

Landesprogramm Digitalisierung historischer Zeitungen in NRW

29.06.2018

Dr. Michael Herkenhoff (ULB Bonn) / Christine Baron (hbz) / Dr. Andrea Ammendola (ULB Münster)

Ministerium für
Kultur und Wissenschaft
des Landes Nordrhein-Westfalen



UNIVERSITÄTS-
UND LANDESBIBLIOTHEK





Warum historische Zeitungen digitalisieren?

- umfangreiches Quellenkorpus
- wichtige Quellen zur politischen Geschichte, zur Kultur-, Lokal- und Regionalgeschichte
- heterogene Überlieferungssituation
- neue Nutzungsmöglichkeiten durch Digitalisate
- „riesige Nutzungspotenziale auch außerhalb der Wissenschaft“ (Hans-Jörg Lieder)
- Erschließung noch nicht nachgewiesener Zeitungen in der ZDB

Zeitungsdigitalisierung ULB Bonn

Beschreibung	Alle Titel	Kalender
<p>Sie sind hier: Startseite → Zeitungen →</p>		
<h2>Zeitungen</h2>		
<p>Handschriften und Autographen (9341)</p> <p>Bibliothek Goussens (980)</p> <p>Translation. Die kognitive Aneignung europäischer Schlüsselkonzepte in asiatischen und nahöstlichen Gesellschaften (1860-1945) (24)</p> <p>Prinz-Georg-Bibliothek (39)</p> <p>Sammlung Klein (18)</p> <p>Rheinische Landeskunde (1393)</p> <p>Zeitungen (193)</p> <p>Bonn (32)</p> <p>Rhein-Sieg-Kreis (16)</p> <p>Bergisches Land (25)</p> <p>Kreis Aachen, Düren, Heinsberg (29)</p> <p>Kreis Ahweiler, Euskirchen, Eifel (19)</p> <p>Kreis Kleve (9)</p> <p>Kreis Neuwied, Westerwald, Mittelrhein (25)</p> <p>Mittelhessen (1)</p> <p>Bad Kreuznach, Rhein-Hunsrück-Kreis, Trier (13)</p> <p>Großraum Düsseldorf (10)</p> <p>Köln (6)</p> <p>Saarland (2)</p> <p>Westliches Ruhrgebiet, Geldern, Mönchengladbach (5)</p> <p>Zentrales Ruhrgebiet (1)</p> <p>Historische Drucke (1073)</p> <p>Universität Bonn (325)</p> <p>Sondersammelgebiet Romanistik (925)</p>	<p>Für die digitalisierten Zeitungen ist eine neue Funktion, die <i>Kalendersicht</i>, verfügbar. Damit können alle in den Digitalen Sammlungen enthaltenen Tagesausgaben zu einem bestimmten Datum parallel aufgerufen werden. Diese Funktion steht sowohl für den Gesamtbestand als auch für regionale Sichten (z.B. Bonn, Bergisches Land etc.) zur Verfügung. Die Kalendersicht bietet zugleich eine alphabetische Liste aller in der Kollektion angebotenen Zeitungen.</p> <p>Alternativ haben Sie weiterhin die Möglichkeit des systematischen Einstiegs.</p> <p>Vor kurzem haben wir damit begonnen, ausgewählte Titel mit einer Optischer Zeichenerkennung (OCR) zu bearbeiten. Bisher besteht aber nur die Möglichkeit, Inhalte über die Detailsuche zu finden. Wir arbeiten bereits an einem komfortableren Sucheinstieg.