

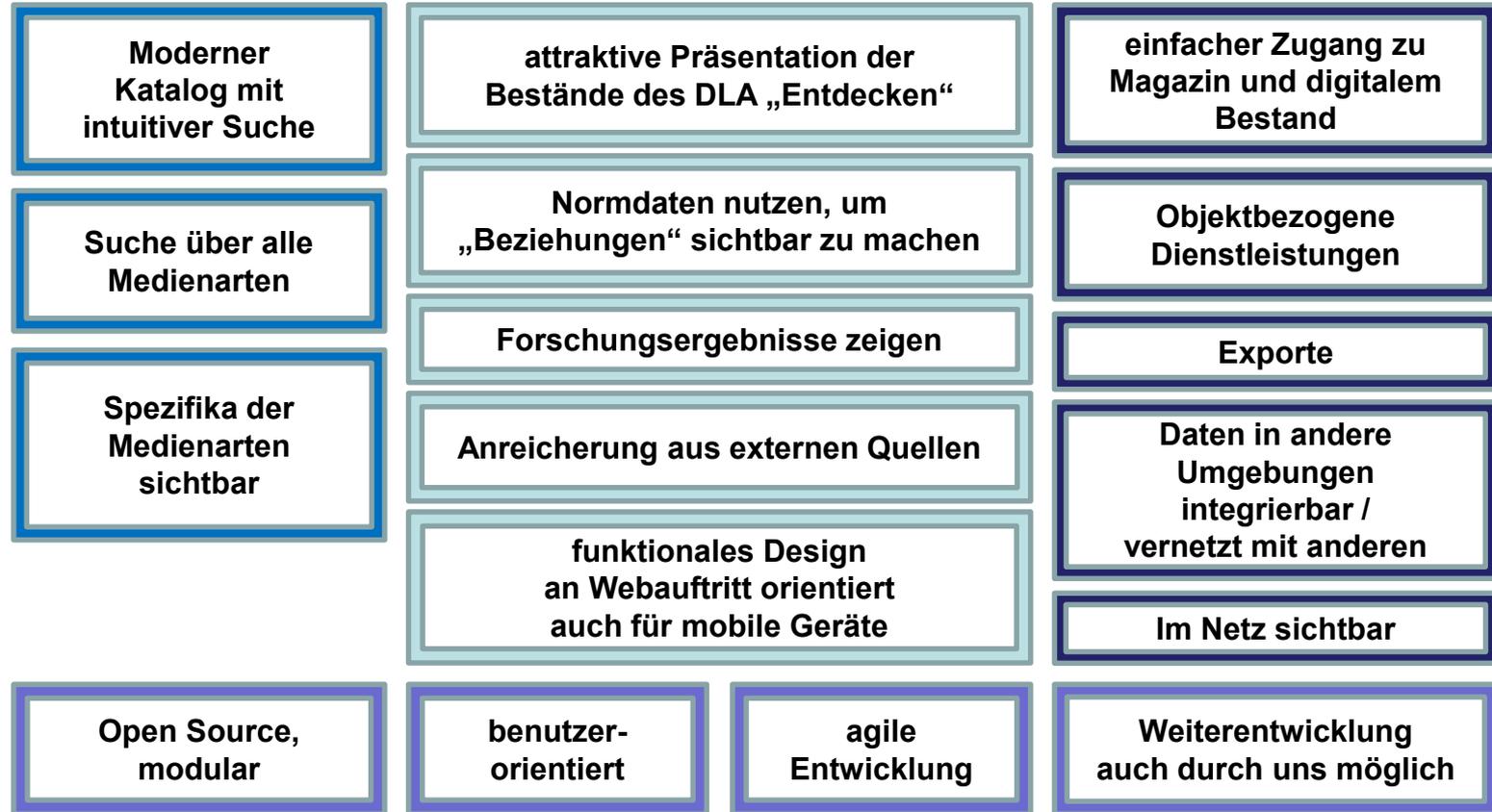
Vom Prototyp zum neuen Onlinekatalog – gute Daten als Schlüssel zu den Quellen der Literaturgeschichte

Felix Lohmeier, Open Culture Consulting

Karin Schmidgall, DLA Marbach

7. Bibliothekskongress, Leipzig 18.3.2019

Erwartungen an einen neuen Onlinekatalog



Was finden Sie?

