



# “Die Grenzboten” on its way to virtual research environments and infrastructures

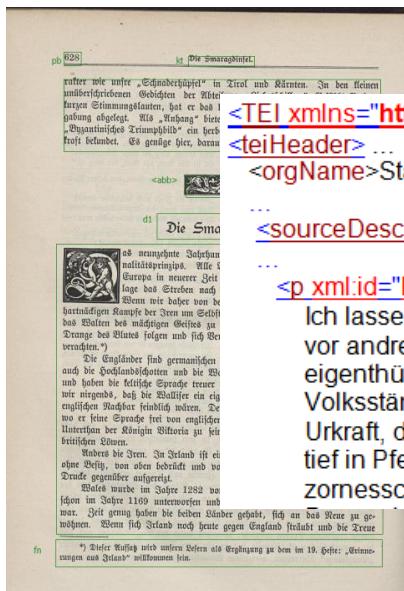
Digitisation  
Project



OCR  
Post Processing



Research  
Infrastructures



```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
<teiHeader> ...
<orgName>Staats- und Universitätsbibliothek Bremen
...
<sourceDesc><bibl type="J">Die Grenzboten.
...
<p xml:id="ID_8" next="#ID_9">
```

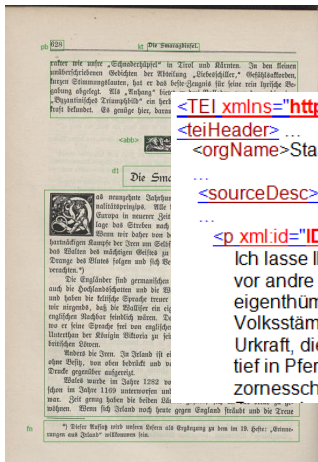
Ich lasse Ihnen das Glauben und Wissen c  
vor andre Richtersthühle gehörig hier zur Se  
eigenthümlichen Stellung dieser drei Männ  
Volksstämmen. Strauß, der Schwabe mit s  
Urkraft, die mit einem Hiebe durch Reiter u  
tief in Pferdes Rücken schlug; Ludwig Feu  
zornesschnaubend und gewandt auch das





# Outline

- The digitisation project – *Die Grenzboten*
- OCR post correction
- Creation of a TEI version
- Working with full texts
  - Digitisation portals
  - Virtual research environments and research infrastructures
- Discussion



```

<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader> ...
  <orgName>Staats- und Universitätsbibliothek (S
  ...
  <sourceDesc><bibl type="J">Die Grenzboten.
  ...
  <p xml:id="ID_8" next="#ID_9">

```

Ich lasse Ihnen das Glauben und Wissen c  
vor andre Richterstühle gehörig hier zur Se  
eigenthümlichen Stellung dieser drei Männ  
Volksstämmen. Strauß, der Schwabe mit s  
Urkraft, die mit einem Hiebe durch Reiter u  
tief in Pferdes Rücken schlug; Ludwig Feu  
zornesschnaubend und gewandt auch das



## The digitisation of the periodical *Die Grenzboten* and the post correction of the OCRRed full texts



## Digitisation Project „Die Grenzboten“

- Historical journal; appeared from 1841 to 1922
  - Subtitle as of 1871: "Zeitschrift für Politik, Literatur und Kunst"
- Founded by Ignaz Kuranda (1811-1884) in Brussels; later on published in Leipzig and Berlin
- Literary and political journal with a national liberal orientation

Studien zur englischen Literatur

1890

Humor und Komik in der griechischen Kunst

Der Kampf um die Armee in Österreich-Ungarn

Die Griechen

1903

Der Vertrag über den Pa...

Ein Wort über das deutsche Theater.

(Eine Vorlesung gehalten im Museum zu Frankfurt.)

1880.

Eine

Das Atelier eines französischen Bildhauers.

atistische Satyre.

1841.

## Digitisation Project „Die Grenzboten“

- Funded by the German Research Foundation (DFG)
- Project duration
  - 18 month – digitisation project
  - 12 month – OCR post correction + TEI version
- Quantity structure
  - 81 years in 270 volumes
  - approx. 187.000 physical pages
  - approx. 33.000 sections
  - approx. 50.000 structure nodes



- Website: <http://brema.suub.uni-bremen.de/grenzboten>



Search Details

Sammlung Grenzboten



Ihre Suche: z.B. Titel, Thema, Person



Title

Search the Journal

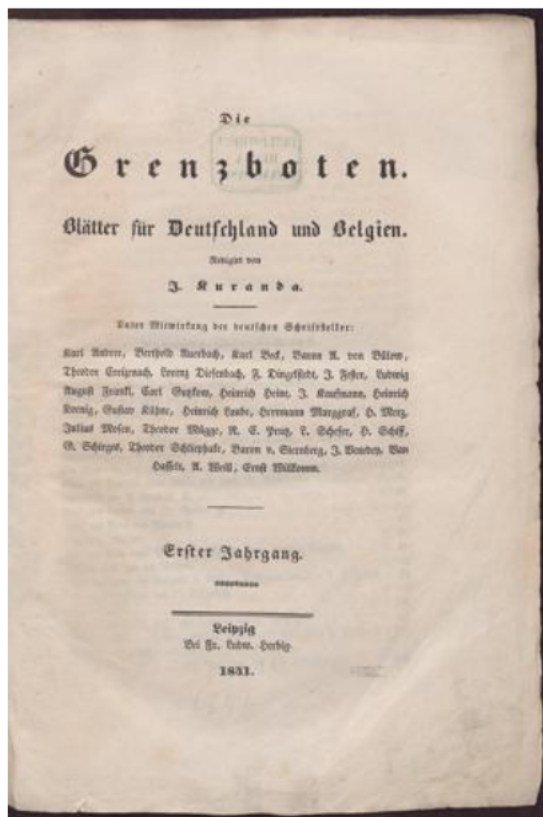
OK

Journal

Title	Die Grenzboten : Zeitschrift für Politik, Literatur und Kunst
Published	Berlin : Dt. Verl, 1841 - 1922
Annotation	Jg. 2.1842 nachträglich als Jg. 1 bez.; 2.1843 auf d. Sem.-Titelbl. fälschlich mit Ersch.-Jahr 1842 angegeben
Electronic Edition	Bremen : Staats- und Universitätsbibliothek, 2012 Berlin : Staatsbibliothek zu Berlin, 1997 - 1999
URN	urn:nbn:de:gbv:46:1-908

Volumes

1 (1841)	29 (1870)	58 (1899)
2 (1842)	30 (1871)	59 (1900)
2 (1843)	31 (1872)	60 (1901)
3 (1844)	32 (1873)	61 (1902)
4 (1845)	33 (1874)	62 (1903)
5 (1846)	34 (1875)	63 (1904)
6 (1847)	35 (1876)	64 (1905)
7 (1848)	36 (1877)	65 (1906)
8 (1849)	37 (1878)	66 (1907)
9 (1850)	38 (1879)	67 (1908)
10 (1851)	39 (1880)	68 (1909)





Ihre Suche: z.B. Titel, Thema, Person ...

1 - 18

Die Grenzboten > 1850 > 9. Jahrgang. I. Semester. I. B > Vorderdeckel > [1]



[1]



[2]



[3]



[4]



[5]



[6]



[7]



[8]



[9] 1



[10] 2



[11] 3



[12] 4



[13] 5



[14] 6



[15] 7



[16] 8



[17] 9



[18] 10



Staats- und  
Universitätsbibliothek  
Bremen

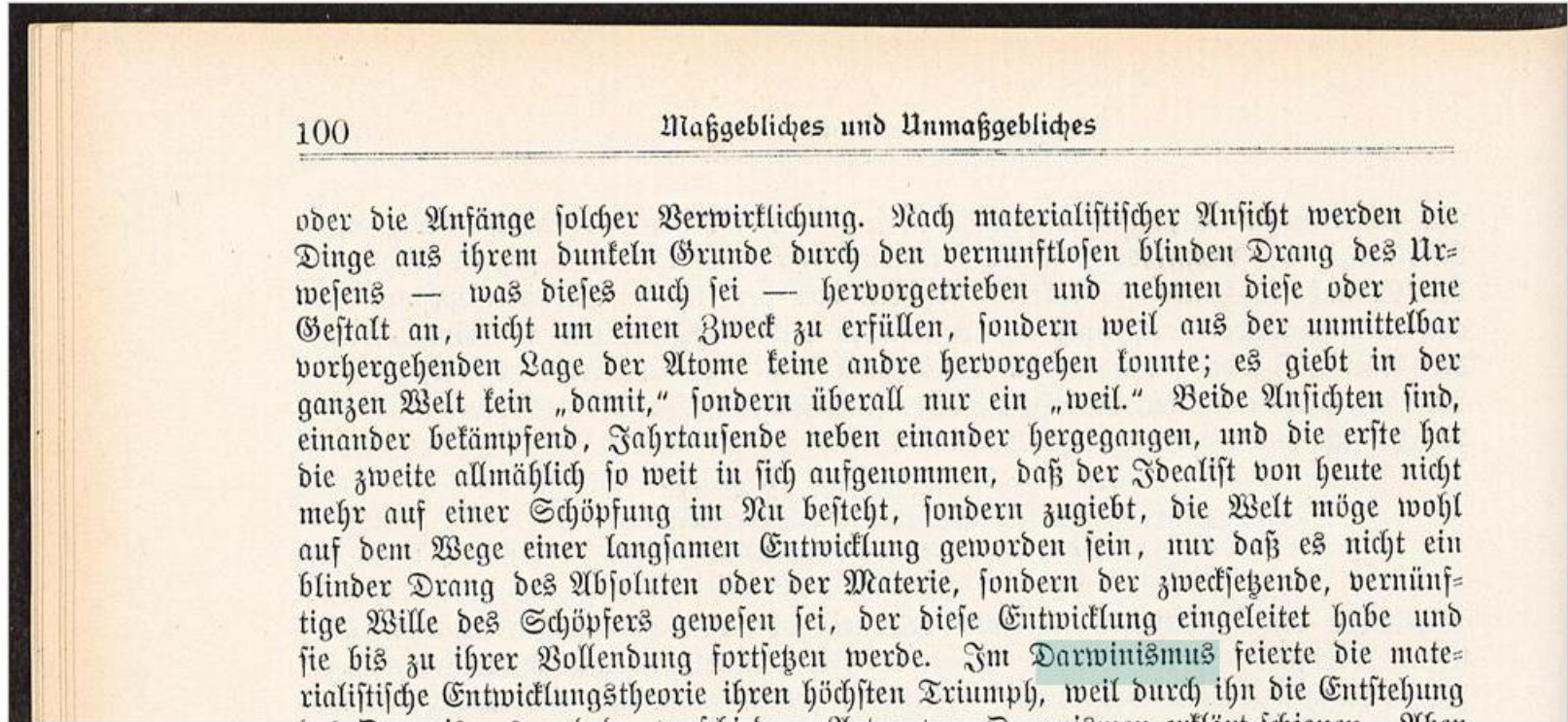
[Über uns](#) [Service & Beratung](#) [Kataloge & Sammlungen](#) [Fachinformationen](#) [Standorte](#) [Kontakt](#) [Literatur verwalten & publizieren](#)

