

# Die SBB in Europa: Europeana 1914-1918, Europeana Newspapers & mehr

Clemens Neudecker  
Staatsbibliothek zu Berlin  
Berliner DH-Rundgang

16. April 2016

[@cneudecker](https://twitter.com/cneudecker)



**Staatsbibliothek  
zu Berlin**

Preußischer Kulturbesitz



# Europeana

- Die europäische Plattform für das digitalisierte kulturelle Erbe aus Bibliotheken, Museen, Archiven
- Aktuell über 52 Mio. digitale Objekte
- 2016: Relaunch als „Europeana Collections“ mit thematischen Schwerpunkten
- [www.europeana.eu/portal/](http://www.europeana.eu/portal/)

# Europeana Collections 1914-1918



- Förderlinie CIP-ICT-PSP (€5.4 Mio. , €2.7 Mio. EU)
- Laufzeit: Mai 2011 – April 2014
- 12 Projektpartner
- 4 Arbeitspakete
- Koordiniert durch Staatsbibliothek zu Berlin
- [www.europeana-collections-1914-1918.eu](http://www.europeana-collections-1914-1918.eu)

# Europeana Collections 1914-1918

- Hauptziele
  - Digitalisierung von 400.000 Objekten zu WW1 unter Gesichtspunkten wie:
    - Hohe Qualität der Digitalisierung
    - Auswahl der Objekte in Kooperation mit Experten und Vertretern relevanter Zielgruppen
    - Abdeckung einer großen Vielfalt von Materialien
    - Seltene und fragile Bestände
    - Zusätzliche Erschließung (z.B. Themengruppen)

# Europeana Collections 1914-1918

- Ergebnisse
  - 425.000 digitalisierte Objekte, darunter:
    - mehr als 10.000 Bücher
    - etwa 1.700 Liedertexte und Notenbücher
    - etwa 100.000 Zeitschriften und Zeitungen
    - 150.000 handschriftliche Objekte (Tagebücher, Briefe)
    - 180.00 Plakate, Fotos und Postkarten
  - Mehrsprachiges, thematisch geordnetes Portal
  - Virtuelle und physische Ausstellungen
  - „Unlocking Sources“ Konferenz



