

Von PUB zu LibreCat - Auf dem Weg zu einem universell einsetzbaren Publikationsmanagement-System



Von PUB zu LibreCat

Agenda

- PUB? LibreCat? Catmandu? – Who is who?
- Auf dem Weg zu LibreCat – erste Schritte
 - Eine Exkursion zur ELAG-Konferenz nach Kopenhagen
- Stand der Dinge
- ToDo
- Zukunft

Who is who? - PUB

PUB – Publikationen an der Universität Bielefeld

- Publikationslisten, Forschungsdaten, Dissertationen, graue Literatur...
- Personen, Projekte, Preise und Auszeichnungen...
- Schnittstellen, Filter, Suche, Importe, Exporte, Formate...
- Verknüpfungen all dieser Punkte untereinander

Who is who? - PUB

Mehr als ein Repository

Weniger als ein Forschungsinformationssystem

PUB ist ein Publikationsmanagement-System

Who is who? - PUB

PUB ist Bielefeld-spezifisch

- Uni-Bielefeld Design, Logos, Farben, Texte
- LDAP, PEVZ, DOI-Registrierung, URN-Registrierung, Zitationsstile, PUB-Theses, Fakultätszuordnung, Projektzuordnung, Rollen, Rechte ...
- Zugeschnitten auf die Anforderungen unserer Bielefelder Wissenschaftler

Who is who? - LibreCat

- Zunächst: Name einer Arbeitsgruppe (UB Lund, UB Gent und UB Bielefeld)
- Das Catmandu-Toolkit löste sich heraus, wird entwickelt in Gent
- Die UB Lund tritt in den Hintergrund, ist nur noch Anwender
- Die UB Bielefeld baut ein komplettes Produktiv-System auf Basis von Catmandu

Einschub: Who is who? - Catmandu

- Catmandu sind Perl-Werkzeuge
 - zur einfachen Datenverarbeitung und –veränderung
 - zum einfachen Datenzugriff
 - für immer wiederkehrende Arbeits- und Programmierabläufe
- Als Perl-Modul über CPAN verfügbar, Bestandteil von Ubuntu 15.04
- "Catmandu vereinfacht 20 Zeilen Code zu 3 Zeilen"
- Catmandu nutzt durchgehend utf-8 (keine Umlautprobleme mehr!)

Who is who? - LibreCat

- Oft werden wir gefragt, wo man die PUB-Software finden kann
- Wir wollen PUB zu github bringen und frei verfügbar machen
- ABER: PUB ist Bielefeld-spezifisch, PUB ist komplex, PUB verwendet eine große Sammlung von Perl-Modulen
- UND: jeder hat andere Anforderungen

Wir möchten PUB verallgemeinern -> LibreCat

Who is who? - LibreCat

Das Ziel ist es, die Frage

Was ist LibreCat?

in Zukunft so beantworten zu können:

LibreCat ist ein Grundgerüst, ein Basispaket, mit dem man sich in wenigen Schritten ein Publikationsmanagement-System wie PUB aufbauen kann.

Who is who? - LibreCat

PUB ist die erste Instanz auf Basis von LibreCat.

- Die SUB Göttingen wird eine weitere, individuelle Instanz von LibreCat aufbauen.
- Die UB Gent noch eine weitere.

Sie werden alle

- unterschiedlich aussehen,
- unterschiedliche Dienste bieten und
- unterschiedliche Komponenten verwenden.

Aber sie werden alle grundlegend gleich funktionieren.

Who is who? - LibreCat

Klingt gut? - MOMENT!

Noch ist das Zukunftsmusik. Wir sind noch nicht (ganz) fertig.

Was ist dafür noch nötig?

- LibreCat muss frei von UniBi Elementen sein
- frei konfigurierbar sein und
- möglichst frei von Bugs sein.

Und

- Jede Institution muss ihre Daten migrieren!

Wir in Bielefeld haben das Pferd von hinten aufgezäumt.

