

WEBINAR

Open Access Statistik II

Ein virtuelles Seminar für potenzielle Teilnehmer

Ulrich Herb (u.herb@sulb.uni-saarland.de)

Andreas Holtz (a.holtz@sulb.uni-saarland.de)

Martin Wanke

Saarländische Universitäts- und Landesbibliothek
Saarbrücken
30.01.2013

Initiiert durch:



Gefördert durch:

Deutsche
Forschungsgemeinschaft

DFG

Ablauf

- **0. Vorab: Begrüßung, Vorstellung Referenten & Ablauf (5 Min.)**

- **1. Open Access Statistik - Hintergründe (25 Min.)**
 - 1.1. Vortrag des ersten Programmpunktes (20 Min.)
 - 1.2. Fragen und Diskussion (5 Min.)

- **2. Die Technik (50 Min.)**
 - 2.1. Vortrag des zweiten Programmpunktes (40 Min.)
 - 2.2. Fragen und Diskussion (10 Min.)

- **3. Rechtliches und Finanzielles zu OAS (25 Min.)**
 - 3.1. Vortrag des dritten Programmpunktes (20 Min.)
 - 3.2. Fragen und Diskussion (5 Min.)

- **4. Abschlussfragen und Verabschiedung (15 Min.)**
~ 120 Min.

Gliederung

- **1. Zum Hintergrund von OA-Statistik**
- 1.1. Wer steht hinter OA-Statistik?
- 1.2. Warum OA-Statistik?
- 1.2.1. Wie es ist: Publizieren Old School
- 1.2.2. Wie es sein sollte: Ziele OAS
- 1.3. OA-Statistik für wen?
- 1.3.1. Entwicklung IR = Herausforderung OAS
- 1.3.2. Welche Vorteile ergeben sich durch OA-Statistik?
- 1.3.3. Wann ist OA-Statistik betriebsbereit?
- **2. Zur Technik und Funktionsweise von OA-Statistik**
- 2.1. Die Technik/der Workflow
- 2.1.1. Data-Provider
- 2.1.2. Service-Provider
- 2.1.3. Software Download, Registrierung
- 2.2. Darstellungsmöglichkeiten: Aufgearbeitete Daten als Basis für Nutzungsstatistiken
- 2.2. Wie kann mitgemacht werden?
- **3. Rechtlicher und finanzieller Rahmen von OA-Statistik**
- 3.1. Welche Vertragsverhältnisse bestehen?
- 3.2. Wodurch finanziert sich OA-Statistik?
- **4. Wo finde ich weiterführende Informationen?**
- 4.1. Wer ist mein Ansprechpartner?

1. Zum Hintergrund von OA-Statistik

IN A NUTSHELL:

OA-Statistik
sammelt Daten,
arbeitet sie auf
und stellt sie bereit.



1. Zum Hintergrund von OA-Statistik ...etwas länger

- Aufbau tech. Infrastruktur zur Ermittlung standardisierter Nutzungsdaten elektronischer Dokumente
- Fehlende Kriterien zur Nutzungsmessung elektronischer Dokumente => Bedarf nach neuen metrischen Verfahren
- Erstellung verlässlicher Nutzungsdaten & eines nutzungszentrierten Impact Maßes
- Ziel des Projekts: Entwicklung & Etablierung eines Standards zur Ermittlung von Zugriffszahlen und Nutzungsstatistiken für Publikationen sowohl in Open-Access-Repositories als auch in konventionellen Angebotssystemen.



1. Zum Hintergrund von OA-Statistik ...etwas länger

- Dokumentenzugriffe werden über Logfiles des eingesetzten Webservers aufgezeichnet.
- In Repositorien implementierte Data-Provider sammeln diese Daten, transformieren sie in OpenURL/Context Objects und stellen sie über eine OAI-PMH-Schnittstelle dem Service-Provider bereit.
- Dieser korrigiert sie standardisiert um Robots, Doppelklickspanne und Deduplizierung. Die gesammelten Nutzungsdaten werden nach den Standards COUNTER, LogEc und IFABC aufbereitet.



1. Zum Hintergrund von OA-Statistik ...etwas länger

- Erstellte Statistiken werden teilnehmenden Repositorien zur Verfügung gestellt, die sie dann den einzelnen Nutzern anbieten können. Also: Serviceprovider = zentraler Data-Provider für die Repositorien.
- Datenerhebung: IP-Adresse, Zugriffsdatum und -uhrzeit, Zeitzone, Webbrowser, Status, Protokolltyp, URL, Dokumentenname und -typ, Get-Parameter, Dateigröße und Referrer. Die dazugehörige Software ist in den Varianten DSpace, WebDoc und OPUS auf der Projektseite downloadbar.

1. Zum Hintergrund von OA-Statistik

Projektphase 1

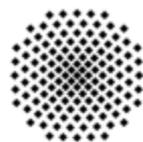
- **Laufzeit:** Mai 2008 - Dezember 2010 (kostenneutrale Verlängerung)
- **Ziele:**
- Aufbau einer dauerhaften Infrastruktur zur Erfassung und Verarbeitung von Nutzungsdaten
- Entwicklung und Etablierung eines einheitlichen Standards zur Ermittlung von Zugriffszahlen und Statistiken
- Anbieten von Serviceleistungen für Repositorien-Betreiber und Nutzer
- Entwicklung von Guidelines zur Implementierung der Services
- Empfehlungen im DINI-Zertifikat

Projektphase 2

- **Laufzeit:** April 2011 - April 2013
- **Ziele:**
- Erweiterung der OAS-Infrastruktur um weitere deutsche Repositorien und Bereitstellung von standardisierten Nutzungsstatistiken
- Erhöhung der Akzeptanz von Open Access bei Autoren und Rezipienten von wissenschaftlichen Publikationen durch exaktere Metriken und Mehrwertdienste
- Kooperationen für international vergleichbare Nutzungsstatistiken
- Anbieten einer nachhaltigen Service-Infrastruktur

1.1. Wer steckt hinter OA-Statistik?

- Das Projekt setzt sich aus diesen Projektpartnern zusammen:



Universität Stuttgart

HUMBOLDT-UNIVERSITÄT ZU BERLIN



NIEDERSÄCHSISCHE STAATS- UND
UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN 
GBV | VZG

Initiated by:



SAARLÄNDISCHE
UNIVERSITÄTS-UND
LANDESBIBLIOTHEK



Funded by:

DFG

1.1. Wer steckt hinter OA-Statistik?

- ...und ist mit diesen Projekten assoziiert:

- ▣ Open Access Statistics



- ▣ DOARC
(Distributed Open Access Reference and Citation Services)



- ▣ Open Access Network



1.1. Wer steckt hinter OA-Statistik?

- ...und kooperiert mit diesen Projekten:



1.2. Warum OA-Statistik?

- Wirkungsanalyse für Open Access-Publikationen ist im Bereich des konventionellen Publizierens bislang nicht möglich/vorgesehen
- Dokumente in Repositorien werden grundsätzlich nicht bewertet
- Dokumente in OA-Journals fallen aus der Bewertung oftmals wegen ihrer relativ kurzen Zitationsgeschichte sowie wegen der Dokumentensprache heraus
- Fehlende Standardisierung & Datenharmonisierung
- Wirkungsanalysen über Zitationen sind stark fehleranfällig, Zugriffsanalysen beheben teils die Defizite
- Damit auch Unterstützung von OA-Publikationen im Allgemeinen
- Deswegen Projektziel: Anwendung resp. Entwicklung einheitlicher Standards zur Erfassung, Analyse und Aufarbeitung von Zugriffszahlen für Dokumente von IRs zunächst im deutschsprachigen Raum

1.2.1. Wie es ist: Publizieren Old School

- Altbekannt: Impact-Messung
- Hintergrund: publish or perish, also:
Publizieren = wissenschaftliches Kapital = Impact = Karriere
- Auf organisatorischer Ebene:
Impact als Basis für Evaluationen
=> Evaluationsergebnisse bestimmen über Mittelverteilung und damit über Forschungsmöglichkeiten (institutioneller Rahmen)
=> institutioneller Rahmen bestimmt über Personal
=> Personal bestimmt Impact

1.2.1. Wie es ist: Publizieren Old School

- Altbekannt: Impact-Messung
- Messung entweder über den Journal Impact Factor (JIF)

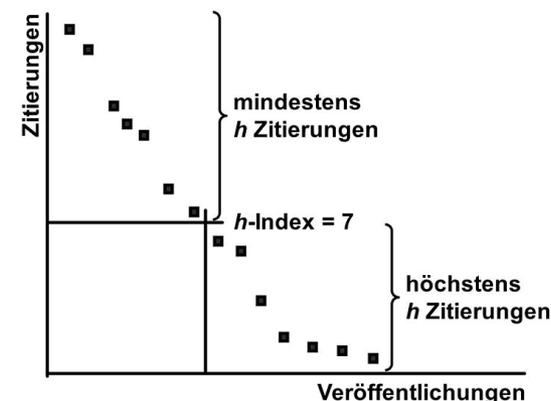
$$\text{JIF} = \frac{\text{Zahl der Zitate im laufenden Jahr auf Artikel (eines Journals) der vergangenen zwei Jahre}}{\text{Zahl der Artikel des Journals der vergangenen zwei Jahre}}$$

Kritik am Journal Impact Factor:

- Ignoranz der Verwertungszyklen in unterschiedlichen Disziplinen: Benachteiligung v. Journals aus Disziplinen mit Verwertungszyklen > 2 Jahren (z.B. Geisteswissenschaften)
- JIF bezieht sich auf Journals, nicht Artikel: i.d.R. führt geringe Anzahl sehr häufig zitierter Artikel zu einem hohen Wert für das Journal
- Vermischung von Popularität & Qualität: begrenzter Anwendungsbereich, damit Ausschluss kompletter Dokumentarten: graue Literatur, Bücher, Großteil der Web-Publikationen
- Berücksichtigung nur der im Journal Citation Report (JCR) indizierten Journals
- Bevorzugung englischsprachiger Journals: niedrigerer JIF v. anderssprachigen Journals

1.2.1. Wie es ist: Publizieren Old School

- Altbekannt: Impact-Messung
...oder über den Hirsch-Index (h-index)
- Hirsch (h-) Index bezieht sich auf Autoren und nicht auf Zeitschriften
- Verhältnis von Zitationen zu Publikationen
- Ein Autor hat einen Index h , wenn h von seinen insgesamt N Veröffentlichungen mindestens jeweils h Zitationen haben



⇒ Ein Autor hat einen h-Index von 7, wenn er 7 Schriften veröffentlicht hat, die jeweils mindestens 7 Mal zitiert worden sind. Ein h-Index von 12 setzt die Publikation von 12 Schriften voraus, die mindestens 12 Mal zitiert wurden.