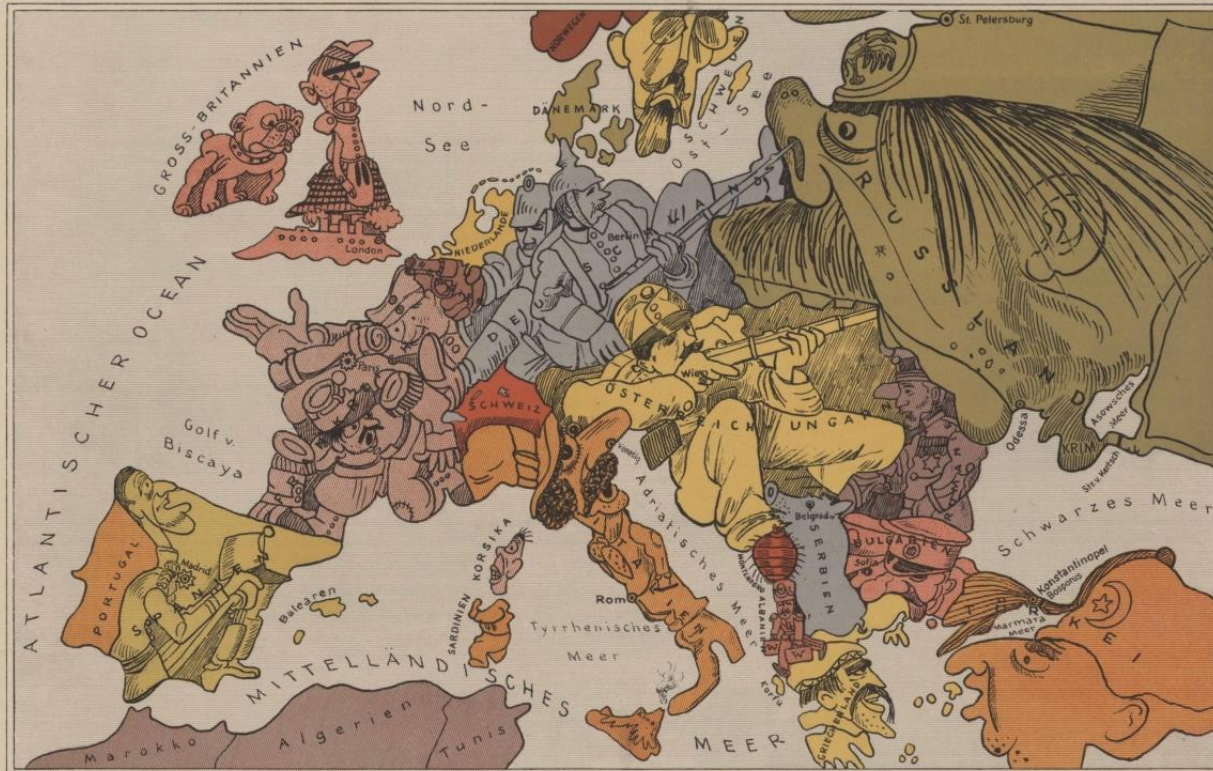
REKLAMEVERLAG ERNST MARX, BERLIN W 8

# KARTE VON EUROPA IM JAHRE 1914

1914 750.

[1]

Gezeichnet von W. Trier



**England** verbringt seine Flotte unter der Heimat Rock, die Dogge **Irland** im Rücken. König Oskar von **Schweden** sieht „gespannt“ an Russland. **Frankreich** zieht sich müde zurück, während sich **Spanien** säuseln Nichtstun hingibt- **Belgien** – hat ihm schon! **Deutschland** und **Oesterreich-Ungarn** teilen gute deutsche Liebe nach allen Seiten aus. In der **Schweiz** herrscht gute „Willt“ **Italien** ist treu zum **Tum** und dem Sieger. **Sicilien** vulkanischer Boden, sonst aber ganz ruhig. **Montenegro** – Lausbabe, **Serbien** – Schweinebändel: Der König von **Durazzo** – **Albanien** – verlässt sein Volk. **Griechenland** und die **Türkei** haben sich zum Fressen gern. **Ferdinand** von **Bulgarien** möchte auch etwas abhaben. **Carolus** von **Rumänien**? **Russland** will das Ganze verschlingen, es wird ihm aber nicht gelingen!

M. R. 32014



PREIS DER DOPPELKARTE 30 PF.

10% dieses Preises erhält das  
**CENTRAL-COMITÉ VOM ROTEN KREUZ.**



1918  
"grace!"  
- fuzg luf  
n gever  
den Uur.  
Wiss  
yger  
ts Writ  
talt-pif  
siele  
yfr scrip  
soll  
yfland  
sue



Europeana  
collections  
1914-1918



Europeana  
1914-1918

Gemeinsames Themenportal:  
[www.europeana1914-1918.eu](http://www.europeana1914-1918.eu)



## Erkunden Sie Europeana 1914-1918 nach Dokumentenarten



Postkarten



Offizielle Dokumente



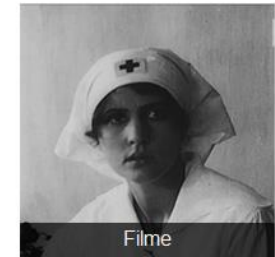
Briefe



Tagebücher



Fotos



Filme

## Themen



Frauen



Kriegsgefangene



Gedenken

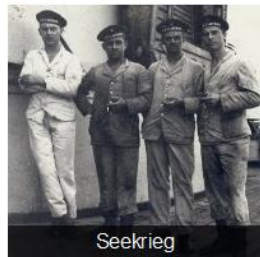


Propaganda



Leben im Schützengraben

## Kriegsschauplätzen



Seekrieg



Luftkrieg



Italienische Front



Heimatfront



Ostfront



Westfront



## Treffer für Ihre Suche

Quelle: aus Archiven, Bibliotheken und Museen

Subject: Propaganda 

Ergebnis Ihrer Suche nach (447 Treffer)

