

# 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)
    - 180.000 Plakate, Fotos und Postkarten
  - Mehrsprachiges, thematisch geordnetes Portal
  - Virtuelle und physische Ausstellungen
  - „Unlocking Sources“ Konferenz

# Der Geekrieg

# KARTE VON EUROPA IM JAHRE 1914

1914 750.

[1]

Gezeichnet von W. Trier



10% dieses Preises erhält das  
"CENTRAL-COMITÉ VOM RÖTEN KREUZ."



England verbringt seine Flotte unter der Heimat Rock, die Dogge Irland im Rücken. König Oskar von Schweden sieht „gespannt“ auf Russland. Frankreich zieht sich mutig zurück, während sich Spanien süßen Nostalgie hingibt. Belgien — hat ihm schon! Deutschland und Österreich-Ungarn teilen gute deutsche Flüche nach allen Seiten aus. In der Schweiz herrscht guter „Will“: Italien ist treu bis zum Tod — dem Sieger. Sizilien vulkanischer Boden, sonst aber ganz ruhig. Montenegro — Lausebandsche — Schweinebandsche! 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!





europeana  
collections  
1914-1918



europeana  
1914-1918

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

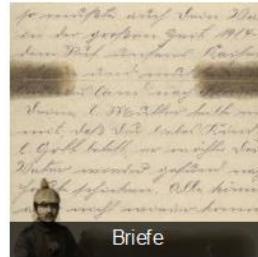
## Erkunden Sie Europeana 1914-1918 nach Dokumentenarten



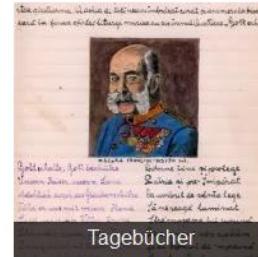
Postkarten



Offizielle Dokumente



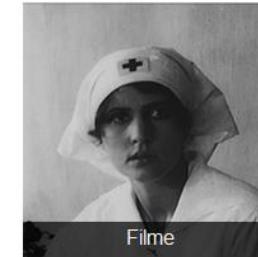
Briefe



Tagebücher



Fotos



Filme

## Themen



Frauen



Kriegsgefangene



Gedenken



Propaganda



Leben im Schützengraben

## Kriegsschauplätzen



Seekrieg



Luftkrieg



Italienische Front



Heimatfront



Ostfront



Westfront

Verfeinern Sie

Schlagwörter hinzufügen

Quelle

Alle Dokumente

von Privatpersonen

Aus archiven

Subject

 Propaganda Krieg und film Patriotismus (2) Karikaturen un Krieg und gese Soldaten (17) Belgien (13) Kriegsanleihen Zivilisten im kri Infanterie (12) Nationalismus Massenmedier Ungarn (10) Grabenkrieg (1) Werbung (10) Militärzeremon Gebäude -- krie Rumänien (9) Serbien (9) Frauen und krie Autonomie- und unabhängigkeitbewegungen (9) Erster weltkrieg 1914-1918 -- denkmäler (9) Mobilisierung der streitkräfte (8) Städtisches leben (8) Russland (7) Kinder und krieg (7)

# Deutsch

## Englisch

[Casimir Legumini](#)[Tasche](#) - [Brasil](#)

Central Institute for the

1941-1945

## Italienisch

[Luoghi di battaglia \(1900\)](#)[Operazioni](#)[Guerra di indipendenza \(1860-1871\)](#)[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 \(774\)](#)

## Französisch

## Niederländisch

## Dänisch

## Serbisch

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)

