

## **Bio (max 100 words)**

Annemieke Romein (Ph.D. Rotterdam, 2016) obtained a Rubicon-grant (round 2016-3) and now works as a post-doctoral researcher at Ghent University's History Department and affiliated researcher of the Erasmus University. Her Ph.D.-thesis is a comparative study on 'fatherland terminology' in German and French political communications (1642-1655). Her current research focusses on creating order and safety in the burgeoning states of Flanders and Holland between 1579-1701, by studying and comparing normative texts and visualising the output in graphs.

Annemieke took part in the Oxford Digital Humanities Summer School ([DHoxSS](#)) 2018 and now actively organises DH-activities: Transkribus-workshops and a [DH-lecture-series](#) *Digitising the Humanities*.

## **2a. Title of the project**

### **Entangled Histories of Early Modern Ordinances. Segmentation of text and machine-learned meta-dating**

## **2b. Abstract (max 250 words)**

Please provide an abstract of your proposal of maximum 250 words. Please note this abstract will be published on the KB Lab website.

This project aims to broaden the scope of the KB-Lab towards the early modern period. It involves the (1) improvement of the currently applied OCR-technique to a much higher recognition-standard with HTR. It (2) enhances readability by systematically segmenting individual texts, recognising text-sections – beginning or end, columns, titles, dates, summaries, the body of the text. In order to improve the searchability, I suggest the (3) application of a standard categorisation (metadata) with a machine-learned algorithm. A categorisation by a machine-learned algorithm will offer ample possibilities to computer-search for similar topics within texts and do content-based longitudinal searches, whereas the actual title may not be so helpful to modern readers.

The KB-library hosts at least 42 digitised *plakkaatboeken* (bundles of normative texts – ca.1540s-1800s) and near 5000 individual *plakkaten* (16<sup>th</sup>-19<sup>th</sup> century). These texts contain indications of how governments of burgeoning states dealt with unexpected threats to safety, security, and order through home-invented measures, borrowed rules, or adjustments of what was established elsewhere. The hundreds of texts within these books are frequently consulted by researchers of various disciplines to unravel rules for controlling society. Having the possibility of a longitudinal search based upon contents rather than the index or title, as well as having an overview based upon several states (e.g. provinces) has been impossible so far. It will disclose the entangled histories of neighbouring provinces, due to synchronic and diachronic comparisons of approximately 15.000 texts on location – allowing a wider search and implementation in other projects Europe-wide.

## **3. Project description (max 1.500 words for sections 3a-c)**

### **3a. Background and research question**

Please give a short introduction on your project, building up to your research question.

Law and order have been key issues in all times and places, whether it be in ancient Rome or present-day threats of public disturbance, crime or terrorism. Order – or the regulation within society to secure what is believed civility – prevents chaos from erupting and protects inhabitants from harm. Protecting public safety from criminality and protecting the social order was the major concern of governments during the sixteenth until the eighteenth century.

Early modern European states struggled for survival, making it impossible to ‘reinvent the wheel’ each time a problem arose. Hence, it was of tremendous importance to copy, adapt and implement normative rules that were already proven successful elsewhere – at least, that is a hypothesis I seek to verify. With this Researcher-in-Residence project, I want to shed light on the entangled history of state-building and governance by using the *plakkaatboeken* from various provinces from the Republic and the Habsburg Netherlands. I would very much like to visualise how problems rose and solutions spread through the Low Countries – and possibly expand that to adjacent German territories.

### **3b. Theoretical background**

Please provide a brief theoretical background on your project

#### **Techniques**

OCR has focussed solely on deciphering individual characters, while HTR also works with the context to create a digitised version of a text. As digitising a text is not the same as improving its searchability, topic modelling is increasingly applied (Leydesdorff 2017) though it seems that it is not that often applied to full texts but rather to small parts.