### Verfeinern Sie

► Schlagwörter hier

#### ▼ Quelle

☐ Alle Dokument

☐ von Privatperso

☒ Aus archiven,

#### ▼ Subject

☒ Propaganda

☐ Krieg und film

☐ Patriotismus (2

☐ Karikaturen un

☐ Krieg und gese

☐ Soldaten (17)

☐ Belgien (13)

☐ Krieganleihen

☐ Zivilisten im kri

☐ Infanterie (12)

☐ Nationalismus

☐ Massenmedien

☐ Ungarn (10)

☐ Grabenkrieg (1

☐ Werbung (10)

☐ Militärzeremon

☐ Gebäude -- krie

☐ Rumänien (9)

☐ Serbien (9)

☐ Frauen und kri

☐ Autonomie- und  
unabhängigkeitsbewegungen (9)

☐ Erster weltkrieg 1914-1918 -- denkmäler (9)

☐ Mobilisierung der streitkräfte (8)

☐ Städtisches leben (8)

☐ Russland (7)

☐ Kinder und krieg (7)

#### ▼ Subject

☐ Battlefields (43

☐ Campaigns (43

☐ Trench warfare

☐ Transportation

☐ Caricatures and

☐ Serbia (1469)

☐ Religious aspe

☐ Literature and t

☐ War and societ

☐ Soldiers (1252)

☐ Aerial operation

☐ Medical care (1

☐ Prisoners of wa

☐ Estonia -- histo  
independence,

☐ Campaigns -- it

☐ Technology (77

#### ▼ Subject

☐ Schlachtfelder (4388)

☐ Feldzüge (4307)

☐ Grabenkrieg (2984)

☐ Verkehr (2169)

☐ Karikaturen und cartoons (

☐ Serbien (1469)

☐ Der krieg in religiöser hinsi  
(1429)

☐ Krieg und literatur (1335)

☐ Krieg und gesellschaft (130

☐ Soldaten (1252)

☐ Luftkampf (1093)

☐ Erster weltkrieg--medizinis  
versorgung (1020)

☐ Kriegsgefangene (948)

☐ Estland -- geschichte --  
unabhängigkeitskrieg 1918  
(918)

☐ Feldzüge -- italien (887)

☐ Technik (774)

Deutsch

Englisch

Italienisch

Französisch

Niederländisch

Dänisch

Serbisch

#### ▼ Subject

☐ Luoghi di battaglia (1886)

☐ Operazioni

☐ Guerra di indipendenza (1887)

☐ Trasporti (21

☐ Caricature e

☐ Serbia (1469)

☐ Ruolo [della] rel

☐ Letteratura -- guer  
1914-1918 (1335)

☐ Società -- effetti [della] guerra  
mondiale 1914-1918 (1300)

☐ Soldati (1252)

☐ Operazioni aeree (1093)

☐ Assistenza sanitaria (1020)

☐ Prigionieri (948)

☐ Guerra di indipendenza estone,  
1918-1920 (918)

☐ Operazioni militari - italia (887)

☐ Tecnologie (774)

Wie Friede wird  
Gewissensfragen und  
Antworten  
F. E. Frankenberg und  
Ludwigsdorf

Staatsbibliothek zu  
Berlin - Preußischer  
Kulturbesitz

# Europeana Newspapers



- Förderlinie CIP-ICT-PSP (€5.2 Mio. , €4.1 Mio. EU)
- Laufzeit: Februar 2012 – März 2015
- 18 Projektpartner + 11 assoziierte Partner
- 6 Arbeitspakete
- Koordiniert durch Staatsbibliothek zu Berlin
- [www.europeana-newspapers.eu](http://www.europeana-newspapers.eu)