Home → [Kataloge & Sammlungen](#) → [Digitale Sammlungen](#) → [Sammlung Grenzboten](#)

aaa

[Titel](#) [Inhalt](#) [Übersicht](#) [Seite](#) [Vollbild](#) [108] 100

Die Grenzboten > 1894 > 53. Jahrgang, Erstes Viertelja > Maßgebliches und Unmaßgebliche > [108] 100



100

Maßgebliches und Unmaßgebliches

oder die Anfänge solcher Verwirklichung. Nach materialistischer Ansicht werden die Dinge aus ihrem dunkeln Grunde durch den vernunftlosen blinden Drang des Urwesens — was dieses auch sei — hervorgetrieben und nehmen diese oder jene Gestalt an, nicht um einen Zweck zu erfüllen, sondern weil aus der unmittelbar vorhergehenden Lage der Atome keine andre hervorgehen konnte; es giebt in der ganzen Welt kein „damit,“ sondern überall nur ein „weil.“ Beide Ansichten sind, einander bekämpfend, Jahrtausende neben einander hergegangen, und die erste hat die zweite allmählich so weit in sich aufgenommen, daß der Idealist von heute nicht mehr auf einer Schöpfung im Nu besteht, sondern zugiebt, die Welt möge wohl auf dem Wege einer langsamen Entwicklung geworden sein, nur daß es nicht ein blinder Drang des Absoluten oder der Materie, sondern der zwecksetzende, vernünftige Wille des Schöpfers gewesen sei, der diese Entwicklung eingeleitet habe und sie bis zu ihrer Vollendung fortsetzen werde. Im Darwinismus feierte die materialistische Entwicklungstheorie ihren höchsten Triumph, weil durch ihn die Entstehung



## „Die Grenzboten“ – Typeset Quality

- Different types of gothic / black letter font



- Different languages and font sizes

esten (τηλόθ' εοῦσαν V,  
rückfahrt). Als später He  
ien, findet er die Nympl

Fern in das Giland  
üft, der gespaltenen, wallte des  
ng mit melodischer Stimm' in



# „Die Grenzboten“ - Quality of Pages

- Quality assurance using the OCR full text



advertisements  
foreign languages

Je crois, Monsieur  
borateur; et pour vou  
de bonne grâce que r  
incontestables de l'All  
contenter de ce qui  
trême fécondité et la  
vains), surtout notre  
à bien des méprises.

foreign  
languages

Arbeitsmarktstatistik.  
Kaiserliche Statistische Amt  
August gegen die Monate  
hauptsächlichen Industrien  
der Saison (Buchdruckergew  
haben aber eine hervortreten  
den Krankenkassen dem An  
August wie im Juli eine weit  
getreten sei. Die Zunahme s  
bei den weiblichen Kassengli  
nicht gezogen werden könn  
gliederzahl von 3,7 Millionen.  
dafür als Maßstab gelten könn  
des allgemeinen Beschäftigungs

Antiqua

da güttheißen,  
Belgien gegen  
ffen Holland.  
Rheinarm verse  
Dwurbeer ihn

bad  
type face



damaged  
pages



translucent  
pages



low contrast



tables

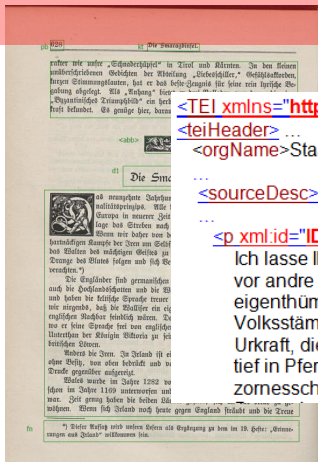
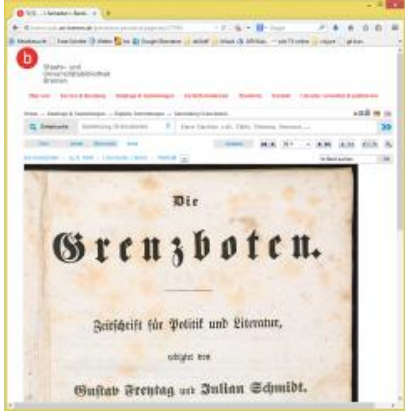


small  
font size



# Outline

- The digitisation project – *Die Grenzboten*
- OCR post correction
- Creation of a TEI version
- Working with full texts
  - Digitisation portals
  - Virtual research environments and research infrastructures
- Discussion



```

<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader> ...
  <orgName>Staats- und Universitätsbibliothek (S
  <sourceDesc><bibl type="J">Die Grenzboten.
  <p xml:id="ID_8" next="#ID_9">

```

Ich lasse Ihnen das Glauben und Wissen c  
vor andre Richterstühle gehörig hier zur Se  
eigenthümlichen Stellung dieser drei Männ  
Volksstämmen. Strauß, der Schwabe mit s  
Urkraft, die mit einem Hiebe durch Reiter u  
tief in Pferdes Rücken schlug; Ludwig Feu  
zornesschnaubend und gewandt auch das

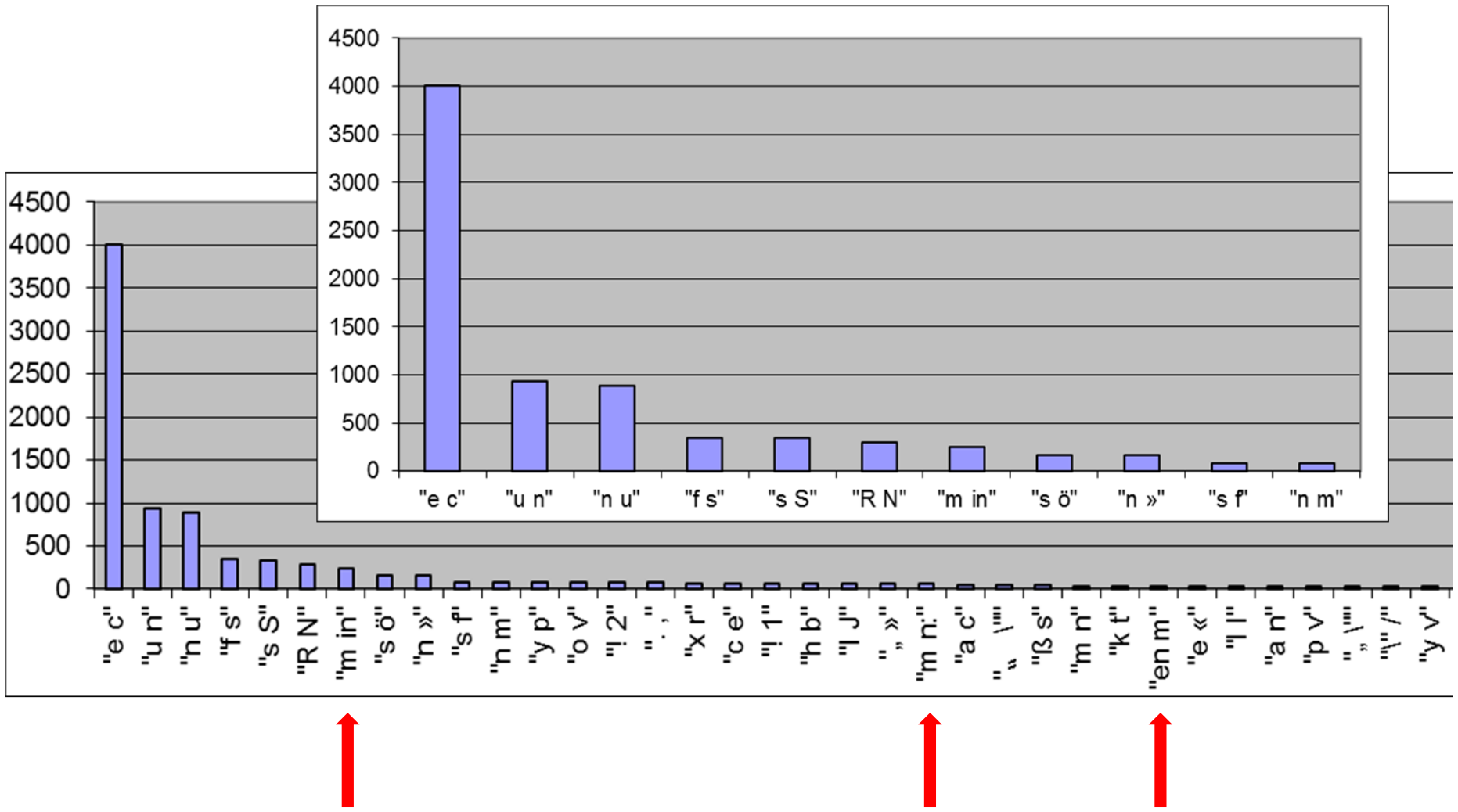


## „Die Grenzboten“ – OCR Post Correction

- Objective: Improvement of the full text quality with respect to the
  - rate of OCR errors (recognition rate: 98,28% → 99,5%)
    - Development of a fully automatic post correction algorithm at the SuUB Bremen
  - structural presentation of the pages  
(a prerequisite for a **TEI** mark-up)
- Cooperation with the German Text Archive, the IMPACT CoC, Wikisource and the Australian company ProjectComputing



# „Die Grenzboten“ – OCR Post Correction





# „Die Grenzboten“ – List of Historical Word Forms

Frequency	Surface Form	Transliteration into the subset of ISO-8859-1 (Latin-1)	Modern Form
278187	und	und	und
239487	der	der	der
233389	die	die	die
...			
28268	ift	ist	ist
14885	fo	so	so
...			
600	Aehnlichkeit	Aehnlichkeit	Ähnlichkeit
322	Säugethiere	Säugethiere	Säugetiere
319	theilt	theilt	teilt
...			
6	ältlicher	ältlicher	ältlicher
1	ältlicher	ältlicher	ältlicher

## „Die Grenzboten“ – OCR Post Correction

- Correction examples

Print Image	Ground Truth	OCR-Text	Correction
	Aber	Ader	Ader
	dem	dein	dein
	Zeitschrift	Zeitschriſt	Zeitschrift
	gewonnen	gcwouneu	gewonnen
	Entwicklung	Eutwickclimg	Entwicklung
	Herz der begeisterten	Herz der begeisterten	Herz der begeisterten
	Sein Wille erfüllte	Sciu Wille ersülte	sein Wille erfüllte

