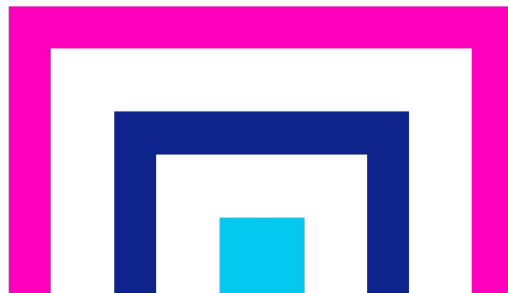


Putting together the wartime propaganda puzzle

Mari Wigham, Marjet Brolsma, Rana Klein, Roeland Ordelman



Clariah



THIS PRESENTATION

- Background: Media War Matching project
- Publishing radio transcripts
- Linking propaganda collections
- Demonstrating added value
- Findings and conclusions



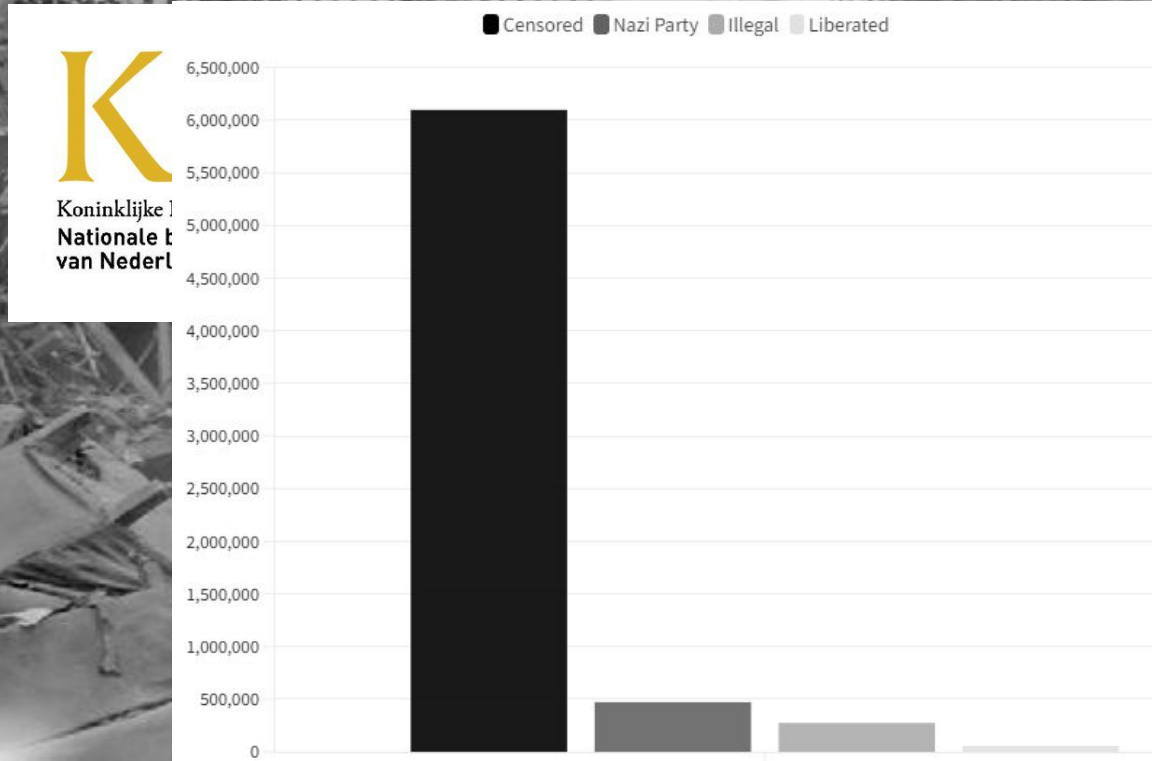
WHY WWII PROPAGANDA?

