

zeit.punktNRW – ein Werkstattbericht zur Zeitungsdigitalisierung in NRW

75. Westfälischer Archivtag
Dülmen
20.03.2024

Dr. Andrea Ammendola
(ULB Münster)

Ministerium für
Kultur und Wissenschaft
des Landes Nordrhein-Westfalen



Warum historische Zeitungen digitalisieren?

- wichtige Quellen zur politischen Geschichte, zur Kultur-, Lokal- und Regionalgeschichte und Ahnenforschung
- umfangreiches Quellenkorpus, Originalbestand teils gefährdet
- verstreute Überlieferungssituation
- neue Nutzungsmöglichkeiten und -potenziale durch Digitalisate
- Erschließung noch nicht nachgewiesener Zeitungen in der Zeitschriftendatenbank (ZDB)
- Landesdigitalisierungsprogramme eine große Chance insb. für Archive (92 der 108 Kooperationspartner sind Archive!)

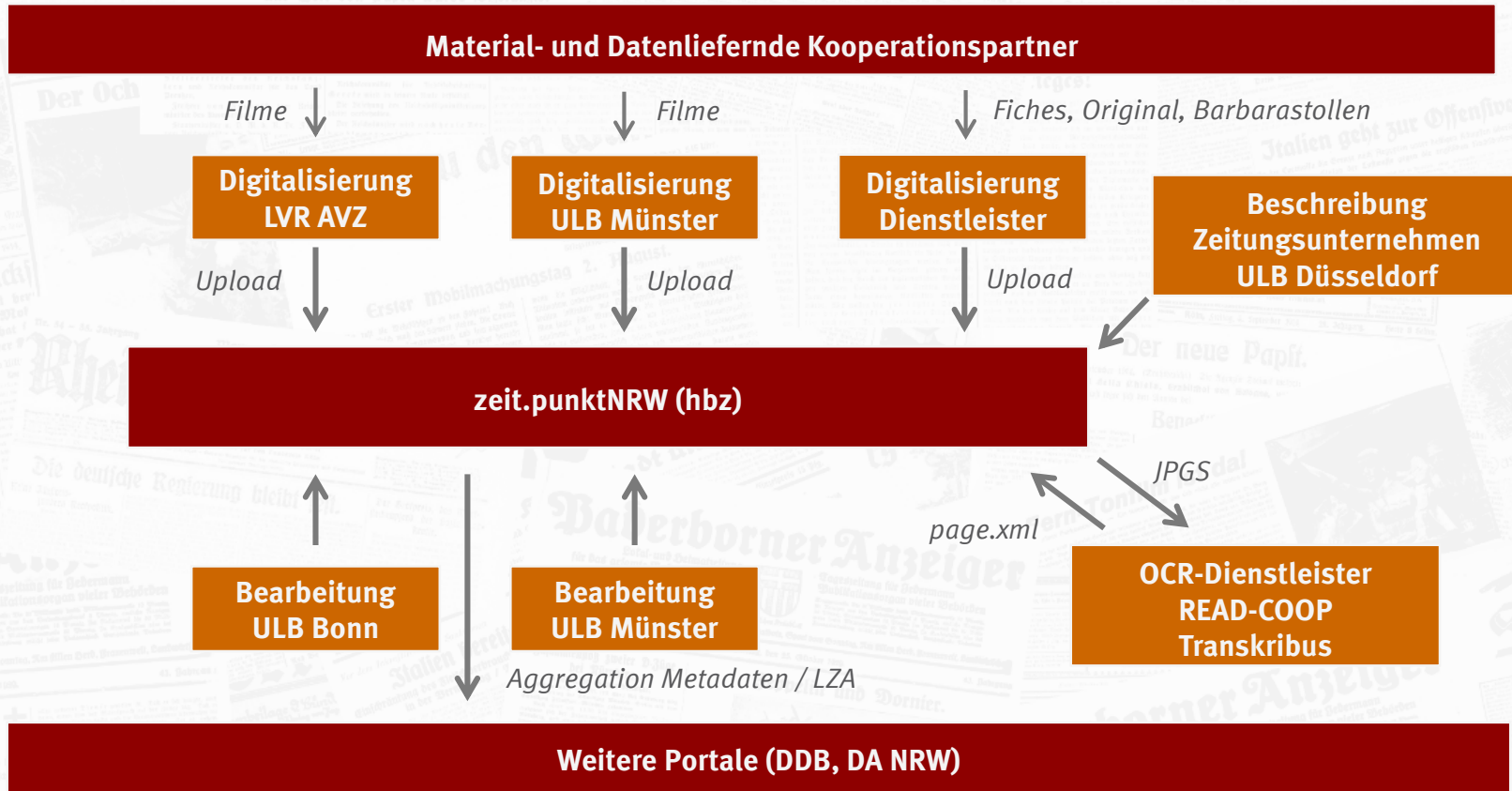
Genese und Projektstart

- Mittelgeber: Ministerium für Kultur und Wissenschaft NRW (MKW)
- 2015 – 2016: Vorbereitungs- und Antragsphase beim MKW
- Idee für ein rheinisches Zeitungsportal der ULB Bonn » ganz NRW!
- Fragebogenaktion, Antragskonzept, technische und personelle Ausstattung planen (insb. mit ULB Bonn und dem Lenkungskreis)
- Nach der Bewilligung: Verträge mit Partnern schließen, Workflow festlegen, Hardware und Software bereitstellen, Transport der Filme
- 2017: Projektstart und Beginn der Digitalisierung im Mai 2017
- Technische Partner: Walter Nagel, semantics (VL), Hardware vom hbz