 Gesamtbestand  Gedrucktes  Handschriften  Bilder & Objekte  Audio & Video  Namen & Werke

Wonach suchen Sie?

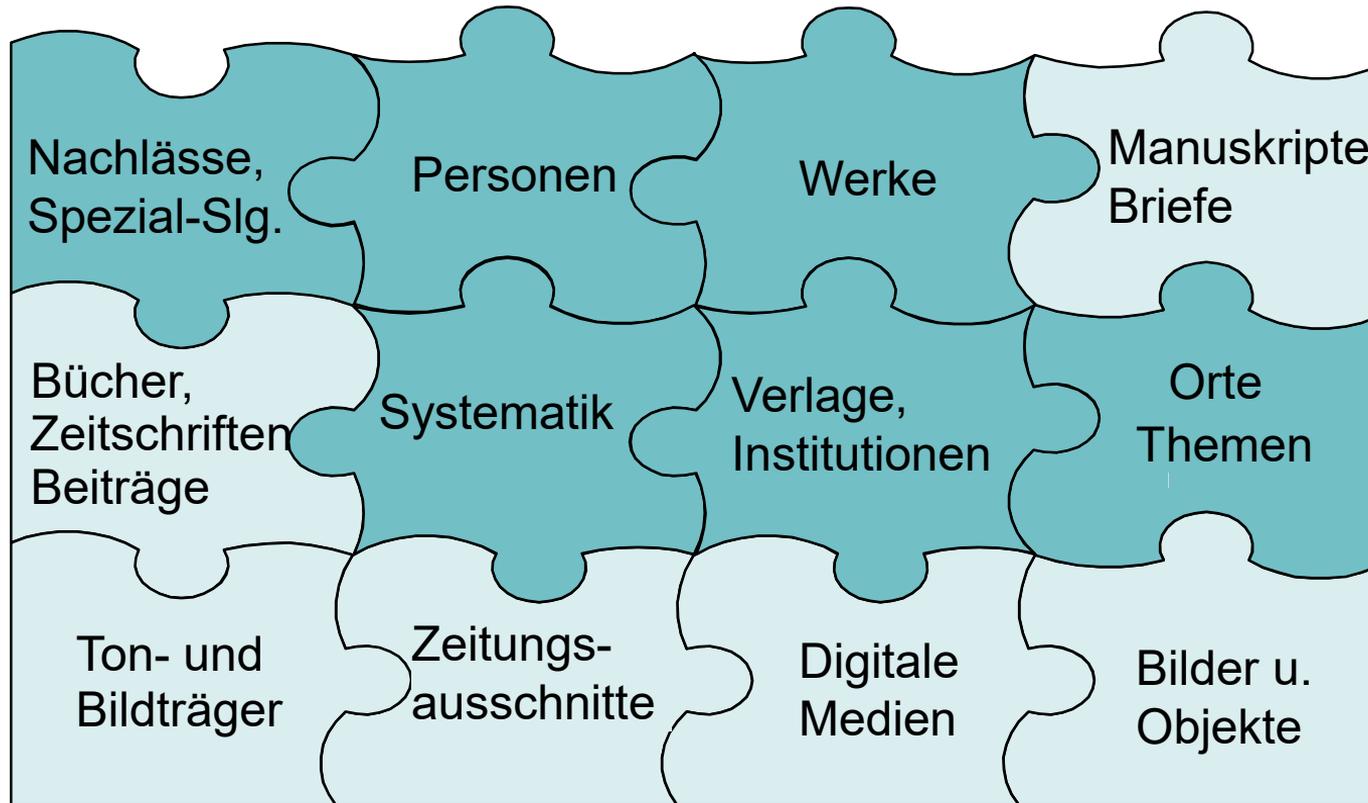


Jetzt suchen

- Literatur vom 19. Jahrhundert bis zur Gegenwart einschließlich Forschungsliteratur in möglicher Vollständigkeit
- literarische Nachlässe, Autographen, Manuskripte, Korrespondenzen, Druckwerke, AV-Materialien, Bildnisse, Erinnerungsstücke
- deutschsprachiger Autoren, Kritiker, Übersetzer, Philosophen, Philologen und anderen Gelehrten
- Dokumente Ihrer Wirkungsgeschichte von 1750 bis zur Gegenwart
- Verlags- und Redaktionsarchive

Datengrundlage des Onlinekatalogs:
Kallias – ein Erfassungs-System für alle Bestände

deutsches
literatur
archiv **marbach**



 **Gesamtbestand**  Gedrucktes  Handschriften  Bilder & Objekte  Audio & Video  Namen & Werke

Wonach suchen Sie?