## OCR Text

b

Aber noch ein zweiter Grund bewegt uns bei unserem Unternehmen, nes ist dieses der Boden selbst aus dem diese Blätter hervorwachsen sollen: Belgien! Als wir den Titel dieser Zeitschrift, die Bezeichnung: "Blätter für Deutschland und Belgien" hinzusetzten, so verhehlten wir uns nicht, daß wir gegen ein gewisses Vorurtheil zu kämpfen haben werden. So poetisch und Interesse erregend der Name Niederland dem Deutschen klingt, so fremdartig und unsicher scheint ihm der Name Belgien. An das Wort Niederland knüpfen sich gar theure Erinnerungen der deutschen Geschichte. Der deutsche Religionszwiespalt hat da seine heißesten Kämpfer gefunden, die deutsche Wissenschaft hat da ihre Grundstümpfe (Erasmus, Justus Lipsius, Grotius, Spinoza, Vesal u. s. w.) gewonnen, die deutsche Kunst hat da ihre kräftigste Ammenmilch gesogen, und die deutsche Poesie hat daher auch diesen Namen zu ihren Lieblingskindern erhoben und Schiller und Göthe haben ihn ins Herz der begeisterten Jugend gelegt, die für Egmont und Posa schwärmt. Der Name Belgien aber - so uralte das Wort auch ist - steht doch anders zu jung und zu fremdartig dem Deutschen gegenüber, um ihm populär zu sein. Wir brauchen nicht erst auf die Ereignisse von 1830 hinzuweisen. Es ist leicht begreiflich, daß Deutschland die Trennung der südlichen Niederlande von den nördlichen mit Unmuth betrachtete, daß es den Kopf schüttelte, da es die germanischen Elemente den gallischen weichen sah. Sein Interesse wendete sich seitdem mit ziemlicher Kälte von Belgien weg, und wenn die politischen Ereignisse es nicht zur Aufmerksamkeit nöthigten, wenn nicht Belgien selbst, durch die glänzende Thätigkeit seiner Eisenwerke ihm die Beachtung abzwang, da blieb es mißmüthig mit dem Rücken ihm zugekehrt. Und wahrlich, es ist nicht gut, daß es so gekommen ist. Belgien hat in diesen zehn Jahren einen riesenhaften Fortschritt gethan, und Deutschland hätte mit mehr Aufmerksamkeit auf die Entwicklung dieses Landes in Kunst und Gewerbe, in socialer und sogar in politischer Beziehung, manche schöne Erfahrung erwerben können. Es ist ein gewöhnlicher Fehler, daß man die französische Revolution von 1830 mit der gleichzeitigen belgischen zusammenkettet, ohne zu betrachten, wie die Folgen beider ganz verschieden sind. Frankreich zielte im Jahr 1830 nach einer Republik und gelangte nur bis zu einer Veränderung der Dynastie. Sein Wille erfüllte sich nur halb, und die andere Hälfte blieb als ein klaffender Riß, als eine eiternde Wunde, welche an dem gesunden Theile des Staates zehrt und ihn nicht zur Ruhe und gesunden Entwicklung kommen läßt. Dieß ist keinesweges mit Belgien der Fall; die Revolution von 1830 zielte hier nur nach einer Loslösung von dem holländischen Mitstaate; sobald dieses geglückt war, und die

## Post Corrected Text

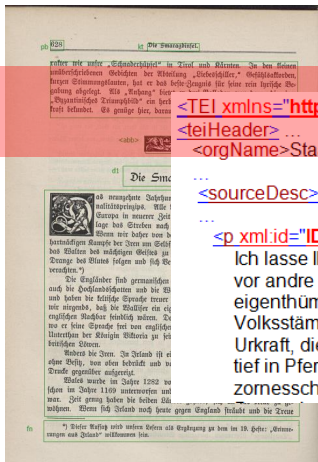
Aber noch ein zweiter Grund bewegt uns bei unserem Unternehmen, nes ist dieses der Boden selbst aus dem diese Blätter hervorwachsen sollen: Belgien! Als wir den Titel dieser Zeitschrift, die Bezeichnung: "Blätter für Deutschland und Belgien" hinzusetzten, so verhehlten wir uns nicht, daß wir gegen ein gewisses Vorurtheil zu kämpfen haben werden. So poetisch und Interesse erregend der Name Niederland dem Deutschen klingt, so fremdartig und unsicher scheint ihm der Name Belgien. An das Wort Niederland knüpfen sich gar theure Erinnerungen der deutschen Geschichte. Der deutsche Religionszwiespalt hat da seine heißesten Kämpfer gefunden, die deutsche Wissenschaft hat da ihre Grundstümpfe (Erasmus, Justus Lipsius, Grotius, Spinoza, Vesal u. s. w.) gewonnen, die deutsche Kunst hat da ihre kräftigste Ammenmilch gesogen, und die deutsche Poesie hat daher auch diesen Namen zu ihren Lieblingskindern erhoben und Schiller und Göthe haben ihn ins Herz der begeisterten Jugend gelegt, die für Egmont und Posa schwärmt. Der Name Belgien aber - so uralte das Wort auch ist - steht doch andererseits zu jung und zu fremdartig dem Deutschen gegenüber, um ihm populär zu sein. Wir brauchen nicht erst auf die Ereignisse von 1830 hinzuweisen. Es ist leicht begreiflich, daß Deutschland die Trennung der südlichen Niederlande von den nördlichen mit Unmuth betrachtete, daß es den Kopf schüttelte, da es die germanischen Elemente den gallischen weichen sah. Sein Interesse wendete sich seitdem mit ziemlicher Kälte von Belgien weg, und wenn die politischen Ereignisse es nicht zur Aufmerksamkeit nöthigten, wenn nicht Belgien selbst, durch die glänzende Thätigkeit seiner Eisenwerke ihm die Beachtung abzwang, da blieb es mißmüthig mit dem Rücken ihm zugekehrt. Und wahrlich, es ist nicht gut, daß es so gekommen ist. Belgien hat in diesen zehn Jahren einen riesenhaften Fortschritt gethan, und Deutschland hätte mit mehr Aufmerksamkeit auf die Entwicklung dieses Landes in Kunst und Gewerbe, in socialer und sogar in politischer Beziehung, manche schöne Erfahrung erwerben können. Es ist ein gewöhnlicher Fehler, daß man die französische Revolution von 1830 mit der gleichzeitigen belgischen zusammenkettet, ohne zu betrachten, wie die Folgen beider ganz verschieden sind. Frankreich zielte im Jahr 1830 nach einer Republik und gelangte nur bis zu einer Veränderung der Dynastie. Sein Wille erfüllte sich nur halb, und die andere Hälfte blieb als ein klaffender Riß, als eine eiternde Wunde, welche an dem gesunden Theile des Staates zehrt und ihn nicht zur Ruhe und gesunden Entwicklung kommen läßt. Dieß ist keinesweges mit Belgien der Fall; die Revolution von 1830 zielte hier nur nach einer Loslösung von dem holländischen Mitstaate; sobald dieses geglückt war, und die





# Outline

- The digitisation project – *Die Grenzboten*
- OCR post correction
- Creation of a TEI version
- Working with full texts
  - Digitisation portals
  - Virtual research environments and research infrastructures
- Discussion



```

<TEI xmlns="http://www.tei-c.org/ns/1.0">
<teiHeader> ...
<orgName>Staats- und Universitätsbibliothek (S
...
<sourceDesc><bibl type="J">Die Grenzboten.
...
<p xml:id="ID_8" next="#ID_9">

```

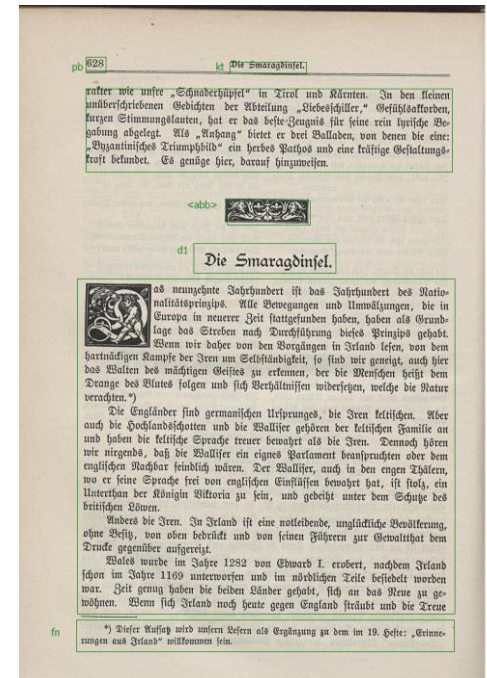
Ich lasse Ihnen das Glauben und Wissen c  
vor andre Richterstühle gehörig hier zur Se  
eigenthümlichen Stellung dieser drei Männ  
Volksstämmen. Strauß, der Schwabe mit s  
Urkraft, die mit einem Hiebe durch Reiter u  
tief in Pferdes Rücken schlug; Ludwig Feu  
zornesschnaubend und gewandt auch das



## „Zoning“ – Adding structural annotation to the full text

- ... page-wise image-clipping at the level of paragraphs
- The semantical function was annotated  
→ it is apparent what parts of the text are titles, running text or footnotes
- Guideline for the annotation: The DTA “base format” (DTABf) according to the TEI-P5 standard
- We used ZOT, a point-and-click image clipping tool developed by the DTA project

<http://www.deutschestextarchiv.de/doku/software#ZOT>



rakter wie unsere „Schnaderhüpfel“ in Tirol und Kärnten. In den kleinen unüberschriebenen Gedichten der Abteilung „Liebeschiller,“ Gefühlstakorden, kurzen Stimmungslauten, hat er das beste Zeugnis für seine rein lyrische Begabung abgelegt. Als „Anhang“ bietet er drei Balladen, von denen die eine: „Byzantinisches Triumphbild“ ein herbes Pathos und eine kräftige Gestaltungskraft bekundet. Es genüge hier, darauf hinzuweisen.

&lt;abb&gt;



d1

## Die Smaragdinsel.



Das neunzehnte Jahrhundert ist das Jahrhundert des Rationalitätsprinzips. Alle Bewegungen und Umwälzungen, die in Europa in neuerer Zeit stattgefunden haben, haben als Grundlage das Streben nach Durchführung dieses Prinzips gehabt. Wenn wir daher von den Vorgängen in Irland lesen, von dem hartnäckigen Kampfe der Iren um Selbständigkeit, so sind wir geneigt, auch hier das Walten des mächtigen Geistes zu erkennen, der die Menschen heißt dem Drange des Blutes folgen und sich Verhältnissen widersetzen, welche die Natur verachten.\*)