</p> <p>Die Universitäts- und Landesbibliothek Bonn verfügt über einen umfangreichen Bestand an historischen rheinischen Zeitungen, der seit 1824 als Pflichtexemplar in die Bibliothek gelangt ist. Wegen des im Vergleich zu heute bis 1945 wesentlich größeren Einzugsbereichs (Pflichtgebiet war die gesamte Rheinprovinz also vom Niederrhein bis zum Saarland) war die Zahl der abgelieferten Titel erheblich. In den 90er Jahren sind knapp 100 Zeitungstitel mit finanzieller Unterstützung des Landes Nordrhein-Westfalen sicherheitsverfälscht worden, um die Informationen dauerhaft zu sichern und die vom Papierverfall bedrohten Bände vor weiterem Verschleiß durch fortgesetzte Benutzung zu schützen. Seit dem Jahr 2009 werden die historischen rheinischen Zeitungen durch Dienstleister vor Mikrofilm in enger Abstimmung digitalisiert. Dabei sind zunächst vor allem Zeitungen aus Bonn und dem Rhein-Sieg-Kreis digitalisiert worden, danach aus dem Bereich Koblenz-Eifel-Ahr. Im Jahr 2012 wurden schwerpunktmäßig Zeitungen aus dem Bergischen Land, 2013 aus dem Kreis Aachen sowie weitere Titel aus dem Regierungsbezirk Koblenz, 2014 aus dem Bereich Köln, Düsseldorf, Niederrhein sowie aus der Region Oberkassel/Siebengebirge gesammelt und online veröffentlicht. Aktuell vervollständigen wir kontinuierlich unser Angebot und werden einige Titel bis ins Jahr 1945 digitalisieren. Zukünftig soll die Zusammenarbeit mit rheinischen Archiven im Fokus stehen, mit deren Hilfe wir unser Angebot kontinuierlich ausbauen wollen.</p> <p>Aus urheberrechtlichen Gründen ist das Jahr 1945 als vorläufige zeitliche Grenze gewählt worden.</p> <p>Um das digitale Angebot inhaltlich abzurunden oder Lücken zu schließen, wurden auch Zeitungen als Leihgabe von anderen Institutionen (z.B. Stadtarchiv Bonn, Stadtarchiv Bergisch Gladbach etc.) digitalisiert und Fremddigitalisate aus anderen Bibliotheken via Links eingebunden.</p> <p>Auf dennoch bestehende Lücken innerhalb einer Zeitung wird in einer Notiz zur Titelaufnahme hingewiesen. Bitte beachten Sie, dass die angegebenen Erscheinungszeiträume denen der Druckausgaben entsprechen. Aktuell haben wir leider nicht die Möglichkeit die Erscheinungszeiträume der Digitalisate anzeigen zu lassen.</p>	<p>Index</p> <ul style="list-style-type: none"> Home Alle Titel Kalender
<p>Listen</p> <p>Alle Titel</p>	<ul style="list-style-type: none"> Bonn (32) Rhein-Sieg-Kreis (16) Bergisches Land (25) Kreis Aachen, Düren, Heinsberg (29) Kreis Ahweiler, Euskirchen, Eifel (19) Kreis Kleve (9) Kreis Neuwied, Westerwald, Mittelrhein (25) Mittelhessen (1) Bad Kreuznach, Rhein-Hunsrück-Kreis, Trier (13) Großraum Düsseldorf (10) Köln (6) Saarland (2) Westliches Ruhrgebiet, Geldern, Mönchengladbach (5) Zentrales Ruhrgebiet (1) 	
Clouds		