Auf dem Weg zu LibreCat – erste Schritte

Eine Exkursion zur ELAG Konferenz nach Kopenhagen

ELAG – European Library Automation Group

<https://elag.org>

ELAG Konferenz

- Europas größte Konferenz zur Anwendung von Informationstechnologien in Bibliotheken
- Veranstaltet seit mehr als 30 Jahren
- Publikum: Bibliotheks- und IT-Spezialisten
- Themen (u.a.): Digitale Bibliothek, Web Services, Langzeitarchivierung, Open Source Software

ELAG 2016

7-9. Juni 2016 - Kopenhagen, Dänemark

<http://elag2016.org/>

Thema 2016:





LibreCat

Transforming an Institutional Repository

ELAG Conference 7-9 June 2016 in Copenhagen, Denmark
Petra Kohorst, Vitali Peil (Bielefeld University Library)

"LibreCat - Transforming an Institutional Repository", ELAG Conference

AGENDA

- ~~INSTITUTIONAL REPOSITORY~~
- TECHNOLOGIEN
- DATEN
- NUTZER (INTERFACE)
- ~~WHAT'S NEXT?~~

"LibreCat - Transforming an Institutional Repository", ELAG Conference

Technologien

AUSSTIEG

- kompliziertes DB Schema, Template Engine, veraltete Module, schlecht zu erweitern

ÜBERGANG

- Entscheidungen: welche Technologien? -> Nutzt Dinge, die ihr kennt!

EINSTIEG

- Open Source, moderne Frameworks, sichere aber schnelle Haupt- und Indexspeicher
- Und für alle Daten, die rein oder raus gehen: use Catmandu;

"LibreCat - Transforming an Institutional Repository", ELAG Conference

Daten

AUSSTIEG

- Verwirrende Daten aufgrund von "geerbter" Datenbank(struktur)
- Viel "Datenmüll", der sich in der DB angesammelt hat
- Überflüssige Felder und Relationen