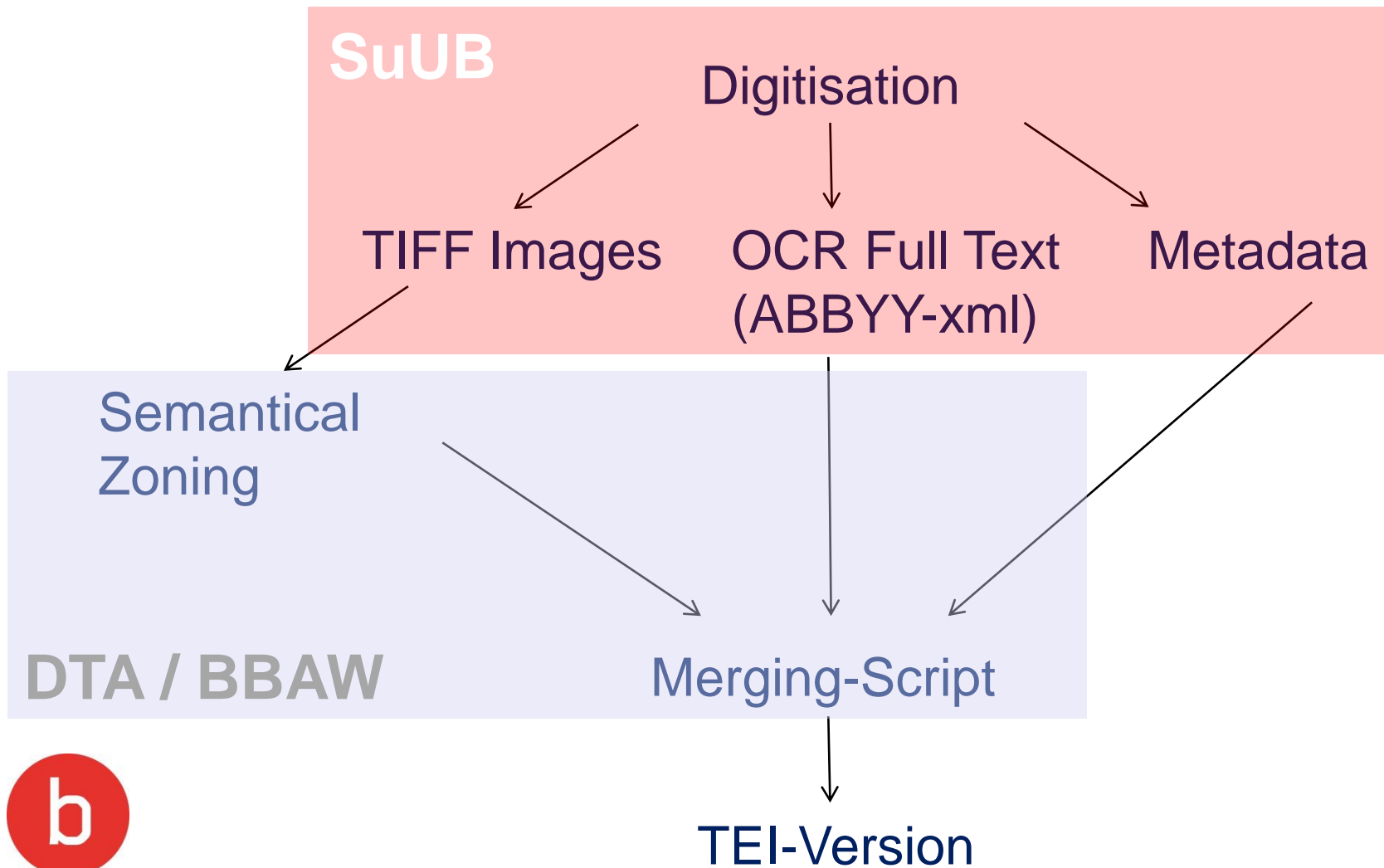
Die Engländer sind germanischen Ursprunges, die Iren keltischen. Aber auch die Hochlandsschotten und die Walliser gehören der keltischen Familie an und haben die keltische Sprache treuer bewahrt als die Iren. Dennoch hören wir nirgends, daß die Walliser ein eignes Parlament beanspruchten oder dem englischen Nachbar feindlich wären. Der Walliser, auch in den engen Thälern, wo er seine Sprache frei von englischen Einflüssen bewahrt hat, ist stolz, ein Unterthan der Königin Viktoria zu sein, und gedeiht unter dem Schutze des britischen Löwen.

Anders die Iren. In Irland ist eine notleidende, unglückliche Bevölkerung, ohne Besitz, von oben bedrückt und von seinen Führern zur Gewaltthat dem Drucke gegenüber aufgereizt.

Wales wurde im Jahre 1282 von Edward I. erobert, nachdem Irland schon im Jahre 1169 unterworfen und im nördlichen Teile besiedelt worden war. Zeit genug haben die beiden Länder gehabt, sich an das Neue zu gewöhnen. Wenn sich Irland noch heute gegen England sträubt und die Treue

\*) Dieser Aufsatz wird unsern Lesern als Ergänzung zu dem im 19. Hefte: „Erinnerungen aus Irland“ willkommen sein.

# Creation of a TEI Version – Process Chain





# Summary – OCR Post Correction and Creation of a TEI Version

- Development of a fully automatic algorithm based on a set of 1,7 million historical word forms
- We eliminated nearly one-third of all character errors achieving a new recognition rate of 98,83%. Reasons for this are:
  - Erroneous and non-complete list of historical words with a hypothetical error-free and complete word list as a theoretical upper performance limit, which is 99,83%, and we were not able to take into account the complex orthographic variations
- In cooperation with the Australian company Project Computing, the service OverProof was enhanced by the ability to correct Fraktur typeface texts
- Results
  - The digital Grenzboten: <http://brema.suub.uni-bremen.de/grenzboten/>
  - Source Codes and primary data were published on GitHub: <https://github.com/suub>
  - All volumes have been converted into a TEI-Version
  - Publication: <https://www.o-bib.de/article/view/2016H1S32-55>

Aufsätze

Aufsätze

Automatische Qualitätsverbesserung von Fraktur-Volltexten aus der Retrodigitalisierung am Beispiel der Zeitschrift Die Grenzboten

Manfred Nölte, Staats- und Universitätsbibliothek Bremen  
Jan-Paul Bultmann, Universität Bremen  
Maik Schünemann, Universität Bremen  
Martin Blenkle, Staats- und Universitätsbibliothek Bremen

Zusammenfassung:

Den Geisteswissenschaften stehen nach und nach mehr computerbasierte Werkzeuge und Infrastrukturen der Digital Humanities zur Verfügung, für die die Existenz und weitere Erstellung von Volltext mit guter Qualität eine unabdingbare Voraussetzung ist. Der Bedarf nach qualitativ hochwertigem Volltext aus Retrodigitalisierungsprojekten steigt daher ständig an. Der zu Frakturschrift berechnete OCR-Volltext hat eine deutlich schlechtere Qualität als von Antiqua-Schrift berechneter. Daher ist für das wissenschaftliche Arbeiten unkorriger und unstrukturierter OCR-Volltext von mehreren Millionen Seiten, Da eine bedarfsgerechte Erzeugung von Volltext in der Größenordnung von mehreren Millionen Seiten in Bezug auf Aufwand und Kosten effizient sein sollte, wird hier eine möglichst weitgehende Automatisierung der Nachbearbeitung von OCR-Volltext vorgestellt. An der Staats- und Universitätsbibliothek Bremen (SuUB) wurde dazu ein Ansatz entwickelt, der sich durch Einfachheit auszeichnet: Eine Liste historischer bzw. dialekt- oder fachspezifischer Wortformen – eine der Voraussetzungen dieses Ansatzes – ist verhältnismäßig leicht erstellbar. Ein effizienter Algorithmus leistet den Abgleich von hier ca. 1,7 Millionen Wortformen gegen bei der Zeitschrift Die Grenzboten knapp 80 Millionen enthaltenen Wörtern und lässt sich auf verständliche und nachvollziehbare Art und Weise parametrisieren, d.h. auf die spezifischen Eigenschaften des jeweiligen Volltextprojektes einstellen. Die erreichbaren Ergebnisse sind stark abhängig von der Ausgangsqualität des Volltextes sowie von dem Umfang und der Qualität der Liste der historischen Wortformen und dem verwendeten Fehlermodell. So können beispielsweise bestimmte Fehler nur mit einem den Kontext berücksichtigenden Ansatz korrigiert werden. Weiterhin wurde zusammen mit der Firma Project Computing mit Sitz in Canberra, Australien, der cloud service overProof<sup>1</sup> um die Funktionalität der Nachkorrektur deutschsprachiger Frakturschrift erweitert. In einem Ausblick werden Bedarfe und Möglichkeiten für die Zukunft aufgezeigt.

Summary:

Gradually, the humanities are provided with a number of computer based tools and scientific infrastructures of the digital humanities. As digital full text is strongly needed for these tools and infrastructures, the demand for high-quality full texts is constantly rising. OCRed full text from Gothic typeface texts is of considerably worse quality than OCRed full text from Antiqua. The value

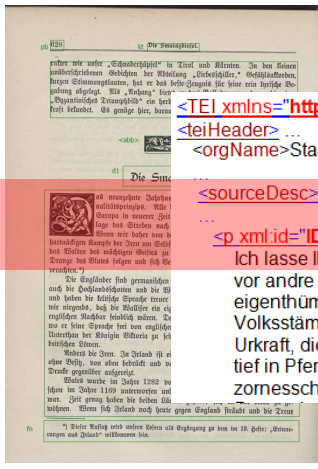
<sup>1</sup> OverProof, zuletzt geprüft am 01.02.2016, <http://overproof.projectcomputing.com/>.





# Outline

- The digitisation project – *Die Grenzboten*
- OCR post correction
- Creation of a TEI version
- Working with full texts
  - Digitisation portals
  - Virtual research environments and research infrastructures
- Discussion



```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
<teiHeader> ...
<orgName>Staats- und Universitätsbibliothek (S
```

```
<sourceDesc><bibl type="J">Die Grenzboten. ...
```

```
<p xml:id="ID_8" next="#ID_9">
Ich lasse Ihnen das Glauben und Wissen c
vor andre Richterstühle gehörig hier zur Se
eigenthümlichen Stellung dieser drei Männ
Volksstämmen. Strauß, der Schwabe mit s
Urkraft, die mit einem Hiebe durch Reiter u
tief in Pferdes Rücken schlug; Ludwig Feu
zornesschnaubend und gewandt auch das
```



## How to get access to the full texts at the SuUB?

- Search feature
  - allows finding different forms of declension of nouns  
Lebensbild → Lebensbildes, Lebensbildern, ...
  - Sample search:  
<http://brema.suub.uni-bremen.de/grenzboten/search/quick?query=Lebensbild>





Detailsuche

Sammlung Grenzboten



Lebensbild



## Suchergebnis

Suche nach: [Lebensbild](#)



### Die Grenzboten

Zeitschrift für Politik, Literatur und Kunst

Berlin : Dt. Verl, 1841 - 1922

576 Treffer im Volltext: [\[Seite 7\]](#), [\[Seite 8\]](#), [339](#), [387](#), [382](#), [517](#), [155](#), [121](#), [226](#), [105](#)

[Übersicht](#) [\[Ausblenden\]](#)

- Jg. 3, 1844
    - II. Semester. I. Band.
      - Inhaltsverzeichnis
        - [\[Seite 7\]](#) ...S. 235. Die Familie Mantini. Ein [Lebensbild](#) aus der neuesten Zeit. Von Frau von... ..S. 291. — Die Familie Mantini. Ein [Lebensbild](#) aus der neuesten Zeit. Von Frau von...
        - [\[Seite 8\]](#) ...[Lebensbilder](#) aus Berlin- Von Albert Fränkel. Ein... ..Hamburg. S. 383\_\_Notizen. S. 385. [Lebensbilder](#) aus Berlin. Bon Albert Fränkel. Ein...
      - Fränkel, Albert: [Lebensbilder](#) aus Berlin : ein Hôtel garni : erste Abtheilung.
        - [339](#) ...[Lebensbilder](#) aus Berlin. Von Albert Fränkel. Ein...
      - Fränkel, Albert: [Lebensbilder](#) aus Berlin : ein Hôtel garni : zweite Abtheilung.
        - [387](#) ...[Lebensbilder](#) aus Berlin. Von Albert Fränkel. Ein...
- Jg. 4, 1845
  - I. Semester.