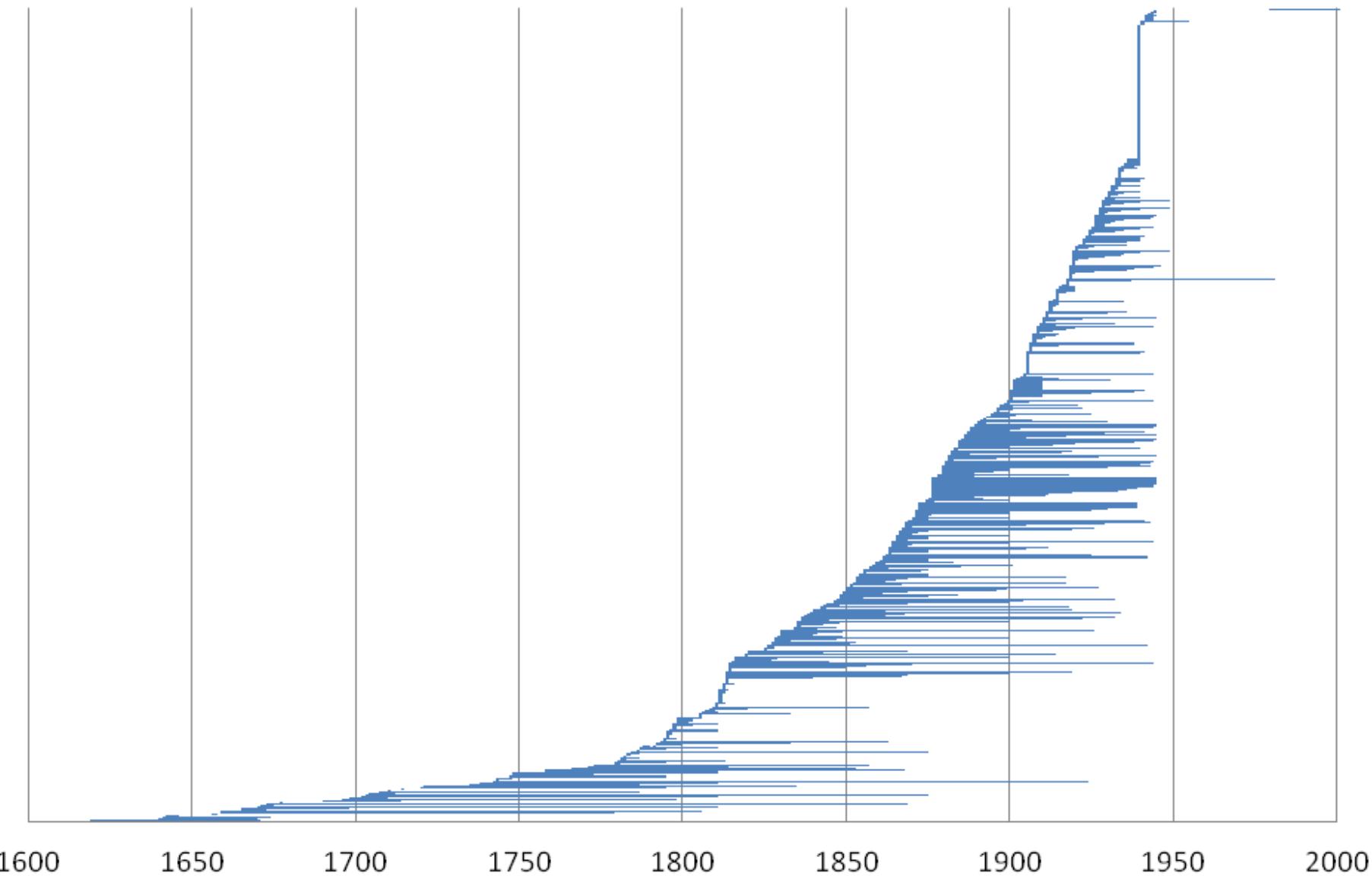
# Europeana Newspapers

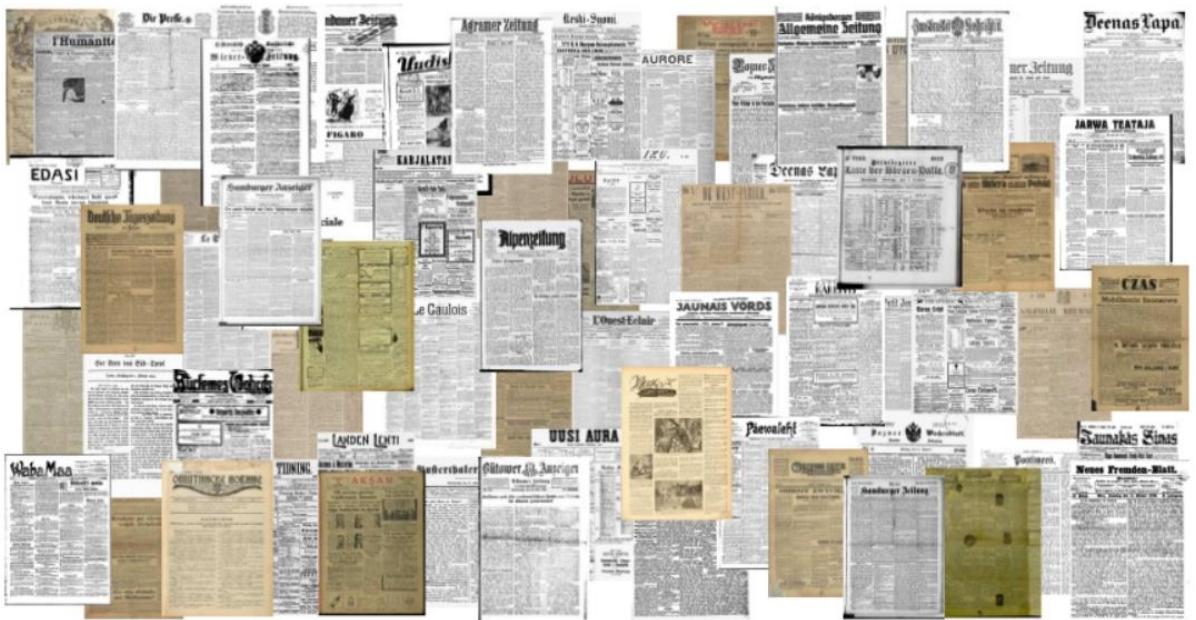
- Hauptziele
  - Aggregation und Transformation zu EDM von Metadaten digitalisierter Zeitungen in Europa
  - OCR/OLR für 10 Mio. Zeitungsseiten
  - Entwicklung eines Webportals für Volltextsuche
  - Design eines Meta- und Strukturdatenmodell
  - Methoden und Werkzeuge für die Evaluation
  - Workshops 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, IIIP, IIIF
  - ENMAP (METS/ALTO) Datenmodell
  - Open Source Software Werkzeuge
  - 3 Datensets zur freien Nachnutzung: Ground truth, Named entities, Text mining

# Years covered per title





Rapport van den Burger *Siderius*, uit naam van de Commissie, op den benoemd en gecommitteerd tot het examineeren van eene Misive van het S den 21 October 1802, houdende eenen Voordragt van eene Verdeeling va ment in Ringen, en een Reglement, bepalende de wyze, waarop het Stem geoeffend, als mede de Gegoedheid der Kiezers; uitgebragt in de Zitting 12 dezer, en als toen dadelijk geconcludeerd.

In de zitting, gisteren door het Wetgevend Lichaam gehouden, is het wetsontwerp tot het verleenen van een crediet van 5 miljoen voor de werklieden der Seine Inferieure, alsmede het buitengewone budget, ingekomen. Morgen zal het eerste wetsontwerp reeds in behandeling worden genomen. Voorts