# Europeana Newspapers

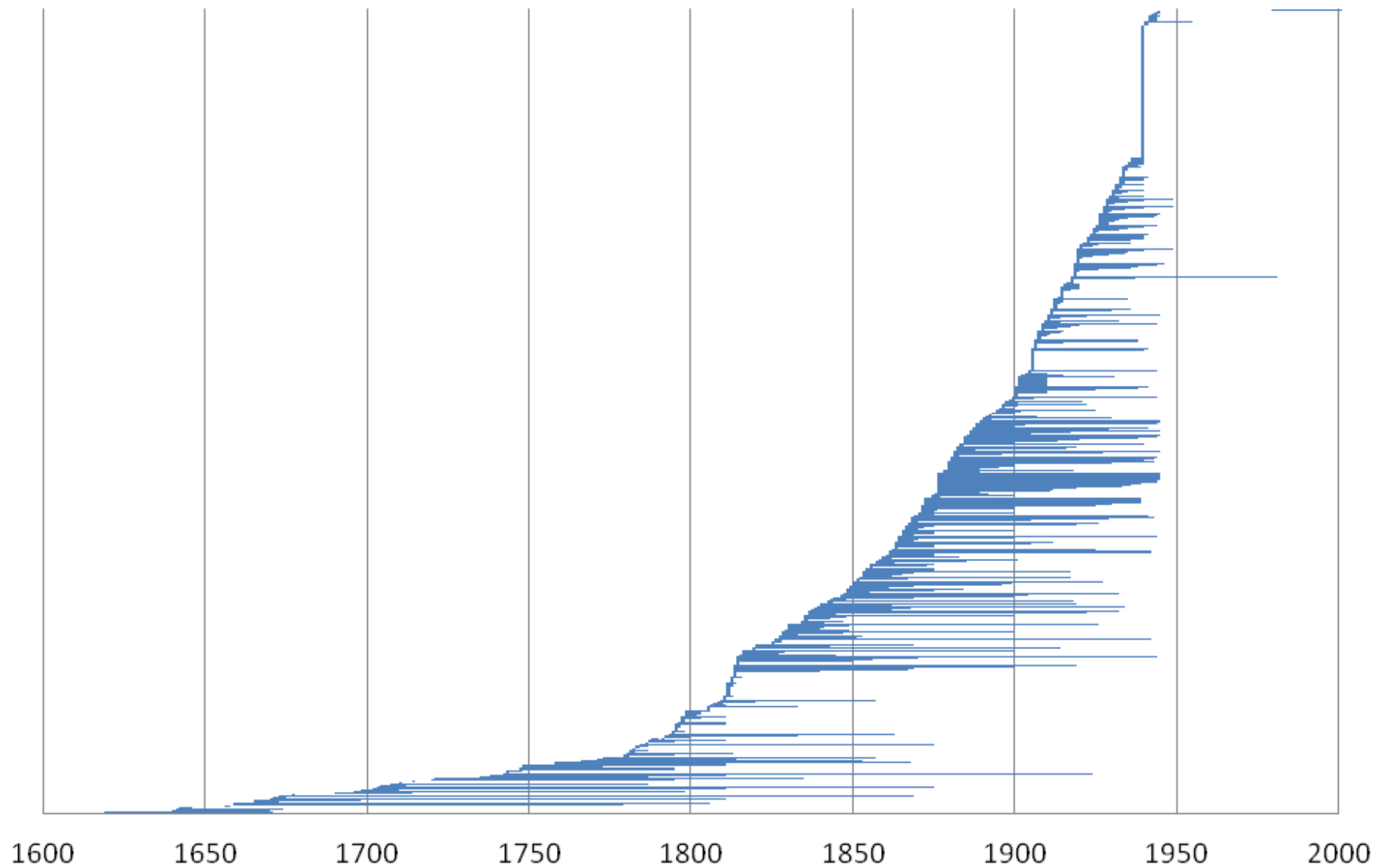
- Hauptziele
  - Aggregation und Transformation zu EDM von Metadaten digitalisierter Zeitungen in Europa
  - OCR/OLR für 10 Mio. Zeitungsseiten
  - Entwicklung eines Webportals für Volltextsuche
  - Design eines Meta- und Strukturdatenmodell
  - Methoden und Werkzeuge für die Evaluation
  - Workshops und Best-Practice Empfehlungen

# Europeana Newspapers

- Ergebnisse
  - Größter Dateningest seit Bestehen der Europeana: 3.6 Mio. Records, >10 TB Daten
  - 10 Mio. Seiten OCR + 2 Mio. Seiten OLR
  - Content-browser basierend auf JP2, IIP, IIIF
  - ENMAP (METS/ALTO) Datenmodell
  - Open Source Software Werkzeuge
  - 3 Datensets zur freien Nachnutzung: Ground truth, Named entities, Text mining



