

EconStor – Umsetzung eines Fach-Repositorys unter DSpace

Wolfgang Riese und Jan B. Weiland

Deutsches DSpace-Anwendertreffen 2016

ZBW – Leibniz-Informationszentrum Wirtschaft

Hamburg, 27.09.2016

EconStor im Überblick (1)

- Fachliches OA-Repository für die Wirtschaftswissenschaften (www.econstor.eu)
- Communities: beteiligte Institutionen
- Collections:
 - Schriftenreihen, Zeitschriften & Konferenzen
 - Self-Archiving, LeibnizOpen
- DB-intern erweitertes Metadaten-Schema:
 - Qualified Dublin-Core
 - mit zusätzlichen Unterfelder (v.a. Aufsatz-Zitationsangaben)
- „On the Fly“-Generierung von PDF-Deckblättern
- DINI-Zertifikat 2010

EconStor im Überblick (2)

- Downloadstatistik „COUNTER-like“ (s.u.)
- Migration auf DSpace-Version 5.5 (Mai 2016)
 - umfassender Relaunch der Webseiten
 - Mobil-Optimierung...(s.u.)
- aktueller Content: ~125.000 PDF-Volltexte
 - ~67% Arbeitspapiere
 - ~20% Zeitschriftenaufsätze
 - ~ 7% Konferenzbeiträge

Import-Service für Forschungsinstitute

- kostenloser „Full Service“
 - pauschale Übertragung von Nutzungsrechten
 - Verwendung verfügbarer Metadaten für den Import:
 - Exporte aus GBV-Verbundkatalog
 - RePEc-Datenbank
 - OAI-Daten (v.a. OJS-Zeitschriften, IR's)
 - Konferenz-Datenbanken
 - Excel-Tabellen...
 - DSpace-Import via „Simple Archive Format“
 - Datenpflege via „Batch Metadata Editing“-Feature, vgl.
 - DuraSpace-Wiki
 - Powerpoint von M. Walsh
-

[Metadaten-Export für Administratoren]

The screenshot shows a search interface with a search bar containing 'EconStor gesamt'. Below the search bar are buttons for 'Suchen' and 'Neue Suche starten'. A filter section is visible with 'Titel' and 'gleich' options. The results section shows 'Treffer/Seite' set to 10 and a button for 'Export metadata' highlighted in yellow. A file opening dialog titled 'Öffnen von search-results.csv' is overlaid on the interface. The dialog shows the file 'search-results.csv' from the URL 'https://www.econstor.eu'. The 'Wie soll Firefox mit dieser Datei verfahren?' section has 'Öffnen mit' selected, with 'LibreOffice Calc Portable (Standard)' chosen in the dropdown menu. Other options include 'DownThemAll!', 'dTa OneClick!', and 'Datei speichern'. The 'Datei speichern' option is also highlighted in yellow. The dialog has 'OK' and 'Abbrechen' buttons at the bottom.

Self-Archiving für Einzelpublikationen

- „*EconStor Direct*“
- Self-Upload durch AutorInnen
- Registrierung beschränkt auf affilierte ForscherInnen (weltweit):
 - Vorab-Prüfung der E-Mail-Adresse!
- 6 Collections mit üblichen Dokumentarten (Aufsätze, Preprints...)
- Standard-Deposit-Lizenzen
- Metadaten-Felder auf wenige „Kerndaten“ beschränkt
 - Upload von BibTeX-Metadaten möglich
- Interne Bearbeitung über DSpace-Submission-Workflows

[BibTeX-Upload]

Beschreiben Heraufladen Überprüfen Lizenzieren Abschließen

Publikation beschreiben

Bitte füllen Sie die erforderlichen Informationen zu dieser Einreichung aus. In den meisten Webbrowsern können Sie die Tabulator-Taste verwenden um zum nächsten Eingabefeld zu gelangen. Zum Speichern müssen Sie jeweils mit der Maus auf den entsprechenden Schalter klicken.

i Hier können Sie eine bereits bestehende BibTeX-Datei hochladen, um die Metadaten Ihrer Publikation in das Formular einzufügen:
!Achtung alle bisherigen Eingaben im Formular werden gelöscht!

Durchsuchen... Keine Datei ausgewählt.

Unicode (UTF-8)

Bitte bestätigen Sie an dieser Stelle nochmals den Publikationstyp des eingereichten Dokuments bzw. wählen einen genaueren Typ aus.

i Publikationstyp *

Zeitschriftenaufsatz / Article
Buchaufsatz / Book Part
Buchbesprechung / Book Review
Zeitungsartikel / Contribution to Periodical

i Geben Sie den Titel der Publikation ein.

Titel *

i Geben Sie den Namen des Autors/des Autors bzw. die Namen des Autors/des Autors ein. (Bitte folgen Nachname, Vorname(s))

Verbreitungsservice (1): RePEc

- international genutzte, frei zugängliche Fach-Datenbank
- direkte Verbreitung in die Fach-Community via
 - IDEAS
 - EconPapers
 - New Economic Papers (Alerting-Dienst)
- international etablierte RePEc-Rankings:
 - Zeitschriften & Schriftenreihen
 - Autorinnen & Autoren
 - Forschungsinstitute
- EconStor gelistet als „Nationaler Input-Service“
- „externes“ PHP-Skript erzeugt RePEc-Metadaten (ReDIF-Format)

Verbreitungsservice (2): LeibnizOpen

- „Mediated Deposit“ durch die Bibliotheken der kooperierenden Leibniz-Institute (DIW, ZEW, IWH...):
 - lokales Monitoring institutioneller Veröffentlichungen
 - Prüfung der Nutzungsrechte
 - „Self“-Archiving von Postprints (mit Autoren-Vollmacht!)
- Einrichtung spezieller Leibniz-Collections auf EconStor:
 - Collection-Admin-Rechte für Leibniz-BibliothekarInnen
 - v.a. für das „Item Mapping“
- „Selective Harvesting“ der Leibniz-Collections (= OAI-Sets) durch LeibnizOpen über eigenes OAI-Metadaten-Format