#### **Casus**

How can a government protect its inhabitants from harm? Upholding social order by preventing crime and violence is arguably a key priority of governments today, especially under the current threat of terrorist attacks. During the sixteenth and seventeenth century it was just as important, as burgeoning states faced threats of increased warfare. Social order can be understood as the mechanism of interlinked institutions, social structures, rules, customs, values and practices that help to regulate the behaviour and activities of inhabitants (Vermeir 2005). In early modern Europe, it was seen as a basic necessity to organise the protection of personal property and safety in order to accomplish a good life, welfare, and strive for the common good. Around the 1500s a shift in political thinking occurred towards a (1) function-oriented interpretation of rules and the role of government. With this shift arose a (2) government ruling over its subjects – slowly replacing the feudal hierarchal traditions within a (3) agreed-upon fixed geographical space (Von Friedeburg 2016). These three developments occurred simultaneously with the introduction of a new type of rules, known as *Politia* (Lat. *Politia*, Fr./Eng. Police, Germ. *Policey*) or *police ordinances*. Such comprehensive normative rules had consequences not only for the former organisational structure of the nobility (and clergy) ruling as lord and master within their own dominion but also bestowed new responsibilities upon secular political-legal institutions. Moreover, the announcements of new rules throughout a jurisdiction grew in importance, which was increasingly amplified due to a rise in literacy; an announcement’s importance had to be clear to the larger audience. Finally, police ordinances were the precursors of a large part of modern legislation, in particular, administrative law: having norms on the environment, production, trade and marketing, product safety, health, building and many more.

Such a general development of secular institutions and normative rules throughout Europe gave rise, initially, to the notion that this phenomenon was triggered by authoritarian-princely governments or small-scale city-governments who were able to coerce their subjects towards an ideal society (Härter). In recent years, however, studies have shown that since the 1500s more large-scale, and non-authoritarian societies, were able to coach their subjects towards a normative society. It may well be that some organisational structures (e.g. small principalities or city-governments) fulfilled a pioneering role within the development of this new normative trend. Unfortunately fragmented, traditional methods of research do not allow us to draw conclusions yet, although several strong indications of cross-border influence have arisen in ongoing research projects (Kotkas 2014; Iseli 2003; Thomas Simon 2004; Pihlajamäki 2017).

### 3c. Methods and techniques

Please explain which methods or approach you will use to successfully complete your project

In this section I would first like to explain the traditional techniques in order to demonstrate why the field of early modern political(-institutional) history and legal history still have a long way to go in digitising their sources and accessibility. Second, I will explain what can be achieved by applying the methods emerging from the field of Digital Humanities (DH).

#### Traditional Techniques

Many of the aforementioned historical developments became apparent through the research of Michael Stolleis (1988) and research carried out at the Max-Planck-Institute für europäische Rechtsgeschichte (Frankfurt a/Main) in the 1990s. Here, a project on [Repertorium der Polizeyordnungen](#), focussing on *Policey* in the early modern era commenced leading to a repertorium of over 200,000 XML files on ordinances in several German areas, the Swiss Confederation and Denmark. The MPIER opted for XML as they could not envision the leap OCR and HTC would make in digitising sources and the ability to machine-read them.

In Belgium, the [Royal Commission for Publication of Ancient Laws and Ordinances](#) has been occupied with transcribing ordinances ever since 1846. These publications have all been compiled in books and it was not until early 2018 that the commission decided to change their policy and make the sources online available.

In the Netherlands, many of the very often used *plakkaatboeken* have been digitised and the Huygens ING's [Resolutions of the Estates General](#) formed an ongoing project since the 1915s, now the [REPUBLIC-project](#). As these projects in the Low Countries stand on their own, little is known about similarities and differences with the MPIER-cases.

Early modern legal history or political-institutional history has waned in popularity over the past decades, mainly due to limited accessibility and searchability of the sources. When digitised, it is often not more than 'digital paper' as it lacks (uniform) metadata. As several projects started over a century ago, the techniques could certainly be improved. This is particularly necessary since the field lacks databases that are available for other types of written publications ([EMLO](#); [STCN](#) or [EEBO](#) or [EDBO](#), [TEMPO](#)). This is striking as its sources – ordinances – are abundantly used in various types of research: legal, historical, linguistic, economics, and philosophy.