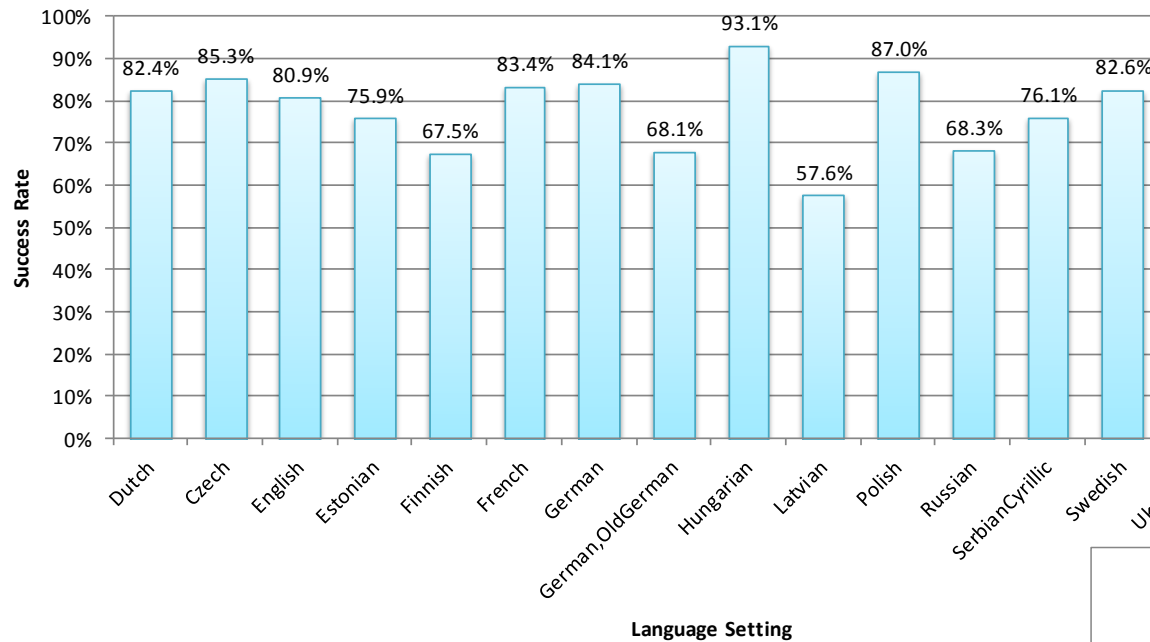
# Years covered per title





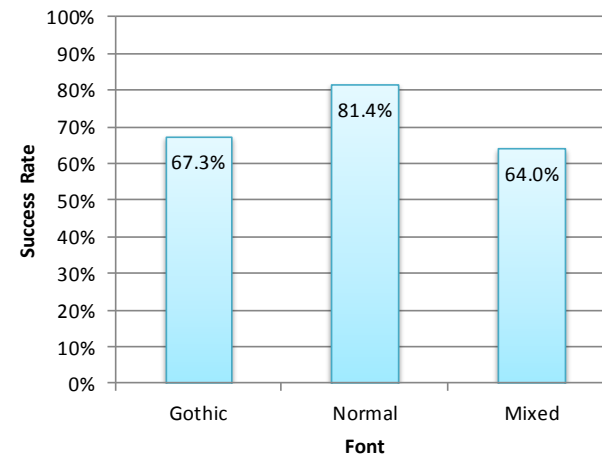
## Bag of Words OCR Evaluation

### Per Language



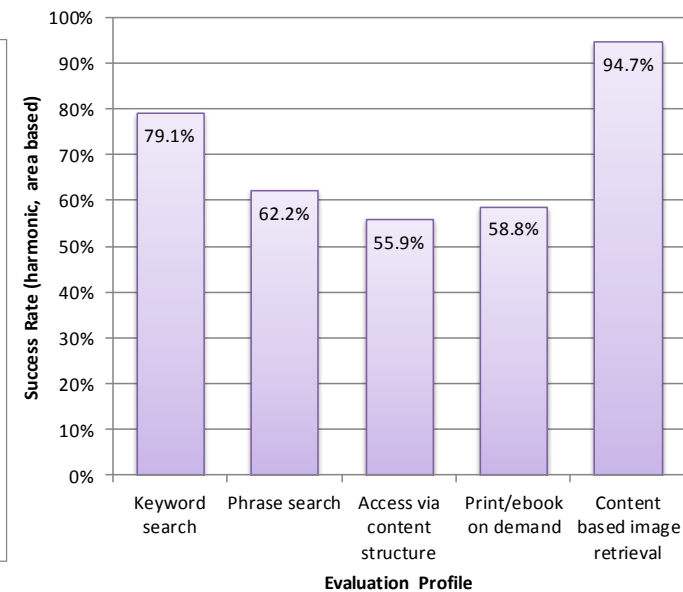
## Bag of Words OCR Evaluation

### Per Font



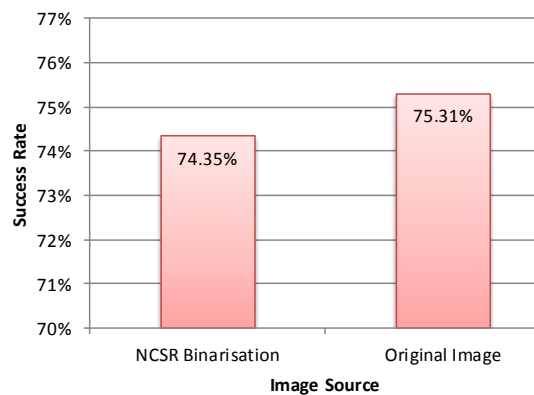
## Layout Analysis Performance

### Per evaluation profile



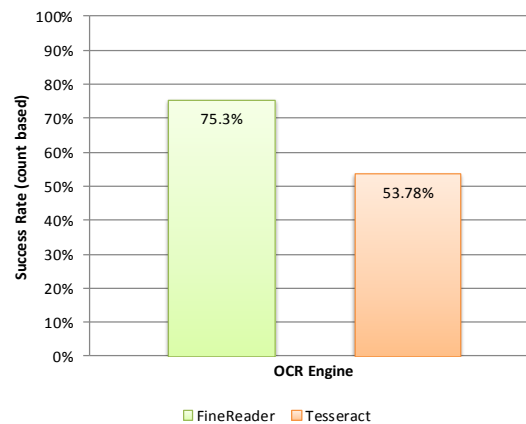
## Bag of Words OCR Evaluation

### Binarised image vs. original image



## Bag of Words OCR Evaluation

### FineReader vs. Tesseract





Select year:  Range: 1618 - 2016

Jan	Feb	Mar	Apr
May	Jun	Jul	Aug
Sep	Oct	Nov	Dec

S	M	T	W
	1	2	3
7	8	9	10
14	15	16	17
21	22	23	24
28	29	30	31

Home → Newspapers Home → Newspapers Gallery → Newspaper "Het nieuws van den dag : kle..." → 07 June 1896

THIS ISSUE PROVIDED BY

NEWSPAPER SOURCE

&lt;&lt; Back to search result

KB

Koninklijke Bibliotheek

See Original at Library...

CONTENT

ZOLA's **ROMÉ**

I Welk een boek! Zevenhonderd-een-en-vijfblad zijden, alle in één deel! L'Assommoir bood indertijd vijfhonderd-negen-en-zestig, Germinal vijf honderd-een-en-negentig, La Débâcle meer dan zeshonderd, en Lourdes vijf honderd-acht-en-negentig. **Romé** is dus het uitvoerigste