 **Jetzt suchen**

 **Expertensuche**

 **Bestandsübersicht**

 **Suchverlauf**

 **Hilfe & Auskunft**



Bestandsübersicht

Nachlässe und Sammlungen



Digitalisate

Friedrich Schiller Manuskripte



Audio-Digitalisate

Lesungen im DLA

Unerwartetes im Katalog

> [Totenmasken](#)

Geburtstage in diesem Monat

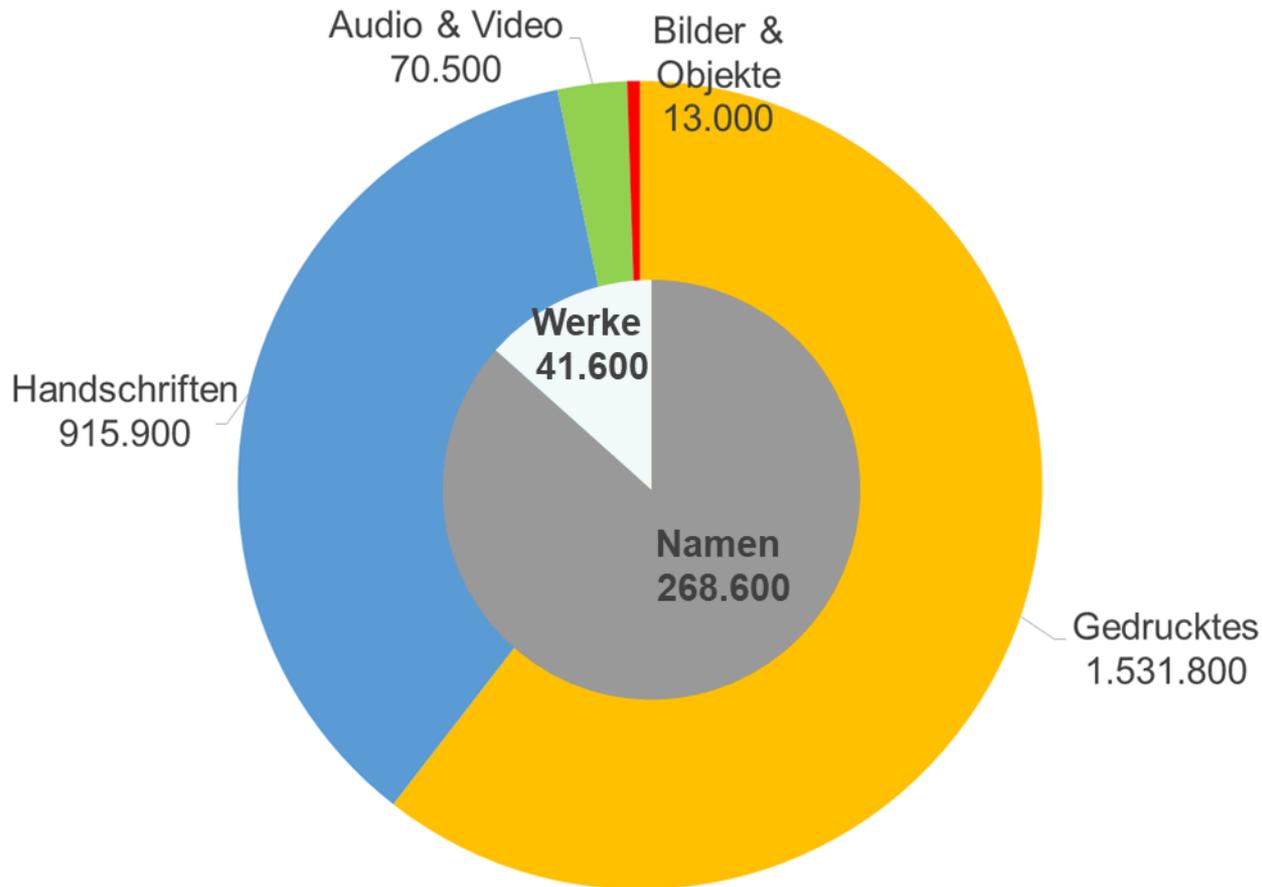
> [Yvan Goll](#)