is het zeer waarschijnlijk dat er eenne groote loterij tot ondersteuning der katoen-arbeiders tot stand komt. Ongelukkigerwijze zijn dezen niet alleen de nooddriftigen. Zij, die in den jongsten tijd de voorsteden St. Antoine of St. Marcel hebben bezocht, zullen moeten erkennen dat de ellende aldaar nimmer zoo groot was en om zoo te zeggen de geheele arbeidende klasse heeft bereikt, dat zijne discussieën over het Adres

De Senaat zal Donderdag zijn discussie over het Aores kunnen beginnen. Ergisteren is reeds voorlezing van het ontwerp gedaan, 't welk even kleurloos als de Troonrede is. Men hoopt nog altijd dat Prins Napoléon het woord over de Romeinsche kwestie zal voeren, doch ik weet dat de heer Pietri in het derde bureau van den Senaat heeft verlaard, dat noch hij, noch *zijne politieke vrienden* die kwestie alleen bespreken.

De consul-generaal van New-Orleans, de graaf Mejean, naarmen vroeger beweerde, van zijn post was ontset, hier aangekomen en heeft reeds een langdurig onderhoud met den heer Drouyn de Lhuys gehad. Heden of morgen oopt hij bij den Keizer te worden toegelaten. Gij weet volgt dat die consul gezegd wordt de ontwerpen van het plan tot bijeenroeping eener vergadering van gedelegeerden te wezen, hoewel de markies de Montholon, onze consul-generaal te New-York, daaraan evenmin vreemd is gebleven.