van alle de werken, door Emile Zola voltooid. Is **Romé** ook het voortreffelijkste? Het antwoord op deze vraag is zeer moeilijk te geven. Het nieuwe kunstwerk heeft ze verschillende eigenschappen. De auteur heeft meer dan één doel te beoogen. Het was voornemen een vervolg op Lourdes te schrijven. In dit opzicht is hij alleen dan geslaagd, wanneer de in beide boeken op tredende hoofdpersoon, de abbé Pierre Froment, het verband tusschen het eerste en tweede kunstwerk voldoende aanwijst. A het slot van Lourdes denkt de abbé Froment aan 'n u n e nouvelle religion' en dit thema wordt in **Romé** vervolgd. Vc worden 'les pères delagrotte' een paar maal te **Romé** genoemd. Aan het sl' "van **Romé** is de hoop van den abbé **For** op 'une nouvelle religion' geheel vervlogen. Hij schijnt daarom zijn hart te willen wijden aan de wetenschap, die de macht der duisternis verzwakt — hij schijnt alleen te vertrouwen op 'Ja rai son.' Aan dit aanknopingspunt zal waarschijnlijk h derde kunstwerk : Paris worden verbonden. Het komt mij voor, dat de eenheid tassol Lourdes en **Romé** beter zal kunnen worden uitgesproken door er op te wijzen beide boeken eene studie bevatten over l'humanité Catholique. Maar dan n