#### Digital Techniques

In order to (re)trace texts among various volumes of ordinances (*plakkaatboeken*) a uniform manner of categorisation seems to be a useful tool. When it comes to early modern texts, then ascribed topics and modern-day keywords (Härter, Kotkas 2014) can differ hugely due to, for example, the development of the language, norms, or by simply having an overview. In fact, the titles of the *plakkaatboeken* have all been spelled differently (e.g. plakkaat; placate; placacet; plakaet, placcaet or boek; boeck), which makes it hard to even find them within the KB catalogue.

This project consists of three steps: 1. Improving the OCR-techniques and 2. segmentate the texts. 3. With [machine-generated metadata](#) or categorisation, based upon the already developed genre classifier (Lonij and Harbers, 2016) that can be fitted to [automated content analysis with the aim](#) to train a computer to recognise the conditions that the categories were based upon. After [training with human-defined](#) examples (from the applicant's ongoing Rubicon research), the computer can then suggest, apply and supplement categories to other texts based on the idea of topic modelling (Leydesdorff 2017). It is assumed that eventually, it will be possible to apply the approach to sources in other languages too. An important new development in this process is to go [from sentence-recognition to article-segmentation](#) of larger texts; this requires that the computer be trained to recognise the beginning and end of texts, either as chapter or as individual text within a compilation of texts. Initially, this can be based upon image/layout but could evolve into integration with HTR, depending on its feasibility. This can be achieved by applying [Transkribus](#), for example. It can not only be used for segmentation of texts, but this Handwritten Text Recognition (HTR) program can also be used to search for similarities in characters, and as such can serve as a substitute for Optical Character Recognition (OCR).

At the Digital Infrastructure of the KNAW Humanities Cluster, a web-based tool under the name of *Pergamon Images* has been developed and applied to early modern texts reaching the precision of approximately 98% in deciphering the characters. Whether or not this feature is available to apply on external projects just yet, it is scheduled for release in December 2018. A self-learning mode within the app to be designed could be able to incorporate the HTR-OCR for future purposes as well.

Due to its significance for such a broadly studied range of sources, the tool should be presented in an RDF (which would allow the use of tools based on the Linked Data paradigm), and stored in [Anansi](#) to allow researchers of other projects to apply the technology to their sources. With the data generated with this project's tool visualisations in [Nodegoat](#), for example, would become possible.

## 4. Outcomes

Please describe what the outcomes of the project will be & how you will use the KB data.

### *Outcomes*

1. Improved performance of the current OCR-recognition of (early modern) texts by incorporating HTR-techniques on printed texts.
2. The possibility to automatically recognise segments in text layout: beginning or end, columns, titles, dates, summaries and the body of the text.
3. Expansion of automatic content analysis based upon segments, rather than on lines or sentences, with a machine-learned algorithm, and applying this with a standardised machine-learned metadata to early modern normative texts. This will allow researchers to search through approximately 15.000 ordinances and resolutions from various provinces in the Low Countries, accessible through uniform search terms.
  - An RDF-compliant tool to enable application in other ongoing and future projects.
  - Integration of the enhanced datasets in the CLARIAH/CLARIN ecosystem

### *Textual output*

- 2 blogposts on researching with DH and early modern political sources.
- I am confident that a journal on legal history or political history will accept an article explaining the use of the tool, next to a possible similar publication in a DH-journal.

### *Wider implications*

The above will also lead to a use-case: the application of the envisaged technology and pipeline to the history of norms throughout the Low Countries. It will show how topics can be selected, when and where normative texts dealing with specific topics occur, and how they can be retraced quickly within the volumes or among the individual texts in order to compare them content-wise.