Deel 1925

Zeg nu maar, dat een winkeber ~~zijn~~  
geld zoo maar verdient.... en wat een  
Nederlandsche taalkennis hij wel moet  
bezitten.

Ongeloofelijk, welke eigenschappen en kennis een winkelier zich moet aankweken om z'n zaak bloeiend te maken.

Op de eerste plaats: goede inkoop, goede etaleur- en reclamekunst en verder een reuzen-geduld en een ontzettend aanpassingsvermogen, om alle klanten gemakkelijk en vlug te bedienen.

van de Reserve den aanval, en de Bajonet beëindigde het lot van dien dag. Nietigenstaande den sterken wederstand werden de Ostionrykers overgeworpen, en tot aan den voet van de Reducties was geen zyde van Caldiero vervolgd.

Wy hebben 3500 gevangenen gemaakt. Het slagsveld is met Ostionrykse bezield. Het aantal huizen doodloos en verwoest.

Oostenrykers bedaaid. Het aantal hunner doden en gekwetsten is ten minsten zon groot als dichenner gevangenen. Prins Karel heeft een stilland van Wapenen doen vragen om de doden te begraven.

Ons verlies is in vergelyking van dat des vyands van weinig belang.

De Maarschalk en Chef kau niet salaten des moed van heit  
Leger openlyk toe te juichen; hy zal inzonderheid die daden  
die den dag merkwaardig maaken, doen kennen; en de names  
dier bravou, die 'er de eer van hebben, onder het oog van  
Z. M. den Keizer en Koning bringen.

De Moniteur van heden schryft onder Parys van gisteren ten  
aanzien van den voorgang der zeggeraende Wapenen der Fran-  
sche Armee in Italië het volgende: „Zes duizend Oostenry-  
kers zyn den 2 November voor de Poorten van Verona zelve  
Krygsgevangen gemaakt. Die Colonne was over de Bergen aank  
gekomen en moest de Fransche Armee in den rug vallen.

Drie honderd Man van het saame Regiment ligie Infanterie hebben dat Colonie een vur lang tegen gehouden; door 1500 Man versterkt geworden synde, hebben sy de vyandelyke Colonie, zes Battalions en een Eskadron Hazareen sterk, genoodzaakt de Wapenen neerleggen."

Behalve die berige geeft een ander Blad (*wascheychelyk* was gien hetzelfde voordeel) als eicht op die volgende tyding uit Milaan van den 3den. "De Maarchalk Maffei had den 2 November de hoogten van Caldiero doen inname, almede den belangryken post van Montebello, waarby 6000 man krygsgevangen waren gemaak. Den 3 November moet die Maarchalk Maffei te Vicenza slapen. De Aartsheer Karel was in voller acht. Binne weinige dagen verwachtte men die vereeniging der Italiaansche Armes met die groote, De Generals Gouvin en St. Cyr beweg nog eene versterking van zoos man aan Hetrurie soos door Spaansche Troeppen bezet wordoen."

Over het overige vindt men in den Munition en in de andere Papieren wagens den Zeelang by Cadix nog nieten. Alleen leeft men in een derzelve een nog zeer overbaard berigt deswegens, behelstende niets anders dan 't geen in ozzen vorigen wordt gemedied; behalven dat er als zeker in wordt opgegeven dat de Vice-Admiraal Rosily, die het hevel van de vereenigde Vloot te Cadix op zich ging menen, aldus drie dagen, na het uitloopen van Villeneuve en Gravina aangekomen is.

