



Staatsbibliothek
zu Berlin
Preußischer Kulturbesitz

DATENPLATTFORMEN AN DER STAATSBIBLIOTHEK UND DER STIFTUNG PREUSSISCHER KULTURBESITZ STATUS QUO UND ZUKÜNFTIGE ENTWICKLUNGEN

**TEMPORAL
COMMUNITIES**

DOING LITERATURE IN A GLOBAL PERSPECTIVE

Dr.-Ing. David Zellhöfer, Informations- und Datenmanagement
Berlin, 25.11.2019

Daten der SBB – Verschiedene Zugänge



**Staatsbibliothek
zu Berlin**
Preußischer Kulturbesitz

Generell frei verfügbar in den folgenden Use Cases

Inspiration

Massen- bzw. automatisierte Downloads

Zitierbarkeit

Daten der SBB – Verschiedene Zugänge



Staatsbibliothek
zu Berlin

Preußischer Kulturbesitz

Generell frei verfügbar in den folgenden Use Cases

Inspiration: [SBB Lab als redaktioneller Hub](#)

Massen- bzw. automatisierte Downloads: [OAI-PMH \(z.B. Digitalisierte Sammlungen\)](#), [SRU \(z.B. Kalliope\)](#)

Zitierbarkeit: [Downloads von aufbereiteten Korpora](#)

Persönliche Anforderung + private Github-Accounts

Staatsbibliothek
zu Berlin
Preußischer Kulturbesitz

SBB-Lab

HOME

DATEN

DEMOS

KONTAKT

Projekte und Prototypen

Experimentelle Anwendungen, Projekte und Prototypen lernen Sie im Demobereich unseres SBB-Lab kennen. Da einige Inhalte allerdings nicht von uns erstellt wurden, leiten wir Sie teilweise auf externe Seiten – vielleicht ja schon bald auch auf Ihre!?

ALTPAPIER
NACHRICHTEN VON VOR-VORGES

Altpapier
Die App Altpapier bringt einige der spannendsten und merkwürdigsten Zeitungsmeldungen des 19. und 20. Jahrhunderts auf die Smartphones.

Berliner Schlagzeilen
Der Bot „Berliner Schlagzeilen“ twittert tagesweise die Schlagzeilen und ein Bild der Titelseite der jeweiligen Ausgabe von vor 100 Jahren.
[Weiterlesen >](#)

Interne Proof-of-Concepts

Deutsch

<https://lab.sbb.berlin/>

Apps aus Wettbewerben

Interne Proof-of-Concepts





Staatsbibliothek
zu Berlin
Preußischer Kulturbesitz

SBB-Lab

[HOME](#) [DATEN](#) [DEMOS](#) [KONTAKT](#) [🔍](#)

Daten

Downloads und Schnittstellen

Entdecken Sie von uns zusammengestellte Datenkollektionen und erfahren Sie, wie Sie sich über Schnittstellen Zugang zu den jeweiligen Inhalten verschaffen können.



Bibliographische Daten aus dem StaBiKat

Der Online-Katalog der SBB (StaBiKat) beinhaltet vor



DIGITALISIERTE
SAMMLUNGEN

Digitalisierte Sammlungen

Die Digitalisierten Sammlungen der SBB präsentieren

<https://lab.sbb.berlin/>

Strategisches Ziel ist die Bereitstellung von Metadaten und Content über Standard-Protokolle wie OAI-PMH, SRU und IIIF in Formaten wie METS/MODS, EAD und JSON

Implementierung über einzelne Dienste (z.B. Kalliope, Digitalisierte Sammlungen oder Handschriftenportal)

Nachnutzung von Verbundstrukturen durch zentralisierte Kataloge (CBS des GBV)

Perspektivisch: Data Hub der SBB und durchgängige Nutzung von Normdaten (mit Hinblick auf Linked Data)



Aufbereitete Korpora (als DOI) zu den Anwendungsfeldern:

OCR Fulltexts of the Digital Collections of the Berlin State Library

Extract from the Berlin State Library's Main Catalog

N-Grams verschiedener chinesischer Korpora

Metadata, Title Pages, and Network Graph of the Digitized Content of the
Berlin State Library (146,000 items)

Bildextrakte der Digitalisierten Sammlungen ...

