

1. Projektplanung

2. Projektablauf

3. Projektergebnis und Lessons learned

Projektplanung

Projektidee auf Direktionsebene bereits seit 2012

Ende 2012: Email an Teilnehmerbibliotheken mit Bitte um
× Titelauswahl + Mengengerüst

2013: OCR-Test mit ULB Innsbruck

Antrag (Herbst 2013) bei HMWK, Projektleitung:
× HEBIS-Verbundzentrale

Teilnehmerbibliotheken
und Titelanzahl

ULB Darmstadt	18 (22)
UB Frankfurt	19
HLB Fulda	11
UB Gießen	9
UB Kassel	0 (2)
UB Marburg	5
HLB Wiesbaden	14



Quelle: Wikimedia Commons


Projektmeetings

- × 6.12.2013 Frankfurt: Kickoff, Aufgabenverteilung
- × 22.1.2014 Frankfurt: Absprachen Ablauf, Basis, grundl. Workflow
- × 11.2.2014 Frankfurt: Absprachen Ablauf, Feintuning

Projektsteuerungsgruppe

Projektkommunikation

- × Mail-Verteiler
- × Dokumentations-Wiki (Confluence)

 Digitalisierung regionaler Zeitungen / Digitalisierung regionaler Zeitungen Startseite
Doku und Projektüberblick
© Erstellt und zuletzt geändert von Roth-Steiner, Roland am Jun 12, 2014

Projektüberblick

Koordination

- Antragsstellung, Projektleitung: Risch (Hebis)
- Steuerungsgruppe: Buchkamp (Giessen), Knepper (Fra), Roth-Steiner (Da)

Dokumentation

- Hebis-Intranet-Plattform
- Mail-Liste HWK1@hebis.de für das Projekt "Digitalisierung regionaler Zeitungen aus der Zeit des 1. Weltkriegs" eingerichtet.
Administration über: <http://dlist.server.uni-frankfurt.de/mailman/listinfo/hwk1>
- Projektwiki bei <https://kid.hebis.de>

Teilnehmerbibliotheken

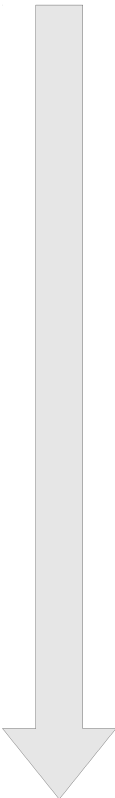
Bibliothek	Kürzel für Masterlist	Bestand in Masterliste eingefügt?	Ansprechpartner	Lokale Plattform	Anmerkungen
ULB Darmstadt	17		Roth-Steiner	dwork	Scanning im eigenen Haus
UB Frankfurt	30		Marko Knepper	VL	
HLB Fulda	66	Ja	Hendrik Wesner	Goobi	Scanning durch externen Dienstleister
UB Giessen	26		Thomas Buchkamp	dwork	Scanning im eigenen Haus
UB Kassel	34		Axel Halle	Goobi	
HLB Wiesbaden	43		Cindy Kathke	VL - Instanz	Scanning durch externen Dienstleister
UB Marburg	4		Hubertus Neuhausen	OPUS 3 + PDFs	

Meetings

- Kickoff 6.12.2013, Frankfurt
- 22.1.2014, Frankfurt
- Di., 11.2., 14 Uhr, UB Frankfurt

Vorgehensweise für alle Teilnehmerbibliotheken

Überblick geplanter Workflow

- 
- × Bestand melden (Masterlist)
 - × Bestand sichten
 - × Hard- und Software (Innsbruck) einrichten
 - × Scanning
 - × Binarisierung
 - × Erschliessung (Ausgaben = Issues)
 - × Qualitätssicherung (QS)
 - × Versand nach Innsbruck auf externer Festplatte
 - × OCR + METS-Erzeugung zentral in Innsbruck (ULB)
 - × Rückversand OCR-Ergebnisse
 - × Einspielung in HEBIS-Portal
 - × Einspielung in lokale Digitalisierungsplattformen