More About This Text

Cite this page

Full-text and image is Public Domain (1.0)

A B C D E

Bade- und Reise-Journal (1930)  
 Badener Bezirks-Blatt (1880-1911)  
 Badener Zeitung (1896-1943) (F)  
 Baldur (1932-1959)  
 Baldur (1868-1870)  
 Baldur (1903-1910)  
 Baldursbrá (1934-1940)  
 Baleares (Palma. 1917) (1917-1918)  
 Balkanski rat u slici i reči (1917-1918)  
 Balss (1905-1905) (Full-text available)  
 Baltija (1907-1907) (Full-text available)  
 Baltijas Vēstnesis (1868-1905)  
 Bandera social (1885-1886)  
 Bankablaðið (1935-1995)  
 Banovac (1888-1912)  
 Baranovičer Kurier (1939-1939) (Full-text available)



Zoom





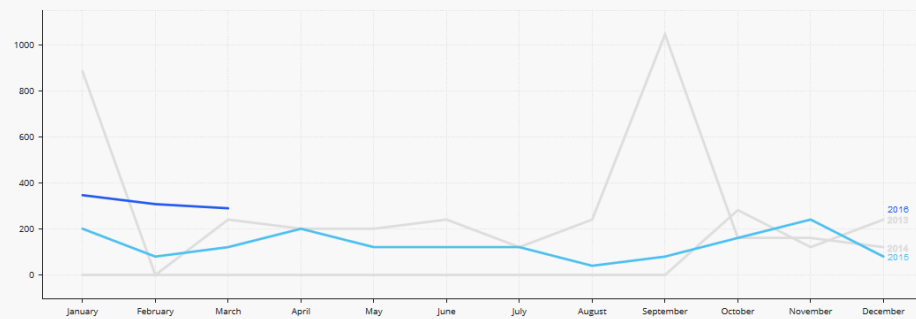
# Europeana DSI

- Seit 2015: Finanzierung der Europeana durch die EU-Kommission als „Digital Service Infrastructure“ im Rahmen des CEF Programm
- SBB ist Projektpartner in Europeana DSI-2 und dort zuständig für die Weiterentwicklung und den Launch einer „Europeana Newspapers Collection“ (voraussichtlich Q1/2017)

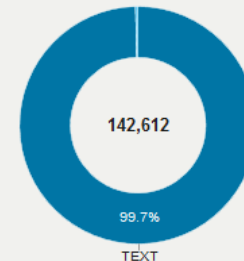
# SBB Daten in Europeana: Nutzung

- Brandneu seit 6. April 2016!

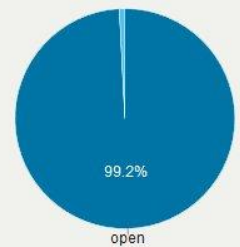
Staatsbibliothek zu Berlin - Preußischer Kulturbesitz Monthly page views



TOTAL NUMBER OF ITEMS



AVAILABLE FOR RE-USE



- Aktuell:
  - 142,612 digitale Objekte der SBB in Europeana
  - 99,2% „available for re-use“ (CC-BY-SA, CC-0)
- [Europeana Statistics Dashboard SBB-PK](#)

# Danke für Ihre Aufmerksamkeit!

## Noch Fragen?

Clemens Neudecker  
Staatsbibliothek zu Berlin  
Berliner DH-Rundgang

16. April 2016

[@cneudecker](#)



**Staatsbibliothek  
zu Berlin**

Preußischer Kulturbesitz