## 25 Titel

### Sortieren

Relevanz

### Autoren / Beteiligte

Speck, Wilhelm	13
Fränkel, Albert	2
Marbach, Hans	2
Boas, Eduard	1
Gerstäcker, Fr.	1

### Zeiträume

1801-1900	11
1901-2000	14

### Dokumenttypen

Beitrag	24
Zeitschrift	1



1 (1841)

Lebensbild  OK  
Im Jahrgang suchen

Ein Wort über das deutsche Theater : (eine Vorlesung gehalten im Museum zu Frankfurt.) → 49

49

ein warnblütiges Lebensbild in kräftigen Zügen vor uns zu sehen? Freilich, wenn das Lustspiel in solcher Bahn fortschreiten soll, muß zuerst noch manche ängstliche politische Rücksicht sich lösen, die auch dem dramatischen Dichter manche freiere Bewegung des Geistes hemmt, und in dieser Hin-

Deutsche Preßsünden. → 285

6) in Hinsicht auf den Styl — im weitesten Sinne genommen — keine poetische Schönheit im Gedanken des Dichters zu unterdrücken,\*) höchstens die für ein Lebensbild nicht geeignete Ueppigkeit der Lyrik zu beschneiden, jeden U) Pathos und Wortschwall zu vermeiden; dagegen einfach und wahr jeden Aufschwung der Gefühle und jede gegebene Gedankenrichtung der Dichtung walten zu lassen. Unvermeidlich ist

2 (1842)

Lebensbild  OK  
Im Jahrgang suchen

Inhaltsverzeichnis → [Seite 9]

Die Pariser Kunstausstellung von 1842. . . . .	419.
Tagebuch: Brief aus London. — Flandereien. . . . .	428.
Der moderne Nomade, ein Pariser Lebensbild, von Philipp H. . . . .	437.
Flandrische Literatur. Der Löwe von Flandern v. H. Conscience. . . . .	443.
Tagebuch: Briefe: aus Leipzig; — aus Brüssel; — aus Teplitz. . . . .	461.
Mozarts erste Reise nach Paris, von F. F. . . . .	469.
Französisch u. Deutsch. Vorlesungen v. Kuranda. 2. Die Literaturen. . . . .	482.

Der moderne Nomade : ein Pariser Lebensbild. → 437

**Der moderne Nomade.**  
Ein Pariser Lebensbild  
von  
Philipp H.\*\*\*

Gehe ich wohl heute Abend aus? Da man nicht auf den Kopf...

Tagebuch. → 201

aus dem Staatsdienste entlassen. Ich kann Ihnen die vollständige Versicherung geben, daß jedes Wort dieser Nachricht eine Lüge ist.  
Die Lebensbilder aus dem Befreiungskriege und die Bemerkungen über den Fürsten Metternich und den seligen Kaiser Franz, die sich darin befinden, haben hier allerdings großen Spektakel gemacht — nichts desto weniger ist von Seiten unseres Cabinets auch nicht mit einer Note Erwähnung hierüber gesehen. Wenn Ihnen der Name \*\*\* eine hinreichende Bürgschaft ist, so

## How to get access to the full texts at the SuUB?

- Search feature
  - allows finding different forms of declension of nouns  
Lebensbild → Lebensbildes, Lebensbildern, ...
  - Sample search:  
<http://brema.suub.uni-bremen.de/grenzboten/search/quick?query=Lebensbild>
- PDF downloads
- Special URLs for an ABBYY-xml version  
<http://brema.suub.uni-bremen.de/periodical/pageview/179392> →  
<http://brema.suub.uni-bremen.de/download/fulltext/fr/179392>
- Special URLs for an ALTO version  
<http://brema.suub.uni-bremen.de/grenzboten/download/fulltext/alto/270882>
- Request access asking a contact person of the SuUB



# How to get access to metadata?

- On the *Title* page <http://brema.suub.uni-bremen.de/periodical/titleinfo/282153> you find:

**Links**

---

Reference	 <a href="#">Staats- und Universitätsbibliothek Bremen ↗</a>
Availability	 <a href="#">In My Library</a>
Metadata	 <a href="#">METS (OAI-PMH)</a> 

**Classification**

---

[Regionalhistorische Werke](#) → [Periodika](#)  
[Nicht Regionalspezifisches](#)

**Terms of Use**

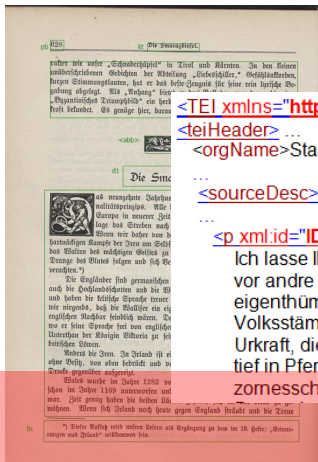
---

Licence	<a href="#">Public Domain Mark 1.0</a>
Usage	Available for free download. In the event of reproduction or exp; accordance with usual scientific practices, to provide: <ul style="list-style-type: none"><li>• information</li></ul>



# Outline

- The digitisation project – *Die Grenzboten*
- OCR post correction
- Creation of a TEI version
- Working with full texts
  - Digitisation portals
  - Virtual research environments and research infrastructures
- Discussion



```

<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader> ...
  <orgName>Staats- und Universitätsbibliothek (S
  ...
  <sourceDesc><bibl type="J">Die Grenzboten.
  ...
  <p xml:id="ID_8" next="#ID_9">

```

Ich lasse Ihnen das Glauben und Wissen c  
vor andre Richterstühle gehörig hier zur Se  
eigenthümlichen Stellung dieser drei Männ  
Volksstämmen. Strauß, der Schwabe mit s  
Urkraft, die mit einem Hiebe durch Reiter u  
tief in Pferdes Rücken schlug; Ludwig Feu  
zornesschnaubend und gewandt auch das

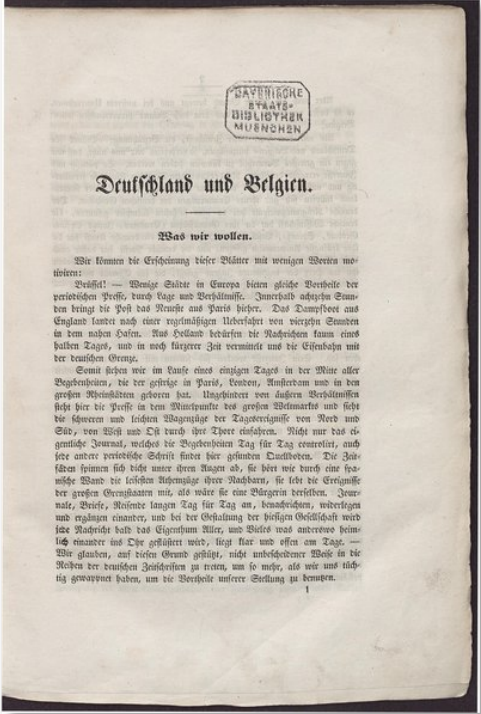


# *Die Grenzboten* – Distribution in the Digital World

- *Die Grenzboten* is integrated in
  - DTAQ : [http://www.deutschestextarchiv.de/dtaq/book/show/grenzboten\\_179382\\_282158c](http://www.deutschestextarchiv.de/dtaq/book/show/grenzboten_179382_282158c)



# Die Grenzboten. Erster Jahrgang. Leipzig, 1841. [automatisch korrigiert]



Cover in hoher Auflösung

[Buch in Korrekturumgebung öffnen ...](#)

## Informationen

**Quelle:** Die Grenzboten  
**Publikationstyp:** Zeitschrift/Zeitung  
**Umfang:** 324 Scans  
 ca. 521542 Zeichen  
 ca. 75641 Tokens  
 /[[[:alnum:]]/  
 ca. 16210 Oberflächentypes

**Schriftart:** Fraktur  
**Genre:** Zeitung  
**im DTA seit:** 2014-11-04 12:27:31  
**zuletzt geändert:** 2014-11-04 13:11:28

**Verfügbarkeit:** Text (TEI-XML-, HTML-, TCF-, E-Book-Fassung): Dieses Werk ist **gemeinfrei**.  
 Weitere Informationen: [Nutzungsbedingungen](#).

**Grundlage dieses Digitalisats:** Staats- und Universitätsbibliothek (SuUB) Bremen: Bereitstellung der Texttranskription. (2013-11-19T17:23:38Z) *Bitte beachten Sie, dass die aktuelle Transkription (und Textauszeichnung) mittlerweile nicht mehr dem Stand zum Zeitpunkt der Übernahme des Werkes in das DTA entsprechen muss.*  
 Bayerische Staatsbibliothek: Bereitstellung der Bilddigitalisate (Signatur Per 61 k-1).

## Ansichten

- [Korrekturumgebung](#)
- *Dieses Werk befindet sich in der Qualitätskontrolle und ist noch nicht auf der DTA-Webseite verfügbar.*
- [Faksimiles \(DFG-Viewer\)](#)
- **Download:**
  - [Text \(UTF-8, Seitenumbrüche als ASCII \014\)](#)
  - [TEI/XML \(mit Silbentrennung\)](#)
  - [HTML \(mit Silbentrennung\)](#)
  - [TEI-Header](#)
  - [CMDI](#)
  - [Dublin Core](#)

## Suche im Buch

DDC nicht verfügbar  
  +

## Inhaltsverzeichnis

- [Dtuschlmech und BtlgitN.](#)
- [Die Seebader in Dstende.](#)
- [Literarische Uebersichten vom Standpuncte der Gesellschaft.](#)
- [Leipziger Skizzen.](#)
- [T a g e b u e h.](#)
- [September Scenen.](#)
- [Ein Wort über das deutsche Theater. \(Eine Vorlesung gehalten im Museum zu Frankfurt.\)](#)
- [Elsaß und Flandern.](#)
- [R e i s e b v i e f e.](#)

# *Die Grenzboten* – Distribution in the Digital World

- *Die Grenzboten* is integrated in
  - DTAQ : [http://www.deutschestextarchiv.de/dtaq/book/show/grenzboten\\_179382\\_282158c](http://www.deutschestextarchiv.de/dtaq/book/show/grenzboten_179382_282158c)
  - <http://kaskade.dwds.de/dstar/grenzboten/>
  - CLARIN-D – 3 edited chapters: <https://vlo.clarin.eu/search?4&q=grenzboten>







# A small part of *Die Grenzboten* was uploaded to the CLARIN Virtual Language Observatory

The screenshot shows the CLARIN VLO search results page for the query 'grenzboten'. The interface includes a search bar, a breadcrumb trail 'VLO / Faceted search / Search results', and a list of search results. On the left, there are faceted search options for Language, Collection, Format, Availability, and Search options. The search results list includes entries like 'Parks', 'Öffentliche Charaktere', and 'Öffentliche Charaktere'. Each result has a document icon and a question mark icon. The page footer contains 'About v4.0.2', 'Service provided by CLARIN', and 'Contact'.