Genese Landesprogramm

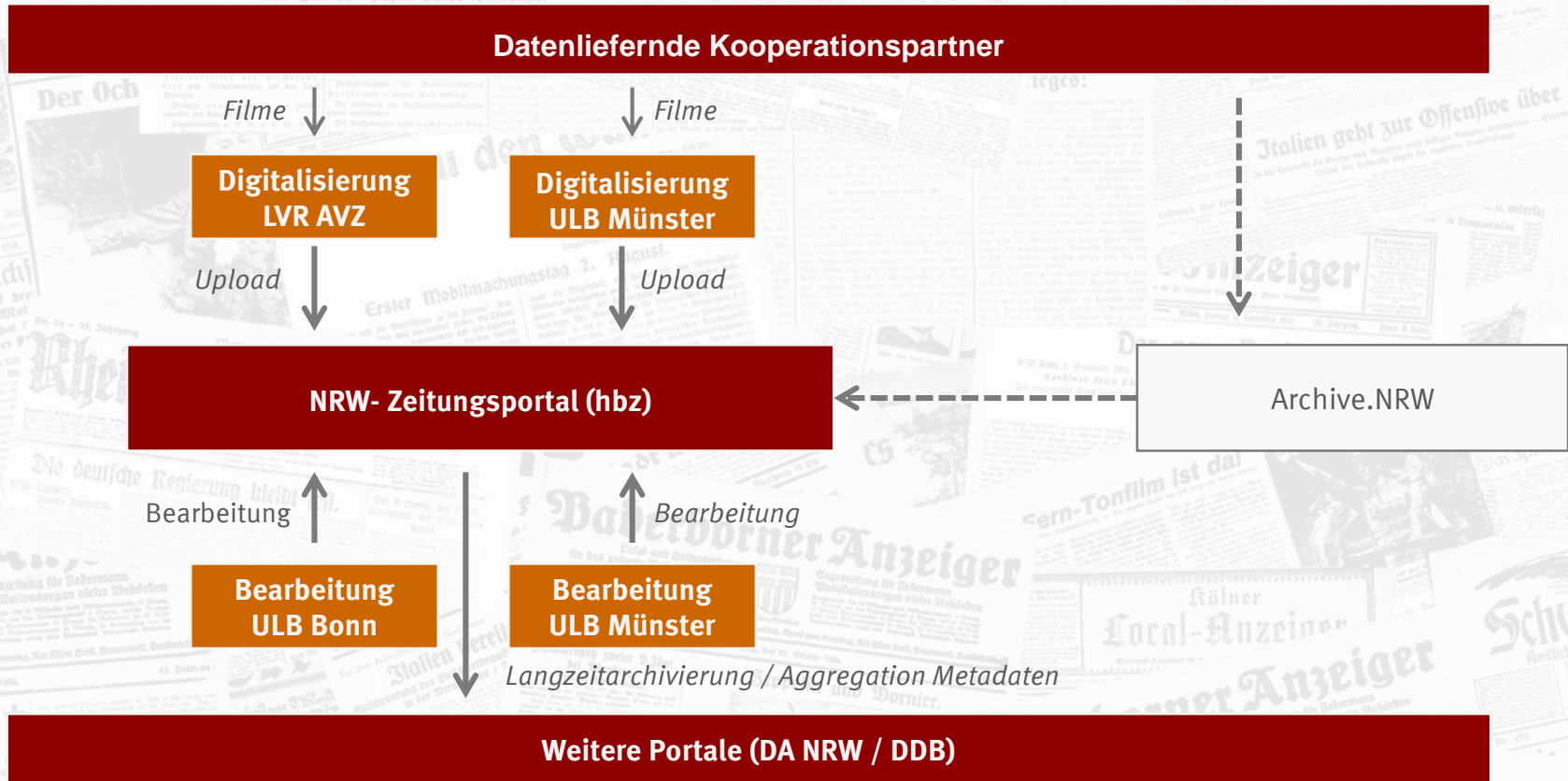
- Mai 2015: Fragebogenaktion im Rheinland (ca. 90 Institutionen befragt)
- September/Oktober 2015: Erstellung eines Konzepts für Zeitungsdigitalisierung im Rheinland (LVR-AFZ, ULB Bonn)
- November 2015: Landesbibliothekenkonferenz stimmt Konzept zu, beschließt Ausweitung auf ganz Nordrhein-Westfalen
- Dezember 2015: Konstituierung eines Lenkungskreises
- Frühjahr 2016: Einreichung eines Förderantrags und dessen Bewilligung durch das Ministerium
- 1. Januar 2017: Projektstart



