

# Retrodigitization and Electronic Representation of the “Vorwärts”

## A Workshop-Report



# Table of Content

## 1. Why do we digitize the „Vorwärts“

## 2. What has been done so far

- The scan-process: digitizing in-house
- The internal structure of the newspaper corpus
- The previous workflow

## 3. What should be done from now on

- Objectives of the web presentation
- Research project: What has been implemented so far
- Research project: Possibilities and challenges for the future

# Why do we digitize the „Vorwärts“

- By far the most important source for social democracy in German Kaiserreich and Weimarer Republik
- Content important for both historical and cultural research
- Currently only accessible via microfilm edition
  - Huge workload for researchers
  - In many cases practically not accessible
- Fits in the narrowest mission of the Library of the Friedrich-Ebert-Stiftung



29.03.2016

Olaf Guercke - Retrodigitization and Electronic Presentation of the "Vorwärts"

# Why do we digitize the „Vorwärts“

Aim:

- a) Complete digitization of the „Vorwärts“ , 1876-1878 ; 1891-1933
- b) Providing it to the public via researchable web-presentation

Extent:

200.000 DIN A3 newspaper-pages

19.000 issues

44 volumes

The „Vorwärts“ and 9 different supplements

Font: German gothic

# Extra-Ausgabe. Vorwärts

**Abonnement-Verfahren**  
Das Abonnement für ein Jahr (12 Hefen) kostet 1,20 Mark. Ein halbes Jahr (6 Hefen) 0,70 Mark. Ein Vierteljahr (3 Hefen) 0,40 Mark. Die Abnahme erfolgt durch Postanweisung an den Verlag. Die Abnahme durch den Postamt ist ebenfalls möglich. Die Abnahme durch den Postamt ist ebenfalls möglich. Die Abnahme durch den Postamt ist ebenfalls möglich.

**Berliner Volksblatt.**  
Zentralorgan der sozialdemokratischen Partei Deutschlands.

Redaktion: 60, 65, Lindenstrasse 69. Sonnabend, den 25. Juli 1914. Gedruckt: 60, 65, Lindenstrasse 69.

## Aufruf!

Noch dampfen die Acker auf dem Balkan von dem Blute der nach Tausenden Hingemordeten, noch rauchen die Trümmer verheerter Städte, verwüsteter Dörfer, noch irren hungernd arbeitslose Männer, verwitwete Frauen und verwaisete Kinder durchs Land, und schon wieder schickt sich die vom öfterreichischen Imperialismus entfesselte Kriegsmaschine an, **Tod und Verderben über ganz Europa zu bringen.**

Verurteilen wir auch das Treiben der großserbischen Nationalisten, so fordert doch die **frivole Kriegsprovokation der österreichisch-ungarischen Regierung** den schärfsten Protest heraus. Sind doch die Forderungen dieser Regierung so **brutal**, wie sie in der Weltgeschichte noch nie an einen selbständigen Staat gestellt sind, und können sie doch nur darauf berechnet sein, den Krieg geradezu zu provozieren.

Das klassenbewußte Proletariat Deutschlands erhebt im Namen der Menschlichkeit und der Kultur flammenden Protest gegen dies verbrecherische Treiben der Kriegsherrn. Es fordert gebieterisch von der deutschen Regierung, daß sie ihren Einfluß auf die öfterreichische Regierung zur Aufrechterhaltung des Friedens ansetze, und falls der schändliche Krieg nicht zu verhindern sein sollte, sich jeder kriegerischen Einmischung enthalte. **Kein Tropfen Blut eines deutschen Soldaten darf dem Machthaber der österreichischen Gewalttäter, den imperialistischen Profitinteressen geopfert werden.**

Parteigenossen, wir fordern Euch auf, sofort in **Massenversammlungen** den unerfüßlichen Friedenswillen des **klassenbewußten Proletariats** zum Ausdruck zu bringen. Eine erste Stunde ist gekommen, erster als irgend eine der letzten Jahrzehnte. Gefahr ist im Verzuge! **Der Weltkrieg droht!** Die herrschenden Klassen, die Euch im Frieden knebeln, verachten, ausnutzen, wollen Euch als Kanonenfutter mißbrauchen. Überall muß den Gewalttätigern in die Ohren klingen:

**Wir wollen keinen Krieg! Nieder mit dem Kriege!**

**Hoch die internationale Völkerverbrüderung!**

Berlin, den 25. Juli 1914.

