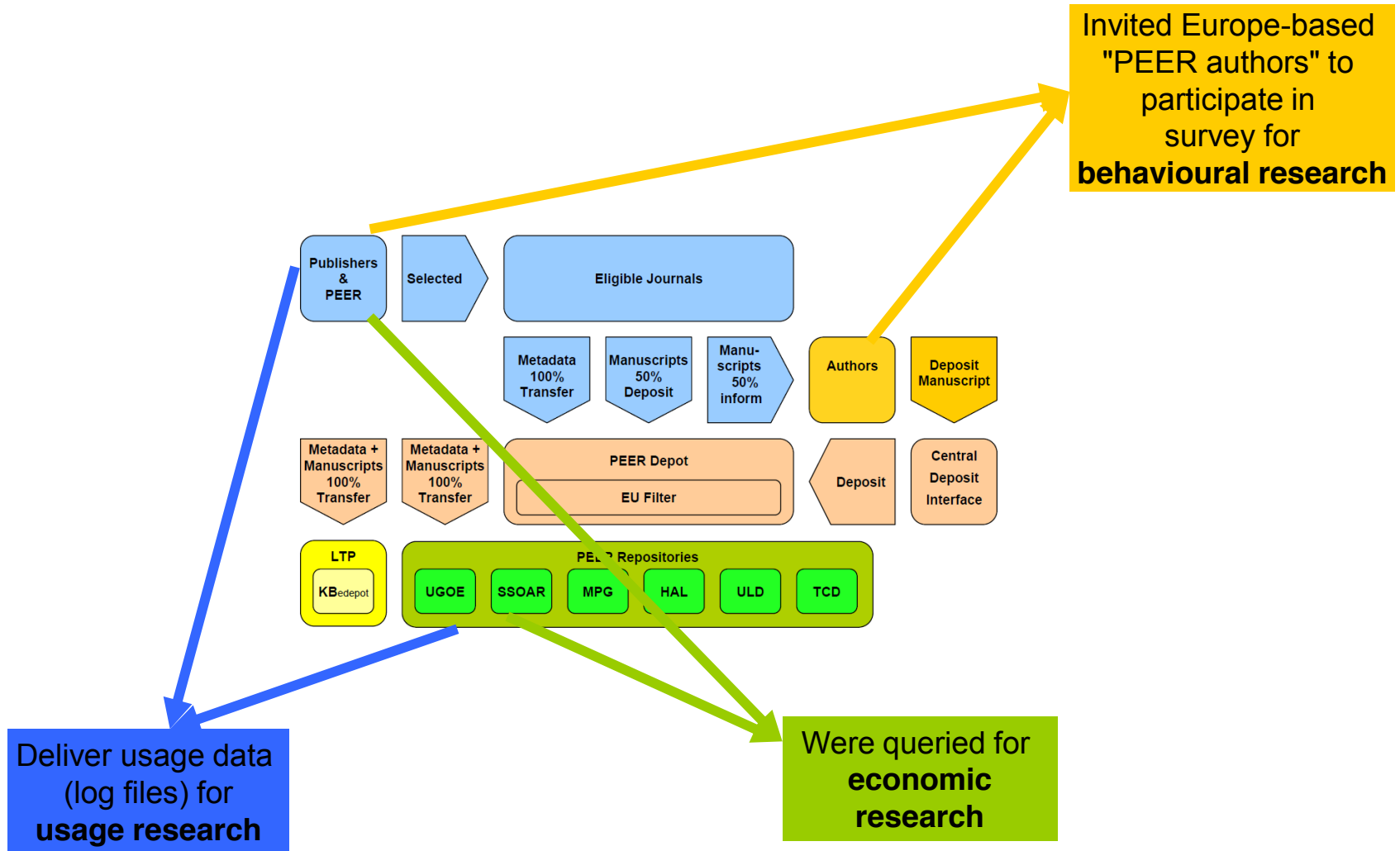
SSOAR is a subject repository accepting only a part of the PEER depot content



Research: Usage, Behaviour, Economics

- **PEER Research Manager:** manage high-quality, credible and transparent research process
 - **Usage research** (CIBER Research Ltd. – David Nicholas and Ian Rowlands)
 - **Behavioural research** (LISU and DIS, Loughborough – Jenny Fry, Claire Creaser and Stephen Proberts)
 - **Economic research** (ASK Bocconi – Paola Dubini)
- **Research Oversight Group (ROG):** independent experts in scholarly communication and publishing
 - Carol Tenopir, University of Tennessee (USA)
 - Cherifa Boukacem, Lille University (France)
 - Tomàs Baiget, *El profesional de la Información*, Barcelona (Spain)
 - Validate research objectives and tenders
 - Advise on methodologies
 - Evaluate the deliverables (e.g. data, conclusion)

A body of evidence



Project solutions and achievements

Publishers and Repositories

- (P) Metadata transfer: format, iterations, embedded in pdf, mandatory requirements
- (P) Manuscript transfer: format, author authentication
- (P) Embargo setting: publication date
- (R) Hosting space: dedicated PEER collection
- (R) Standardization: PEER standards
- (R) SWORD protocol: Implementation, automatic transfer