Quelle: <http://de.wikipedia.org/wiki/Hirsch-Index>

1.2.1. Wie es ist: Publizieren Old School

- Altbekannt: Impact-Messung
...oder über den Hirsch-Index (h-index)

Kritik am Hirsch-Index

- vgl. JIF-Kritik: Vernachlässigung von Dokumentengattungen und nicht- englischer Publikationen, Vermischung von Popularität und Qualität, ...
- Trennscharfe Autorenidentifikation in der Datenbasis (z.B. Web of Science) ist nicht sichergestellt
- h-Index hängt von Forschungsdisziplin und Alter des Autors ab
- Vorteil im Vergleich zum JIF: Zitationen einer einzigen, vielzitierten Veröffentlichung schlagen sich nicht nieder

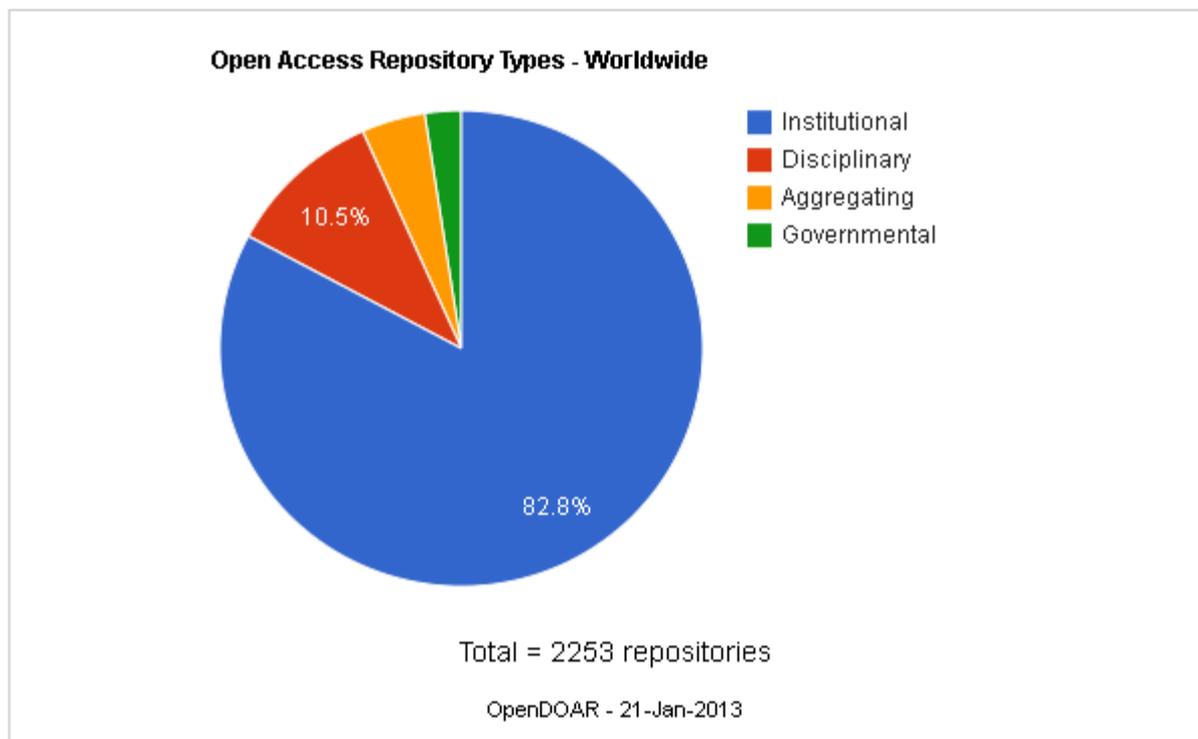
1.2.2. Wie es sein sollte: Ziele OAS

- Gemeinsame Standards zum Austausch von Nutzungsdaten
- Aufbau dauerhafte Infrastruktur zur Erfassung und Verarbeitung von Nutzungsdaten
- Aufarbeitung der Nutzungsdaten nach den Standards von COUNTER, LogEc und IFABC
- DINI-Empfehlung
- Verbesserung des internen Monitorings für Repositorien
- Anbieten von Serviceleistungen für Repository-Betreiber und Nutzer
- Guidelines zur Implementierung der Services
- Klärung der juristischen Standards und Lizenzen
- Erweiterung der OAS-Infrastruktur um weitere deutsche Repositorien und Bereitstellung von standardisierten Nutzungsstatistiken
- Erhöhung der Akzeptanz von Open Access bei Autoren und Rezipienten von wissenschaftlichen Publikationen durch exaktere Metriken und Mehrwertdienste
- Kooperationen für international vergleichbare Nutzungsstatistiken
- Angebot einer nachhaltigen Service-Infrastruktur

1.3. OA-Statistik: Für wen?

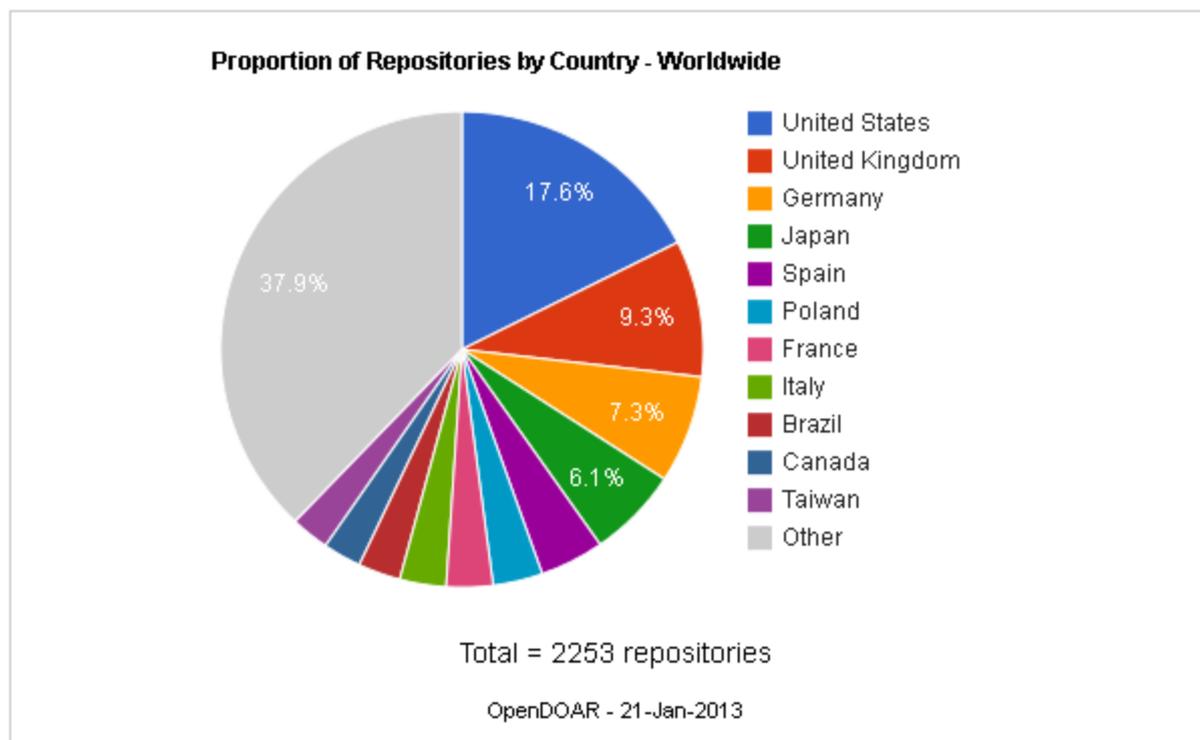
- Grundsätzlich richtet sich OA-Statistik an Sie, also die Repositorienbetreiber
- Mittelbar dient der Service aber auch Autoren, Institutionen/Organisationen und ganz allgemein auch den Nutzern.