Neu im Katalog

> [Bibliothek Ernst Jünger](#)

Katalog in Zahlen

deutsches
literatur
archiv **marbach**



Autoren und ihre Werke, Nachlässe und Autorenbibliotheken,
Provenienzen – im Zentrum des Sammelns und der Katalogsuche

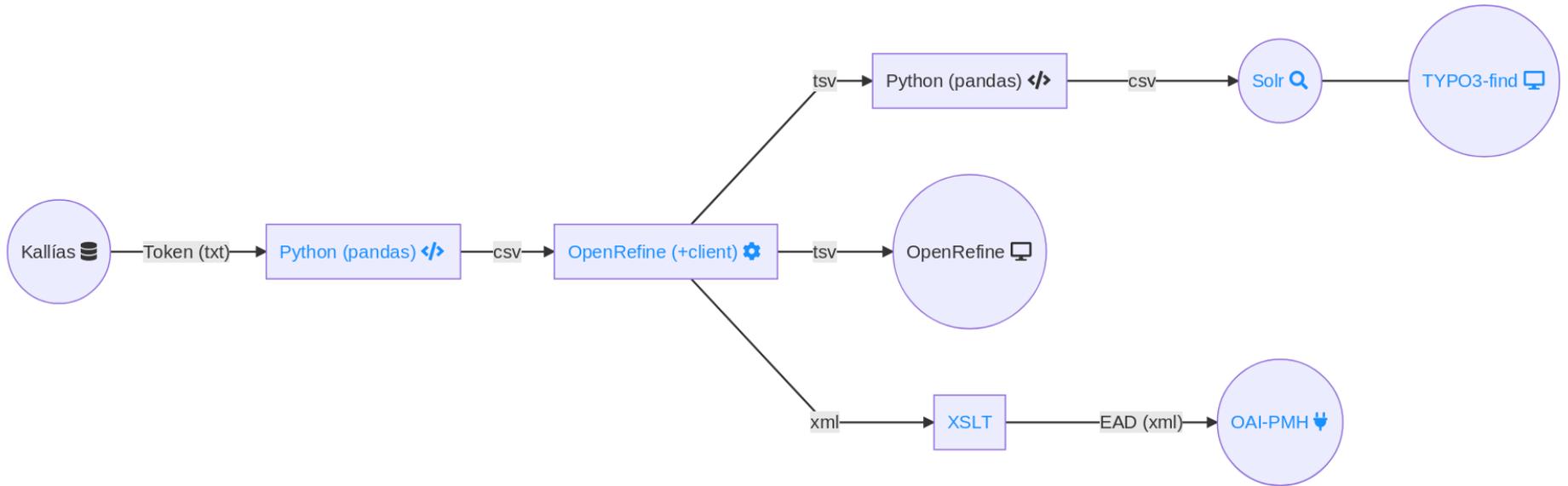
1. Autovervollständigung im Suchschlitz auf Basis von Normdaten (Namen und Werke)
2. Normdaten in der Trefferliste
3. Normdaten als „Wissensraum“: eine Detailanzeige die Beziehungen sichtbar macht und zu den Beständen führt
4. Integration und Vernetzung mit Normdaten-Informationen aus anderen Quellen (z.B. Wikidata, Deutsche Biographie, ...)
5. Entdecken im Katalog – Normdaten zur Kontextualisierung – Metadaten für die Beschreibung eines Nachlasses – funktional auch ein „Normdatensatz“

Einblick in die Werkstatt der Datenprozessierung

- ✓ Einsatz von Open-Source-Komponenten
- ✓ Modulare Systemarchitektur
- ✓ Weiterentwicklung „auch durch uns möglich“
- ✓ Daten in andere Umgebungen integrierbar