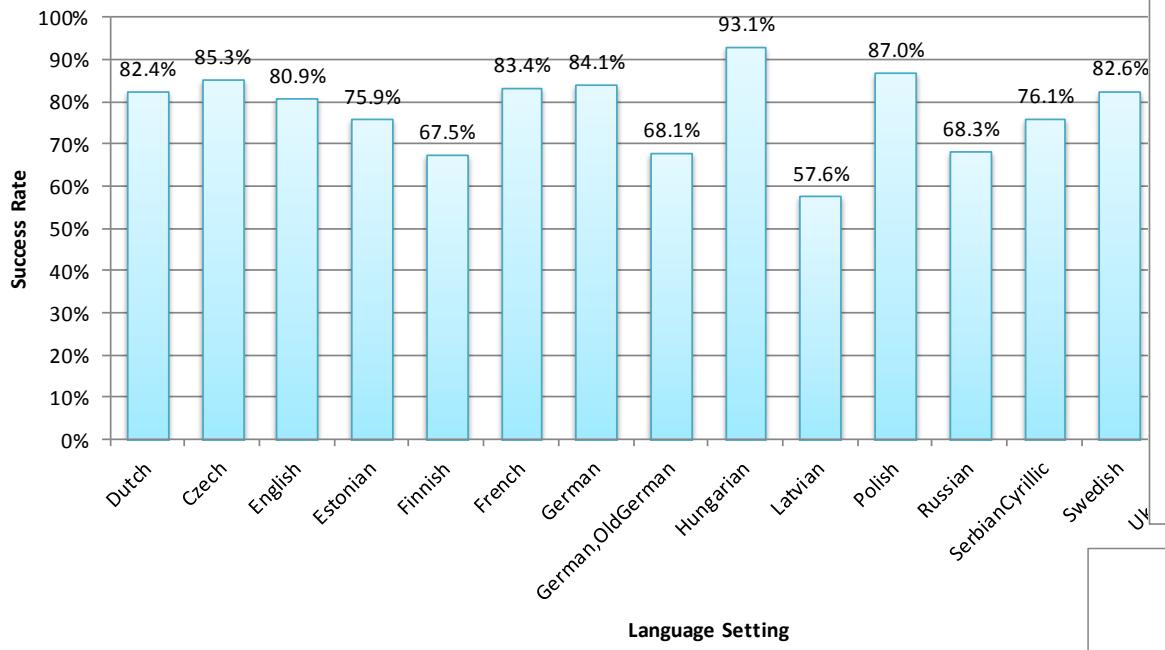
De Fondien staan 60 Fr. De Bank-Actionen 1180 Fr.

**BATAAFSCHE REPUBLIEK.**  
AMSTERDAM den 10 November. Den 11 deszelfs in Tessa op Breewyk ten anker gekomen één Americanae Bril, onbekend; en een zynthegezel N. Mycen na Baltimore, C. F. O. H. Harenberg na St. Thomas, D. Hall na St. Ubes, L. Evenen na Tonsberg, A. H. Kiaras Dramme; de Wind Z. Oost.  
Den 10 deszelfs is in 't Vlie nieten binden gekomen; maar uitzig  
zei L. Netzel na Koningsbergen, en D. Simons na Leer; de  
Wind Z. Oost.  
Den 10 deszelfs is te Terfchelling binnen gekomen L. Hoek van