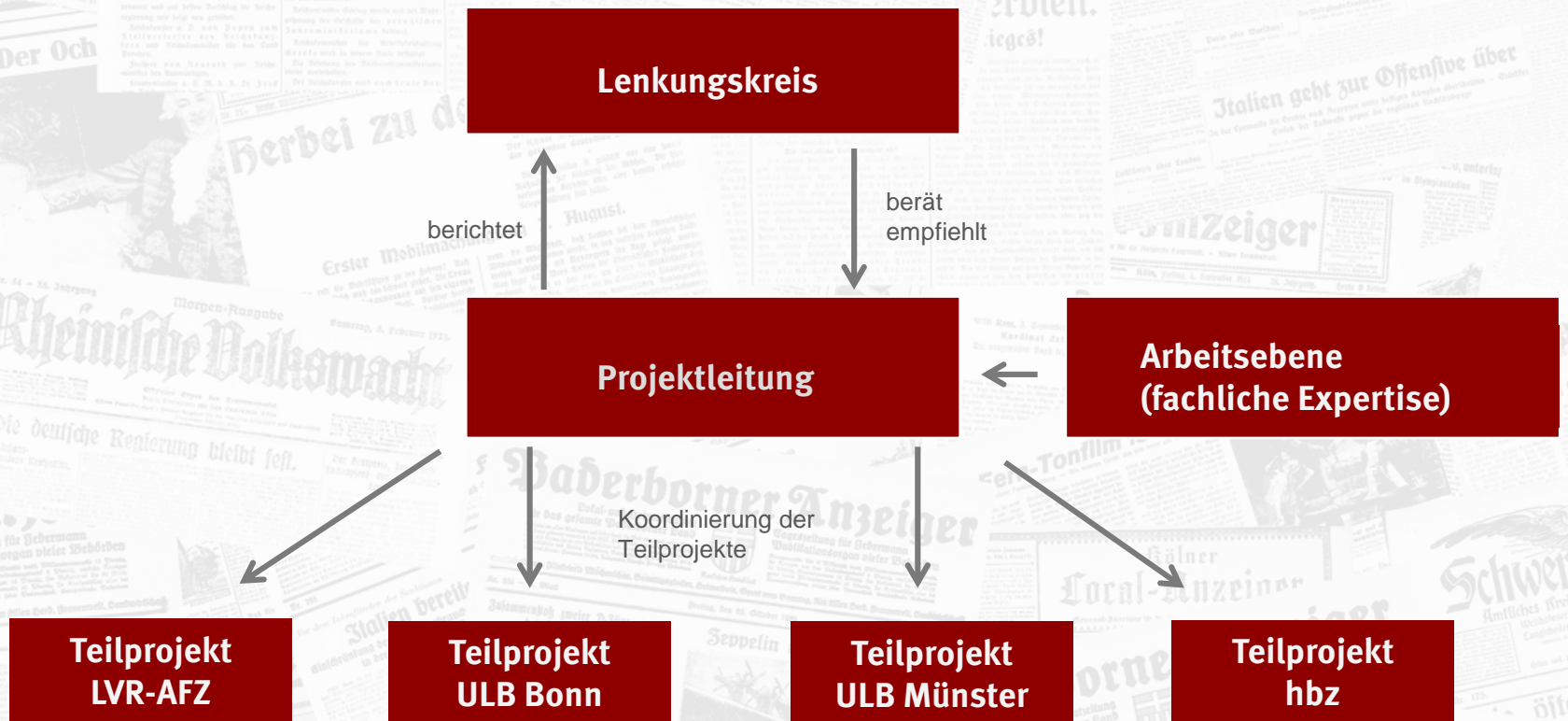
Inhaltliche Festlegungen

- Digitalisierung von Mikroformen (Mikrofilm, Mikrofiche)
- Übernahme bereits vorhandener Digitalisate
- Lokale, Kreis- und Regionalzeitungen allgemeinen Inhalts
- Orientierung an den Grenzen des heutigen Bundeslandes Nordrhein-Westfalen
- Zeitraum 1801 – 1945
- technische Features: Kalendersicht, „Visualisierungskomponente“, Download von Zeitungsausgaben als PDF-Dateien, keine OCR-Erkennung

Workflow



Projektstruktur





Technischer Aufbau – hbz

Christine Baron
baron@hbz-nrw.de



Scantoweb hosted by hbz

- Im Einsatz: Visual Library von semantics
- Partner Walter Nagel GmbH & Co. KG und semantics Kommunikationsmanagement GmbH
- Entwicklung, Installation, Betrieb und Wartung durch semantics
- Hardware-Hosting durch das hbz



Die Zahlen

20.000.000 = 20 Millionen Zeitungsseiten

- 1 gut ausgerüsteter (virtueller) Server
- 500 Terabyte Plattenplatz für die Scanns
- 150 Terabyte Plattenplatz für die Präsentation

Wir starten mit:

- 100 Terabyte Archivspeicher
- 50 Terabyte Cache für die Präsentation

Ein Vergleich

100 Gigabyte

1970



1995



2018





Rechenzentrum hbz

- Speichersysteme mit insgesamt 700 Terabyte Festplattenplatz (= 716.800 Gigabyte)
- 80 Server
- 3 Virtualisierungssysteme
 - Basierend auf 20 Servern
 - 100 virtuelle Server

Virtualisierung



Rosetta
Scan2Web
ELi:SA



Virtualisierung



Rosetta
Scan2Web
ELi:SA
Zeitungsportal





Das erste Projektjahr 2017

Vorproduktionsphase

- Dezember 2016 – April 2017:

- Leihverträge mit den ersten Kooperationspartnern
- Transport der ersten Filme nach Münster und Brauweiler
- Einstellen von Personal
- Planung und Abstimmung von technischen Workflows



Das erste Projektjahr 2017

Produktionsphase

- Mai 2017: Start der Digitalisierung in Münster und Brauweiler
- Juli 2017: Start der Strukturierung in Münster und Bonn