Systemarchitektur

deutsches
literatur
archiv **marbach**



Erläuterungen auf der Webseite <https://felixlohmeier.de/dla/systemarchitektur.html>

```
felix@tux:~  
ext_lohm@neo:~/opac-labor/logs> tail -n 19 20190308_222811.log  
  
=== Statistik ===  
  
Beginn und Laufzeiten der einzelnen Prozessierungsschritte:  
  Beginn Gesamtprozessierung Fr 8. Mär 22:28:11 CET 2019 (00:00:00)  
  Arbeitsverzeichnis leeren Fr 8. Mär 22:28:11 CET 2019 (00:01:54)  
  Vorprozessierung aDIS Token Fr 8. Mär 22:30:05 CET 2019 (00:13:34)  
  Transformation aDIS Token Fr 8. Mär 22:43:39 CET 2019 (01:05:50)  
  Kopplungen und Bestandspezifisches Fr 8. Mär 23:49:29 CET 2019 (01:00:27)  
Zusammenfassung und Ergebnisprüfung Sa 9. Mär 00:49:56 CET 2019 (00:00:30)  
  Neuindexierung Solr 1 Sa 9. Mär 00:50:26 CET 2019 (01:31:38)  
  Neuindexierung Solr 2 Sa 9. Mär 02:22:04 CET 2019 (01:09:50)  
  csv2ead 1/2 OpenRefine Sa 9. Mär 03:31:54 CET 2019 (00:31:41)  
  csv2ead 2/2 XSLT Sa 9. Mär 04:03:35 CET 2019 (00:04:01)  
  OpenRefine Demo aktualisieren Sa 9. Mär 04:07:36 CET 2019 (00:11:45)  
  Ende Gesamtprozessierung Sa 9. Mär 04:19:21 CET 2019 (00:00:00)  
  
Gesamtlaufzeit: 05:51:10 (hh:mm:ss)  
  
ext_lohm@neo:~/opac-labor/logs> echo "hello #bibtag19" █
```

Quelldaten:
4,3 Mio. Datensätze
mit bis zu 115 Feldern
(7,1 GB txt)

Suchindex:
4,2 Mio Datensätze
mit 531 Feldern
(5,1 GB Solr)

Server:
2 vCPUs
64 GB RAM

vollautomatisiert mit
Bash-Scripts

Facet / Filter [Undo / Redo](#) 0

491963 rows

Extensions: [VIB-Bits](#) ▾

Refresh [Reset All](#) [Remove All](#)

Show as: [rows](#) [records](#) Show: [5](#) [10](#) [25](#) [50](#) rows [Sort](#) ▾

« first < previous **1 - 5** next > last »

PI814 [change](#)

8258 choices Sort by: [name](#) [count](#)

- [Schriftsteller](#) 18057
- [Schriftstellerin](#) 5139
- [Autor](#) 4834
- [Journalist](#) 4675
- [Historiker](#) 3665
- [Übersetzer](#) 3411
- [Maler](#) 3198
- [Germanist](#) 3168
- [Schauspieler](#) 3023
- [Autorin](#) 2867
- [Komponist](#) 2848

PD814 [change](#)