**But:**

**I don't want to emphasise this.  
This talk is about transferring large  
amounts of "just" OCRed,  
post corrected and structured  
full texts.**

# *Die Grenzboten* – Distribution in the Digital World

- *Die Grenzboten* is integrated in
  - DTAQ : [http://www.deutschestextarchiv.de/dtaq/book/show/grenzboten\\_179382\\_282158c](http://www.deutschestextarchiv.de/dtaq/book/show/grenzboten_179382_282158c)
  - <http://kaskade.dwds.de/dstar/grenzboten/>
  - CLARIN-D – 3 edited chapters: <https://vlo.clarin.eu/search?4&q=grenzboten>
  - CLARIN crawler is harvesting the DTAQ ++ CLARIN Collection is planned
  - The integration into DARIAH-DE is intended but not yet accomplished
- Used as sample data
  - in PoCoTo workshops at the Center for Information and Language Processing (CIS, LMU München)
  - for a DARIAH topic modeling tool: <https://github.com/dariah-de/topics>
  - within the “OCR-D coordination project”: <http://ocr-d.de/>
- Used in a digital humanities seminar at the University of Ghent
- Several scientist directly contacted the SuUB Bremen to ask for the full text and metadata

b

Staats- und  
Universitätsbibliothek  
Bremen

## Virtual Research Environments and Infrastructures



CLARIN / CLARIN-D - *Common Language Resources  
and Technology Infrastructure*

DARIAH / DARIAH-DE / Textgrid

DARIAH-VL Virtual Research Environment  
Service Infrastructure (VRE-SI)



DTA – Deutsches Textarchiv  
at the BBAW in Berlin

DTAQ

D\*-Korpusinfrastruktur

FuD, Trier –

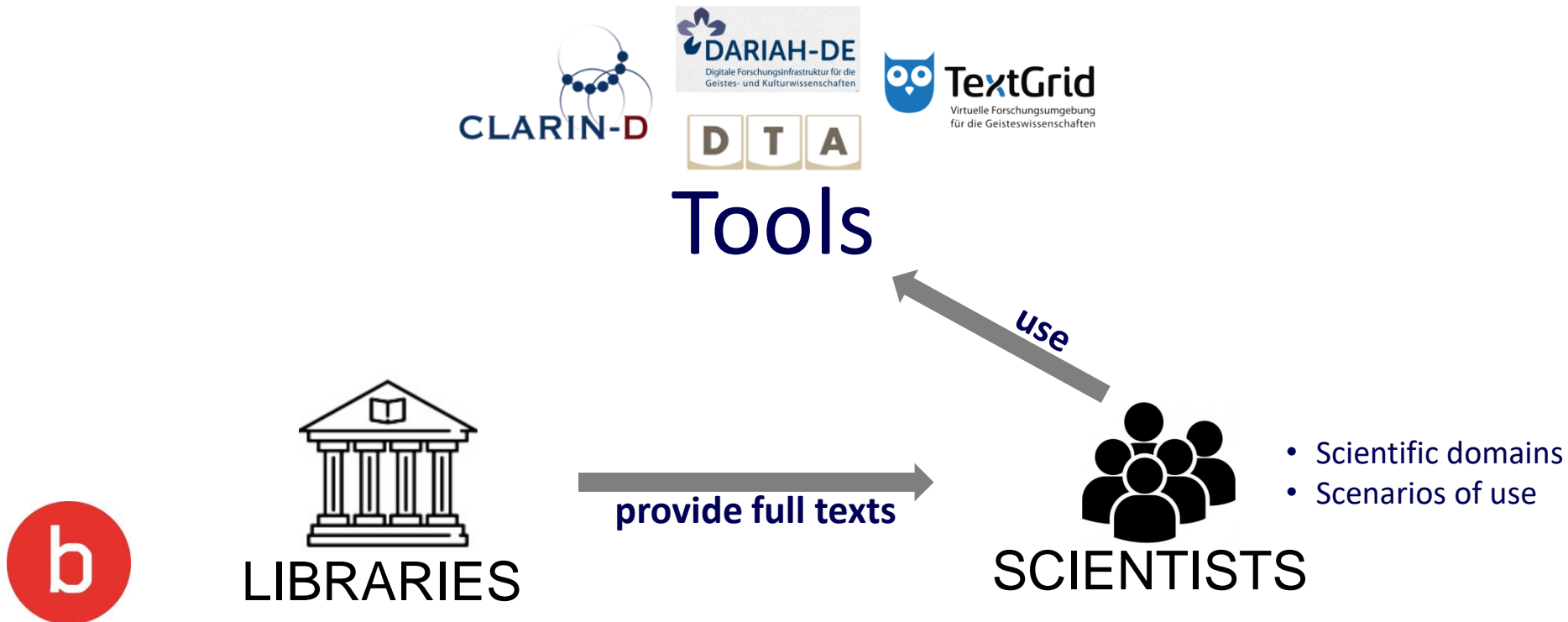


**FuD**

Eine virtuelle Forschungsumgebung  
für die Geisteswissenschaften

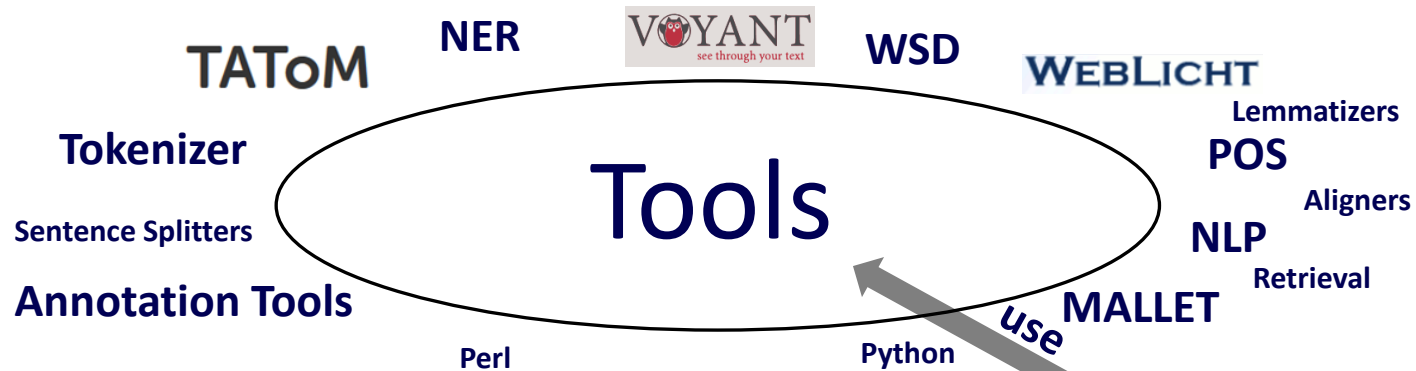
# Virtual research environments and research infrastructures

- Providing full texts
  - Scientific domains
  - Scenarios of use
  - Tools
  - ... and requirements given by VREs/RIs themselves



# Virtual research environments and research infrastructures

- Providing full texts there might be different requirements by different
  - Scientific domains
  - Scenarios of use
  - Tools
  - ... and requirements given by VREs/RIs themselves



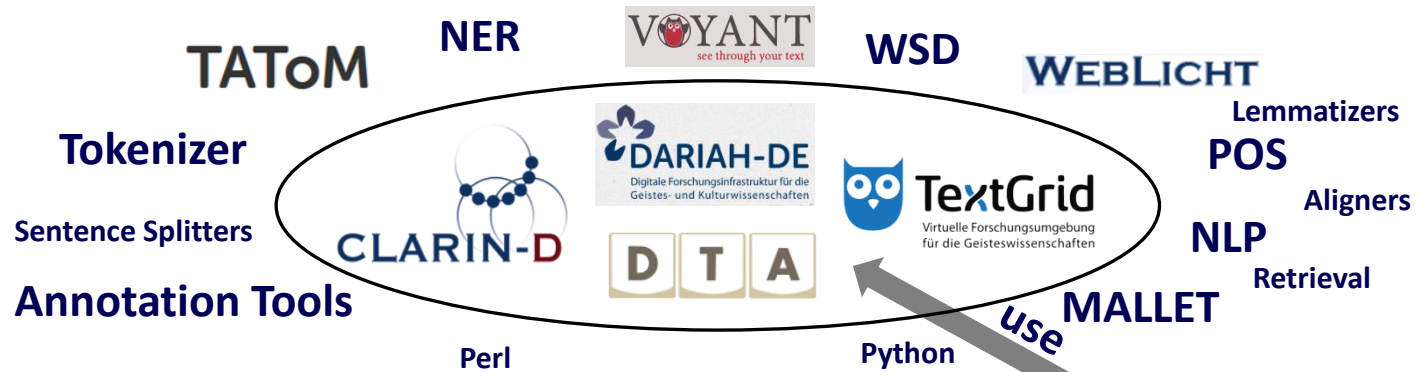
provide full texts



- Scientific domains
- Scenarios of use

# Virtual research environments and research infrastructures

- Providing full texts there might be different requirements by different
  - Scientific domains
  - Scenarios of use
  - Tools
  - ... and requirements given by VREs/RIs themselves

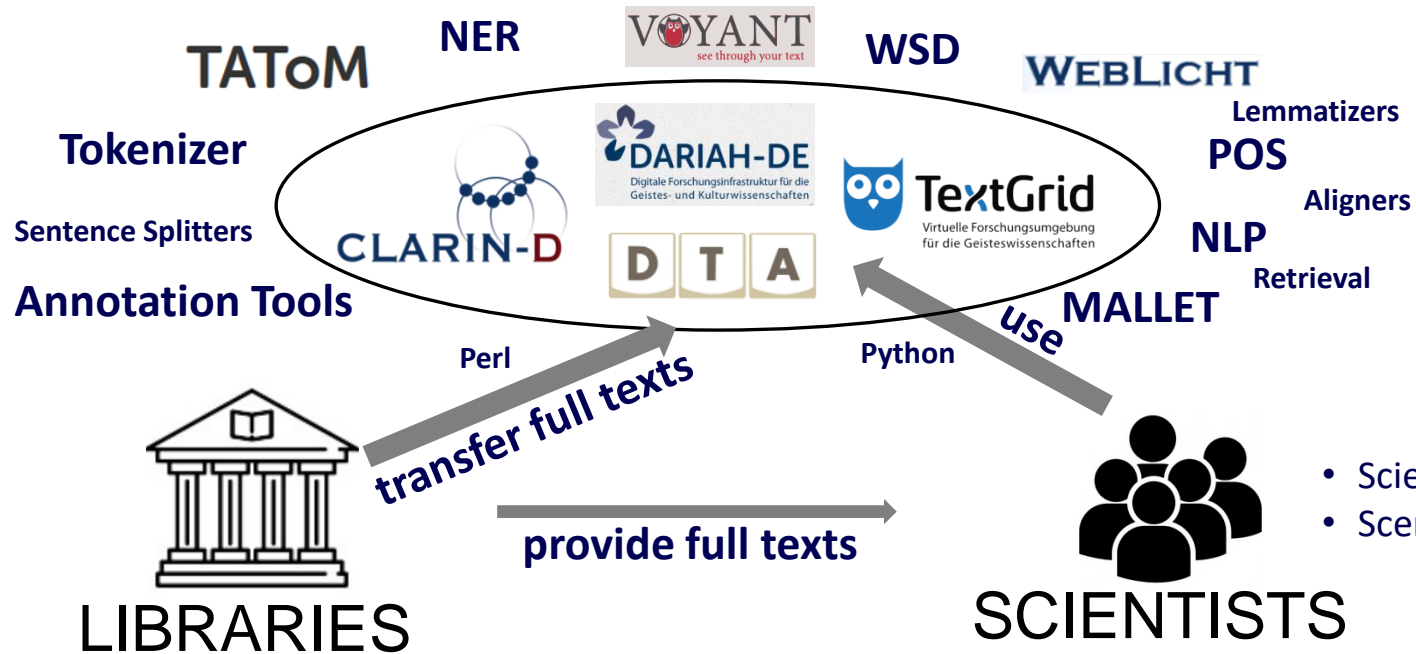


provide full texts



# Virtual research environments and research infrastructures

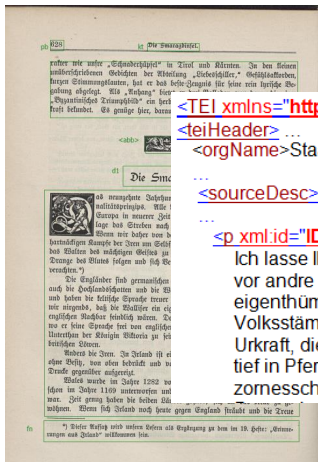
- Providing full texts there might be different requirements by different
  - Scientific domains
  - Scenarios of use
  - Tools
  - ... and requirements given by VREs/RIs themselves





