THE ENHANCEMENT OF THE SEARCH-ABILITY OF OCRED TEXTS

HTR+, segmentation and meta-dating early modern ordinances

KB - NATIONAL LIBRARY OF THE NETHERLANDS (DEN HAAG)¹, GHENT UNIVERSITY², ERASMUS UNIVERSITY ROTTERDAM³ CHRISTEL ANNEMIEKE ROMEIN¹²³, SARA FLOOR VELDHOEN¹, AND MICHEL DE GRUIJTER¹

Hypothesis: When problems arose, small 'states' had to act swiftly. Governments may have adopted successful legislation from neighbouring areas. Hence 'entangled histories'. The project uses printed books of ordinances from the Low Countries (1500-1800s) to answer this.

Using Computer Technologies to read and follow Early Modern **Political Legislation**



Early modern ordinances were affixed to 'known places' and proclaimed by the city crier. (Paalhuis c. 1660).

Provincial governments decided to bundle the most important rules in books of ordinances (in Dutch: plakkaatboeken). These are not complete, but do give an overview of what government officials thought to be the most important rules that had been cast. Not all are digitised.

The set used comprises:

Steps within the project

en dat op Velwwe nyemant fal mogen huyahanden handen , ten maere die tavee voorft

clauwen tottet leste lit toe In affgeknypt.

By den Keyler.

die men by ellick Stadt, Scholtus und Richter van

peene van vier Heeren ponden, soo dick und

fal werden hier tegen doende, van welcke vier

ponden wy die twee, und die aenbrenger der

overtrederen deles gebodts, die andere twee

hebben fullen. Und fal dit ghebot vast und

werden, ter tijt toe anders daer inne verordent

fal werden. Beyelen derhalven u Wilhem van

Scherpenzeel Droft op Veluwen, Johan Mom

Richter tor Arnhem und Velawen-Zoom, und

voirt allen anderen Scholriffen , Richteren und

Officieren, dat ghy ernitelick opfihen hebben, dit onfe gebodt vast stede und onverbreecklijck

gehalden worde, und foe derhalven by u ec-

nigh gebreek bevonden wurde, fullen wy veroir-

facekt worden fullix aen n te laeten verhaelen,

Gegeven onder onffen op gedruckt fecreet Segel

tho Arnhem, den xxIV. dach Septemb. xv*

xLv11. In margine stons, Dit is noch vernieuwt

Placaet en ordonnantie op die Tol-

Philips van Lalaingh, Grave van

Hooichstraten, &c.

Y doen kondt allen den ghenen die dit

den, Dat wy van wegen Key. Maj". oofes allerge-

nedichsten Heeren, den Tolner, Besiehre ende Tol-

schrijver van siiner Majis, groten Gelderschen tol tor

Nymegen mit allen ernst belast ende bevolen heb-

ben, belaften ende bevelen by defen, dat fy in 't

vertollen van aflen schepen, koopmanschappen en-

tegenwoordich Placaer gethoent fal wer-

gewood av. Marty 1548.

mehmael toe verboeren als ymants bevonden

1. Within 'Entangled Histories' we apply the P2PaLA-tool of the University of Valencia to recognise sections in the text.

Groot Gelders

van Gelre ende Grave van Zurphen gelaft ende

ven ende aenteykenen alle goeden die fy verrol-len fullen, in wyns schip die goeden sijn, wiese toekommen, in wat Steden off plactsen die schip-

pers en coopluyden off eygenaers van den goeden

de die opvaerende goeden eick by lich felven,

Van gelijken fullen die tolfchryvers haer con-

trebouck alleen halden, en daer inne op 't aengeven ende in bywefen des belienders aenteykenen. den goeden die in eleke fehip bevonden fullen

werden, wiese toecommen, waer die schippers en coopluyden off eygenzers dez goeden woen-

achtich fim, ende hoe viele yeder den tollenaer

van eleker parthien van goede beraelt heeft, fon-

der den tollenaers haer bouck ofte opschrijvon-

gen to thonen ende openen, ofte dat een bouck

bouck van alle 't gunt voor den felven tolle ge-

Teden ende verrolt fal fijn , den commiffarien ofte an-

dere des laft-hebbende toegesloten aver seynden

fullen, gereyckent ende geattefteert mit huer felfs

handt en de handt des besienders in deser ma-

Wy A. tolfchrijver, ende B. beliender van den

groten Gelderschen tolle tot Ysselssort, affirmeten

ende getuygen by onten eeden ende confeientien,

dat geen ander, noch meer onvrije goeden voor

desen tolle gepasseert noch vertolt sim geweest,

ende dat oick niet meer binnen den tijt voirfe.

daer van ontfangen is, dan hier voren in onfen

quoijer aengereyckent staet, t'Oirkoade onfer beyder handrichrifte hier ondergestelt, den dach

Item dat die tollenaers alle maenden, off des by

den commissarien off anderen bevel daer van hebbende vermaent fijnde overfeynden fullen, alle die

pennongen by henluyden ontfangen, ende haer re-

keninge macken uyt haer felfs aentevekenongen off boucken, ende niet uyt des tolfehrijvers boucken,

omme die selve huere rekenonge tot gesinnen des commissarien te commen presenteren, en doen vol-

Item foe men bevindt, dat in den Jaere xys. xxxij.

by wielen Heer Kaerlen van Egmont, in fijnen tije

regierende Furst der Landen van Gelre en Zurphen.

een ordenough gemaeckt is geweelt, waer nae die

Tchipperen en coopluyden by den Amptluyden van

de tollens vertolt fouden werden, ende hoe die fel-

vige Amprluyden fich dair inne halden follen; Dat

gens huer verschrijvonge

de waeren voir-by den vuriz tol neder off opwaerts | men oick onderricht is , hoe beven op den Rhijn,

vaerende ende pafferende, fich regulieren nae in- nyet meer dan vier enekel voeder wyns voer een haldt der ordinantie, op den vuriz tol, van we- tolbaer voeder geleden wordt, Ende die olde boue-

gen Key. Majes opgricht ende gemacekr, ende van ken vermelden dat van olts nyet meer dan thien off

Item dat die tolfchrijvers t' allen tijden des vermaent wefende haer uytfchrijvongen ende contre-

na des anders bouck te richten.

2. Creating "Handwritten Text Recognitionmodels" with Transkribus.

- language (Dutch/French)
- font (roman/gothic/variations)
- character Error Rate <5%

This data we can export as XML/TEI files (but also as PDF/A or Word-files).

- 3. Machine-learning categorisation through the use of ANNIF through the GhentHPC: - hierarchical categories created by the MPI für europäische Rechtsgeschichte (1800 in total)
- per individual law
- based on manually labelled data (pre-trained)

Annif (annif.org)

lesults:

- searchable data.
- visualisations through plotting
- on maps (Palladio for example). data will be connected
 - through Linked Data and URI's with the dataset in Frankfurt.



e.g. Palladio -Stanford

108 Books:

- font:

- language:

words:

- Dutch 67
- French
- Latin
- Mixed 14
- 75.372 pages: ± 550 mln.

Project-team

Michel de Gruijter – Project Advisor

Annemieke Romein – Primary Investigator

Sara Veldhoen Research Software Engineer

Contact

Annemieke.Romein@ugent.be https://research.flw.ugent.be/nl/annemieke.romein www.caromein.nl



in Christel Annemieke Romein