Verbreitungsservice (3): Google Scholar

- „Google Scholar Indexing Guidelines“ implementiert:
 - Ausgabe der Metadaten in normierten „HTML Meta Tags“
- bis Juni 2016:
 - hohe Abdeckung der EconStor-Volltexte im Google-Scholar-Index
 - lt. Referrer: ~50% der Volltext-Zugriffe aus Google Scholar!
- seit Juni 2016:
 - Weiterleitung aller EconStor-URLs auf *https*-Adressen
 - von Google Scholar interpretiert als „*PDF not available*“!
 - ➔ EconStor-Titel aus GS-Trefferlisten entfernt!
 - ➔ deutlicher Rückgang der EconStor-Zugriffe

DNB-Pflichtablieferservice

- optional für Wiwi-Institutionen aus Deutschland
 - v.a. außeruniversitäre ohne eigenes Repository
- OAI-Format „xMetaDissPlus“ implementiert
- jede Community als eigenes OAI-Set abrufbar:
 - Ablieferservice kann pro Institution eingerichtet werden!
 - evtl. aus „fremden“ Communities gemappte Titel nicht im OAI-Set enthalten!
- systematische Nachpflege vorhandener, „externer“ urn:nbn-Identifizierer in der EconStor-DB notwendig

Relaunch 2016 (1): Ziele

- Update auf aktuellen Programmcode (v1.4.2 auf v5.5)
 - Anschluss an aktuelle DSpace-Entwicklung
 - Verbesserte Wartbarkeit – Monolithisches Projekt (v1.4.2) vs. Modulare Projekte (JSPUI, API, SOLR, OAI, usw. - v5.5)
- Umstellung der Dokumentsuche in EconStor auf das ZBW-eigene Produkt EconBiz
- Mobiloptimierung der Weboberfläche (Bootstrap)
- Positive Effekte beim Google-Ranking (Mobiloptimierung, moderneres HTML der DSpace-Oberfläche)

Relaunch 2016 (2): Durchführung I

1. Ende Nov. 2013 bis Ende Dez. 2013

"Einbau" unserer DSpace Version 1.4 in einen lokalen Git-Klon und Update auf Version 1.5 (Maven basierter Build-Prozess – vorher Ant).

Zeilen Code in Git: +739.020, -347.491

2. Dez. 2014 bis Ende März 2015

Nachziehen von Updates und Fixes des Produktivsystems und stückweises „mergen“ der Major-Versionen bis DSpace 5.1

Zeilen Code in Git: +79.351, -54.195

Relaunch 2016 (3): Durchführung II

3. September 2015 bis Ende März 2016

Zusammenführen unserer Anpassungen mit dem aktuellen DSpace JSPUI via Maven Overlay, mit anschließender Implementierung des neuen EconStor-Layouts und „mergen“ der jeweils aktuellen Version bis DSpace 5.4

Zeilen Code in Git: +167.820, -211.110

4. April 2016 bis Launch am 2016-05-02

Anpassungen der Suche, des Layouts und der englischen Texte, sowie Mergen der aktuellen stabilen Version 5.5 und Relaunch von EconStor

Zeilen Code in Git: +15.058, -24.488

Relaunch 2016 (4): Durchführung III

5. Fine-Tuning bis Ende August 2016

Layout-Optimierungen, kleinere Bug-Fixes, Optimierungen für Google (Kombinieren und Minimieren der Java-Scripte und CSS-Dateien, aktivieren der serverseitigen Kompression von textuellen Daten wie HTML, CSS, JavaScript – erzwingen von https für die gesamte Webseite)

Zeilen Code in Git: +34.420, -27.533

Geänderte Zeilen Programmcode im Git-Repository seit Beginn des Projekts „EconStor Relaunch“ (DSpace 1.4x auf 5.x) insgesamt:

Neu hinzugekommen oder Ersetzt: +1.035.669

Entfernt oder Ersetzt: -664.817

Downloadstatistik nach COUNTER (1)

- Apache-Logfile-Analyse (combined Format)
- Eigenentwicklung als DSpace Modul mit PostgreSQL-Backend
- Monatliche Push-E-Mail mit eigener Statistik je DSpace-Community
- Übersichtsseiten je Dokument / Collection / Community / Site
- Seiten mit Vergleich aller Collections bzw. aller Communities

Downloadstatistik nach COUNTER (2)

- Implementiert die 10/30 Sekunden Doppelklick Regel für HTML-Views/ PDF-Downloads von COUNTER
- Verschärft auf ein Download je Dokument + IP pro Tag. Alle nachfolgenden Downloads des Dokuments von dieser IP werden für den Tag ignoriert
- Verschärft auf max. 3.000 Downloads je IP und Monat, bei Überschreiten werden die Downloads der IP in dem entsprechenden Monat verworfen
- Erweiterte COUNTER-Filterliste für Bots anhand des Useragents
- Manuel erstellte IP-Liste von auffällig gewordenen IP-Adressen
- Nur erfolgreiche Downloads (HTTP Response Code 200 oder 302)

Downloadstatistik - Links

- <https://www.projectcounter.org/>
- <http://www.irus.mimas.ac.uk/>
- <https://www.coar-repositories.org/activities/repository-interoperability/usage-data-and-beyond/>
- <https://www.gbv.de/Verbundzentrale/serviceangebote/oas-service/open-access-statistik-service>
- <http://www.slideshare.net/eservice/icanhazrobot-improved-robot-detection-for-ir-usage-statistics>