These techniques can, therefore, be applied to the repository that is to be created with the Huygens ING project *Republic* (digitising the Resolutions of the Estates General), which has recently received the NWO Groot Grant. One of the entanglements that may then come to light, will be the implementation of normative texts (ordinances/resolutions) of the Estates General within provinces of the Republic.

With the Belgium Royal Commission's gradual shift to go digital in the (near) future, an implementation there can be foreseen too. Furthermore, the Max-Planck-Institute für europäische Rechtsgeschichte (Frankfurt a/Main) has an interest in undertakings within this field too.

Being able to recognise categories and topics of ordinances more quickly, my research will benefit from the rapid expansion of accessible, searchable sources. It will no longer be 'digital paper', but meaningful research material that will allow me to unravel the entangled history of institutions and norms within the Low Countries – which will potentially be a stepping stone for future research projects. It also fits neatly into the current COST Action-proposal 'dMEANOR' (OC-2018-2-23156: Early ModErn goverAnce and NORmative Rules) of which I am the main applicant. It aims to connect researchers throughout Europe on governance, institutions and normative rules to one other in order to come to terms with the used terminology and unravel – based upon (mainly) analog in-depth research of cases studies – the entangled history of early modern good governance.

### *KB data*

The KB has a large collection of digitised *plakkaatboeken* as well as several relevant individual texts that may, or may not be part of these books. With such a large collection there should be enough material to train a computer to recognise categories. There are a minimum of 42 *plakkaatboeken* digitised and stored under various titles – and this could be even more – covering the early modern period. Furthermore, there are approximately 4958 individual *plakkaten* that were published during the 16<sup>th</sup>-19<sup>th</sup> century (stored in [Delpher](#)).

## **5. Link to the KB Research agenda**

This project broadens the scope of the KB-Lab towards earlier periods (early modern period) and widens it, as article segmentation and categorisation is an enrichment of its earlier developed *genre classifier* and *frame generator*. The technique will be applicable to any type of text, and in the future even to handwritten texts.

This project will, therefore, contribute to the KB Research Agenda ‘Access and Sharing’, because of the ability to search segmented texts better and apply automatic meta-dating.

Finally, better OCR-techniques applied throughout the collection will enable the accessibility of texts for blind and visually impaired users – regardless of when the texts were written/printed.

## **6. Workplan and timetable**

Please describe: 1) how you will work together with the KB team, 2) where or if you would need assistance, and 3) a short overview of the work per 2 months.

### **Work with the KB team**

In the past years, many very useful tools have been created by the KB-lab. The programmers, for example, Willem Jan Faber, are good at breaking down problems into small units that can then be used within a sequence to solve the bigger research questions. Hence, I am looking forward to working with Mr. Faber and his colleagues in 1) improving the OCR-quality of early modern printed texts; 2) segmenting the texts; 3) mining the texts with a machine-learned algorithm. I believe that these bits will not only work to solve the big research question that I am posing but will in themselves improve the accessibility of the KB’s services and possibly also have an impact outside the KB.

Working in a stimulating environment such as the KB will allow me to develop a different point of view on a research library, acknowledging there is much more going on behind the scenes than just lending books and preservation. I am looking forward to exchanging thoughts, inspiring conversations and stimulating debates with people that may hold an entirely different view of “my” sources that will refine my views and my research. I believe that working with people from various academic disciplines enriches research and may open up opportunities for further research and growth that I cannot yet imagine.

Moreover, learning more about tools and methodology that can be developed by using DH is an enrichment for research itself. I truly believe that if it wasn’t for a project such as the KB Researcher-in-Residence, the gap between early modern political history would become unbridgeable.

### **Assistance and co-operation**

As an early modern historian, I do not have a wealth of experience with programming. However, I do, however, bring a lot of knowledge regarding specific (digital) sources and have developed the ambition to fully engage in digital scholarship. With this knowledge and a general notion of what could happen if I were able to have an overview of publications by geographical area (province), I believe to be able of explaining the above plan in great detail.