"LibreCat - Transforming an Institutional Repository", ELAG Conference

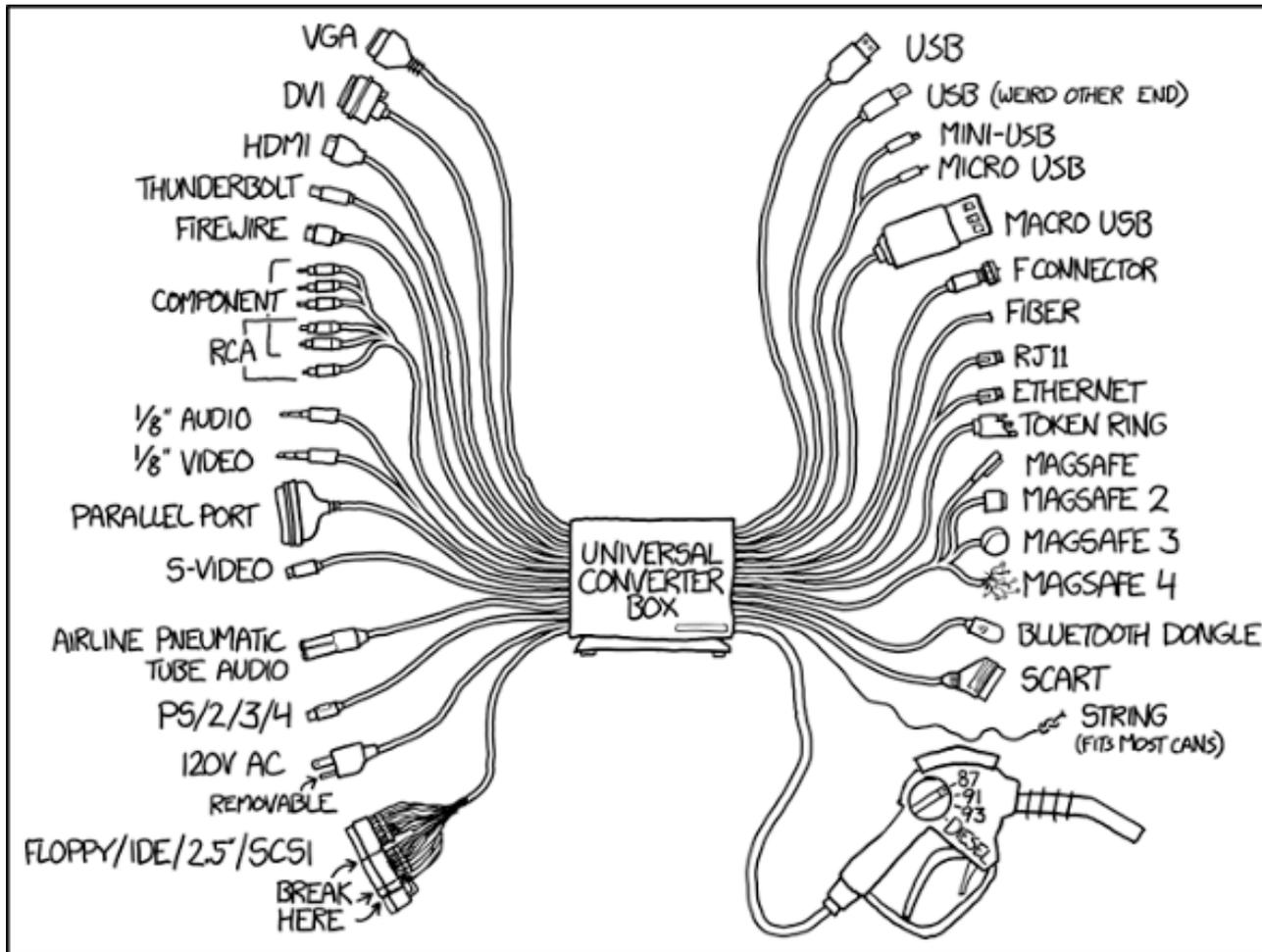
Daten

ÜBERGANG

- Datentransformation mit Catmandu
 - technisch einfach, inhaltlich sehr kompliziert
 - Felder **erstellen**, **umbenennen**, **neu strukturieren**, **löschen**
 - Daten in den Feldern aufräumen
- Neue Struktur des FileStore
 - Aufgabe: sicheren Weg finden, Dateien/Volltexte zu migrieren
- Überraschungen finden wir heute noch!

"LibreCat - Transforming an Institutional Repository", ELAG Conference

Datentransformation fühlt sich an wie der Versuch, dies hier zu bauen...



[CC BY-NC 2.5]
<https://xkcd.com/1406/>

"LibreCat - Transforming an Institutional Repository", ELAG Conference

Daten

EINSTIEG

- MySQL als Hauptdatenspeicher - nur json-Blobs, also reiner Text
 - Nach Datentyp (publication, person, project) auf Tabellen verteilt
 - Einfache Struktur, leicht zu warten
- Elasticsearch Suchindex
 - Gleiche Struktur wie Hauptdatenspeicher, aber indexiert
 - Beim Indexieren mit weiteren Daten anreichern
- Leichter Zugriff auf alle Daten Dank Catmandu

"LibreCat - Transforming an Institutional Repository", ELAG Conference

Nutzer Interface

AUSSTIEG

- Veraltetes Design, nicht responsiv, kompliziert

ÜBERGANG

- Wie geht "intuitiv bedienbar"?
- Übergang? **Keiner!** Wir haben einfach umgeschaltet und die Nutzer mit dem neuen System konfrontiert.

EINSTIEG

- Schnell, responsiv, benutzerfreundlich (hoffentlich), intuitiv (hoffentlich), leicht(er) zu warten

Zurück nach Deutschland...

Stand der Dinge

- LibreCat ist auf github <https://github.com/LibreCat/LibreCat>

Wir haben:

- Installationsanleitung
- automatische Perl-Modul-Installationen (via cpan-file)
- Dummy-Datensätze für den Schnellstart mit Datenbanken
- Unterstützung für verschiedene Betriebssysteme
- Konfigurierbarkeit in einigen Bereichen, aber noch längst nicht allen

ToDo

- Austauschbare File- und MainStores
- Jobserver zur Lastverteilung
- CLS-Engine modularisieren
- Validierung vor dem Speichern
- Schema für Daten festlegen
- Design verallgemeinern
- Design (etc.) lokal anpassbar machen
- Dokumentation, Dokumentation, Dokumentation
- Trennung von Backend und Frontend durch API

Zukunft

- LibreCat in
 - Gent und Göttingen als Publikationsmanagement-System
 - Lund als Oberfläche für kleinere Datenbanken
 - Bielefeld als Basis für PUB
- LibreCat in Bielefeld und Göttingen als Grundlage für ein System, das in Richtung "Forschungsinformationssystem" geht.

"LibreCat - Transforming an Institutional Repository", ELAG Conference

JOIN US!

- <http://librecat.org> - home page
- <https://librecatproject.wordpress.com> - blog
- <https://github.com/LibreCat/LibreCat> - code
- librecat-dev@lists.uni-bielefeld.de - mailing list

Vielen Dank!