**Der Parteivorstand.**

## Protest des österreichischen Proletariats!

Arbeiter, Parteigenossen!

Die Forderung stellt sich wieder mit aller Schärfe vor uns: Die Österreicher sollen sich nicht an dem Weltkriege beteiligen, sondern sich für den Frieden einsetzen. Die Österreicher sollen sich nicht an dem Weltkriege beteiligen, sondern sich für den Frieden einsetzen. Die Österreicher sollen sich nicht an dem Weltkriege beteiligen, sondern sich für den Frieden einsetzen.

Wir fordern Euch auf, sofort in Massenversammlungen den unerfüßlichen Friedenswillen des klassenbewußten Proletariats zum Ausdruck zu bringen. Eine erste Stunde ist gekommen, erster als irgend eine der letzten Jahrzehnte. Gefahr ist im Verzuge! Der Weltkrieg droht!

Die Österreicher sollen sich nicht an dem Weltkriege beteiligen, sondern sich für den Frieden einsetzen. Die Österreicher sollen sich nicht an dem Weltkriege beteiligen, sondern sich für den Frieden einsetzen. Die Österreicher sollen sich nicht an dem Weltkriege beteiligen, sondern sich für den Frieden einsetzen.

Redaktion: 60, 65, Lindenstrasse 69. Sonnabend, den 25. Juli 1914. Gedruckt: 60, 65, Lindenstrasse 69.

# Vorwärts

**Berliner Volksblatt**  
Zentralorgan der sozialdemokratischen Partei Deutschlands.

Redaktion: 60, 65, Lindenstrasse 69. Sonnabend, den 25. Juli 1914. Gedruckt: 60, 65, Lindenstrasse 69.

Vorwärts-Verlag G. m. b. H.

Vertrieb: 60, 65, Lindenstrasse 69. Sonnabend, den 25. Juli 1914. Gedruckt: 60, 65, Lindenstrasse 69.

# am 1. August! Wer mit dem Faschismus spielt, der spielt mit Deutschlands Untergang!

Die sozialdemokratische Partei Deutschlands hat sich gegen den Faschismus ausgesprochen. Sie hat sich gegen den Faschismus ausgesprochen. Sie hat sich gegen den Faschismus ausgesprochen. Sie hat sich gegen den Faschismus ausgesprochen.

Die sozialdemokratische Partei Deutschlands hat sich gegen den Faschismus ausgesprochen. Sie hat sich gegen den Faschismus ausgesprochen. Sie hat sich gegen den Faschismus ausgesprochen. Sie hat sich gegen den Faschismus ausgesprochen.

Die sozialdemokratische Partei Deutschlands hat sich gegen den Faschismus ausgesprochen. Sie hat sich gegen den Faschismus ausgesprochen. Sie hat sich gegen den Faschismus ausgesprochen. Sie hat sich gegen den Faschismus ausgesprochen.

## Bomben im Arbeiterviertel

Schweigsame Polizei — Die SA. als Täter erkannt

Die sozialdemokratische Partei Deutschlands hat sich gegen den Faschismus ausgesprochen. Sie hat sich gegen den Faschismus ausgesprochen. Sie hat sich gegen den Faschismus ausgesprochen. Sie hat sich gegen den Faschismus ausgesprochen.

Die sozialdemokratische Partei Deutschlands hat sich gegen den Faschismus ausgesprochen. Sie hat sich gegen den Faschismus ausgesprochen. Sie hat sich gegen den Faschismus ausgesprochen. Sie hat sich gegen den Faschismus ausgesprochen.

# What has been done so far

## Scan process: Challenges of in-house digitization

- Advantages
  - Complete control over production process with sensitive, partly damaged material
- Challenge 1: Suitable devices
  - Extensive market research
  - In our case: Scanner: „Zeutschel OS 12000 A1“, Software: „OS 12“
- Challenge 2: Manpower
  - 1 Librarian (25 hours) and 2 scan operators (23 hours in total)
  - 55.000 images scanned and postproduced since June 2015
  - Support your scan-operator!!