AP[WORLD](#) [U.S.](#) [ELECTION 2024](#) [POLITICS](#) [SPORTS](#) [ENTERTAINMENT](#) [BUSINESS](#) [SCIENCE](#) [FACT CHECK](#) [ODDITIES](#) [BE WELL](#)[POLITICS](#) | [UNITED STATES OF AMERICA](#)

Biden says Trump using Hitler's language with 'Reich' post

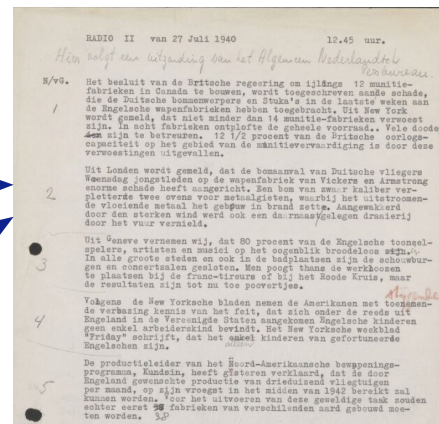
Trump says Biden is running a 'Gestapo' administration. It's his latest reference to Nazi Germany

WHY IS IT A PUZZLE?



d
ar,
d
ies

68,986

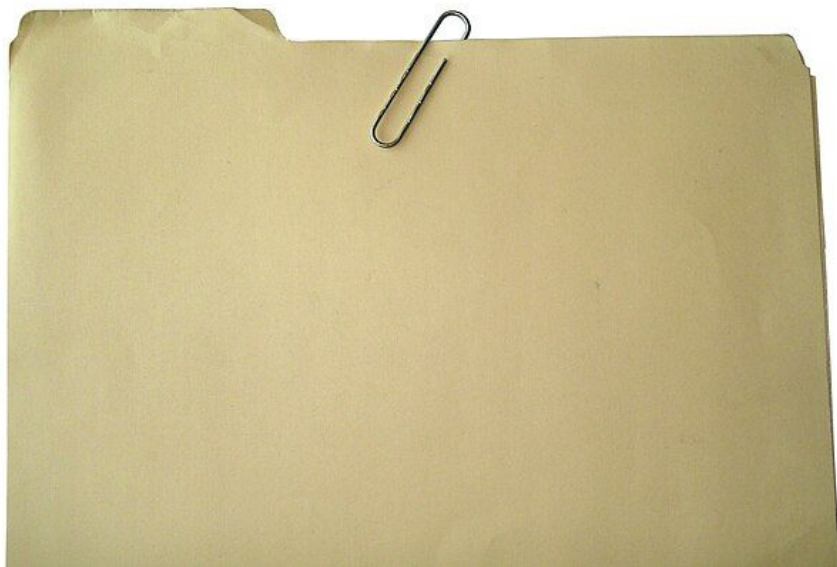


- Publish radio transcripts
- Link propaganda collections
- Demonstrate added value

PUBLISHING TRANSCRIPTS

The ideal

328 1941, 13-17 november



13th November
1941 - Radio 1



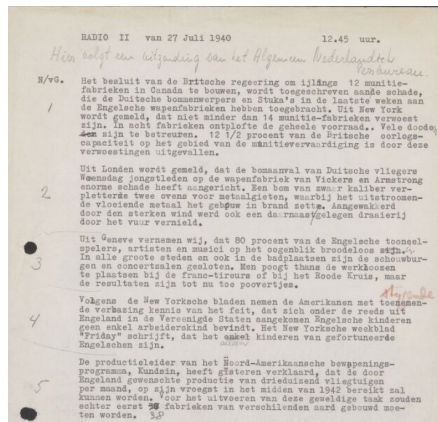
13th November
1941 - Radio 2



14th November
1941 - Radio 1

PUBLISHING TRANSCRIPTS

The proposed method



RADIO II van 27 Juli 1940 12.45 uur. -C^az-
/ct'/jövjöÖxxaÖ £xaz. ttvu J^^JCA^Lou^cL/t- L
vG. Het besluit van de Britsehe regeering om
ijlffings 12 munitie fabrieken in Canada te
bouwen, wordt toeges