<https://zenodo.org/communities/stabi>

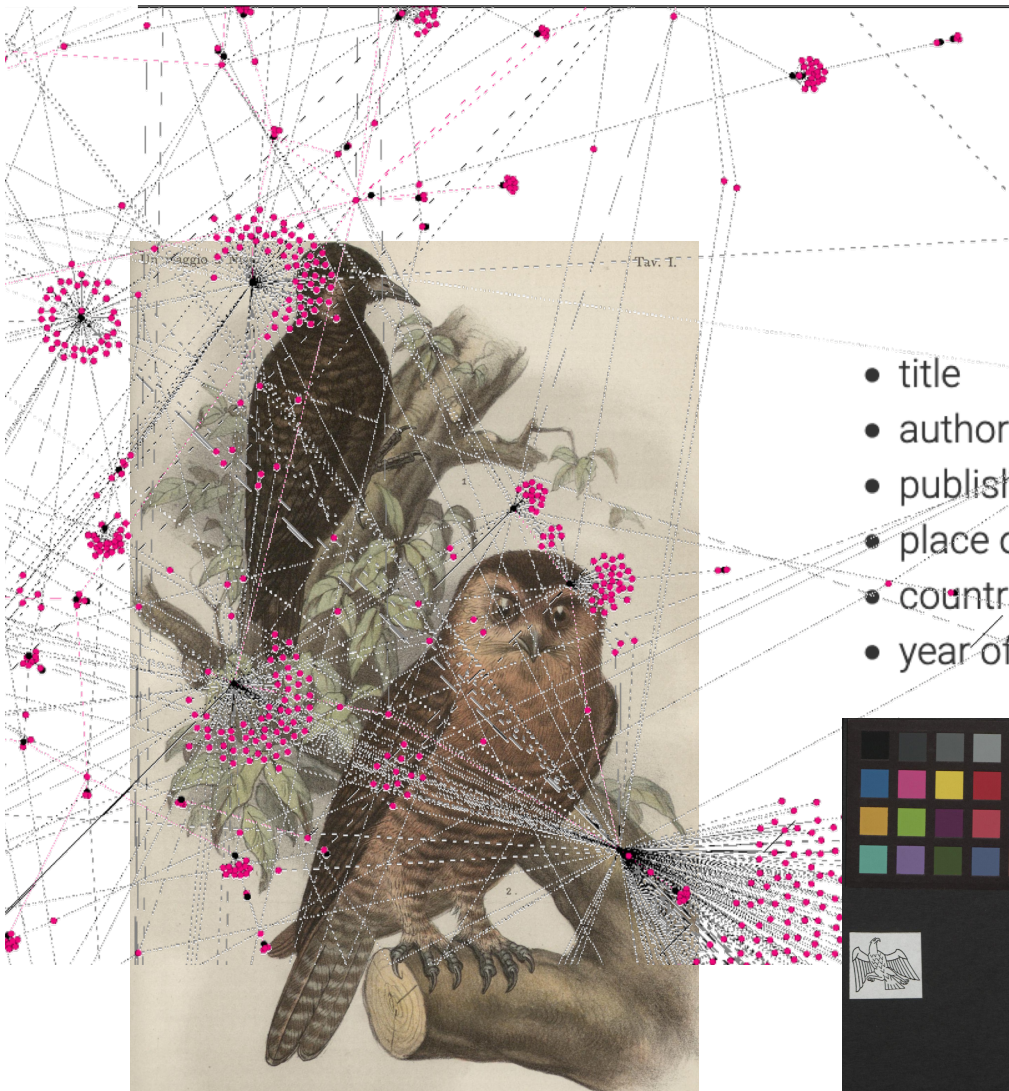


Korpora-Downloads



Staatsbibliothek
zu Berlin

Preußischer Kulturbesitz



- title
- author (+ optional GND ID)
- publisher
- place of publication
- country of publication
- year of publication



- 3. dut
- 1. cze
- 2. dan
- 3. dut
- 4. eng
- 5. fre
- 6. fry
- 7. ger
- 8. gre
- 9. ice
- 10. ita
- 11. lat
- 12. nor
- 13. pol
- 14. por
- 15. rum