Den 9. deser is te Terschelling binnengekomen J. Hoek van

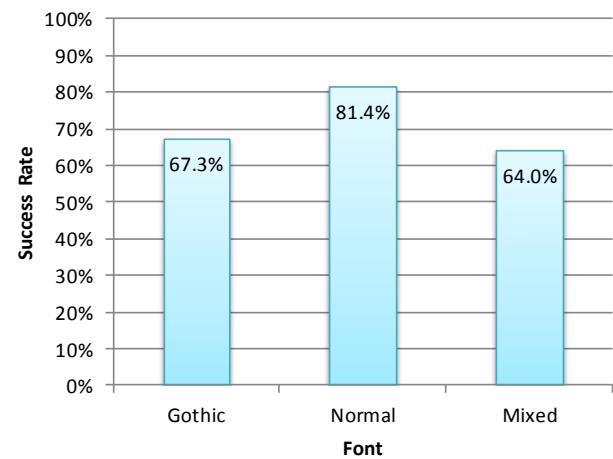
## Bag of Words OCR Evaluation

### Per Language



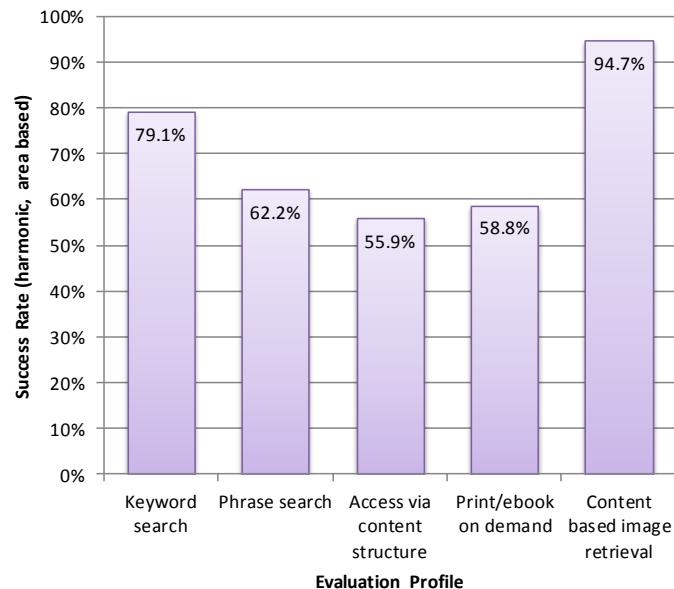
## Bag of Words OCR Evaluation

### Per Font



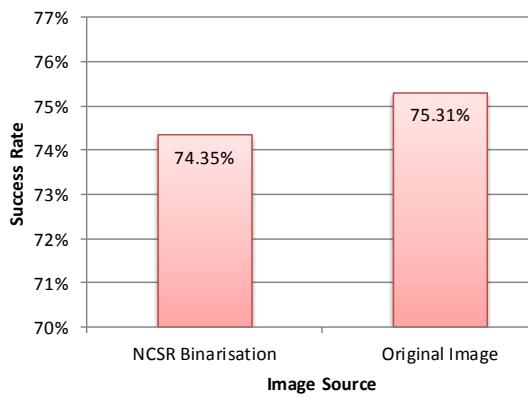
## Layout Analysis Performance

### Per evaluation profile



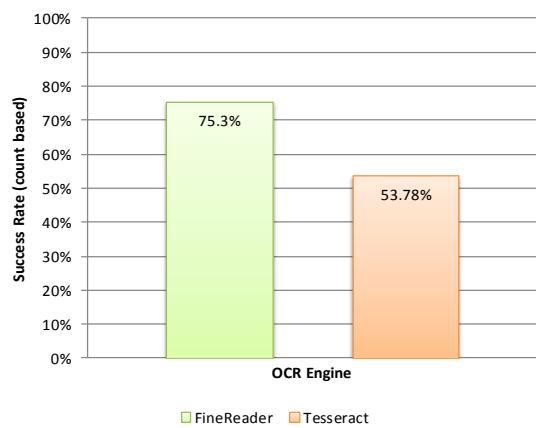
## Bag of Words OCR Evaluation

### Binarised image vs. original image



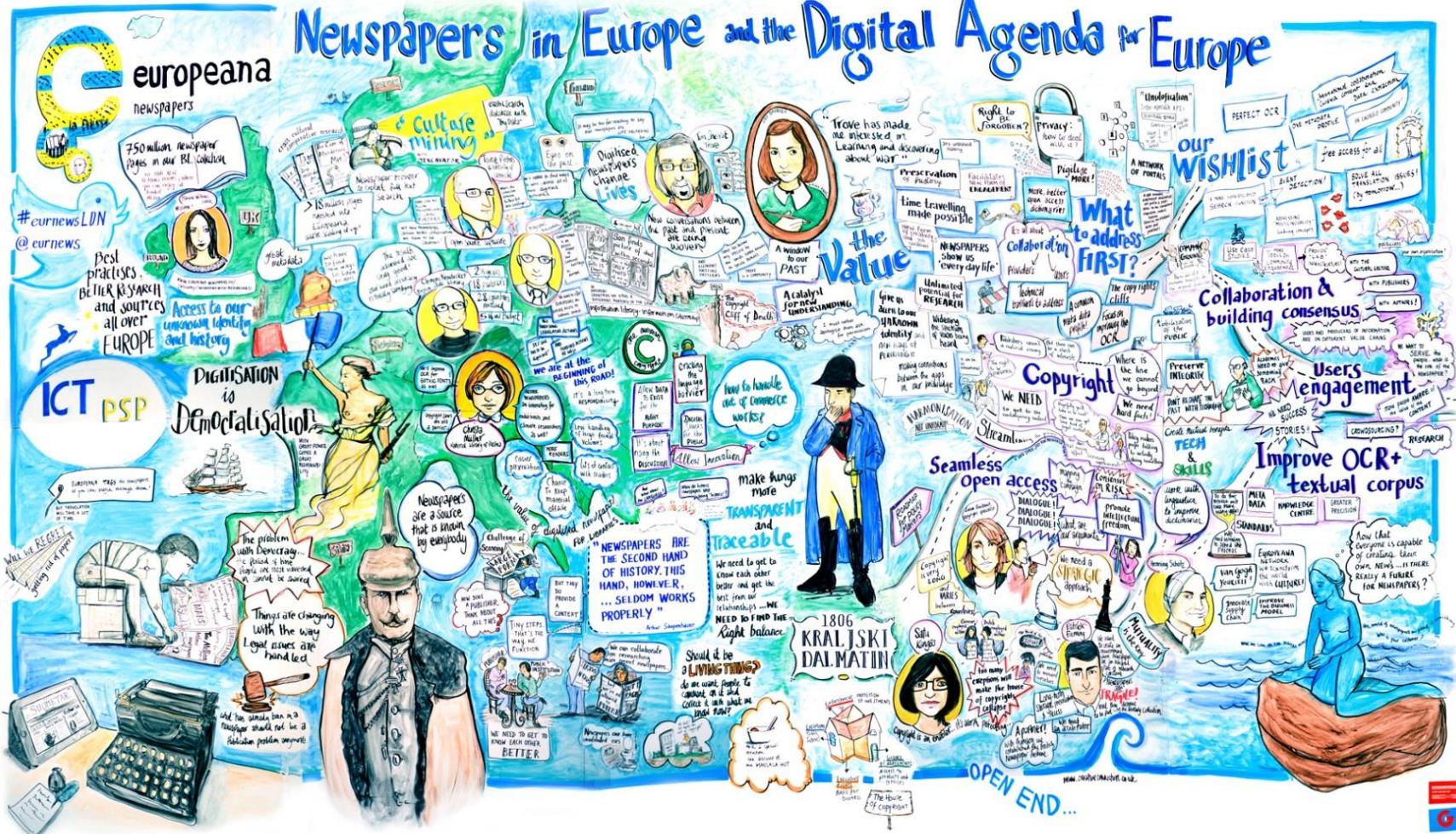