Festlegungen der Projektphasen 1 und 2

- Entscheidung für ein kooperatives Massendigitalisierungsprojekt
- Digitalisierung von Mikroformen (Mikrofilm, Mikrofiche) in Münster für Westfalen und Brauweiler für das Rheinland
- Übernahme bereits vorhandener Digitalisate
- Lokale, Kreis- und Regionalzeitungen allgemeinen Inhalts im Zeitraum von 1801–1945, zunächst keine sog. Spartenzeitungen
- Orientierung an den Grenzen des heutigen Bundeslandes NRW
- Technische Features: Kalendersicht, Georeferenzierung, Zeitungsbeschreibungen, PDF-Download, OCR-Erkennung seit 2021

Workflow

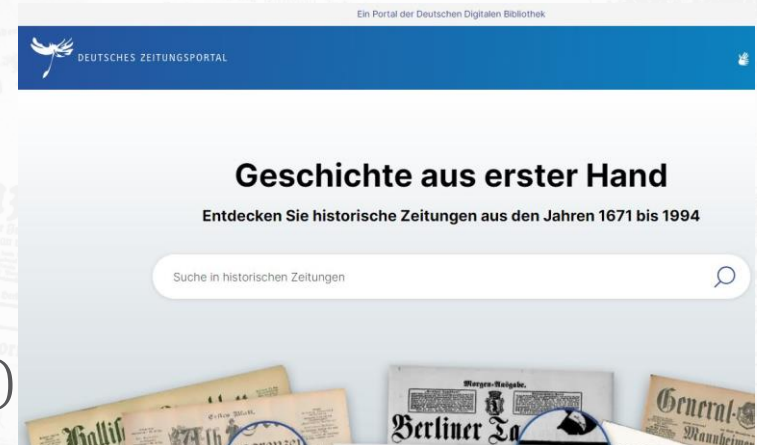


Workflow im Schnelldurchgang

- Kontaktaufnahme mit Archiven/Bibliotheken + Kooperationsverträge
- Abholung der Filme
- Digitalisierung der Filme + Nachbearbeitung der Seiten
- Hochladen der Seiten in die Visual Library
- Segmentierung der Titelblätter +
- Strukturierung der Zeitungen (beides für die Kalenderfunktion)
- Katalogisierung der Titel in der ZDB + Freigabe im Portal
- Rückgabe der Filme

Erfolge der Projektphasen 1 und 2

- Aufbau der Projektstrukturen
- Freischaltung von zeit.punktNRW
- Erreichen der Zielmarken (trotz Corona)
- Digitalisierung zahlreicher Mikrofiches, v.a. aus Westfalen
- Nachweis im Deutschen Zeitungportal
- OCR-Texterkennung mittels Transcribus
- Pilotprojekt zur Direktdigitalisierung
- Etablierung des Projektes (Vorträge etc.)