QURATOR – Curation Technologies <https://qurator.ai/>

Maschinelles, multimodales Lernen & Normdaten/NER

SoNAR - Interfaces to Data for Historical Social Network

Analysis and Research <https://sonar.fh-potsdam.de/>

Netzwerkanalysen und Linked Data

OCR-D <http://www.ocr-d.de/>

Verbesserung OCR-Qualität historischer Quellen

Private Github-Accounts



Staatsbibliothek
zu Berlin

Preußischer Kulturbesitz

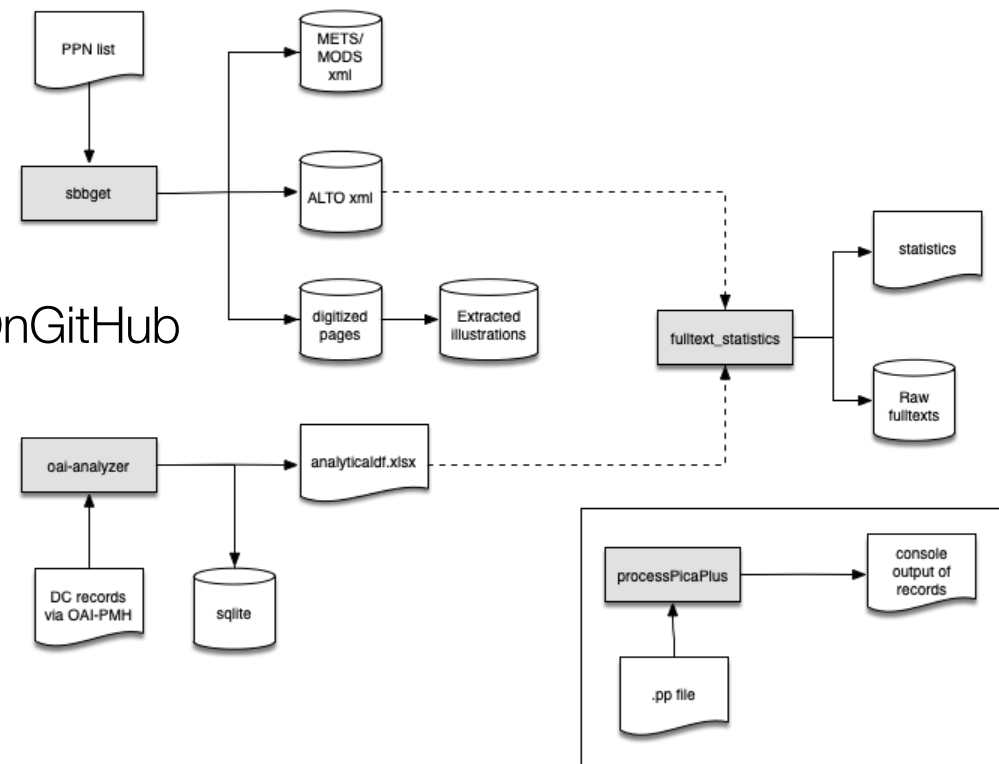
Zum Beispiel:

<https://github.com/hbunke/BibsOnGitHub>

<https://github.com/cneud>

<https://github.com/elektrobohemian>

i.d.R. Python/Perl-Skripte für
typische Harvesting- und
Pre-Processing-Tasks



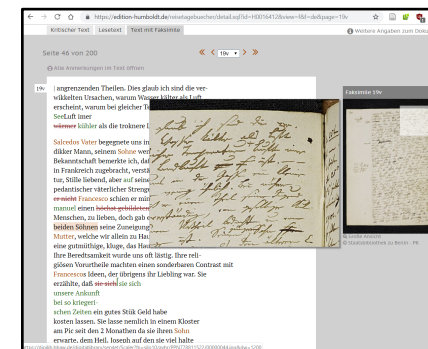
STIFTUNGSPERSPEKTIVE

Über einfache Online-Suche und -Anzeige hinausgehen und mit einer technischen Plattform Daten-Nachnutzung und -Verknüpfung ermöglichen.

Daten anreichern und über Sammlungen/Einrichtungen hinweg verknüpfen.



Welche Textpassagen
behandeln die Herkunftsregion
der Mais-Göttin Chicomecoatl?



Förderung öffentlichkeitswirksamer Datennachnutzung (u.a. Interaktion mit neuen Zielgruppen)

Leistung konzeptioneller Vorarbeiten für die Datenplattform

Inventur und Dokumentation von Daten und Schnittstellen

Rechteklärung und Lizenzierung der Daten bzw.
Unterstützung bei der Klärung

Organisation interner und externer Veranstaltungen (inkl. Wettbewerbe)

Ansprache externer Zielgruppen und Etablierung von Partnerschaften

z.B. „Datenbesuche“ – Workshops in Institutionen mit vorhandenen Daten

Interne Information über die Aktivitäten, Kolleg*innen gewinnen und einbeziehen.

TEMPORAL COMMUNITIES

DOING LITERATURE IN A GLOBAL PERSPECTIVE



Staatsbibliothek
zu Berlin
Preußischer Kulturbesitz