PEER Depot

- Deposit channels: each publisher, the authors
- Extraction routines: GROBID, matching and mapping, manual intervention
- Validation processes: integrity of metadata and file, match metadata and author submissions
- TEI conversion
- Filtering: journals, European content, valid research articles
- Distribution: release to repositories, embargo management, recall and correction

Usage research: Logfile analysis

- Objectives
 - ✓ Determine usage trends at publishers and repositories;
 - ✓ Understand source and nature of use of deposited manuscripts in repositories (so called Green Open Access) via usage data provision
 - ✓ Track trends, develop indicators and explain patterns of usage.
- Method
 - ✓ First large-scale and comparative collection of item level usage at publishers and repositories
 - ✓ At Publishers: 241 participating journals (items with EU authors), 200 control journals
 - ✓ At repositories: embargo-expired items

Usage research: towards the final report

- High volume of content in the project:
 - February 2011: Crucial milestone of 9000 embargo-expired items in February 2011
- Ongoing: measure activity over 12 months, starting March 2011
- Log file collection until 31 August 2011 (analysis of first six months), then 30 November 2011 and 28 February 2012
- Interim confidential reporting: Nov 2011 (after 6 months), Jan 2012 (after 9 months)
- Public report: May 2012
 - **“Final Report on usage research“**: May 2012 available via <http://www.peerproject.eu/reports/>

Scholarly attitudes (...from the final report)

Method: two representative surveys, preceded by focus groups and followed by a workshop

- Scholars have a conservative set of attitudes and do not desire fundamental changes in the way research is currently disseminated and published.
- Researchers who associated Open Access with 'self-archiving' were in the minority.
- Authors tend to be favourable to Open Access and receptive to the benefits of self-archiving... with the caveat that the pivotal role of the published article is not compromised.
- Researchers have concerns about the authority of article content and the extent to which it can be cited if the version accessed is not the final published one.
- Open Access repositories are perceived by scholars as complementary to, rather than replacing, current forms for disseminating and publishing research.
- There appears to be confusion amongst scholars about the distinction between OA journals, which may have a journal impact factor, and OA repositories, which do not.

Behaviour of authors and users (...from the final report)

- Broad similarity of behaviour in the Physical sciences (&M) and the Social sciences (&AH)
 - Shared culture of depositing and downloading pre-prints
 - Open access more likely to be associated with self-archiving
 - If self-deposit is voluntary, authors are more likely to choose a subject-based repository
- Medical sciences stand out
 - PubMed Central plus publisher deposit shape behaviour and expectations of authors and users
 - Open access least likely to be associated with self-archiving
- Copyright issues and the embargo period may be barriers to authors self-archiving, though authors do not generally experience much difficulty.
- Google and GS are the most likely access routes to Open Access repositories.
- Readers' ability to distinguish different versions of an article is critical as to how the quality of repository content is perceived.

Cost structures of publishers and repositories

(...from the final report)

Method: twenty-two case studies (e.g. scale and scope of activity, costs, salaries, revenue).

- Content certification per article: Average cost \$250 in this sample, with no significant economies of scale, highly dependent on rejection rate and complexity of process.
- Publishing cost (including metadata): ranges from \$170 to \$400 in this sample, influenced by journal size and make/buy decisions, with an advantage resulting from outsourcing to low-wage countries.
- Digital platform (content management): Very hard to calculate, with high variance in terms of investment, size and service. Maintenance costs may be estimated (ranging from \$170k to \$400k in this sample). Open source platforms possibly cheaper, but costs differentials to proprietary platforms not quantifiable.
- Open source software has reduced barriers to self-organization and market entry in journal publishing and repositories benefit from open (shared) platform development.
- Content uploading to repositories: Highly variable costs, dependent on make/buy decisions
- Repositories have high level of sunk costs, and coordination and transaction costs are not accounted for.

Cost and competition: state of play

(...from the final report)

- Opening up of a competitive space
- Strategies of individual players matter more than ‘the model’
- Repository growth is slow compared to overall growth of scholarly publications
- Key challenges of the decade
 - ✓ Further rapid growth of documents (&data)
 - ✓ More efficient preservation strategies
 - ✓ Increased competition for attention
 - ✓ Aftermath of the financial crisis and reduced resources

Green OA environment issues encountered by PEER

- Non uniformity of publisher outputs
- Varying requirements by repositories
- EU & article type filtering of content
- Embargo management
- Author authentication for deposit
- Non uniformity of log files
- Format problems with back-content files
- Technical & financial challenges for repository participation (non PEER Partner repositories)

Announcement: Final PEER CONFERENCE: May 29th 2012 Brussels



For details please visit our Webpage: <http://www.peerproject.de>

Read more soon:

<http://www.peerproject.eu/reports/>

or

e-mail: peer@stm-assoc.org

PEER cumulated usage

March to August 2011