# What has been done so far

## Scan process: File formats

- Tiff uncompressed - ca. 55 MB per Image
- Jpeg 100% - ca. 18 MB per Image
- Jpeg 80% - ca. 3 MB per Image
- Pdf (Jpeg 80%) of whole issues
  
- All formats are generated automatically by the scanner-software

# What has been done so far

## Internal structure of the newspaper-corpus

- Periodicity: 4 issues per week - daily appearance - 2 issues per day
- Numbered sequentially during each volume
- Additional numbering for morning/evening-issues
- 9 different supplements with independent numbering
- Various special cases





# What has been done so far

## Mirroring the structure in filenames and folders

- Filename (example): vw-1923-12-25-325\_b165-004.tiff
- Translation: „Page 4 of the Vorwärts issue nr. 1923.325, evening-edition nr. 1923.165, released in December 25th“
- Folder-structure: file-format/newspaper-title/volume/release-date/issue
  
- Filenames and folders are automatically generated by the scan-software

# What has been done so far

## Workflow:

- Paper originals bound in quarterly folios are examined and documented with excel-spreadsheet (volume, date, number, pages per issue, conservation-status)
- Folios are opened carefully by a bookbinder
- Originals are scanned and stored in archive-boxes
- Quality control and upload of the scans
- Missing or heavily damaged pages are replaced with help of ULB-Bonn
- All these steps are performed parallel



# What should be done from now on

## Objectives of the web-presentation

- 2/3 of actual client inquires: Researchers and students
- Client request:
  - For the preparation and design of a multiple-semester seminar at the historian faculty of his university, a docent needs access to reports from the "Vorwärts" about left-wing labor-unrest in Halle in the years of 1919, 1920 and 1921.*
- Solution:
  - Keyword-search on basis of OCR, ideally to be narrowed down to certain timeframes
  - Browsing approach via volume, release date and issue

# What should be done from now on

Additional important objectives:

- Traceability via URNs at a page-level
- Managable mechanics to block content with copyright-related issues at a page level



# What should be done from now on

Research project: Library of the FES – Imageware Components GmbH (IWC)

Aim:

- Presenting the „Vorwärts“ in the web with the presentation-software MyBib eL®
- Developing the software to meet the requirements of such a presentation with the input of expertise from both sides
- Web presentation will be hosted by hbz

# What should be done from now on

Research project: What has been implemented so far

- Viewer / display:
  - Pages can be scaled continuously
  - Keyword search / Truncation, / Search for phrases and multiple words
  - Hit-lists / highlighting of hits
- OCR
  - Fraktur-OCR: ABBYY vs. Tesseract
  - Pretest: ABBYY provides much better results with much faster performance

# What should be done from now on

Research project: Possibilities and challenges for the future

- Configuration of the „advanced search“
  - Search in certain time-periods
  - Search with all Boolean operators
  - Fuzzy search
- Search entry „browsing“
  - Structurizing the browsing opportunity to three levels: volume, date of appearance (calendar-function), issue



# What should be done from now on

## Research project: Possibilities and challenges for the future

- Option to block content at a page level
  - Extremely important requirement to cope with copyright issues
  - Has to be developed on base of the access management functions of MyBib el





Im wunderschönen Monat Mai  
 Wenn alle Knospen spriessen  
 Hol ich mein schönstes Kleid herbei  
 Den Lenz geschmückt zu grüssen.

Und setz ich mich in's grüne Gras	Ich lache - mache keinen Mucks
Und kriegt mein Kleidlein Flecken	Und springe auf mein Zimmer
Und pfeffelt mich der Osterhas-	Drei Löffel Seifenflocken Lux-
Ich lass mich nicht erschrecken.	Schon hat's den alten Schimmer

Sunlicht<sup>®</sup> Mannheim



511 mal wurde ein Leinenstück  
 mit Persil gewaschen unter fach-  
 amtlicher Aufsicht! Das ge-  
 waschene Stück wurde auf seine  
 Festigkeit untersucht, und es ergab sich  
 im Hinblick auf ein so oftmaliges Waschen

ein ausserordentlich günstiges Er-  
gebnis für die Schonung des Gewebes!