1.3.1. Entwicklung IR = Herausforderung OAS



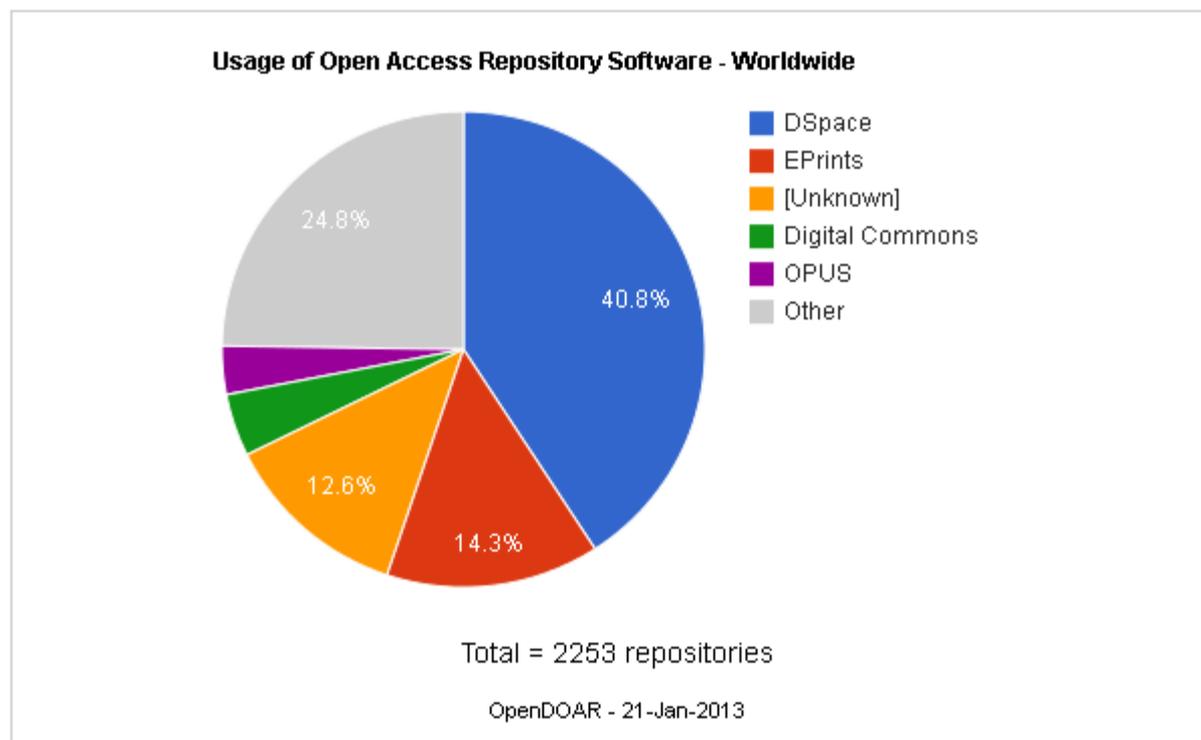
<http://www.opendoar.org>

1.3.1. Entwicklung IR = Herausforderung OAS



<http://www.opendoar.org>

1.3.1. Entwicklung IR = Herausforderung OAS



<http://www.opendoar.org>

1.3.1. Entwicklung IR = Herausforderung OAS

- IR-Software sind in Deutschland und Europa nicht homogen
- Heterogenität erschwert Vergleichbarkeit der Nutzungsinformationen
- Es fehlt sowohl in Europa als auch in Deutschland an einer einheitlichen Integration von Repositorien
- Angebot einer solchen Vereinheitlichung im Rahmen von einheitlichen Standards bildet Anreiz zur verstärkten Nutzung von Repositorien

1.3.1. Entwicklung IR = Herausforderung OAS

- Situation in Deutschland heterogen
- Unterschiede in Design, Größe und Anwendungsbereich
- Verbreitete Anwendung von OAI-PMH
- Dennoch Defizite in Standardisierung und Datenharmonisierung trotz Erfolgen von DINI (www.dini.de) und Open Access-Netzwerk (www.dini.de/projekte/oa-netzwerk/)

1.3.2. Welche Vorteile ergeben sich durch OA-Statistik?

- Genaue Kenntnis der Zugriffe auf die einzelnen Texte
- Verbesserung des internen Monitorings für Repositorien
- Verbesserung der Vergleichbarkeit
- Allgemeine Unterstützung alternativer Publikationen
- Ranking der Texte
- Graphische Aufarbeitung der Nutzungsstatistik
- Aktuelle Auswertung
- Aufarbeitung der Nutzungszahlen für Weiternutzung durch Deutsche Bibliotheksstatistik (DBS) und für den Bibliotheksindex (BIX) => Einsparen von Mehraufwand, da Daten bereits gesammelt

1.3.2. Welche Vorteile ergeben sich durch OA-Statistik?

- Statistiken geben Auskunft über ...
 - Relevanz eines Dokuments, Rankingelement (Leser)
 - Nutzung/Sichtbarkeit/Impact eines Repositories (Leser, Autor, Admin, Verwalt.)
- Statistiken geben Auskunft über ...
 - Relevanz eines Forschungsthemas (Autor, Verwaltung)
 - Autor/Lehrstuhl/Forschungseinrichtung: Dokumentnutzung gibt Feedback über Publikations- und Forschungsthemen
 - Interpretation erfordert Sensibilität und Beachtung der Rahmenbedingungen, wie...
 - ...Technikaffinität der Disziplin
 - ...Publikations- und Rezeptionspräferenzen der Disziplin
 - ...Alter und Verwertungszyklus der Dokumente
 - ...Dokumenttyp

1.3.2. Welche Vorteile ergeben sich durch OA-Statistik?

- Statistiken geben Auskunft über ...
 - Relevanz eines Forschungsthemas (Autor, Verwaltung)
 - Autor/Lehrstuhl/Forschungseinrichtung: Dokumentnutzung gibt Feedback über Publikations- und Forschungsthemen
 - Verwertung durch Verwaltung (Benchmarking) erfordert Sensibilität und Beachtung der Rahmenbedingungen, wie...
 - ...Technikaffinität der Disziplin
 - ...Publikations- und Rezeptionspräferenzen der Disziplin
 - ...Alter und Verwertungszyklus der Dokumente
 - ...Dokumenttyp
 - Impact/Qualität eines Dokuments (Leser, Autor, Verwaltung)
 - Zitationsbasierte Maße: autorenzentriert, Messung erst in folgenden Publikationsgenerationen möglich, Messung erfolgt i.d.R. auf Journal- oder Autorenebene, Impact eines Textes/Objekts wird nicht abgebildet
 - Nutzungsbasierte Maße: leserzentriert, Messung live und fortlaufend möglich, Impact eines Textes/Objekts wird abgebildet, stark automatisierte Messung möglich

1.3.3. Wann ist OA-Statistik betriebsbereit?

- **ab sofort!**

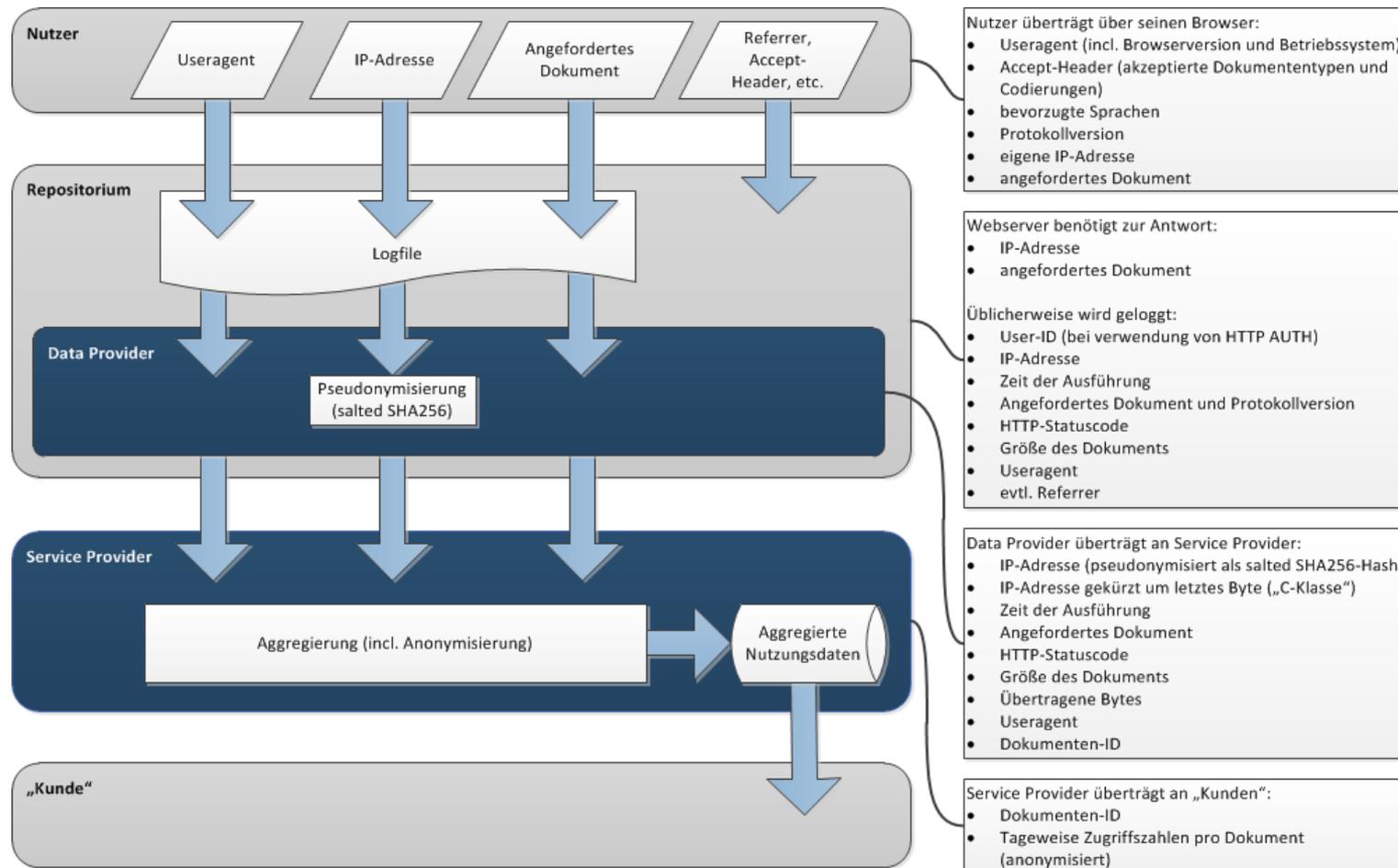
2. Wie funktioniert OA-Statistik?

Zum Hintergrund:

- Jeder der teilnehmenden Institutionen verfügt über ein eigenes Datenpool
- Zusammenführung aller Nutzungsdaten in einem zentralen Service Provider
- Dieser erstellt die jeweiligen Dienste auf Basis der Nutzungsdaten
- Aufbereitete Nutzungsdaten werden an die teilnehmenden Institutionen transferiert

2.1. Die Technik/der Workflow

kurz und knapp:



2.1. Die Technik/der Workflow

...und etwas ausführlicher:

Schritt 1) Nutzer eines Repositoriums übertragen beim Aufruf eines Dokumentes Informationen.

Fordert ein Nutzer eines Repositoriums ein Dokument an, so überträgt sein Browser automatisch Daten. Jeder Nutzer überträgt über seinen Browser Informationen wie seinen Useragent (inkl. Browserversion und Betriebssystem), ACCEPT-Header (akzeptierte Dokumententypen und Codierungen), bevorzugte Sprachen, Protokollversion, eigene IP-Adresse und den Namen des angeforderten Dokuments.

Schritt 2) Nutzungsdaten der Repositorien werden in Logfiles gespeichert.

Die Zugriffe der NutzerInnen auf die Dokumente eines Repositoriums werden über Logfiles des eingesetzten Webservers aufgezeichnet.

Aus datenschutzrechtlichen Gründen wird vom Repositorium die Möglichkeit eines Opt-Out angeboten. Wird dieser vom Nutzer gewählt, werden seine Daten nicht weiterverarbeitet.