Überblick tatsächlicher Workflow

- x ~~Bestand melden (Masterlist)~~
- x Bestand sichten, Hard- und Software einrichten
- x Scanning
- x ~~Binarisierung~~
- x Erschliessung mit Farb-Images (Größenunterschied Faktor 10)
- x (QS)
- x Versand nach ~~Innsbruck~~ Frankfurt inkl. Masterlist
- x ~~OCR + METS~~-Erzeugung zentral in Innsbruck (ULB)
- x Rückversand OCR-Ergebnisse
- x Katalogisierung in ZDB
- x Einspielung in HEBIS-Portal
- x ~~Einspielung in lokale Digitalisierungsplattformen~~ manuell in Eigenregie

OCR (Texterkennung + daraus resultierende Volltextsuche)
konnten (zunächst) nur für frühzeitig eingelieferte Scans realisiert werden

Projektergebnis Portal

<http://morus.hebis.uni-frankfurt.de/>

Testplattform, das offizielle Portal wird anders heißen (zur Buchmesse).
Realisiert mit Visual Library (VL).

Einige Features:

Navigationsmöglichkeiten
Bibliotheksprofile
Suche, Treffer mit Highlighting

Beispiele:

Detailsuche Volltext „burghausen“
→ fuzzy-Suche + Relevanzranking




Suche nach „franz ferdinand“ (als Phrase), dann Filter nach Bestand aus HLB Fulda → 9 Titel → Bonfatiusbote / Beilage, 1914
→ OCR-Erkennungsfehler (insgesamt aber sehr zufrieden stellendes OCR-Ergebnis für Fraktur)




The screenshot shows the HeBis website interface. At the top, there is a search bar and navigation links. The main content area displays a search result for the newspaper 'Centralanzeiger für den Odenwald'. The newspaper's masthead is prominently displayed in a decorative, gothic-style font. Below the masthead, there is a small illustration of a person sitting under a tree. The text on the page includes the newspaper's name, its founding year (1824), and its publication details. The main headline of the newspaper is 'Der Krieg zwischen Oesterreich und Serbien.' The page is dated 'Erbach, Donnerstag den 30. Juli 1914'.

Projektergebnis in Zahlen

In Zahlen (Scans = Seiten):

Standort	Antrag	Innsbruck/OCR	22.7.2014 gesamt
Darmstadt	20.010	0	21.231
Frankfurt	11.031	12.210	12.210
Fulda	15.047	6.399	6.399 
Gießen	3.564	6.835	6.835
Kassel	326	0	0
Marburg	4.845	4.137	4.137 
Wiesbaden	47.700	0	23.261 
	102.523	29.581	74.073

Darmstadt: als Farbbilder knapp 1 TB

 Scanning durch externen Dienstleister

Projektergebnis lokal

<http://tudigit.ulb.tu-darmstadt.de/sml/Regionalzeitungen>

Digitale Sammlungen **Universitäts- und
Darmstadt Landesbibliothek**

**Der Zeitspiegel: illustrierte
Unterhaltungsbeilage [der
Lampertheimer Zeitung]**
1914.1914 19140620_24_00001
Bibliographische Information
Startseite des Bandes (PDF, 166 MB)
Alle zugehörigen Bände
Digitale Sammlungen Darmstadt -
Regionalzeitungen (Universitäts- und
Landesbibliothek Darmstadt)

Sprung zur Seite (z. B.: IV,
145, xii)

- Inhaltsverz.**
- [Ausgabe 21]
 - 30.05.1914 (Samstag),
[Ausgabe 22]
 - 06.06.1914 (Samstag),
[Ausgabe 23]
 - **20.06.1914 (Samstag),**
[Ausgabe 24]
 - 27.06.1914 (Samstag),
[Ausgabe 25]
 - 04.07.1914 (Samstag),
[Ausgabe 26]
 - 11.07.1914 (Samstag),
[Ausgabe 27]
 - 18.07.1914 (Samstag),
[Ausgabe 28]
 - 25.07.1914 (Samstag),
[Ausgabe 29]
 - 01.08.1914 (Samstag),
[Ausgabe 30]
 - 08.08.1914 (Samstag),
[Ausgabe 31]
 - 15.08.1914 (Samstag),
[Ausgabe 32]
 - 22.08.1914 (Samstag),
[Ausgabe 33]
 - 12.09.1914 (Samstag),
[Ausgabe 34]
 - 19.09.1914 (Samstag),
[Ausgabe 35]
 - 26.09.1914 (Samstag),
[Ausgabe 36]
 - 02.10.1914 (Samstag),
[Ausgabe 37]
 - 10.10.1914 (Samstag),
[Ausgabe 38]
 - 17.10.1914 (Samstag),
[Ausgabe 39]
 - 24.10.1914 (Samstag),
[Ausgabe 40]