Gibt es einen besseren Beweis für die hochwertige Beschaffen-  
 heit eines Waschmittels?

**Denken Sie: 511 mal gewaschen!**

Das entspricht bei regelmässigem vierwöchentlichem Waschen  
 einer Gebrauchsdauer von mehreren Jahrzehnten!

**Persil schont und pflegt die  
 Wäsche in ausserordentlichem Masse!**



Eine jede Bibliothek hat ihre Stammgäste. Seltsame Leute sind das, die Tag für Tag in den Besejalen herumtun und nichts weiter machen als lesen, am frühen Morgen, am hellen Mittag und am späten Abend, die abendrein aber noch ärgerlich sind, wenn die Bibliothekare Feierabend bieten. „Nachts müßten Sie aufhören, mindestens bis zwölf“, röhnierten neulich die Bücherwürmer in der Kunstgewerbebibliothek, „dann ist es wenigstens ruhig zum Lesen.“ Kaum, daß diese Menschen eiligen Schritts um die Ecke in eine Koltzküche gehen und etwas Kohlrabi essen! Gleich sind sie wieder da, putzen ihre Brillengläser und klappen die Bücher auf, um zu lesen. Was sollten sie auch weiter tun? Eigentlich müßten alle Gelehrte fein oder langsam welche werden, aber das stimmt nicht: Man eh einer schnappt nur über. Dann wehe der Bibliothek. Eima so:

### „Was halten Sie vom Awiensstil?“

„Haben Sie ein Buch über Rekutadnezar?“ fragt Frä. S. den Bibliothekar. Frä. S., das ist die Dame mit dem abgebrochenen Regenschirm. „Bedaure sehr“, antwortet der Beamte, „dieses Buch fällt nicht in unsere Sammlung.“ „Das werde ich dem Minister fragen!“ „Ja, was wollen Sie denn vom Minister?“ „Ich will ihm sagen, daß er das Buch endlich anschaffen soll.“ „Aber das können Sie doch auch unserem Direktor fragen.“ „Nein, ich gehe zum Kultusminister, ich habe sowieso bei ihm zu tun.“ Das ist Frä. S., immerhin noch ein harmloser Fall.

„Kann ich ein Buch haben vom Einzug Christi in Jerusalem?“ kommt ein Mann. „Bitte sehr.“ Eine ganze Bibliothek wird auf den Kopf gestellt, Fahrtrüble voll Bücher werden in den Besejalen geschleppt, aber nichts gefällt dem Mann. „Ach nein, das ist alles nichts, ich möchte eine Plastik haben“, quängelt dieser Qualgeist. „Dann sagen Sie doch in Zukunft gleich, was Sie haben

immer nur Awiensstil. Das könnten ebenjogut Brafsartoffeln auf Rand genöht sein, man schlägt hier nach und sucht dort, eine geschlagene halbe Stunde schon. Großer Gott, dämmert es schließlich: „Sie meinen Möbel im Queen-Ann-Stil (Admign-Anna-Stil)?“ „Ja, die meine ich“, strahlt der Mann und bekommt sein Buch über diese Art englischer Möbel.

### Fabrik für Reiseberichte.

Ob es draußen Backsteine regnet oder der Asphalt in der Sonne schmilzt, die großen Besejale der Bibliotheken sind immer vollbesetzt. Man erwischt gerade noch einen Platz zwischen zwei Stammgästen. Der dicke Fünfsziger zur Linken ist eingeknickt müde werden die Leute ja auch einmal von dem vielen Lesen, er schnarcht immer lauter wird das, man sieht sich den Schläfer genauer an, ein roter Haarplauum kraucht ihm aus den Armen, der Kranz um die Gläse ist auch rot wie die fleischigen Ohren, mit einemmal klappert der Dicke die kleinen Augen auf, und, ohne sie zu bewegen, liest er die angefangene Zeile zu Ende. Aus ist das Schnarchen. Es vergeht keine Minute, da hält er einem ein Buch unter die Nase und sagt: „Lesen Sie das doch mal, so etwas kann nur Theodor Storm schreiben!“ Man weiß im ersten Moment gar nicht, was man sagen soll, erst schnarcht der Mann, dann hält er einem seine Bücher unter die Nase, und tatsächlich, wie man noch auf das Gebiät starrt, fängt man auch schon zu lesen an, vier Strophen sind das immerhin, und nebenaan sitzt der Dicke, man fühlt, wie seine kleinen Augen fragen: „Na, wie finden Sie das Gebiät.“ Alles, was man herausbringen kann, ist ein Schwärzernes „hm“, bloß nicht mit dem Mann einlassen, aber der