2.1. Die Technik/der Workflow

...und etwas ausführlicher:

Schritt 3) Die Logfiles werden mit dem Logfile-Parser in OpenURL Context-Objekts umgewandelt.

- OAS Logfile-Parser wandelt Logfiles d. Webservers in OpenURL Context-Objects um und speichert sie im Datenformat XML ab. Dabei: Infos aus Logfiles wird ein Identifier eines Dokuments mitgegeben (v.a. auf Grundlage der abgerufenen URL). Individualisiertes Vorgehen: Anpassung bei Installation im LogfileParser.
- Irrelevante Daten werden zwecks Laufzeitverbesserung ausgefiltert.
- Datenschutz: IP-Adressen werden mit *Salt* versehen und anschließend mit *Hashverfahren SHA-256* bearbeitet. *Hash* = Pseudonymisierung der von einer IP-Adresse erfolgten Nutzungsmuster, ohne die IP selbst zu verwenden. *Salt* = weitere Schutzmaßnahme zur Verhinderung der Rückübersetzung des Hashwertes. Dazu noch: *Salt* und *Hash* des C-Klasse-Netzes der IP-Adresse (3/4 Zahlen der IP-Adresse).
- Datenformat OpenURL ContextObjects wurde angepasst und mit anderen europäischen Projekten abgestimmt und dokumentiert. Komplexer, aber kompatibel.

2.1. Die Technik/der Workflow

...und etwas ausführlicher:

Schritt 4) Die OpenURL ContextObjekts werden durch einen lokalen oder einen zentralen Data-Provider angeboten.

- Data-Provider (OAS-DP) basiert auf OAI Standard: Es werden keine bibliografischen Metadaten sondern Nutzungsdaten ausgetauscht. Deswegen: Installation eines modifizierten OAI-Data-Providers.
- OAS-DP nimmt Informationen aus Datenbank auf und stellt sie als OpenURL Context-Objects über OAI-Protokoll OAS-Service-Provider (OAS-SP) zur Verfügung.
- Ohne eigene Installation des OAS-DP: Logfiles können alternativ auch verschlüsselt zum zentralen OAS-Data-Provider transferiert werden. Umwandlung in ContextObjects und Bereitstellung für den OAS-Service-Provider erfolgen dann dort.

Schritt 5) Datentransfer

Der OAS-Service-Provider (angepasster OAI-PMH Service-Provider) holt Daten vom OAS-DP (5 Min oder je nach Verarbeitungsdauer). Im Anschluss werden die Daten auf dem OAS-DP gelöscht.

2.1. Die Technik/der Workflow

...und etwas ausführlicher:

Schritt 6) Aggregation der Daten auf dem OAS-Service-Provider

- Aufbereitung gesammelter Nutzungsdaten auf OAS-SP.
- Auswertung nach verschiedenen Standards wie COUNTER möglich.

Schritt 7) Bereitstellung der Statistiken

- Dokumenten-IDs & anonymisierte Zugriffszahlen pro Tag/Dokument werden teilnehmenden Repositorien und für Mehrwertdienste zur Verfügung gestellt
- Bereitstellung tagesgenauer Zugriffszahlen v. Dokumenten für teilnehmende Repositorien durch:
 - a) Abruf gesamter Auszüge aus Datenbank des OAS-SP über eine Schnittstelle
 - b) Angebot eines iFrame, das in das Repository eingebunden werden kann. Als Parameter muss der Identifier des Dokuments mitgegeben werden.

2.1.1. Data Provider



2.1.1. Data Provider

Der OAS Data Provider besteht im Wesentlichen aus zwei Elementen:

- Der Logfile-Parser analysiert die Server-Logfiles, wertet ergänzende Metadaten des Repositoriums aus und bündelt sie als OpenURL Context Objects. Diese werden lokal gespeichert.
- Der Data-Provider stellt diese OpenURL Context Objects über ein standardisiertes OAI-Interface zur Verfügung.

Der OAS-Data-Provider für DSpace, Webdoc und andere Repositorien ist über eine MIT-Lizenz zugänglich. Einige Elemente sind anders lizenziert. Mehr Infos finden Sie in der Datei license.txt.

2.1.1. Data Provider

Exkurs: Open Context Objects:

Ein Context Object kann optional im Fall von OpenAccess-Statistik folgende Informationen enthalten:

- Referent: genutztes Dokument (aktueller Fokus)
- ReferringEntity: Herkunftsdocument
- Requester: eindeutige und anonyme Besucheridentifikation
- ServiceType: Art der Nutzung (Volltext, Abstractseite)
- Resolver: benutzter Linkresolverserver (falls zutreffend)
- Referrer: Erzeuger des Context Object

2.1.1. Data Provider

Systemvoraussetzungen für den OAS Data Provider:

Der OAS-Data-Provider arbeitet unter einem „LAMP“-System, d.h. ein Linux Server mit dem Apache Web Server, MySQL und PHP unter diesen Voraussetzungen:

- PHP-Version 5.2.x oder eine neuere Version
 - Der Logfile-Server benötigt PHP mit pcntl und CLI (bitte beachten Sie, dass für Windows keine pcntl-Erweiterung zur Verfügung steht)
 - Beide Elemente benötigen die DOM-Erweiterung.
- Eine Datenbank zum Speichern der OpenURL Context Objects und Metadaten wird benötigt. Diese muss über PHP Data Objects (PDO) angesprochen werden können (z.B. MySQL). Darüber hinaus ist ein Datenbankzugang erforderlich, mit dem Tabellen angelegt werden können.

2.1.1. Data Provider

Installation:

- Entpacken Sie die beiden Elemente Logfile-Parser und Data-Provider.
- Der Logfile-Parser kann an jedem beliebigen Ort gespeichert werden, allerdings sollte er aus Sicherheits- und Datenschutzgründen nicht von außen zugänglich sein. Der Data-Provider sollte zugriffsbeschränkt. Der Data-Provider sollte zugriffsbeschränkt auf dem Webserver liegen und muss von außen zugänglich sein. Zum Zugriffsschutz nutzen Sie bitte .htaccess.

2.1.1. Data Provider

Konfiguration:

- Grundsätzlich: Bereitstellung angepasster Klassen für **DSpace, Webdoc und Opus**
Weg: Editieren Sie die Konfigurationsdateien im Hauptverzeichnis *Logfile-Parser*, bestehend aus *config.php & config-<name>.php*
- Wenn Sie einen anderen Server benutzen, müssen Sie diese Klassen ggf. selbst erstellen und konfigurieren. => *oasparser-webserver-standard.php* erstellt OpenURL ContextObjects. Diese stellen keinen ServiceType & keinen Identifier außer URI bereit. Deshalb muss *logfile-parser/lib/oasparser-webserver-<name>.php* durch repositorienspez. Klasse in *logfile-parser/lib/oasparser-webserver.<name>.php* erweitert werden.
- Diese braucht Programmbibliothek (*logfile-parser/lib/identifiers/lib-<name>.php*), um Metadaten zu erfassen, die nicht durch das Logfile ermittelt werden können sowie eine Konfigurationsdatei namens *logfile-parser/config-<name>.php*.
- Binden Sie die Datei *lib-<name>.php* mit *require_once* in *logfile-parser/lib/oasparser-webserver-<name>.php* ein und legen Sie in *config-<name>.php* auf die gleiche Weise Ihren oasparser fest.

2.1.1. Data Provider

Konfiguration:

- Bitte legen Sie eine Datenbank & Nutzer für Ihre ContextObjects an und gewähren Sie diesem Nutzer Schreibrechte auf die Datenbank (mit *grant*).
- Danach: Datenbankinitialisierung zur automatischen Erstellung der Tabelle für ContextObjects. Dazu Aufruf von *log2ctx -O*. Die Konfigurationsdatei *config.php* wird standardmäßig verwendet, nutzen Sie *log2ctx-c* als benutzerdefinierte Konfigurationsdatei. Der Parameter *-i* wird benutzt, um den ContextObjects eine Zeitangabe zuzuweisen, über die sich veraltete Einträge identifizieren und löschen lassen.
- Weitere Infos in *readme.txt*.
- Zwei *Unix-Shell-Scripts* zur Zusammenarbeit mit dem Logfile-Parser:
 - a) In *update-script.sh* wird beschrieben, wie Metadaten gesammelt und Webserver-Logfiles konvertiert werden. Passen Sie diese Datei für Ihren OAS Data Provider an.
 - b) Passen Sie bitte auch *cleanup-script* an und verwenden Sie diese, um veraltete ContextObjects zu entfernen.

