

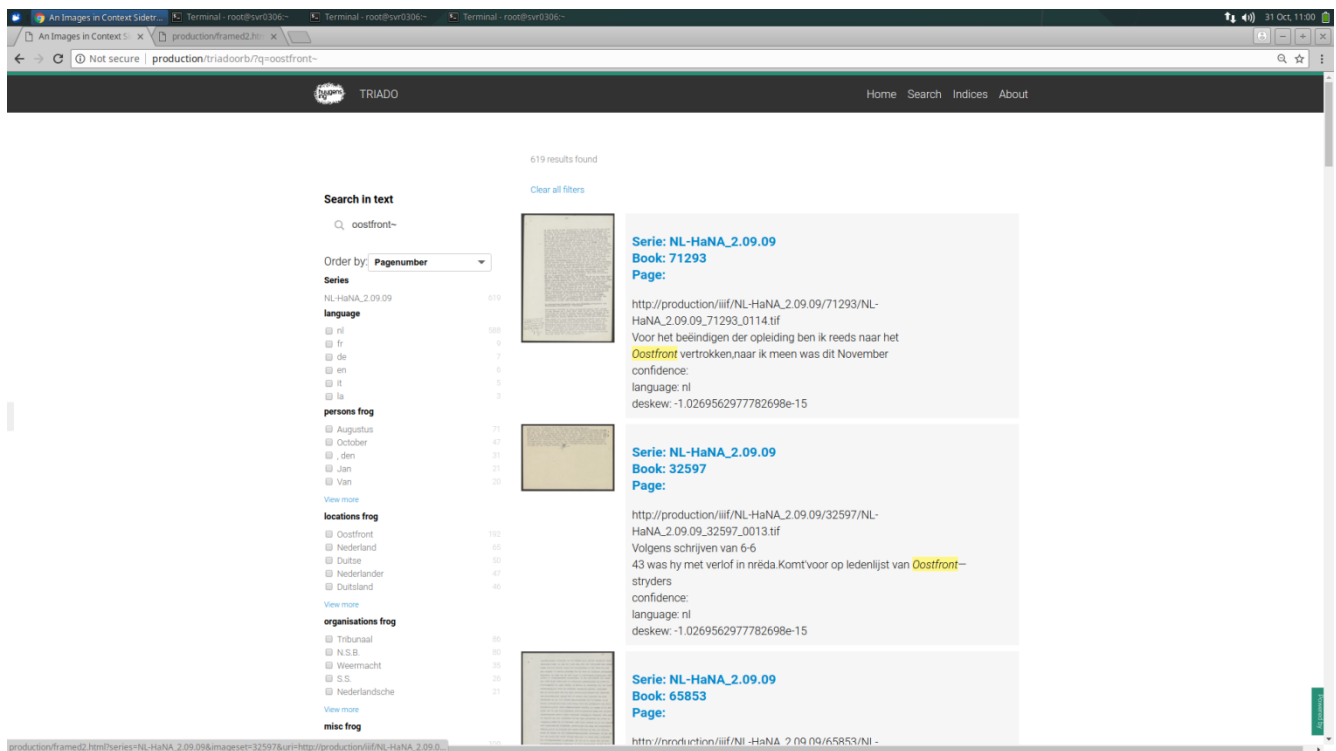
Bijlage D: TRIADO demonstrator

Amsterdam, 7 maart 2019

Achtergrond

Voor de verrijkingfase van TRIADO is in een streng beveiligde omgeving een demonstrator gebouwd waarmee door de 13,8 meter gedigitaliseerde steekproef kan worden gezocht. De onderliggende machineleesbare tekst bestaat uit verschillende lagen: de resultaten van Abbyy Finereader (Abbyy CLI OCR 11), Tesseract (Tesseract 4.0.9 rc4) en Tesseract op images die vooraf bewerkt zijn ('darkened'). Aan de linkerhand zijn facets gemaakt voor inventarisnummer, taal, personen (FROG), locaties (FROG), organisaties (FROG), producten (FROG), namen in NDVS en CABR-database. De zoekfunctie staat toe om fulltext door het gehele corpus te zoeken met bijvoorbeeld Levenshtein-distance. De resultaten worden gehighlight in fragmenten van de machineleesbare tekst.

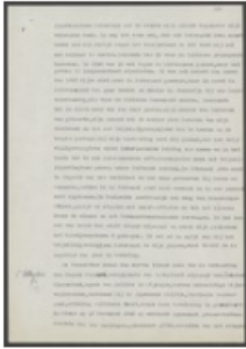
Bijvoorbeeld hieronder in deze screendumps.¹ Zoeken op het woord 'Oostfront' levert 619 resultaten op:



The screenshot displays the TRIADO search interface. At the top, the search bar contains the query 'oostfront-'. The results are sorted by 'Pagenummer'. The interface shows a list of facets on the left, including 'Series', 'language', 'persons frog', 'locations frog', 'organisations frog', and 'misc frog'. The main content area displays three search results, each with a thumbnail image of a document page and a snippet of text. The first result is for 'Serie: NL-HaNA_2.09.09', 'Book: 71293', and 'Page:'. The snippet reads: 'http://production/iif/NL-HaNA_2.09.09/71293/NL-HaNA_2.09.09_71293_0114.tif Voor het beëindigen der opleiding ben ik reeds naar het Oostfront vertrokken, naar ik meen was dit November confidence: language: nl deskew: -1.0269562977782698e-15'. The second result is for 'Serie: NL-HaNA_2.09.09', 'Book: 32597', and 'Page:'. The snippet reads: 'http://production/iif/NL-HaNA_2.09.09/32597/NL-HaNA_2.09.09_32597_0013.tif Volgens schrijven van 6-6 43 was hy met verlof in nreda Komt voor op ledenlijst van Oostfront- stryders confidence: language: nl deskew: -1.0269562977782698e-15'. The third result is for 'Serie: NL-HaNA_2.09.09', 'Book: 65853', and 'Page:'. The snippet reads: 'http://production/iif/NL-HaNA_2.09.09/65853/NL-'. The interface also shows a '619 results found' message and a 'Clear all filters' link.

Deze resultaten zijn afkomstig uit de door de machineleesbare tekst. Het fragment waar het resultaat is gevonden, wordt kort in een zogenaamde 'snippet' getoond.

¹ Vrijgegeven door het NA, gecontroleerd op privacygevoelige informatie.



Serie: NL-HaNA_2.09.09

Book: 65853

Page:

http://production/iiif/NL-HaNA_2.09.09/65853/NL-HaNA_2.09.09_65853_0059.tif

kwam het

vrijwilligerslegioen geheel onder Duits ehe leiding. In Februari i944 werd

ik ingezet aan het **Oostfront**

confidence:

language: nl

deskew: -1.0269562977782698e-15

De link van dit resultaat verwijst door naar de scan van deze pagina binnen de context van de andere documenten zoals hieronder (ander archief, hetzelfde principe)

1920 x 1048 x 24 BPP Not a file 88 % Not a file 5,76 MB Not a fi

Aan de linkerkant worden de filters zichtbaar en tussen haakjes het aantal treffers per zogenaamde 'facet' aangegeven.

organisations frog

<input type="checkbox"/> Tribunaal	86
<input type="checkbox"/> N.S.B.	80
<input type="checkbox"/> Weermacht	35
<input type="checkbox"/> S.S.	26
<input type="checkbox"/> Nederlandsche	21

[View more](#)

[View more](#)

Soort document

<input type="checkbox"/> Proces-verbaal	330
<input type="checkbox"/> Dagvaarding	90
<input type="checkbox"/> Staat van Inlichtingen	33
<input type="checkbox"/> Uitspraak	30
<input type="checkbox"/> Besluit	30

[View more](#)

Met deze filters kunnen combinaties worden gemaakt en binnen de gevonden resultaten eigen selecties worden gemaakt. De zoekresultaten aan de rechterkant in het scherm passen zich aan

locations frog

- Oostfront 192
- Nederland 65
- Duitse 50
- Nederlander 47
- Duitsland 46

[View more](#)

organisations frog

- Tribunaal 86
- N.S.B. 80
- Weermacht 35
- S.S. 26
- Nederlandsche 21

[View more](#)

misc frog

- April 109
- Maart 94
- Aan 45
- Augustus 28
- Tribunaalbesluit 22
- Vaderland 22
- Commissaris 20
- Duitschers 17
- Politie 13
- Voorzitter 13

[View more](#)

product frog

- Mei 157
- Volk 44
- Zoo 25
- Van 24
- Dossier 21
- October 19
- Nederlands 7
- Dagblad 6
- Ter 5
- Verwundetenaandekens 5

[View more](#)

cabrsimplematches

- Kleijn 1
- R. van Koert 1
- Gerrit 1
- Edwin Kleijn 1
- Naarden 1
- Vaderland 1
- Scherer 1

http://production/iiif/NL-HaNA_2.09.09/32597/NL-HaNA_2.09.09_32597_0013.tif

Volgens schrijven van 6-6
43 was hy met verlof in nrëda Komt'voor op ledenlijst van **Oostfront**-stryders
confidence:
language: nl
deskew: -1.0269562977782698e-15



Serie: NL-HaNA_2.09.09
Book: 65853
Page:

http://production/iiif/NL-HaNA_2.09.09/65853/NL-HaNA_2.09.09_65853_0059.tif

kwam het
vrijwilligerslegioen geheel onder Duits ehe leiding.In Februari i944 werd ik ingezet aan het **Oostfront**
confidence:
language: nl
deskew: -1.0269562977782698e-15



Serie: NL-HaNA_2.09.09
Book: 57267
Page:

http://production/iiif/NL-HaNA_2.09.09/57267/NL-HaNA_2.09.09_57267_0034.tif

Voor Leider, Volk en
Vaderland viel aan het
OostfrontOostfront den
hedendood heeft gevonden.,Voor Leider, Volk en
Vaderland viel aan bet
Oostfront
Kameraad
ff Sturmman
geboren 28 Feb,Voor Führer, Leider,
Volk en Vaderland viel
aan het **Oostfront** onze