Open Access Statistics
From a Project to a Service

Workshop “Usage Statistics and Beyond”
22-23 April 2013, Berlin
Daniel Beucke | beucke@sub.uni-goettingen.de
Open Access Statistics

OAS 1
- May 2008 – December 2010

OAS 2
- April 2011 – July 2013
Objectives

Internationally **standardised usage statistics**

Sustainable **infrastructure** for collecting and evaluating usage statistics

Increase of the **acceptance** of Open Access among authors and users of scholarly publications
Standardised Usage Statistics

- Uniform **standards** for the gathering, exchange, and analysis of usage data
  - OpenURL ContextObjects
  - OAI-PMH
  - COUNTER

- **Tools**
  - Guidelines (OAS, KE)
  - Robot list
Infrastructure

User
... downloads a document from a repository.

Repository
... records the information on document usage in a log file.

... publishes the internationally standardised usage statistics.

... gathers the log files and extracts the usage data.

... processes the usage data and transforms it into an XML format.

OAS Data Provider
... makes the processed data available via an OAI interface.

OAS Service Provider
... retrieves the anonymised usage data.

... processes the data according to international standards.

... sends the aggregated and standardised usage data back to the repository.

... opens the Open-Access-Statistik.
Infrastructure

... downloads a document from a repository.

... publishes the internationally standardised usage statistics.

... sends the aggregated and standardised usage data back to the repository.

... processes the data according to international standards.

... retrieves the processed data via an OAI interface.

... gathers the log files and extracts the usage data.

... records the information on document usage in a log file.

User

Repository

OAS Service Provider

OAS Data Provider

Workshop “Usage Statistics and Beyond” | Open Access Statistics - From a Project to a Service | 22-23 April 2013 | Daniel Beucke
Infrastructure

Combined Log Format

```
LogFormat "%h %l %u %t "%r" %>s %b " %{Referrer}i" " %{User-agent}i"
combined
```

- downloads a document from a repository.
- publishes the internationally standardised usage statistics.
- records the information on document usage in a log file.

```
Infrastructure
Combined Log Format

LogFormat "%h %l %u %t "%r" %>s %b " %{Referrer}i" " %{User-agent}i"
combined
```

... sends the aggregated and
... records the information on
don usage statistics.
... gathers the log files and extracts
... available via an OAI interface.
... OAS Data Provider

Workshop “Usage Statistics and Beyond” | Open Access Statistics - From a Project to a Service | 22-23 April 2013 | Daniel Beucke
Infrastructure

- Filter pdf and html views
- Link with local identifier
- Salt and hash IP

... publishes the internationally standardised usage statistics.

User... downloads a document from a repository.

Repository... records the information on document usage in a log file.

... gathers the log files and extracts the usage data.

... processes the usage data and transforms it into an XML format.

... makes the processed data available via an OAI interface.

OAS Data Provider
Infrastructure

Data formats of aggregated usage data
- JSON
- XML
- CSV
- iFrame

... sends the aggregated and standardised usage data back to the repository.
... processes the data according to international standards.
... retrieves the anonymised usage data.
Infrastructure

Repository

... publishes the internationally standardised usage statistics.

User

... downloads a document from a repository.

For internal or external use

... records the information on document usage in a log file.
Infrastructure

User

- downloads a document from a repository.

Repository

- records the information on document usage in a log file.

OAS

... publishes the internationally standardised usage statistics.

... sends the aggregated and standardised usage data back to the repository.

... processes the data according to international standards.

... retrieves the anonymised usage data.

OAS Service Provider

... gathers the log files and extracts the usage data.

... processes the usage data and transforms it into an XML format.

... makes the processed data available via an OAI interface.

OAS Data Provider
Acceptance of Open Access

Usage data as a value-added service

- 24% of repositories offer usage statistics.
- 48% offer checksums.
- 45% offer RSS feeds.
- 11% offer social networking.


5) Bibliographic export (at least one format, e.g. RIS) is available on item or collection level.
6) Usage statistics (e.g. downloads, views) are available for unregistered users on item.
7) Checksums (e.g. MD5, SHA1) of full text publications are available on item level.
8) A functioning RSS feed is available on the home or browsing page.
9) Social bookmarking (at least one service, e.g. connotea) is available on item level.
10) Social networking (at least one service e.g. facebook, twitter or AddThis button) is available on item level.

The bigger a repository, the more likely it offers usage statistics.
Lessons Learned
Lessons Learned I

Implementation of infrastructure is expensive
- amount of data
- heavy traffic
- permanent maintenance

Data privacy act
- our assumption: IP address is personal data
- strict data privacy
- salt and hash
Lessons Learned II

Standards for analysis
  – COUNTER ≠ COUNTER – depending on the robot list for example

Robots
  – 30% - 50% of all users are robots
Outcomes
Infrastructure

OAS Data Provider
- Demo OAS Data Provider
- Software

OAS Service Provider

Involved partners
- Digizeitschriften
- Econstor
- EDOC HU Berlin
- OJS - CeDis Berlin
- Pedocs
- Psydoc SULB Saarbrücken
- PUB Bielefeld
- Qucosa
- GoeScholar
- Universitätsverlag Göttingen
- VZG
Guidelines

Guideline for participation in OA Statistics

*Specification: Data Format and Exchange for OA Statistics*

Cooperation with KE Usage Statistics Guidelines
*Guidelines for the aggregation and exchange of Usage Data*

Recommendations in
*DINI Certificate – Document and Publication Services 2010*
Sustainable Infrastructure

Feasibility study

Data privacy report

OAS Validator
Project Publications

Project publications on project website

- **Publications**
- **Presentations**
- **Poster**
- **Workshops**

In summer 2013

- *Usage Statistics and Beyond*  
Version 2.0  
also available in English
From a Project to a Service

Repository manager indicates interest, service will be offered by the VZG from summer 2013

... Marco Recke
Thank you for your attention.

Questions?

E-Mail: beucke@sub.uni-goettingen.de

Website: http://www.dini.de/projekte/oa-statistik/english/

License: [Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)](http://creativecommons.org/licenses/by-sa/4.0/)