## Bag of Words OCR Evaluation

### FineReader vs. Tesseract





# Newspapers in Europe and the Digital Agenda for Europe



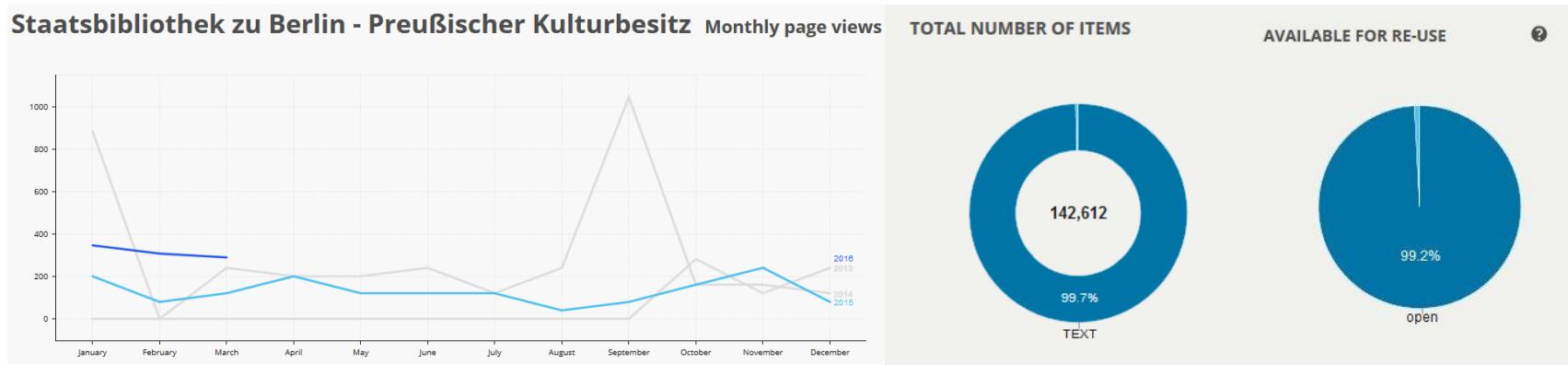
<http://europeananewspapers.github.io>

# 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!



- 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](https://twitter.com/cneudecker)



**Staatsbibliothek**  
**zu Berlin**  
Preußischer Kulturbesitz