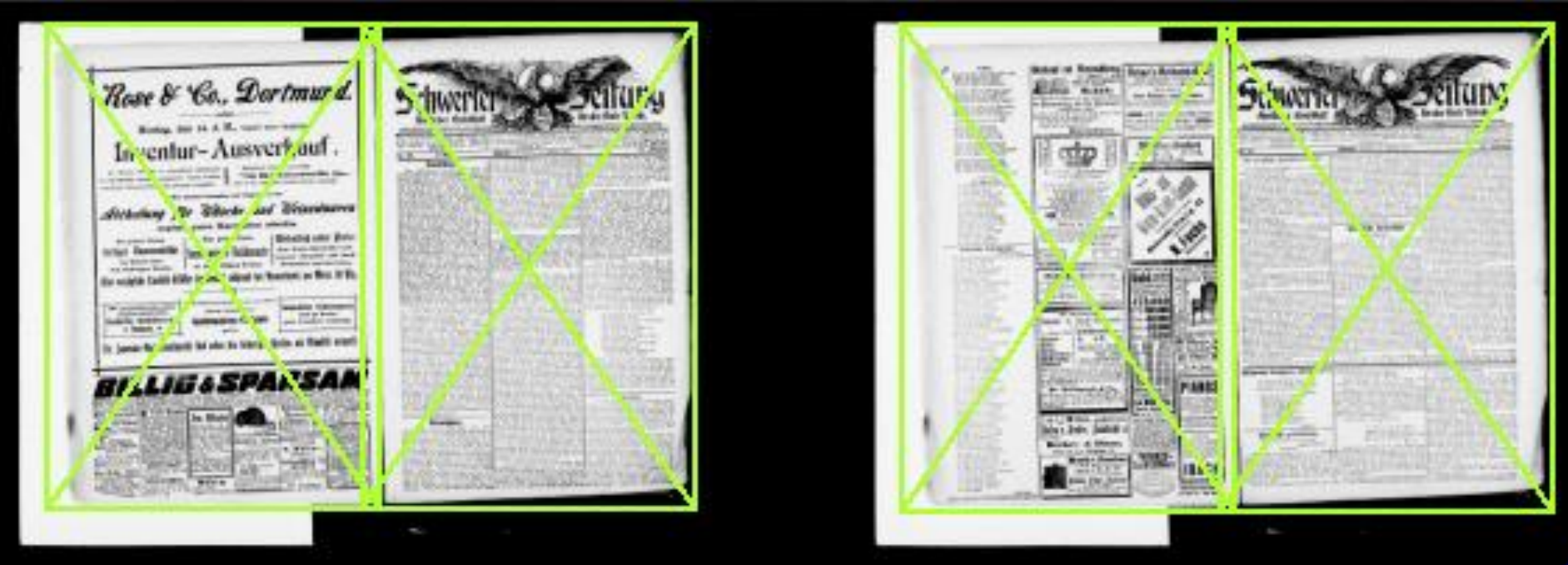
Das erste Projektjahr 2017

Workflow im Schnelldurchgang:

- Die Mikrofilme werden digitalisiert und die Seiten so nachbearbeitet, dass sie gut lesbar und nicht verschoben oder verdreht sind.

Das erste Projektjahr 2017

Workflow im Schnelldurchgang:



Das erste Projektjahr 2017

Workflow im Schnelldurchgang:

- Anschließend werden die Seiten in die Visual Library (VL) hochgeladen; dort werden die Titelblätter segmentiert, Lücken werden notiert und die Zeitungen strukturiert.
- Bei der Strukturierung wird ein Strukturbaum erstellt, mit dessen Hilfe sich die späteren Nutzer über eine Kalenderfunktion im Portal die verschiedenen Jahrgänge bis zur jeweiligen Tagesausgabe anzeigen lassen können.

Das erste Projektjahr 2017





Das erste Projektjahr 2017

- Die Metadaten werden aus der ZDB in die VL eingespielt. Vorher werden dort die digitalen Sekundärformen katalogisiert.
- Falls noch keine Metadaten in der ZDB vorhanden sein sollten, katalogisieren wir den betreffenden Zeitungstitel der Primärausgabe.