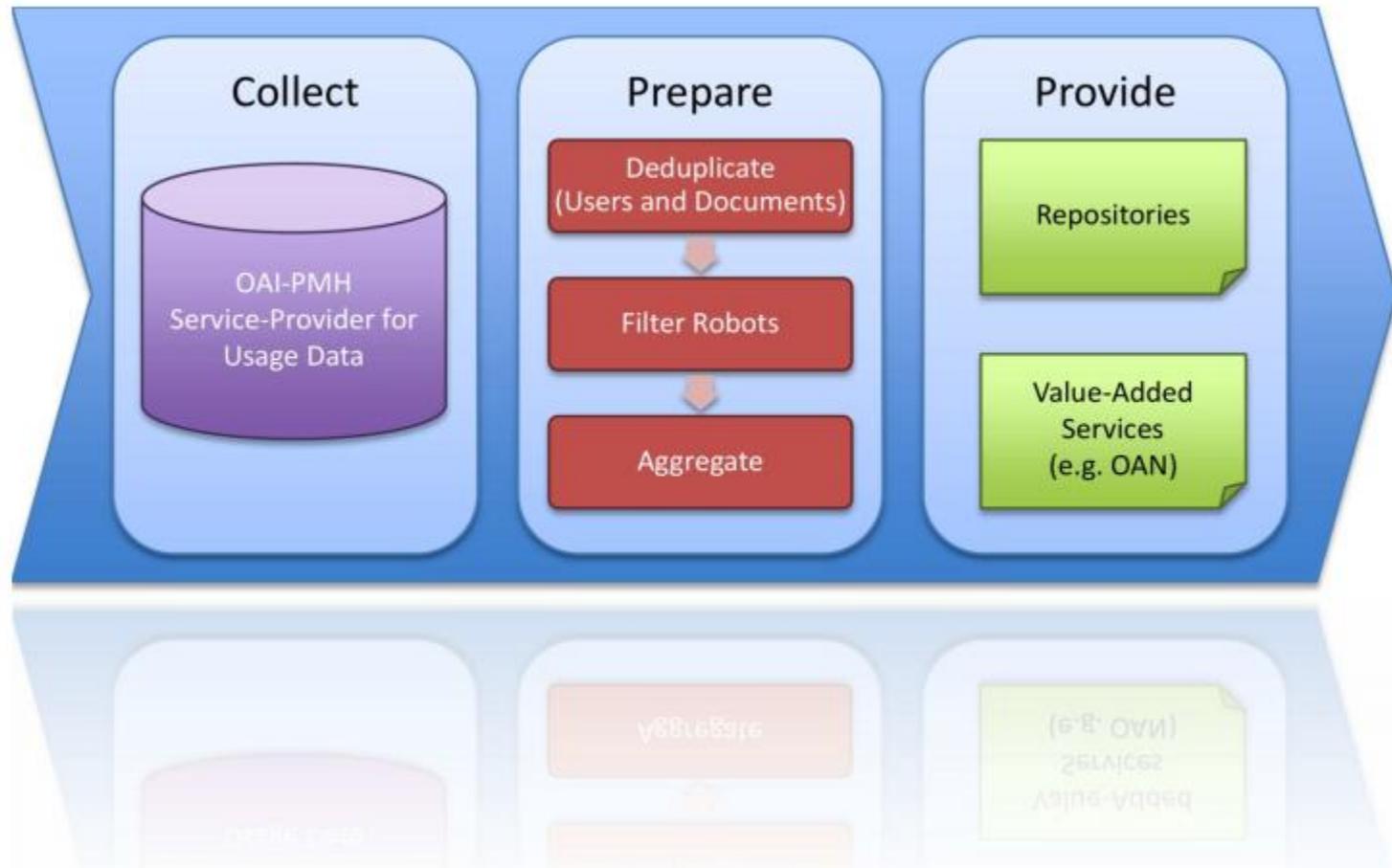
Beide Skripte sollten regelmäßig mit Hilfe eines Cronjobs aufgerufen werden.

2.1.1. Data Provider

Datenübertragung/Bearbeitung der Logdaten:

- Pseudonymisierung
 - Abschneiden des letzten Oktetts der IP
 - Hashing mit monatlich wechselndem SALT
- Isolation der Dokumentidentifikation
 - „OPUS Ident-Nummer
 - „URL
 - „URN
- Optional:
 - „Geolocation
 - ~~GeoIP~~ Extension (für PHP)
 - ~~GeoIP~~ Database (z.B. Maxmind GeoIP Database)

2.1.2. Service Provider



2.1.2. Service Provider

Was ist überhaupt ein Service Provider?

- System, das einen Service bereit stellt, der auf OAI basiert
- verwendet die über das OAI-Protokoll gesammelten Metadaten
- besteht aus zwei Komponenten
 - => Harvester (gemäß OAI-PMH)
 - => nach „außen“ zur Verfügung gestellter Service
- Harvester: kann mit Data Providern kommunizieren, also
 - => Anfragen stellen und/oder Antworten auswerten
- Service-Modul kommuniziert mit „Nutzer“

...und was zeichnet ihn aus?

- Einheitliches Berechnungsmodell:
 - Zeitspannen bei Mehrfachzugriffen
 - Crawler-/Robot-Listen
- Kann bei globalem Salt mit zunehmender Zahl teilnehmender Data Provider serviceübergreifende Effekte berücksichtigen
- Zentrale Implementation notwendiger Schnittstellen/Formate (z.B. SUSHI)

2.1.2. Service Provider

Wie funktioniert der Service Provider?

- Aggregation:
 - regelmäßiges Einsammeln neuer Nutzungsdaten aller Data Provider
 - Einstellen der Daten im Relational Database Management System
 - Verarbeitung: Auswertungslauf
 - Für auszuwertende Dokumente (definiert durch Identifier) dann Berechnungsschritt
 - Bereinigung des Datenbestands
 - „Ausmisten“ von Alt-Daten
 - Bereitstellung/Rücktransfer zum Repository
 - Schnittstellen zur Nutzung der generierten Daten
 - Transportprotokolle: OAI-PMH, RSS
 - Syntax: XML oder JSON.
 - Zeitliche Auflösungsebene: Monate
 - Objektebene: Dokumente (Volltexte)
 - Rückgabe aggregierter Daten in Formaten XML, JSON, iFrame Graphik;
- => Formatauswahl liegt bei Repositorien

2.1.2. Service Provider

Wie funktioniert der Service Provider? Aggregation:

‡

- Gemäß OAI-PMH-Standard

- Verarbeiten der Context Objects:

- „ - Parsen des XML, Kontrolle der Validität

- „ - Neue/geänderte Daten werden in eine relationale Datenbank gespeist

- „ - Herausforderungen:
 - Komplexes Datenformat, keine einfache Speicherung in einer einzigen Tabelle

- ‡
 - Anzahl der Datensätze: Beispiel „Webdoc“-Repository der SUB Göttingen (ca. 3000 Dokumente, überwiegend Dissertationen): ca. 50.000-70.000 Einzelzugriffe (HTTP) pro Tag – betrifft nötigen Rechenaufwand und Speicherplatz

- Geschwindigkeit der Datenübertragung

2.1.2. Service Provider

Wie funktioniert der Service Provider? Verarbeitung:

‡

- Für relevante Identifier (solche von Dokumenten, für die gemessen werden soll) wird 1x täglich Auswertungslauf gemacht
- Herausforderungen: Rechenaufwand, da Umsetzung der Zählalgorithmen rechenintensiv ist
- Verarbeitung unter folgenden Voraussetzungen:

Nutzerdeduplizierung

Ausfilterung von Mehrfachzugriffen durch denselben Nutzer innerhalb eines bestimmten Zeitfensters

Robotfilterung

Ausfilterung von automatisierten Zugriffen durch Robots, Spider und Crawler mittels Blacklist und anderer Verfahren

Speicherung

Tageweise aggregierte Nutzungszahlen werden zentral gespeichert und sind jederzeit abrufbar

2.1.2. Service Provider

Wie funktioniert der Service Provider? Bereitstellung:

‡

- Bereitstellung der Statistiken
 - „ - Fertige Grafiken oder Teil-Webseiten, die einfach in Repositories o.ä. eingebunden werden können
 - Darstellung folgt weiter unten
- „Reines Zahlenmaterial

2.1.2. Service Provider

Wie funktioniert der Service Provider? Datenrückführung zum Repository:

‡

1. ‡ransportprotokolle
 - OAI-PMH
 - RSS
2. ‡yntax
 - XML
3. ‡eitliche Auflösungsebene:
 - Monate
4. ‡bjektebene
 - Dokumente (Volltexte)

2.1.3. Software Download, Registrierung

Die zur Installation notwendige Software steht hier zum Download bereit:

<http://www.dini.de/projekte/oa-statistik/die-projektergebnisse/software-download/>

The screenshot shows a Mozilla Firefox browser window displaying the DINI website. The address bar shows the URL: www.dini.de/projekte/oa-statistik/die-projektergebnisse/software-download/. The page content includes:

- Navigation Menu (Left):** Startseite, Über DINI, Mitgliedschaft, Arbeitsgruppen, **Projekte**, DINI Zertifikat, Dokumente, Veranstaltungen, Service, Kalender, English, Intern.
- Main Content:**
 - Breadcrumb: DINI / Projekte / OA-Statistik / Die Projektergebnisse / Software Download
 - Language: deutsch | English
 - Software Download**
 - Data-Provider für DSpace und WebDoc**: Hier finden Sie den Quellcode des Data-Providers für DSpace- und WebDoc-Repositorien einschließlich Dokumentation sowie Beispielen für Shell-Skripte. Download (ZIP-Archiv, 110 KB)
 - Data-Provider für OPUS**: Hier finden Sie den Quellcode des Data-Providers für OPUS-Repositorien. Download (ZIP-Archiv, 112 KB)
- Übersicht (Right Sidebar):**
 - Der Dienst
 - Die Technik
 - Registrierung als Data Provider
 - Das Projekt
 - Die Projektergebnisse
 - Publikationen
 - Vorträge
 - Workshops
 - Poster & Flyer
 - Software Download
 - Die Ansprechpartner
 - Die Kooperationspartner

2.1.3. Software Download, Registrierung

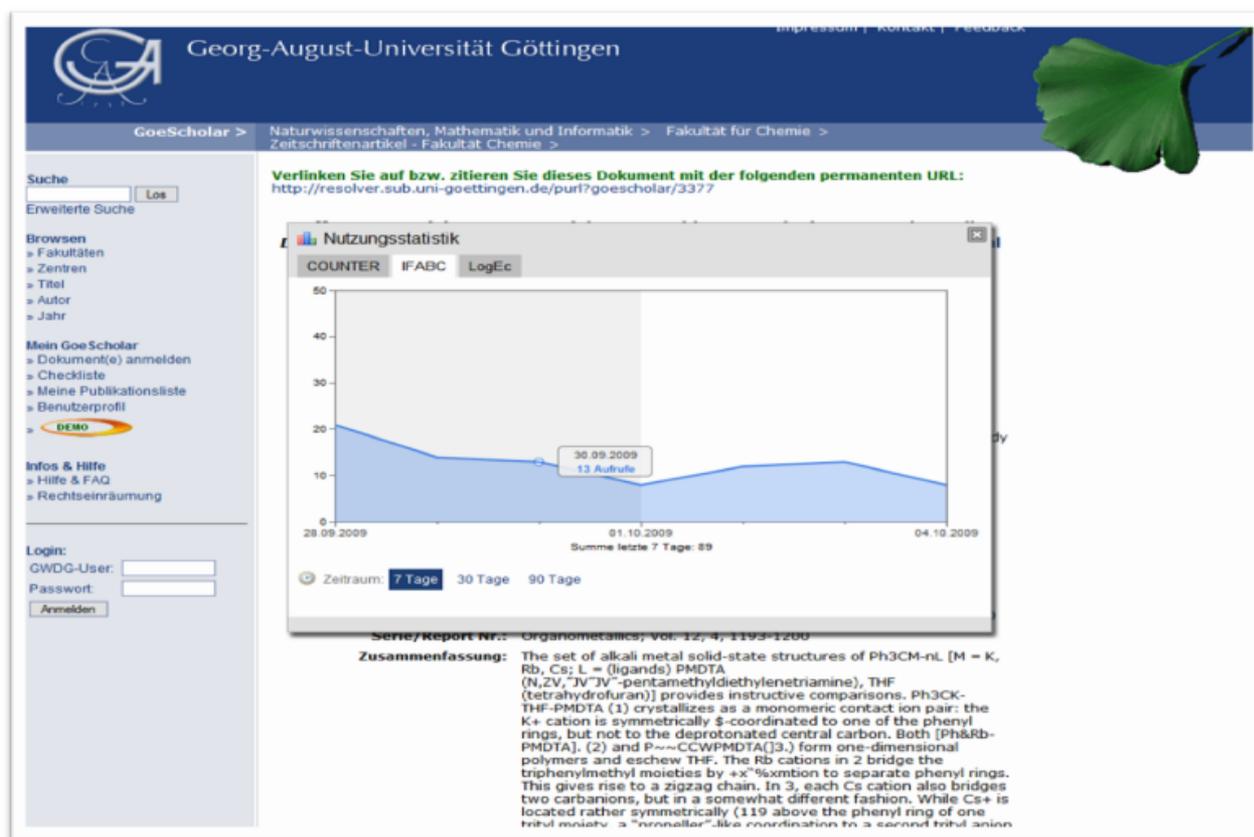
Die Registrierung erfolgt einfach über:

<http://www.dini.de/projekte/oa-statistik/die-technik/registrierung/>

The screenshot shows a web browser window displaying the registration page for becoming a Data-Provider on the DINI website. The page is titled "DINI: Registrierung als Data-Provider" and is part of the "DEUTSCHE INITIATIVE FÜR NETZWERKINFORMATION E.V." website. The navigation menu on the left includes "Startseite", "Über DINI", "Mitgliedschaft", "Arbeitsgruppen", "Projekte", "DINI-Zertifikat", "Dokumente", "Veranstaltungen", "Wettbewerbe", "Service", "Kalender", "English", and "Intern". The main content area is titled "Registrierung als Data-Provider" and includes a brief introduction and a registration form. The form is divided into three sections: "Allgemein", "Institution", and "Ansprechpartner". The "Allgemein" section includes fields for "Data Provider URL", "Name des Repositories", and "DINI Zertifikat". The "Institution" section includes fields for "Name", "Anschrift", and "Web-Adresse". The "Ansprechpartner" section includes fields for "Anrede" (Herr/Frau), "Titel", "Name", "Telefon", and "Email". A "Senden" button is located at the bottom of the form. A sidebar on the right titled "Übersicht" lists various project categories such as "Der Dienst", "Die Technik", "Das Projekt", "Die Projektergebnisse", "Die Ansprechpartner", and "Die Kooperationspartner".

2.2. Darstellungsmöglichkeiten: Aufgearbeitete Daten als Basis für Nutzungsstatistiken

Säulen-/Balkendiagramm, Liniendiagramm, evtl. Histogramm. Dabei Beschränkung auf für den Nutzer nützliche Daten und Darstellungen



2.2. Darstellungsmöglichkeiten: Aufgearbeitete Daten als Basis für Nutzungsstatistiken

• Ranking

Statistik für: psydok.sulb.uni-saarland.de
Zuletzt aktualisiert: 10.03.2010 - 04:05
Zeitraum:

[Zurück zur Hauptseite](#)

Filter: **Exklusiv-Filter:**

Seiten-URL		
Filter pdf: 1058 Unterschiedliche Seiten		
	Zugriffe	durchschnitt. Größe
/volltexte/2006/583/pdf/pdf15.pdf	1376	1.08 MB
/volltexte/2004/321/pdf/BoBe1997_1.pdf	732	107.13 KB
/volltexte/2004/151/pdf/schriftenreihe_42.pdf	562	213.19 KB
/volltexte/2005/447/pdf/Herzog_Brunner_Text04.pdf	491	206.17 KB
/volltexte/2007/981/pdf/e_Buch_MENSCHENBILDER_J._Fahrenberg_2007...	489	1.90 MB
/volltexte/2005/567/html/KapA5.pdf	462	246.91 KB
/volltexte/2004/309/pdf/FAGS-AMSE_Fragebogen.pdf	430	114.69 KB
/volltexte/2003/22/pdf/Funke_2000_Kreativitaet.pdf	425	367.81 KB
/volltexte/2009/2358/pdf/deubel.pdf	425	929.76 KB
/volltexte/2005/448/pdf/Miklautz_Jenull.pdf	412	156.75 KB
/volltexte/2004/383/pdf/Statistik_Skript_2004.pdf	392	456.39 KB
/volltexte/2006/572/pdf/4-Fachtagung-AO-Programmheft-Abstractban...	372	2.63 MB
/volltexte/2006/772/pdf/ANDROGYN.pdf	351	327.95 KB
/volltexte/2004/350/pdf/report_psychologie_06-2003_1.pdf	331	284.99 KB
/volltexte/2004/315/pdf/Diplomarbeit.PDF	319	888.55 KB
/volltexte/2004/102/pdf/02_194.pdf	299	1.12 MB
/volltexte/2007/1007/pdf/Fb1033xvvt.pdf	289	1.09 MB
/volltexte/2009/2363/pdf/SPF_Artikel.pdf	273	212.59 KB
/volltexte/2006/775/pdf/Verkehrstherapie_01_2005.pdf	258	6.63 MB
/volltexte/2004/153/pdf/beitraege_is_21.pdf	248	837.87 KB
/volltexte/2007/985/pdf/Studieren_lernen.pdf	247	884.85 KB

2.2. Darstellungsmöglichkeiten: Aufgearbeitete Daten als Basis für Nutzungsstatistiken

- Ranking

Ranking

Bitte wählen Sie das gewünschte Rankingkriterium im Menü rechts.

Die Daten zu Nutzungs- und Zitationsrelevanz stammen jeweils aus sämtlichen verfügbaren Quellen: lokal, national*, international und Verlagsdaten**

* Beteiligte Betreiber: Berlin, DNB, Göttingen, Saarbrücken, Stuttgart
 ** Beteiligte Verlage: BioMed
 *** In einem zweiten Schritt werden Sie gebeten die zugrundezulegende Metrik auszuwählen.

#	Titel	
1.	Schallehn, Volker (2007) 4 Jahre E-Books an der LMU München - Erfahrungen und Perspektiven	9.165
2.	Schäffler, Hildegard (2007) Checkliste für Standards aus bibliothekarischer Sicht : Vorstellen der Ergebnisse der AG E-Books im Bibliotheksverbund Bayern	9.053
3.	Giebenhain, Sabine (2007) E-Book-Angebot in der Universität Stuttgart - Zwischenbilanz nach 1 1/2 Jahren	

Nutzungsrelevanz*** ▼

Nutzungsrelevanz***
 Zitationsrelevanz
 Downloads***

2.2. Darstellungsmöglichkeiten: Aufgearbeitete Daten als Basis für Nutzungsstatistiken

- Relevanz

Statistik für: scidok.sulb.uni-saarland.de
 Zuletzt aktualisiert: 10.03.2010 - 04:20
 Zeitraum:

[Zurück zur Hauptseite](#)

Filter : Exklusiv-Filter :

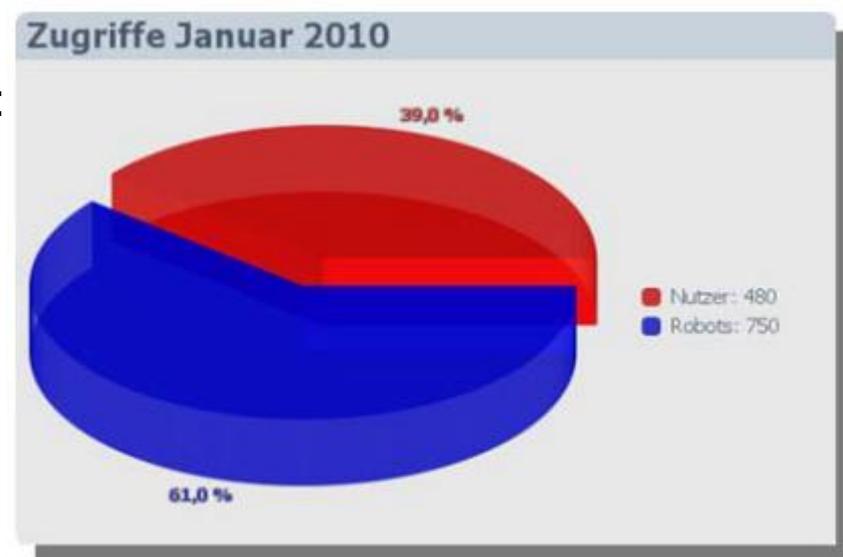
Seiten-URL			
Seiten-URL	Zugriffe	durchschnitt. Größe	
Filter /102/ : 3 Unterschiedliche Seiten			
/volltexte/2003/102/pdf/markus_becker.pdf	860	1.56 MB	
/volltexte/2003/102/	56	8.36 KB	
/volltexte/2003/102/index.html	13	15.60 KB	
Sonstige	60451	2.18 MB	

Advanced Web Statistics 6.5 (build 1.857) - Created by awstats

2.2.1. Probleme bei der Darstellung

Z.B. Relationen bestimmter Zugangstypen:

- Campuszugriffe vs. Gesamtzugriffe
- optional Robots vs. Nutzer
- Abstract vs. Volltext
- Abstrakte Zahlen oder Verlauf



2.3. Wie kann mitgemacht werden?

9 Schritte zum Glück:

1. Softwarepakete von der OAS-Projektwebseite herunterladen
2. Anpassung des OAS-Data-Providers an das eigene Repository
3. Nutzung der Datei Robots.txt
4. Geschützter Ort für den OAS-Data-Provider mit .htaccess
 - 4a. Geschützter Ort für Ihre Logfiles mit .htaccess/
geschützter Transfer zum zentralen OAS-Data Provider
5. Apache Konfiguration: Log-Konfiguration
6. Implementation Opt-out
7. Registrierung Ihres OAS-Data-Providers
8. Lizenzbedingungen akzeptieren
9. Anbieten der Ergebnisse in Ihrem Repository

3. Welche Bedeutung hat der Datenschutz

Grundsätzlich vorweg: Open Access Statistik berücksichtigt den Datenschutz über die gesetzlichen Rahmenbedingungen hinaus.