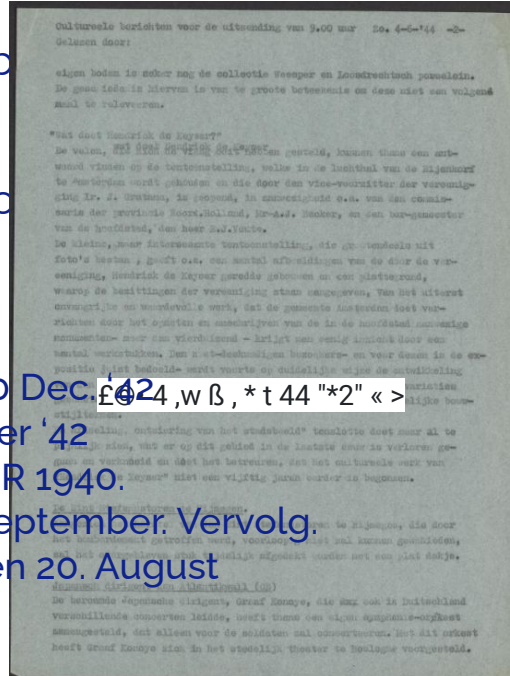
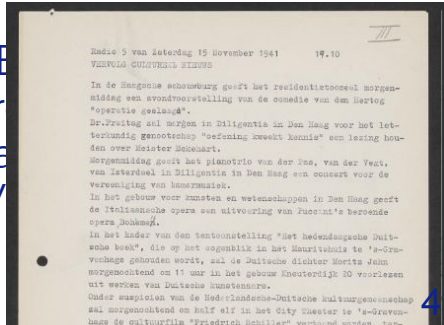
NER

Radio 2
27th July 1940

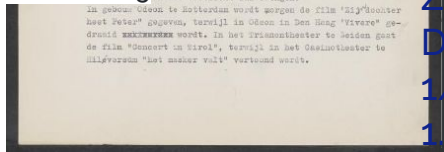
PUBLISHING TRANSCRIPTS

The challenges

- NE
 - Br
 - Da
 - W
- ates and b
hilar
rieval algo



NIOD Amsterdam, archief 103 Nederlandse
Radio 5 van Zaterdag 15 November 1941
VERVOLG CULTUREEL NIEUWS 1^10 In de
Haagsche schouwburg geeft het
residentietoneel morgen middag een
avondvoorstelling



4-6-'44
20 mei 1941
Maandag 20 Dec. 1941
20 December '42
7 DECEMBER 1940
Zondag 7 September. Vervolg.
Dienstag den 20. August
1/1 2/1941
11.'41

PUBLISHING TRANSCRIPTS

Date extraction

RADIO II van 27 Juli 1940 12.45 uur. -C^az-
/ct'/jövjöÖxxaÖ £xaz. ttvu J^^JCA^Lou^cL/t- L/
vG. Het besluit van de Britsehe regeering om
ijlffings 12 munitie fabrieken in Canada te
bouwen, wordt toeges

27th July 1940

Manual annotation for 3
complete folders, 100 pages per
folder for 2 folders per year

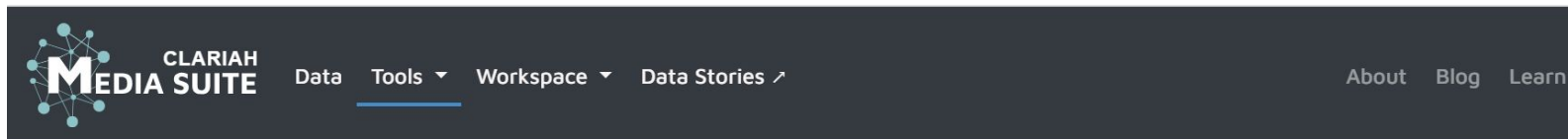
90-100% accuracy
10-50% coverage

For 62% of missed pages the
OCR was unreadable for a human



PUBLISHING TRANSCRIPTS

The reality




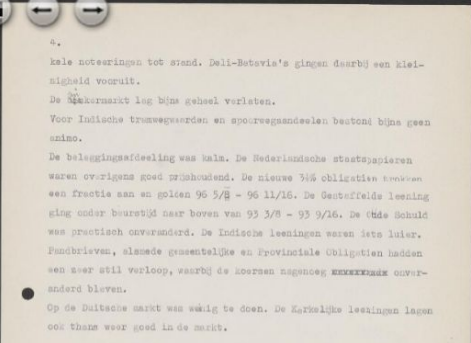
Resource Viewer

328 1941, 13-17 november - page 127

i Playlist
 Page 126 ▾

< >

Metadata *i*
 Content Annotations *i*
✕

Title

328 1941, 13-17 november - page 127

Date *i*

1941-11-13 - 1941-11-17

Description

kele not geringen tot stand. Deli-Batavia's gingen daarbij een klei nigheid vooruit. De t^hkermarkt lag bijna geheel verlaten. Voor Indische tramweg@aarden en spoorwegaandeelen bestond bijna geen anim (...)