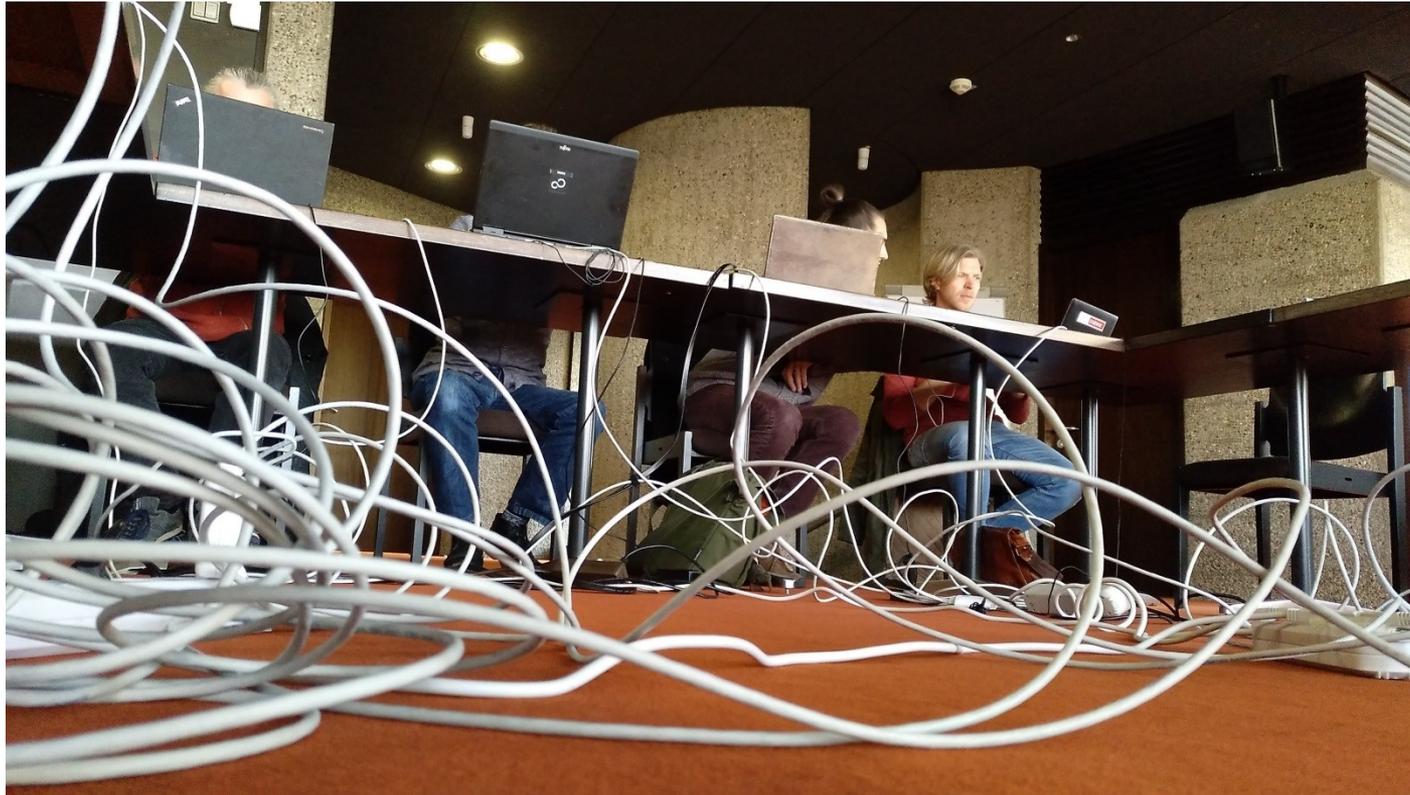
7118 choices Sort by: [name](#) [count](#)

- [Berlin](#) 2381
- [München](#) 1468
- [Wien](#) 1379
- [Paris](#) 858
- [Stuttgart](#) 722
- [Hamburg](#) 642
- [Zürich](#) 451
- [London](#) 424
- [Leipzig](#) 419
- [Dresden](#) 355
- [Heidelberg](#) 268

▼ All	▼ id	▼ entity_score	▼ autocomplete_pay	▼ autocomplete_term	▼ facet_time	▼ listview_additions	▼ listview_title	▼ helper_alive	▼ XX
☆ 🗨 477427.	PE00000863	29401	PE00000863	Goethe, Johann Wolfgang von (1749-1832)	Goethe, Johann Wolfgang von (1749-1832) \ Goethe, Johann Wolfgang \ Gete, Iogann W. \ Goethe, Johann W. \ Gête, Iogann V. \ Goethe, Wolfgang von \ Goethius, ... \ Goethe, Volfgango \ Goethe, Johan Volphgang \ Gete, Johan Volfgang \ Goethe, Johann Volfgang \ Göthe, J. W. \ Goethe, Johann W. von \ Gete, Iogann V.	[1749-08-28 TO 1832-03-22]	Schriftsteller, , Politiker, Naturwissenschaftler, ,	Goethe, Johann Wolfgang von (1749-1832)	false
☆ 🗨 218406.	PE00039252	28874	PE00039252	Unsold, Siegfried (1924-2002)	Unsold, Siegfried (1924-2002) \ Ulmer, Peter \ Unseld, Karl Siegfried \ Unseld, Siegfried \ Unseldt, Siegfried \ Unsöld, Siegfried	[1924-09-28 TO 2002-10-26]	Verleger, , , , , , ,	Unsold, Siegfried (1924-2002)	false
☆ 🗨 24639.	PE00001233	28701	PE00001233	Jünger, Ernst (1895-1998)	Jünger, Ernst (1895-1998) \ Junior, Richard \ Junger, Ernst	[1895-03-29 TO 1998-02-17]	Schriftsteller, , , , Entomologe, Offizier, , , ,	Jünger, Ernst (1895-1998)	false