- Szenario: Zugriffe auf Dokumente eines Repositoriums werden über Logfiles des eingesetzten Webservers aufgezeichnet.
- Aus datenschutzrechtlichen Gründen wird die Möglichkeit eines **Opt-Out** angeboten. Wird diese gewählt, werden die Zugriffe von der entsprechenden IP nicht weiterverarbeitet.
- Für die Nutzer, die den Opt-Out nicht wählen gelten die folgenden Erläuterungen:

3. Welche Bedeutung hat der Datenschutz

Akteure:

Dienstanbieter (OAS, vertreten durch die Verbundzentrale Göttingen)
Dienstnehmer (Repositorien)
Nutzer
Etwaige externe Dienste/ Projekte, welche die „Produkte“ laut Lizenzbedingungen nachnutzen

Dimensionen:

Datenschutz
Verträge (zwischen OAS und Repositorien bzw. Nutzern)

Bereiche:

Datenschutz der Nutzer
Rechtsverhältnis OAS – Repositorien
Rechtsverhältnis OAS – Nutzer
(kommerziell/nicht kommerziell)

3. Welche Bedeutung hat der Datenschutz

Ziel: Nutzungsstatistiken

Methode: Auswertung der vom Webserver in Form von Logfiles protokollierten Nutzungsereignissen eines Repositoriums

Erhoben werden:

IP-Adresse, Zugriffsdatum/zeit, Zeitzone, angewandte Webbrowser, HTTP-Status und der Protokolltyp, URL, Dokumentname/typ, Get-Parameter, Referrer, Dateigröße

=> Obgleich unter Juristen konträr diskutiert, besteht für OAS ein Personenbezug, so dass die **deutschen datenschutzrechtlichen Vorgaben Anwendung finden.**

3. Welche Bedeutung hat der Datenschutz

Notwendigkeit: Juristisch zu prüfen ist also, ob die einzelnen Datenfragmente zur technischen Erstellung des Dienstes notwendig sind. Mit eigenen Mitteln kaum möglich, deswegen Unterstützung durch die



Zentrale Datenschutzstelle der baden-württembergischen Universitäten (ZENDAS, www.zendas.de) und



Schnittstelle zwischen IT und Recht

RA Michael Weller, Fachjurist der Europäischen EDV-Akademie des Rechts (EEAR, www.ear.eu)

Ergebnis: Alle erfassten Datenfragmente sind zur Erstellung des „Produkts“ technisch **notwendig**. Erhebung ist nach § 15 I TMG zulässig.

3. Welche Bedeutung hat der Datenschutz

- Neben dem Grundsatz der **Notwendigkeit** nach § 15 I TMG ist für die Speicherung auch der Grundsatz der **Erforderlichkeit** gemäß § 15 III TMG zu beachten: Ist die Speicherung der Daten technisch erforderlich?
- Ergebnis: Eine Speicherung ist unter bestimmten Voraussetzungen möglich.
- Diese Voraussetzungen bestehen darin, dass die Daten zwecks **Pseudonymisierung** einem HASH- und einem SALT-Verfahren unterworfen werden und dass die personenbezogenen Daten unmittelbar nach Beendigung des Nutzungsvorgangs zu löschen sind (§ 13IV2 TMG), unmittelbar \triangleq wenige Minuten
- Grund: Erfordernisse des § 3 VIa BDSG – statt der klaren Daten werden Identifikationsmerkmale genutzt mittels derer Bestimmung der Betroffenen ausgeschlossen ist.

3. Welche Bedeutung hat der Datenschutz

- **Pseudonymisierung 1:** Zwar kann ein DP als verantwortliche Stelle über die Kenntnis der Zuordnungsregel eine Person identifizieren. Dritte wie der SP kennen diese Zuordnungsregel hingegen nicht (oder können sie nur mit einem unverhältnismäßig großem Aufwand an Zeit, Geld und Arbeitskraft erfahren), so dass die Daten dort anonymisiert eintreffen.
- **Pseudonymisierung 2:** Codierung der Daten durch Hashverfahren SHA-256. Verhinderung Rückcodierung durch SALT => doppelte Codierung, die datenschutzrechtlichen Vorgaben sind für den SP nicht mehr zu beachten.
- Voraussetzung: DP und SP müssen räumlich getrennt sein (auch der zentrale DP), DP dürfen untereinander keine Daten austauschen können, das SALT wird regelmäßig und **nicht** durch den SP geändert.
- Vorsichtshalber (automatisierte Zugriffe !) wird auch das C-Klasse-Netz der IP-Adresse (die ersten drei der vier Zahlen der IP-Adresse) gesalzen und gehasht.
- **Damit ist es möglich, Nutzungsinformationen aufzuzeichnen, ohne Nutzerinformationen zu verwenden.**

3. Welche Bedeutung hat der Datenschutz

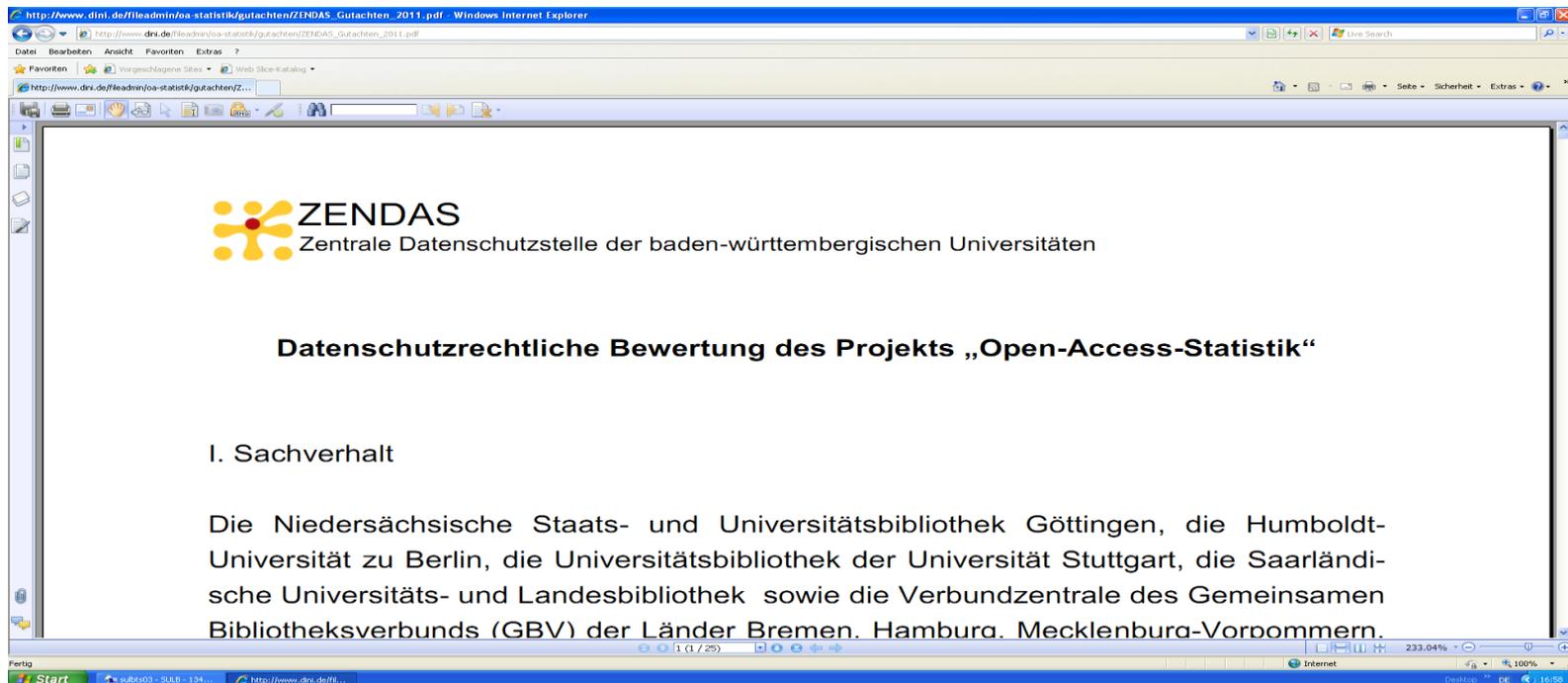
Noch einmal kurz zum Überblick:

- Hash: dient Pseudonymisierung und damit der Identifikation der von einer IP-Adresse erfolgten Nutzungsmuster, ohne die ursprüngliche IP bei der Datenverarbeitung zu verwenden.
- Salt: Beschreibt eine weitere Schutzmaßnahme zur Verhinderung der Rückübersetzung des Hashwertes.
- Dazu 1: Für eine spätere Berücksichtigung von Roboterzugriffen und zur Auswertung der Daten wird zusätzlich auch das C-Klasse-Netz der IP-Adresse (die ersten drei der vier Zahlen der IP-Adresse) gesalzen und gehasht.
- Dazu 2: Räumliche Trennung zwischen DP und SP
- Dazu 3: Kein Datenaustausch zwischen den DP
- Dazu 4: Regelmäßige Änderung des SALT (NICHT durch SP!)