# Outline

- The digitisation project – *Die Grenzboten*
- OCR post correction
- Creation of a TEI version
- Working with full texts
  - Digitisation portals
  - Virtual research environments and research infrastructures
- Discussion



```

<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader> ...
  <orgName>Staats- und Universitätsbibliothek (S
  ...
  <sourceDesc><bibl type="J">Die Grenzboten.
  ...
  <p xml:id="ID_8" next="#ID_9">

```

Ich lasse Ihnen das Glauben und Wissen c  
vor andre Richterstühle gehörig hier zur Se  
eigenthümlichen Stellung dieser drei Männ  
Volksstämmen. Strauß, der Schwabe mit s  
Urkraft, die mit einem Hiebe durch Reiter u  
tief in Pferdes Rücken schlug; Ludwig Feu  
zornesschnaubend und gewandt auch das



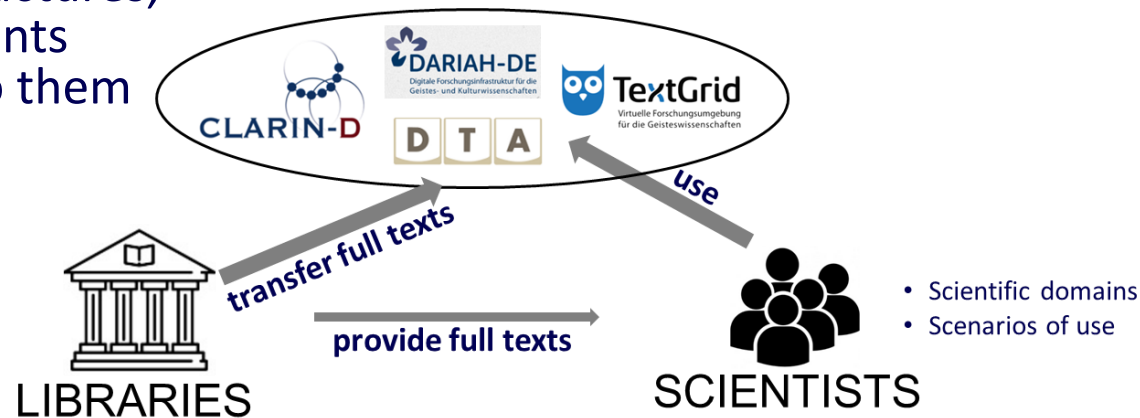


## Discussion

- Requirements to discuss
  - Quality limits: Minimum character recognition rate: 99 +/- x %
  - Formats: plain text, ABBYY, Alto, TEI, DTABf, ...
  - Line breaks: Wikisource vs. DTA/SuUB
  - Strictness of character transcription:

ift ist ältlicher ältlicher

- Annotation guide lines / subsets
  - Back link to the images (with URN granular persistent URLs)
- Should digitising libraries rethink the services they offer?
- Again: So it seems a good deal for libraries to TALK to research infrastructures, coordinate some requirements and transfer the full texts to them





Staats- und  
Universitätsbibliothek  
Bremen

# Thank you for your attention!

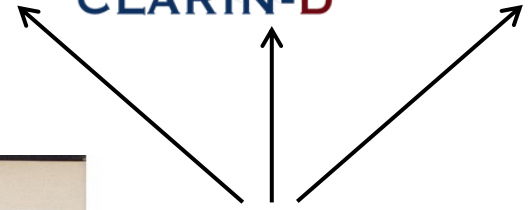


Professional Imaging and Archiving  
**H+K**  
Herrmann und Kraemer

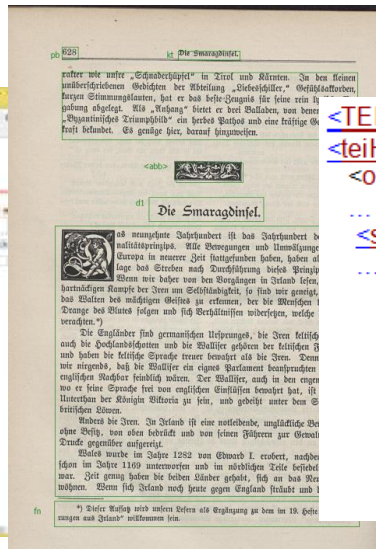
**D T A**



**CLARIN-D**



**ABBYY**



```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
<teiHeader> ...
<orgName>Staats- und Universitätsbibliothek (S
...
<sourceDesc><bibl type="J">Die Grenzboten.
...
<p xml:id="ID_8" next="#ID_9">
```

Die jüngste päpstliche Bulle und die künftige Papstwahl.  
Bedenken über die Sprachveränderung.  
Eine neue Sentenz.  
Das Atelier eines französischen Bildhauers.  
Die jüngste Papiere.  
Der Vertrag über den Papst.  
Bedenken über die Sprachverbesserung.

Ich lasse Ihnen das Glauben und Wissen vor  
andere Richterstühle gehörig hier zur Se  
eigenthümlichen Stellung dieser drei Männ  
Volksstämmen. Strauß, der Schwabe mit s  
Urkraft, die mit einem Hiebe durch Reiter u  
tief in Pferdes Rücken schlug; Ludwig Feu  
zornesschnaubend und gewandt auch das



Staats- und  
Universitätsbibliothek  
Bremen

# Backup Slides

## „Die Grenzboten“ – Correction Algorithm

- For each word (roughly 80 Mio.) do the following
  - Check the word against the list of 1.7 Mio. historical word forms
    - If there is a match: do nothing
    - If there is no match:
      - Create modified words using the list of typical OCR errors (candidates)
      - Sort these candidates by the distance to the initial word
      - An evaluation function takes this sorting together with criteria like frequency of the candidate word and the frequency of the OCR character error
      - Take best candidate with respect to the evaluation function
- Example „gelden“: **gelten**, gelben, gelden, gelder, gelbem, gelteu, gelber, getten, geiten
- Runtime: 4 hours and 15 minutes



WIKISOURCE

- Hauptseite
- Systematischer Einstieg
- Themenübersicht
- Autorenindex
- Zufällige Seite

- Mitmachen
- Letzte Änderungen
- Neuer Artikel
- Korrekturen des Monats
- Gemeinschaftsporta I
- Skriptorium
- Auskunft
- Hilfe
- Spenden

- Drucken/exportieren
- Buch erstellen
- Als PDF herunterladen
- Druckversion

- Werkzeuge
- Links auf diese Seite
- Änderungen an verlinkten Seiten
- Datei hochladen
- Spezialseiten
- dauerhafter Link
- Seiteninformationen

Sprachen

Benutzerkonto erstellen Anmelden

Quellentext Diskussion Lesen Bearbeiten Versionsgeschichte Suchen

# Die Grenzboten (1841/1842)

## 1. Jg., Band 1, 1841 [Bearbeiten]

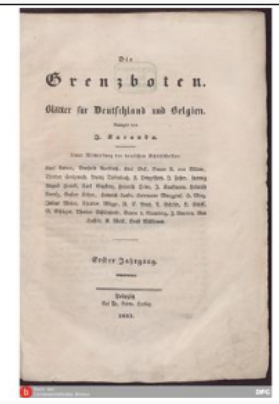
### Seitenindex

- Titel. **S. i**
- Inhaltsverzeichnis. **S. I-II**
- I. Kuranda: *Deutschland und Belgien*. S. 1–10
- D. e. t.: *Die Seebäder in Ostende*. S. 11–18
- A. Baron: *Literarische Uebersichten vom Standpuncte der Gesellschaft 1*. S. 19–25
- Anonymus: *Leipziger Skizzen 1841*. S. 26–29
- *Tagebuch*. S. 30–36
  - P.: *Paris*. S. 30–31
  - L.: *Brüssel*. S. 31–33
    - Anonymus: *Göthe's Egmont als Flamänder. Selbstmörder auf der Eisenbahn. Geimüller als Maltheser. Gervinus. Gutzkow. Lewald. Ein gerechtes Urtheil. Laube als Dramatiker. Gute Nachrichten aus Stuttgart*. S. 33–36
- J. F.: *September-Scenen*. S. 37–45
- Theodor Greizenach: *Ein Wort über das deutsche Theater*. S. 46–50
- Sch.: *Elsaß und Flandern*. S. 51–55
- A. Weil (=Alexander Weill): *Reisebriefe 1-2*. S. 56–59
- B. B.: *Briefe aus London 1*. S. 60–64
- P.: *Briefe aus Paris*. S. 65–66
- *Tagebuch*. S. 67–68
- E. R.: *Deutsche Eisenbahnen und belgische*. S. 69–77
- Alexander von Württemberg: *Mohrenrache*. S. 78–81
- Anonymus: *Alphons Karr und die Wespen*. S. 82–93
- Ludwig August Frankl: *Eine Begegnung mit Paganini in Genua*. S. 94–96
- C.: *Briefe aus Frankfurt 1*. S. 97–99
- *Notizen*. S. 91–91
- I. Kuranda: *Börne in der letzten Zeit*. S. 101–112
- J. Cr.: *Zur Würdigung der neuesten nationalen Bewegungen in Deutschland*. S. 113–117
- A. Baron [?]: *Literarische Uebersichten vom Standpuncte der Gesellschaft 2*. S. 118–123
- Jobard: *Ueber den Bau eiserner Häuser in Belgien*. S. 124–127
- A. Weill: *Reisebriefe 3*. S. 128–130
- *Tagebuch*. S. 131–132
- P.: *Schauspiel und Schauspieler*. S. 133–140
- Anonymus: *Lacordaire*. S. 141–151
- A. Lebermuth: *Lelewel auf dem Marktplatze zu Brüssel*. S. 143–143
- Anonymus: *Carl de Brouckère über die industriellen Wünsche Belgiens*. S. 153–158
- *Literatur*. S. 159–161
- *Tagebuch*. S. 162–164
- I. K...da (= Ignaz Kuranda): *Auch ein Wort über die orangistische Verschwörung*. S. 165–168
- Th. Schliephake: *Die Abdankung Kaiser Heinrichs des Vierten*. S. 169–185
- Anonymus: *Eine politische Satyre*. S. 186–194
- *Tagebuch*. S. 195–196
- Th. Schliephake: *Deutsche Taschenbücher 1*. S. 197–204