Hackathon Februar 2019

deutsches
literatur
archiv **marbach**



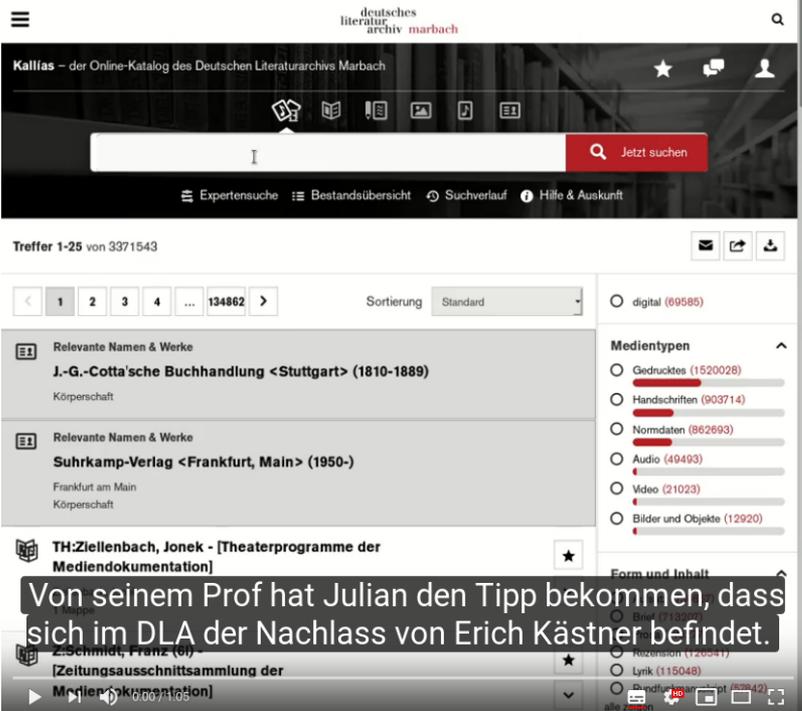
Konzept	Umsetzung
Normdatenbasierter Suchschlitz	<ol style="list-style-type: none">1. Felder mit OpenRefine generiert: „autocomplete_term“ (Texte und IDs), „autocomplete_payload“ (Pseudonyme) und „entity_ids“ (Normdaten-IDs);2. Solr Suggester auf Basis u.a. von „autocomplete_term“ mit Gewichtung;3. Abfrage des Solr Suggester und Aufbereitung mit TYPO3-eID-Script;4. Darstellung im Suchschlitz mit jQuery Tokeninput
Normdaten in der Trefferliste	Darstellung der zwei TOP-Normdaten-Treffer für die jeweilige Suche mit angepasstem TYPO3-find Template (u.a. Facette auf „entity_ids“); Für die Sortierung der Trefferliste im Sucheinstieg Namen & Werke wird mit OpenRefine „entity_score“ (Anzahl der Links auf Normdaten) generiert
Normdaten als „Wissensraum“	Auf den Detailseiten werden mehrere komplexe Suchen mit TYPO3-find konfiguriert, welche auch die ersten fünf Ergebnisse direkt anzeigen
Integration und Vernetzung von Normdaten	Auf Basis der vorhandenen GND-IDs werden mit OpenRefine Daten aus Wikidata und Entity Facts angereichert; interessante externe Daten sind beispielsweise das Bild der Unterschrift und Links auf externe Datenbanken
Normdaten zur Kontextualisierung	Normdatensätze sind eigene Objekte im Suchindex

Live-Vorführung

- Wir versuchen nah an den Wünschen unserer Benutzer zu entwickeln und arbeiten mit Methoden aus der Usability-Forschung.
- Wir führen Thinking Aloud Tests durch, bei denen Testpersonen mit dem prototypischen Katalog Aufgaben lösen und aufgefordert werden laut mitzudenken.
- Keine Sorge wir werden jetzt nicht einen Probanden aus dem Publikum auf die Bühne bitten. Denn unser Projekt begleiten 6 Personas.
- Das sind fiktive, archetypische Nutzer die Ziele und Bedürfnisse ihrer Zielgruppe vertreten. Diese dürfen verschiedene Szenarien durchspielen.