3. Welche Bedeutung hat der Datenschutz

- Die genauen datenschutzrechtlichen Ausführungen sind in der datenschutzrechtlichen Bewertung der Zentralen Datenschutzstelle der baden-württembergischen Universitäten nachlesbar:

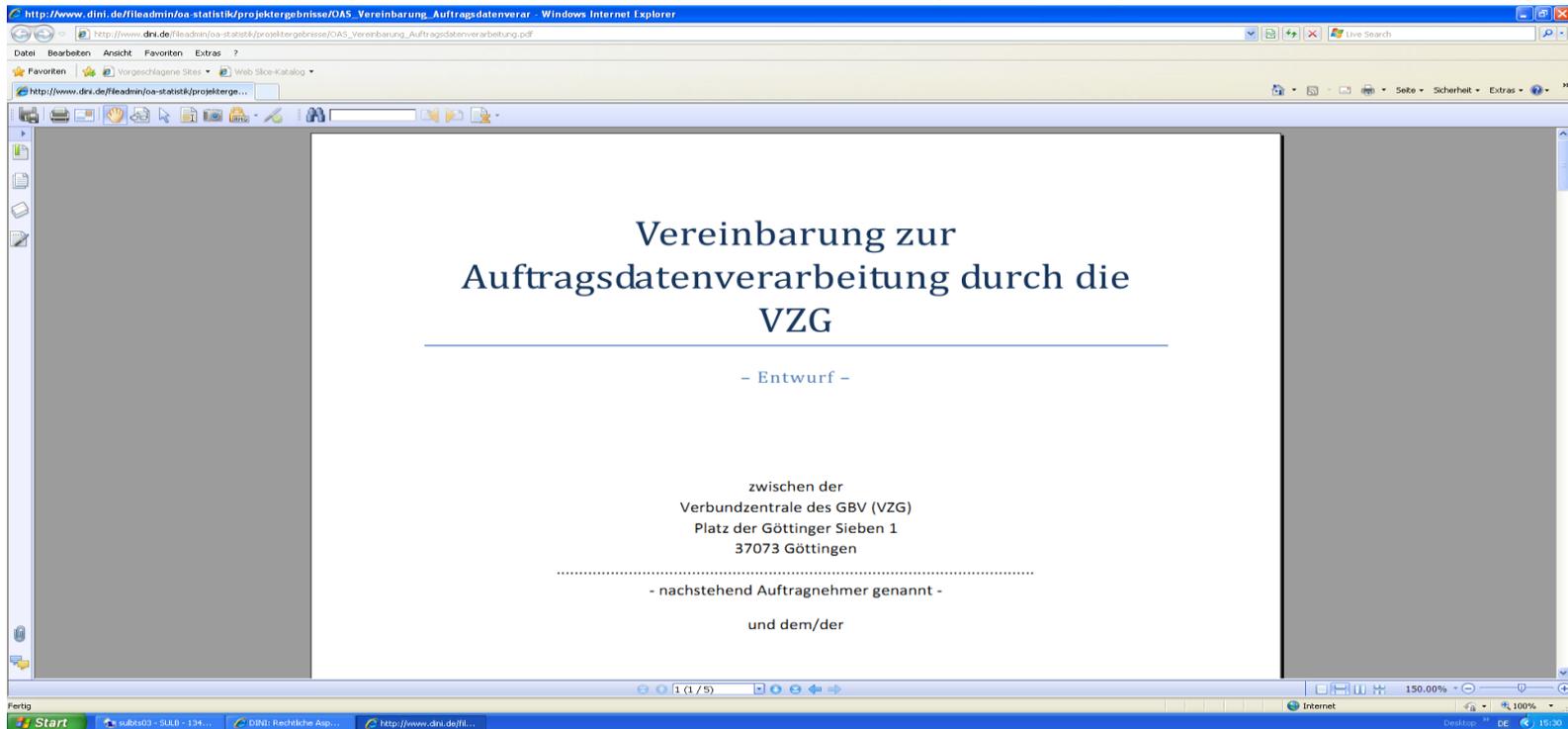
http://www.dini.de/fileadmin/oa-statistik/gutachten/ZENDAS_Gutachten_2011.pdf



3.1. Welche Vertragsverhältnisse bestehen? Repositorien

- Der Wortlaut für den vertraglichen Rahmen zwischen der VZG als Auftragnehmer und Repositorien als Auftraggeber sind hier nachlesbar:

http://www.dini.de/fileadmin/oa-statistik/projektergebnisse/OAS_Vereinbarung_Auftragsdatenverarbeitung.pdf



3.1. Welche Vertragsverhältnisse bestehen? Nutzer

- Ausgangspunkt: Creative Commons BY-NC-SA-Lizenz 3.0 (Deutschland)
- Problem: Werk = Urheberrecht (=CC-Lizenzen), Datenbank = Leistungsschutzrecht. Welcher Bereich der Schutzfähigkeit liegt hier vor?
- Urheberrechtlicher oder leistungsrechtlicher Schutz hängt davon ab, welche Art der Bereitstellung der Informationen gewählt wurde.
- Deswegen Modifizierung des Lizenztextes CC BY-NC-SA auf die Bedürfnisse von OAS über den Verzicht auf Schutzrechte an Datenbanken/Datenbankwerken
- Mit diesem neuen Lizenztext ist es möglich, die Daten unter Beachtung der Schranken des deutschen Urheberrechts auch in Form einer Datenbank zur Verfügung zu stellen.
- Langfassung der Lizenz unter: http://www.dini.de/fileadmin/oa-statistik/projektergebnisse/OAS_Lizenz.pdf

3.1. Welche Vertragsverhältnisse bestehen? Nutzer

LIZENZZUSAMMENFASSUNG

SIE DÜRFEN:

-  die Daten vervielfältigen, verbreiten und öffentlich zugänglich machen,
-  Abwandlungen und Bearbeitungen der Daten anfertigen,

WENN SIE:

-  die Herkunft der Daten durch die Angabe *hier die gewünschte Bezeichnung des Datenanbieters eintragen* kenntlich machen;
-  die Daten nicht zu Zwecken Ihres Geschäftsbetriebes gegen Entgelt vervielfältigen, verbreiten, öffentlich zugänglich machen, Abwandlungen und Bearbeitungen davon anfertigen;
-  die Daten sowie Abwandlungen und Bearbeitungen hiervon unter den Bedingungen der zugrunde liegenden Lizenz weitergeben.

HIERBEI GILT STETS:

Eine über die in der Lizenz genannte Nutzung hinausgehende Verwendung ist durch Individualvereinbarung mit *hier die gewünschte Bezeichnung des Datenanbieters eintragen* möglich.

Gemeinfreie oder nicht schützbar Daten werden von der Lizenz nicht erfasst.

Die Lizenz beschränkt den Nutzer nicht in Rechten, die ihm aufgrund der urheberrechtlichen Schrankenbestimmungen.

Unveräußerliche Rechte, die dem *hier die gewünschte Bezeichnung des Datenanbieters eintragen* aufgrund eines etwaigen Urheberpersönlichkeitsrechts zustehen.

Im Falle der Verbreitung müssen Sie anderen alle Lizenzbedingungen mitteilen, die für die Daten gelten.

3.2. Wodurch finanziert sich Open Access Statistik?

- Die Kosten werden durch die VZG Göttingen als Rechnungsstelle erhoben.
- Neben einer Grundfinanzierung (in der Projektphase ist die Teilnahme **kostenlos!!**) durch die Projektteilnehmer und die DFG tragen die teilnehmenden Repositorien zur nachhaltigen Finanzierung bei.
- Grundsätzlich gilt, das OAS nicht gewinnorientiert arbeitet, sondern lediglich eine Kostendeckung anstrebt.
- Bei der Berechnung der Kosten ist von zwei Modellen auszugehen:
 - a) Lokaler OAS-Data Provider,
 - b) Zentraler OAS-Data-Provider

3.2. Wodurch finanziert sich Open Access Statistik?

- **Kostenbeispiel Lokaler OAS-Data Provider:**
- Annahme für das Beispiel: 10 Teilnehmer, Dauer der Teilnahme 5 Jahre:

Jahr (einmalig):	Fixe Stückkosten:	1.200,- € ($\hat{=}$ 2 Tagessätze á 600.- € für Standardaggregation pro Repositorium und Institution)
Kosten pro Jahr: (Folgejahr)	Fixe Gesamtkosten:	2.250,- €
	Variable Stückkosten:	225,- € pro Jahr bei 10 Teilnehmern
	Dauer 5 Jahre	1.125,- € ($= 225,- € * 5$)
	Fixe Stückkosten:	150,- € pro Jahr ($\hat{=}$ 3 Stundensätze á 50,- € für Softwarepflege)
Kosten 1. Jahr insgesamt:		1.200,- € (inkl. 375,- €)
Kosten 2. – 5. Jahr pro Jahr insgesamt:		<u>375,- €</u>
Gesamtkosten/Teilnehmer nach 5 Jahren bei 10 Teilnehmern:		2.700,- € ($= 1.200 + 4 * 375$)

3.2. Wodurch finanziert sich Open Access Statistik?

Kostenbeispiel Zentraler OAS-Data Provider:

- Kostenbeispiel **zusätzlich** zu den Kosten für den lokalen Data-Provider
- Annahme für das Beispiel: 10 Teilnehmer, Dauer der Teilnahme 5 Jahre:

Jahr (einmalig):	Fixe Stückkosten:	1.800,- € ($\hat{=}$ 2+1 Tagessätze á 600.- € Standardaggregation pro Repositorium und Institution)
Kosten pro Jahr:	Fixe Gesamtkosten:	3.750,- € ($\hat{=}$ 2.250 + 1.500,- Betriebskosten ZOAS-DP)
(Folgejahr)	Variable Stückkosten:	375,- € pro Jahr bei 10 Teilnehmern
	Dauer 5 Jahre $\hat{=}$	1.925,- € (= 375 € * 5)
	Fixe Stückkosten:	300,- € pro Jahr ($\hat{=}$ 3+3 Stundensätze á 50,- € für Softwarepflege)
Kosten 1. Jahr insgesamt:		1.800,- € (inkl. 675.- €)
<u>Kosten 2. - 5. Jahr pro Jahr insgesamt:</u>		<u>675,- €</u>
Gesamtkosten/Teilnehmer nach 5Jahren bei 10 Teilnehmern:		4.500,- € (1.800 + 675 * 4)

4. Wo finde ich weitere Informationen?

Weiterführende Informationen sind auf unserer Projektwebsite unter

<http://www.dini.de/projekte/oa-statistik/>

zu finden.

Darüber hinaus lohnt sich ein Blick auf das Informationsportal CARPET unter:

<http://www.carpet-project.net/>

Natürlich können Sie uns auch unter oas@dini.de immer direkt kontaktieren!

4.1. Wer ist mein Ansprechpartner?

Für Fragen zu Open Access Statistik schreiben Sie bitte eine Mail an:

oas@dini.de

Vielen Dank für Ihre Aufmerksamkeit!

Initiiert durch:



Webinar, OA Statistik, Herb/Holtz/Wanke

Gefördert durch:

Deutsche
Forschungsgemeinschaft

30.01.2013 73 **DFG**