OCR-Volltexte plus Exportmöglichkeiten

573 TREFFER FÜR ZEITUNGEN = KÖLNISCHE ZEITUNG UND VOLLTEXT = ALAAF

ALLE DETAILS

Treffer pro Seite: 20
Reihenfolge: aufsteigend

- KÖLNISCHE ZEITUNG (15.2.1824) 27**
Köln : DuMont Schauberg, Nro. 50 (1. Jänner 1803)-Nro. 182 (2)
Details
- KÖLNISCHE ZEITUNG (6.2.1825) 22**
Köln : DuMont Schauberg, Nro. 50 (1. Jänner 1803)-Nro. 182 (2)
Details
- KÖLNISCHE ZEITUNG (4.2.1826) 20**
Köln : DuMont Schauberg, Nro. 50 (1. Jänner 1803)-Nro. 182 (2)
Details

Aufruf an die Kölner
zur
Ehren-Garde der VENETIA.

Alaaf! alaaf eer köllsche Jungen!
Eer sollt han en grose Ehr;
En der Schmierstrohs trefft eer ungen
Mich em Werfhuhs em Quartehr;
Sin se auch ens knatsch versunken,
Stont se doch, de köllsche Funken,
Durch üch vun den Duhden op,
Un sint widder bovven drop.

Nit zo hauen, nit zo stechen,
Wet gehöt dem tapfern Held,
Geinem auch den Hals zo brechen,
Treckt eer nu nit en et Feld,
Auch nit op de Fühlinger Heiden: —
Ekkers well ich üch vereiden
Ei Paar Däg lang als Vasall.
Für den Helden Carnaval

PDF
txt
XML
jpeg

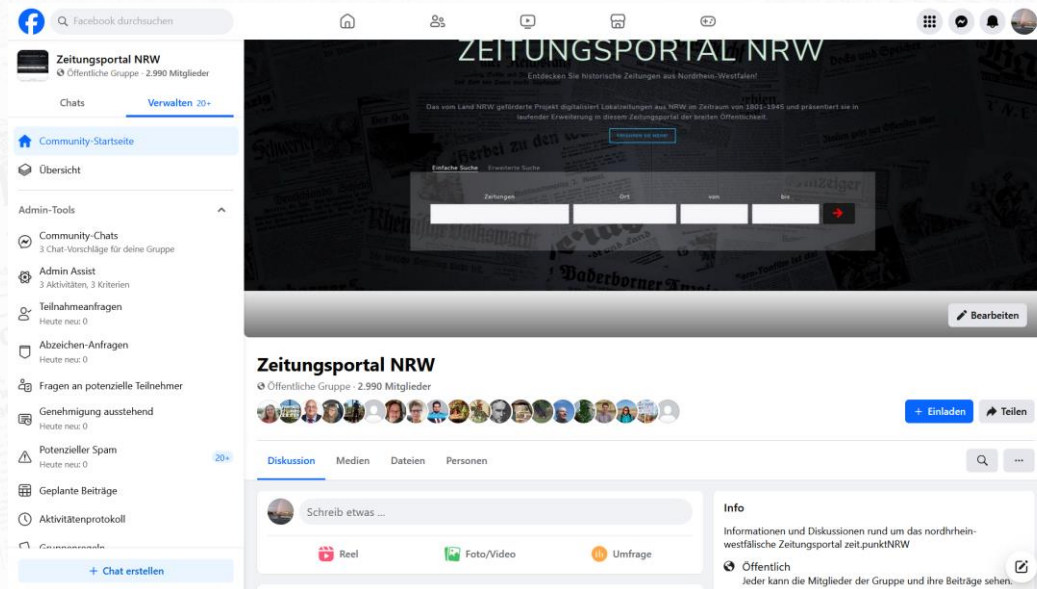
Ziele der 3. Projektphase (2023–2025)