Persona Julian Kulak

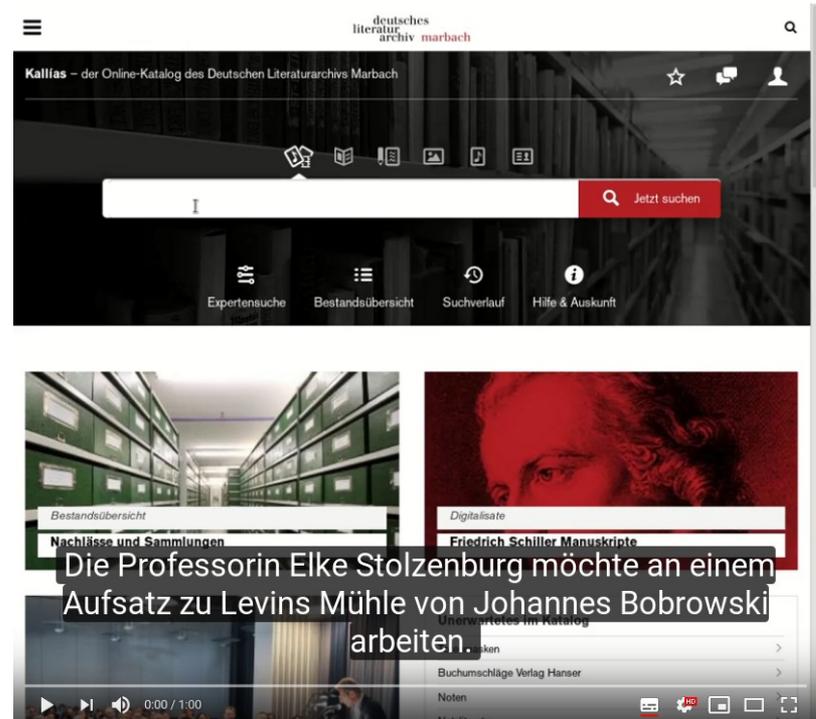
- Sie dürfen hier den Studenten Julian Kulak bei seiner Arbeit kennenlernen.
- Er ist 22 Jahre alt und studiert im 4. Semester Germanistik an der Uni Stuttgart.
- Er ist Single und wohnt in einer 3er-WG.
- Als Digital Native hat er das Smartphone immer griffbereit und ist ständig erreichbar.



The screenshot shows the homepage of the German Literature Archive Marbach (DLA). The header includes the logo and navigation icons. Below the search bar, there are filters for 'Treffer 1-25 von 3371543' and 'Sortierung Standard'. The main content area displays search results for 'Relevante Namen & Werke', including 'J.-G.-Cotta'sche Buchhandlung <Stuttgart> (1810-1889)' and 'Suhrkamp-Verlag <Frankfurt, Main> (1950-)'. A text overlay at the bottom of the screenshot reads: 'Von seinem Prof hat Julian den Tipp bekommen, dass sich im DLA der Nachlass von Erich Kästner befindet.'

Persona Elke Stolzenburg

- Sie dürfen hier die Professorin Elke Stolzenburg bei ihrer Arbeit kennenlernen.
- Sie wurde 1953 in Stuttgart geboren. Sie kommt aus dem Bildungsbürgertum und hat eine humanistische Schulbildung genossen.
- Sie promovierte zum Werk von Christoph Meckel. An der Universität Frankfurt lehrte sie viele Jahre und ist inzwischen emeritiert, aber publiziert und forscht weiterhin.
- Sie war schon oft in Marbach und schätzt die Arbeitsatmosphäre und die Dichte der Bestände.



Während des Bibliothekskongresses steht die Testumgebung frei im Netz. Wir freuen uns über Feedback!



<https://www-test.dla-marbach.de>

(verfügbar bis 22.3.19)

Fragen zum Projekt beantworten gerne:

Felix.Lohmeier@opencultureconsulting.com

Karin.Schmidgall@dla-marbach.de

Beteiligte Dienstleister:

[Open Culture Consulting \(OpenRefine\)](#)

[effective WEBWORK GmbH \(TYPO3\)](#)

[Markus Mandalka \(Solr\)](#)

[Kai Mertens \(Design\)](#)