(Stadtbibliothek). Aber was baut der Mann alles um sich auf, Atlanten, große und kleine, Reisebeschreibungen, echte natürlich, wirtschaftsgeographische, historische Literatur, Statistiken, vier Bände vom Brockhaus und drei vom Großen Meier noch dazu, schön muß sich so ein Reisebericht von ihm lesen, wo er doch im Schweiß seines Angesichts gewissenhaft errechnet hat, wie lange man wohl von Ingolstadt bis Regensburg laufen muß. Nur ins Handwerk läßt er sich weher pfuschen noch gucken, sein Vis-a-vis, ein armer Hiob ohne Kragen, der von früh bis spät alles liest, was er über Jesus von Nazareth ergattern kann, hat sich nur einmal ein wenig verschaukelt, indem er das Auge träumend auf den bunten Farbenkleben der Atlanten ruhen läßt, da leiht der Reiseberichte-Fabrikant auch schon los: „Mein Herr, Sie fixieren mich!“ Der arme Hiob schlägt demütig die Augen nieder. Wenn der Wanderer durch den Marfall mühte, was ich mir eben notiert habe!

### Buddha und der Marfall.

Im vorgangenen Sonnabend begann der große Auszug aus dem Marfall, denn der Bibliothek der Stadt Berlin hat man die Marmorstiege gegenüber dem Schloß gekündigt. Es war wie norm Schloßgehen, der Regen ließ sich nicht stören und klatschte unaufhörlich gegen die Scheiben, manchmal klappte ein Buchdeckel auf den kunstseidüberzogenen Tisch, dann blätterte jemand zwei Seiten um, die Lichter an den Wänden waren schon ausgelöscht, nur oben von der Kuppel herab brannte noch der große Leuchter. Eduard W., der irgendwo mal Bürgermeister gewesen sein soll, erzählt man sich, und heute ein achtzigjähriger Greis ist, hatte vorgekört und alles herbeigeschleppt, was an Literatur über den Buddha zu haben ist, und das ist nicht wenig. Keine Zeit und keine Stunde kümmerte ihn. „Ich schreibe ein großes Werk“, hatte er sich dem Garderobenfräulein annertraut, ganz auf die Nasenspitze war ihm die Brille gerutscht, und was er aufschrieb auf das vergilbte Papier, das buchstabierte er erst leise vor sich hin. Am Grunde genommen schreibt er gar nicht, sondern malt alles, bedächtig und behutsam, rauf, runter, rauf, runter, Buddha, Asien, Buddha, Asien...

# What should be done from now on

## Conclusion:

- We are deep inside the project but there still is a long way to go.
- Now is not the time to make conclusions but to ask questions to be answered in the future:
  - How good is the result of the OCR precisely and are there possibilities to optimize it?
  - Is it possible, to make use of the OCR-generated text in other research-projects?
  - Could the function of flexible content blocking be implemented successfully
  - Which desirable user features could be implemented at what expense?
  - What aspects of the now imminent workflow conversion brought real benefits?
  - How and to what extent has the user interest evolved?

# Thank you for your kind attention!

## Flug ins Unendliche

### Professor Piccard zum Stratosphärenflug gestartet

Muggsb., 27. Mai. (Eigenbericht.)