- Filmdigitalisierung in Eigenleistung (ca. 2.5 Millionen Seiten)
- Filmdigitalisierung Barbarastollen (ca. 1.5 Millionen Seiten)
- Direktdigitalisierung durch Dienstleister (ca. 1 Millionen Seiten)
- **Ende 2025: 20–21 Millionen OCR-erkannte Images im Portal**
- Weiterentwicklung des Portals: Suchmöglichkeiten, Funktionalitäten
- Zeitungsbeschreibungen erweitern
- Stärkere Einbindung der Community (Digital Humanities-Bedarfe)
- Langzeitarchivierung im Digitalen Archiv NRW (DA NRW) starten

Zahlen von zeit.punktNRW

- **16.023.963** Images freigegeben, davon 14.831.108 mit OCR (92%).
Im Backend: 18.470.274 (Stand 15.03.24)
- **Größtes regionales Zeitungsportal Deutschlands!**
- Speicherplatz (TB): Images: 411 (600) | Cache 91 | Volltext: 13
- Besuche von Einzelnutzern seit 2018
 - 2018: 46.174 | 2019: 113.948 | 2020: 160.042
 - 2021: 256.213 | **2022: 525.901** | **2023: 815.262**
- Große Facebook-Community: Mitgliederzahl 2.996 (Stand 15.03.24)

Facebook-Community



Hans-Dieter Proff
Dank dem NRW Zeitungsportal, Informationen zum gefallenen Bruder meiner Großmutter gefunden, durch die Todesanzeige seiner Ehefrau Sibilla, haben wir jetzt auch das Todesdatum. Gefunden in der Oberkasseler Zeitung.



Gefällt mir Antworten Teilen 1 Wo. Bearbeitet

Christa Beschmitt
Auch ich konnte aufgrund der Zeitungen aus Köln Geschwister meiner Großmutter finden, von denen ich nichts wusste. Schönes Projekt.

Gefällt mir Antworten Teilen 1 Wo.

Sabine Hartmann
Habe gestern, Dank des Projektes, den Verbleib meines Ur-Urgroßvater nach 10 Jahren Ahnenforschung klären können! 😊

Gefällt mir Antworten Teilen 1 Wo.

Andreas Janik
Ihr habts mehr als verdient. Viel Arbeit, viele Erkenntnisse und eine zukunftsweisende Plattform. Dankeschön und weiter so 🙌🙌🙌

Gefällt mir Antworten Teilen 1 Wo.

Ich befasse mich mit Ahnenforschung und was für ein Glück das "Der Graftschafter" einsehbar ist. Jemand hat mir den link zum Zeitpunkt NRW gesendet und ich konnte somit, zum erstmal ein Bild von meiner Ururgroßoma sehen. Bin überglücklich!!!

Grenzen des Projekts

- Es geht nicht Alles
- Es geht nicht Alles gleichzeitig
- Teils schwierige Überlieferungssituation
- Mangelhaftes Material
- (Urheber)rechtliche und/oder vertragliche Probleme
- Begrenzte Fördermittel

Ausblick

- Diverse Vorträge 2024: VL-Anwendertreffen, Westfälischer Archivtag, BiblioCon, KPF.NRW, etc.
- Antrag für weitere Projektphase ab 2026 bereits in Planung.
- Bei Projektende 2029 sollen über **27.000.000** Seiten im Portal sein
- Weitere Fachtagungen zum Projekt geplant in 2025/2026
- Neues Format: Webinar „Sie fragen, wir antworten“ als offene Sprechstunde zu zeit.punktNRW, erstmalig am 15.04.24 (18–20 Uhr)
- Stay tuned auf der Facebook-Seite des Portals oder:
- auf der Mitteilungsseite des Portals oder bei Fragen einfach per Mail