Textdaten	
Autor:	Verschiedene <span style="float:right">&gt;&gt;&gt;</span>
Titel:	<b>Die Grenzboten</b>
Untertitel:	<i>Blätter für Deutschland und Belgien</i> , später: <i>Zeitschrift für Politik, Literatur und Kunst</i>
Herausgeber:	<b>Ignaz Kuranda, Gustav Freytag, Julian Schmidt, Georg Cleinow</b>
Erscheinungsdatum:	1841–1842
Verlag:	Herbig
Erscheinungsort:	Leipzig
Quelle:	Band 1: SUUB Bremen  = Commons
Kurzbeschreibung:	1. Jahrgang (in drei Bänden)

Artikel in der Wikipedia

### Bild



Bearbeitungsstand

unvollständig



WIKISOURCE

- Hauptseite
- Systematischer Einstieg
- Themenübersicht
- Autorenindex
- Zufällige Seite

Mitmachen

- Letzte Änderungen
- Neuer Artikel
- Korrekturen des Monats
- Gemeinschaftsportal
- Skriptorium
- Auskunft
- Hilfe
- Spenden

Drucken/exportieren

- Buch erstellen
- Als PDF herunterladen
- Druckversion

Werkzeuge

- Links auf diese Seite
- Änderungen an verlinkten Seiten
- Datei hochladen
- Spezialseiten
- dauerhafter Link
- Seiteninformationen

Sprachen

Links hinzufügen

Index Diskussion

Lesen Bearbeiten Versionsgeschichte Mehr Suchen

# Index:Die Grenzboten 1-1841.pdf



Autor: Verschiedene  
 Titel: [Die Grenzboten \(1841/1842\), 1. Jahrgang, Band 1](#)  
 Verlag: Herbig  
 Erscheinungsjahr: 1841  
 Erscheinungsort: Leipzig  
 Quelle: [SUUB Bremen](#) = Commons

### Seiten

- [Titel i](#)
- [Inhaltsverzeichnis I II](#)
- I. Kuranda: Deutschland und Belgien. [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#)
- D. e. t.: Die Seebäder in Ostende. [11](#) [12](#) [13](#) [14](#) [15](#) [16](#) [17](#) [18](#)
- A. Baron: Literarische Uebersichten vom Standpuncte der Gesellschaft 1. [19](#) [20](#) [21](#) [22](#) [23](#) [24](#) [25](#)
- Anonymus: Leipzig Skizzen 1841. [26](#) [27](#) [28](#) [29](#)
- Tagebuch. [30](#) [31](#) [32](#) [33](#) [34](#) [35](#) [36](#)
- J. F.: September-Scenen. [37](#) [38](#) [39](#) [40](#) [41](#) [42](#) [43](#) [44](#) [45](#)
- Theodor Creizenach: Ein Wort über das deutsche Theater. [46](#) [47](#) [48](#) [49](#) [50](#)
- Sch.: Elsaß und Flandern. [51](#) [52](#) [53](#) [54](#) [55](#)
- A. Weil: Reisebriefe 1–2. [56](#) [57](#) [58](#) [59](#)
- B. B.: Briefe aus London 1. [60](#) [61](#) [62](#) [63](#) [64](#)
- P.-: Briefe aus Paris. [65](#) [66](#)
- Tagebuch. [67](#) [68](#)
- E. R.: Deutsche Eisenbahnen und belgische. [69](#) [70](#) [71](#) [72](#) [73](#) [74](#) [75](#) [76](#) [77](#)
- Alexander von Württemberg: Mohrenrache. [78](#) [79](#) [80](#) [81](#)
- Anonymus: Alphons Karr und die Wespen. [82](#) [83](#) [84](#) [85](#) [86](#) [87](#) [88](#) [89](#) [90](#) [91](#) [92](#) [93](#)
- Ludwig August Frankl: Eine Begegnung mit Paganini in Genua. [94](#) [95](#) [96](#)
- C.: Briefe aus Frankfurt 1. [97](#) [98](#) [99](#)
- Notizen. [100](#)
- I. Kuranda: Borne in der letzten Zeit. [101](#) [102](#) [103](#) [104](#) [105](#) [106](#) [107](#) [108](#) [109](#) [110](#) [111](#) [112](#)
- J. Cr.: Zur Würdigung der neuesten nationalen Bewegungen in Deutschland. [113](#) [114](#) [115](#) [116](#) [117](#)
- A. Baron (?): Literarische Uebersichten vom Standpuncte der Gesellschaft 2. [118](#) [119](#) [120](#) [121](#) [122](#) [123](#)



WIKISOURCE

- Hauptseite
- Systematischer Einstieg
- Themenübersicht
- Autorenindex
- Zufällige Seite

- Mitmachen
- Letzte Änderungen
- Neuer Artikel
- Korrekturen des Monats
- Gemeinschaftsportal
- Skriptorium
- Auskunft
- Hilfe
- Spenden

- Drucken/exportieren
- Buch erstellen
- Als PDF herunterladen
- Druckversion

- Werkzeuge
- Links auf diese Seite
- Änderungen an verlinkten Seiten
- Datei hochladen
- Spezialseiten
- dauerhafter Link
- Seiteninformationen

- Sprachen
- Links hinzufügen

Seite Diskussion Scan

Lesen Bearbeiten Versionsgeschichte Mehr Suchen

# Seite:Die Grenzboten 1-1841.pdf/10

verschiedene: Die Grenzboten, 1. Jahrgang



Fertig. Dieser Text wurde zweimal anhand der Quelle korrekturgelesen. Die Schreibweise folgt dem Originaltext.

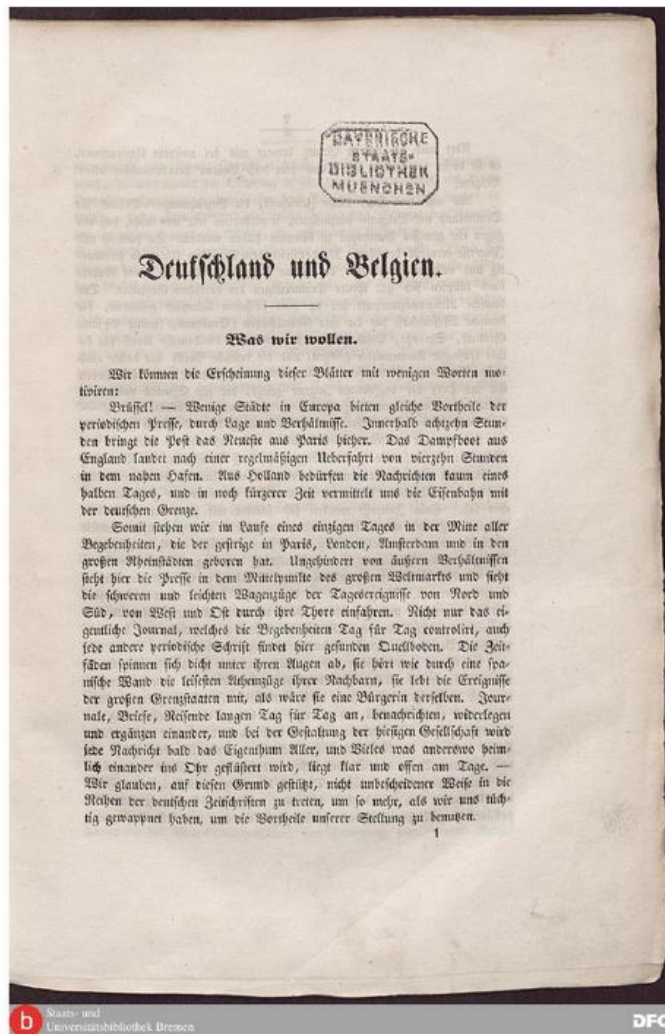
## Deutschland und Belgien.

### Was wir wollen.

Wir könnten die Erscheinung dieser Blätter mit wenigen Worten motiviren:

Brüssel! – Wenige Städte in Europa bieten gleiche Vortheile der periodischen Presse, durch Lage und Verhältnisse. Innerhalb achtzehn Stunden bringt die Post das Neueste aus Paris hieher. Das Dampfboot aus England landet nach einer regelmäßigen Ueberfahrt von vierzehn Stunden in dem nahen Hafen. Aus Holland bedürfen die Nachrichten kaum eines halben Tages, und in noch kürzerer Zeit vermittelt uns die Eisenbahn mit der deutschen Grenze.

Somit stehen wir im Laufe eines einzigen Tages in der Mitte aller Begebenheiten, die der gestrige in Paris, London, Amsterdam und in den großen Rheinstädten geboren hat. Ungehindert von äußern Verhältnissen steht hier die Presse in dem Mittelpunkte des großen Weltmarkts und sieht die schweren und leichten Wagenzüge der Tagesereignisse von Nord und Süd, von West und Ost durch ihre Thore einfahren. Nicht nur das eigentliche Journal, welches die Begebenheiten Tag für Tag controlirt, auch jede andere periodische Schrift findet hier gesunden Quellboden. Die Zeitfäden spinnen sich dicht unter ihren Augen ab, sie hört wie durch eine spanische Wand die leisesten Athemzüge ihrer Nachbarn, sie lebt die Ereignisse der großen Grenzstaaten mit, als wäre sie eine Bürgerin derselben. Journale, Briefe, Reisende langen Tag für Tag an, benachrichten, widerlegen und ergänzen einander, und bei der Gestaltung der hiesigen Gesellschaft wird jede Nachricht bald das Eigenthum Aller, und Vieles was anderswo heimlich einander ins Ohr geflüstert wird, liegt klar und offen am Tage. – Wir glauben, auf diesen Grund gestützt, nicht unbescheidener Weise in die Reihen der



## Sample Searches and Analyses

- „Image and Situation of Women in the 19<sup>th</sup> century“
- Linguistically motivated searches
  - E.g. specific word forms or uses of words
  - Collocation analysis
  - Stylometry
- “Bildprosa” (,Zeit-, Welt-, Lebensbilder’, ,Reise- und Kulturbilder’, ,Charakterbilder’ und ,Sittengemälde’, ,Umriss’, ,Skizzen und Studien’)  
<http://www.fb10.uni-bremen.de/germanistik/fachgebiete/literatur/aga/forschung/bildprosa.aspx>
- Search for combinations of classes of words like geological names, personal names, food, weather or mentality
- Rhetorical figures like metaphors



## Context Information

- Miscellaneous
  - [DARIAH-BE: https://youtu.be/I7EumUAh-J4](https://youtu.be/I7EumUAh-J4) <http://hdl.handle.net/2268/203796>
  - <http://dhd-blog.org> [www.youtube.com/user/dhdkanal](http://www.youtube.com/user/dhdkanal)
  - 2/2016 at GhebtCDG: Dr. Martin Reynaert: From text images or files to online corpus:  
<http://www.flw.ugent.be/sync/activiteiten/1418/a/martinreynaert.pdf>
  - <https://www.scads.de/en/blog-en/223-connecting-digital-humanities-with-the-clarin-infrastructure-2?showall=1> Section “Technical Variety” ← very near to my point of view!!!

**Last Slide**