Zahlen und Ziele

Mai 2017 – Dezember 2019:

Soll

- Digitalisierung und Strukturierung von **4800** Filmen in
Münster und Brauweiler/Bonn

Ist (Stand 2.6.2018)

- Digitalisierung – Soll: 1950. Ist: 2067 (6% über dem Soll)

- Strukturierung – Soll: 1080. Ist: 1532 (40% über dem Soll)



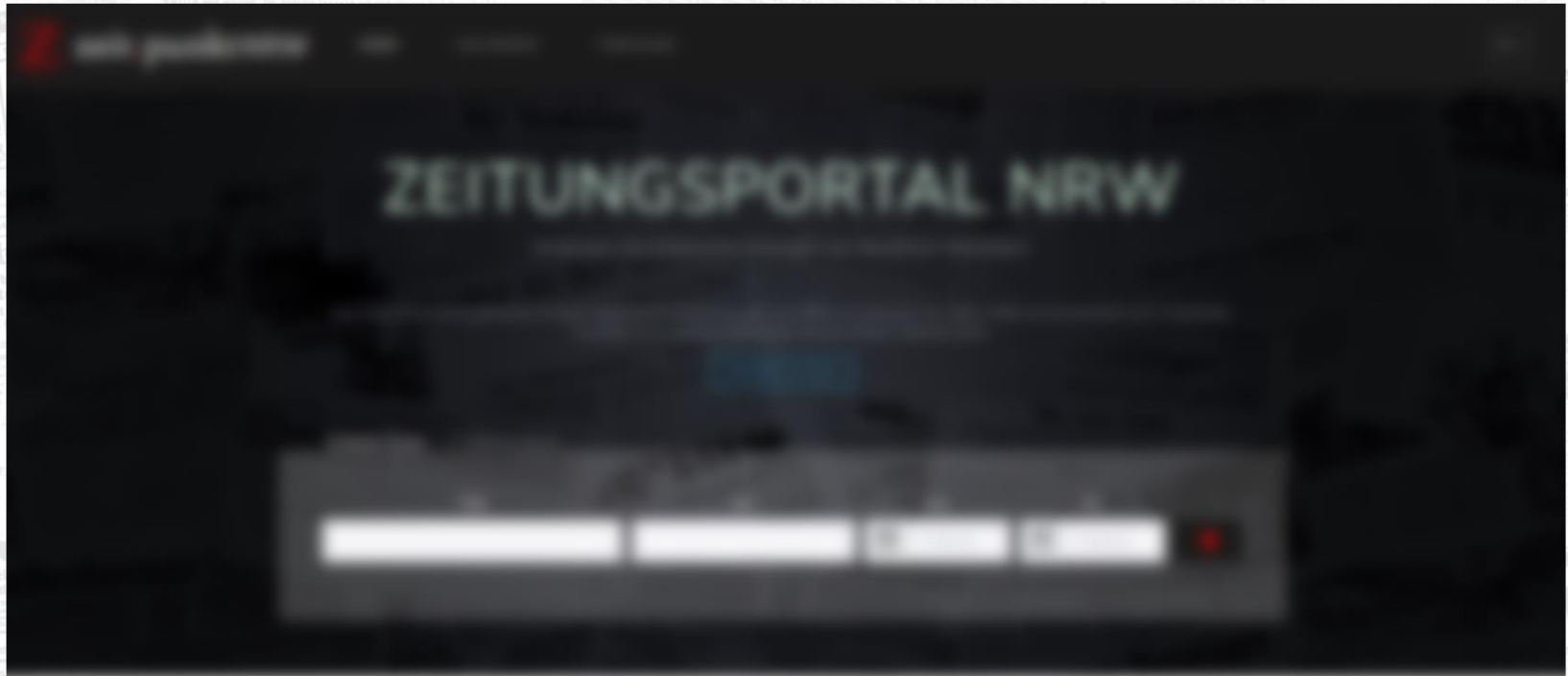
Bis Projekthalbzeit

- Mai 2017 bis Juni 2018: Planung der heutigen Fachtagung
- Herbst 2017 bis Juni 2018: Entwicklung und Testphase des NRW-Zeitungsportals mit semantics
 - Erstellen eines Anforderungskataloges, Prüfung durch semantics und Aufsetzen eines Release Candidate
 - Umfangreiche Testphase + Einspielen der Images



Landesprogramm Digitalisierung historischer Zeitungen in NRW

Freischaltung des Portals: zeit.punktNRW



Beate Möllers (Land NRW), 29.06.2018

Ministerium für
Kultur und Wissenschaft
des Landes Nordrhein-Westfalen