Am Mittwochmorgen wurde der wiederholt verschobene sensationelle Ballonaufstieg des schweizerischen Professors Piccard zur Fahrt in die Stratosphäre Tatsache. Punkt vier Uhr startete der kühne Forscher mit seinem Assistenten Dr. Kipfer in einem eigens für diesen Zweck konstruierten Ballon von dem Gelände der Muggsburger Ballonfabrik Niedinger aus, um in die bis heute unerforschten Regionen vorzustoßen. Trotz der bis zum letzten Tage geübten Geheimhaltung



Prof. Piccard

hatten sich viele hundert Menschen eingefunden, um Zeuge des bedeutsamen Momentes zu sein. Die kühnen Piloten erhoffen von ihrer Fahrt außerordentlich wichtige wissenschaftliche Entdeckungen. Der Aufstieg vollzog

aus Schupoleuten und Arbeitern der Ballonfabrik bestehende Startmannschaft waren um die Gondel beschäftigt. Man sah Professor Piccard im grünen Sportanzug und mit einer Zylinderhaube auf dem Kopf die letzten Anweisungen erteilen. Um 3 1/2 Uhr schlopfte er mit Ingenieur Kipfer in die

#### Gondel, die alsbald hermetisch geschlossen

wurde. Um 3.55 Uhr erfolgte kurze Kommandos. Unmittelbar darauf erhob sich der Ballon unerwartet schnell vor den Augen der ziemlich überfüllten Zuschauer in die Lüfte. Erst als er über dem Fabrikgelände schwebte, erfolgte lautes Händeklatschen. Der Ballon schlug zunächst nordwestliche Richtung ein, drehte sich dann aber in etwa tausend Meter Höhe nach Südosten und glänzte wie eine weiße Kugel in der Morgenionne. Er blieb lange den unbewaffneten Augen sichtbar und schwebte etwa anderthalb Stunden nach dem Start am südöstlichen Horizont in schätzungsweise bereits 5000 Meter Höhe und ungefähr drei Kilometer Entfernung. Der Ballon fliegt wiederum unter Schweizer Flagge.

#### Ballon außer Sicht.

Muggsb., 27. Mai.

Nach Schätzung von ballonfachverständiger Seite hat Professor Piccards Ballon, soweit man seinen sichtbaren Flug verfolgen konnte, die Grenze der Cirruswolken bereits überschritten und damit eine Höhe von 10 000 bis 12 000 Meter erreicht.

Professor Piccard hat kurz vor seinem Aufstieg die Absicht geäußert, spätestens um 11 Uhr zu landen. Die Landung dürfte nach seiner ungefähren Berechnung zwischen Basel und Freiburg im Breisgau erfolgen.

München, 27. Mai.

Von den Vertretern einer Muggsburger Zeitung, die den Flug Piccards im Kraftwagen verfolgten, erhält B.Z.B. um 11.50 Uhr folgenden Bericht:

Unsere Fahrt ging von Muggsb. zunächst nach Krumbach, hier wurde der Ballon um 6 Uhr gelichtet. Um 7 Uhr erliefen er über Kaufbeuren, zwischen 8 und 8.30 Uhr über Kempten, dann schlug er westliche Richtung ein. Es ist anzunehmen, daß er in der langen Zeit zwischen 7 und 8.30 Uhr, die der Ballon zur Zurücklegung der kurzen Strecke zwischen Kaufbeuren und Kempten brauchte, in die Stratosphäre aufgestiegen ist. Dann ging der Flug über Remminger, Leutkirch, Ravensburg und Friedrichshafen. Vorausichtlich hat der

Ballon nunmehr die Wolkenschicht, die sich nach Schätzung Sachverständiger in einer Höhe von 10 000 bis 15 000 Meter befindet, durchstoßen. Zur Zeit ist der Ballon außer Sicht; er treibt vermutlich in westlicher Richtung auf das Rheintal zu, so daß die Landung wahrscheinlich dort oder in den Ausläufern der Hohefen stattfinden wird. Nach Angabe Piccards sollte er zwischen 11 und 1 Uhr landen, doch erklärte er noch kurz vor dem Aufstieg, daß er, wenn die Verhältnisse günstig seien, drei Stunden sich in der Stratosphäre aufhalten wolle.



Der Stratosphären-Ballon