PUBLISHING TRANSCRIPTS

Data/tool criticism

 Dataset

 Groups

 Activity Stream

 Manage

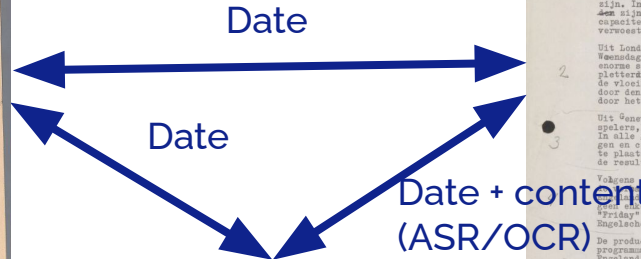
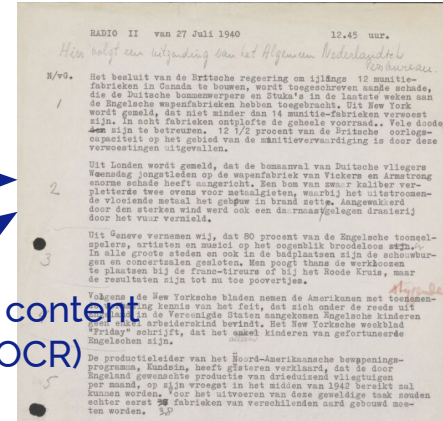
Manual page date

The dates in the Manual page date (ocrDateManual) field were added in 2023 by a researcher who is an expert in wartime propaganda. The dates were determined by reading the top of the scanned page to obtain the date. Pages without dates were assumed to have the same date as the previous dated page. Obvious errors, such as a page labelled Saturday 13th February 1944 within a group of pages labelled Saturday 12th February 1944, were corrected. Where two dates were listed on a page, the top date was used. No attempt was made to verify dates by checking the page content. Manual dates are available for 6% of the total collection, focused on specific folders.

(Berichtendienst Nederlandsche Omroep). The BNO operated under the aegis of the occupation regime in the Netherlands during the Second World War and actively made propaganda to promote national-socialism in its news broadcasts. The collection is provided by the NIOD (the Dutch Institute for War, Holocaust and Genocide Studies), [collection 103](#), [Nederlandsche Omroep](#).

The collection consists of images of the scanned transcripts. The transcripts have been processed by the NIOD using OCR

LINKING ITEMS



The challenges - insurmountable
(for now)



"Man met een man af.
Ik beleg en de uit."



ADDED VALUE

Comparison of search results



Query #1: Radio Europa ■

COLLECTION: Sound And Vision Archive
QUERY TERM: Europa

Query #2: Nazi Kranten Europa ■

COLLECTION: Newspaper Collection
QUERY TERM: Europa

Query #3: Bno Europa Page Date ■

COLLECTION: Wartime Radio News Bulletins: Berichtendienst Nederlands (BNO)
QUERY TERM: Europa

Number of items containing the word 'Europa'

ADDED VALUE

Analysis

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-------------------|--------|--------|---------|--------|--------|--------|---------|-----------|----------------|-----------|
| Anti-nazi kranten | west | oost | oorlog | rijk | geheel | landen | bevrijd | heel | vaste- land | nieuwe |
| Pro-nazi kranten | nieuwe | geheel | strijdt | nieuw | west | oost | landen | oorlog | midden | vasteland |
| BNO | nieuwe | geheel | oost | west | landen | oorlog | nieuw | vrijheid | vasteland | heel |
| ASR | heel | oost | west | landen | we | zuid | binnen | Nederland | vanuit | rest |

Top 10 most frequently occurring words before 'Europa'

FINDINGS & CONCLUSION

Challenges

- Poor source quality
- Large variability in the radio transcripts

... so couldn't achieve dream of linking it all together



I, Luca Galuzzi, CC BY-SA 2.5

FINDINGS & CONCLUSION

Achievements

- Transcript collection searchable
- Transcript collection browsable
- Partial dating enables partial linking
- Partial dating enables comparison
- ASR and OCR enable analysis

Lessons learned

- Quality starts with the source
- Data/tool criticism vital
- Cooperation between data engineers and historical experts essential!



Niklassletteland, CC BY-SA 3.0