Persistente URL: <http://tudigit.ulb.tu-darmstadt.de/show/Za-160a-1914/0183>

Seitenansicht Vorschau



Der Zeitspiegel
Nr. 24. Illustrierte Unterhaltungs-Beilage. 1914.

Margarete.
Roman nach englischen Original bearbeitet
von A. Geisel.

Dreundzwanzigstes Kapitel.

Im dritten November gegen elf Uhr vormittags trat Noel Warren in das Frühstückszimmer der Villa Balliol bei Dumfries in Schottland, und einen unzufriedenen Blick auf den gedienten Tisch werfend, kintete er und befahl den eintretenden Dienstmädchen, frischen Tee zu bringen.

Als das Mädchen bald darauf wieder erschien und außer der Teeanne eine Platte mit heißen Knetlecken nebst weichen Omeu auf den Tisch gestellt hatte, fragte der Hausherr mitrühlich: „H Frau Warren abgereist?“

„Jawohl, Herr Warren.“

„H Luise dahine!“

„Hat Ihre Herrin keinen Auftrag für mich hinter lassen?“

„Nein, Herr Warren.“

„Hm — sagte sie, wann sie wiederkommen werde?“

„Ja, Herr Warren, — in spätestens acht Tagen.“

„War Frau Warren better, als sie abreiste?“

„Nein, Herr Warren — Frau Warren sah sorglos voll aus.“

„Sagte Frau Warren, sie wolle nach London reisen?“

„Ja, Herr Warren.“

„Und was badien Sie sich dabei, daß Frau Warren ohne Begleitung reise, he?“

„Ich nehme mir's nicht heraus, etwas darüber zu denken, Herr Warren. Frau Warren ist meine Gesteierin, und was sie tut, erscheint mir recht.“

„So — na, das klingt fast, als ob's wahr wäre — wie ist das Wetter heute — kalt oder mild?“

„Wilt, Herr Warren, im Garten ist's jaung und warm.“

„So will ich hinausgehen, bringen sie mir Nebelzucker und Gut.“

„Reib darauf schritt Noel Warren in den hübschen Hausgarten auf und ab, trotz des Sonnenheims erheiterte sein Gesichtsausdruck sich nicht, und seine Stirne auf die niedere Blanke stützend, welche den Garten von der daranstoßenden Wiege schied, leuzte er laut auf:

„Stillsam in Grundierung dieses Sturzfers laute jetzt eine Stimme dicht hinter Noel Warren: „Zu

Features u.a.:

Binarisiert

Eingebaut in Sammlung
„Regionalzeitungen“

Erschliessung übernommen

Langzeitarchivierung

Keine OCR

Erstes gemeinsames Digitalisierungsprojekt der Hessischen Universitäts- und Regionalbibliotheken

Zu knappe Planung (zeitlich + Ressourcen)

- × keine fachliche Begleitung (Auswahl Zeitungstitel)
- × Ressourcen für Scanning + Erschliessung unscharf geplant
- × Datenmanagement im Terabyte-Bereich

Erstens kommt es anders...

- × Umsteuerungen im Projektverlauf (z.B. Farbe statt SW, TIF-LZW → TIF-baseline)
- × Pannen und Risiken – jede Bibliothek kämpft mit ihren eigenen Dämonen (ULB Darmstadt: dpi-Zahl in TIF-Header, Personalausfall, konservatorische Vorgaben, ...)

Aber trotzdem ...

- × z.Zt. wird das Mengengerüst für 1915 – 18 eingeholt

Vielen Dank für die Aufmerksamkeit...



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Zeit für Fragen, Diskussion ...

... weiterhin viel Spass bei der Tagung.

Sie finden die Präsentation auch bei **tuprints**,
dem ePublishing-Service für die TU Darmstadt:
<http://tuprints.ulb.tu-darmstadt.de/4147/>