I am still new to Digital Humanities, meaning that I am still searching for the correct DH-terminology. Or as some would say: I may be able to hum the tune, but I do not know the words (yet). Hence, I believe that the ‘interpreters’ (e.g. Lotte Wilms and Steven Claeysens) will be of unequalled importance to help translate this plan to a ‘programmers’-mindset. However, it should be stressed that a DH-project can

only succeed when programmers are at work, so I am surely looking forward to working with Willem Jan Faber and other programmers of the KB-library.

Pip Willcox and Dave De Roure said at the DHoxSS2018 that “those within humanities ask the good questions” – and those questions can be used for programming. This does not merely apply to the research one wants to do, but also to questioning other scholars’ ideas and motivations. I would be more than happy to assist those at the KB-lab to unravel the mysteries of early modern ordinances.

### Timetable

Months	Activities
1-2	Re-OCR-HTR texts and article segmentation: training a computer with Transkribus (text segmentation tool).
3-4	Metadating texts: training with prepared samples (Gelderland, Flanders, Holland).
5-6	Searching through other plakkaatboeken and improving the research-tools. Creating RDF, Linked-Data and generating comparable, visual output.

## 7. References

Please add a list of references cited in the proposal, with full bibliographic details

Websites have been indicated within the text with hyperlinks.

### Technique

- Digital Infrastructure - KNAW Humanities Cluster (Amsterdam), *Pergamon Images* (2018). (to be released around 1 December 2018).
- Leydesdorff, L., and A. Nerges. (2017) "Co-word maps and topic modeling: A comparison using small and medium-sized corpora (N< 1,000)." *Journal of the Association for Information Science and Technology* 68.4: 1024-1035.
- Lonij, J., Harbers, F. (2016), *Genre classifier*. KB Lab: The Hague <http://lab.kb.nl/tool/genre-classifier>
- Lonij, J., Eijnatten, J. van. (2016), *Frame Generator*. KB Lab: The Hague <http://lab.kb.nl/tool/frame-generator>

### Casus

- Iseli, A. (2003), “Bonne Police”: Frühneuzeitliches Verständnis von der guten Ordnung eines Staates in Frankreich. (Epfendorf/Neckar).
- Härter, K.; Stolleis, M.; Schilling, L. (eds.) (1996A), *Policey im Europa der Frühen Neuzeit* (Frankfurt a/Main).
- Härter, K. and M. Stolleis (1996B), *Introduction to Repertorium der Policeyordnungen der frühen Neuzeit, Band 1. Deutsches Reich und geistliche Kurfürstentümer (Kurmainz, Kurköln, Kurtrier)*, edited by K. Harter, (Frankfurt a/Main), pp. 1-36.
- Härter, K. (1999) “Social Control and the Enforcement of Police-Ordinances in Early Modern Criminal Procedure.” In: *Institutionen, Instrumente und Akteure sozialer Kontrolle und Disziplinierung im frühneuzeitlichen Europa/Institutions, Instruments and Agents of Social Control and Discipline in Early Modern Europe*, edited by H. Schilling (Frankfurt a/Main) pp. 39–63.
- Härter, K. (2005), “Edition, Repertorium oder Datenbank? Die Erschließung frühneuzeitlicher Policeygesetzgebung durch das Repertorium der Policeyordnungen.” In: *Vom Nutzen des Edierens: Akten des internationalen Kongresses zum 150-jährigen Bestehn des Instituts für Österreichische Geshichtsforschung Wien, 3.–5. Juni 2004*, edited by B. Merta, A. Sommerlechner and H. Weigl (Wien) pp. 131–144.
- Härter, K. (2010) “Security and ‘Gute Policey’ in *Early Modern Europe: Concepts, Laws and Instruments*.” *Historical Social Research* 35, pp. 41–65.
- Härter, K. (2010), “Die Verwaltung der ‘guten Policey’: Verrechtlichung, soziale Kontrolle und Disziplinierung.” In *Herrschaftsverdichtung, Staatbildung, Bürokratisierung. Verfassungs-, Verwaltungs- und Behördengeschichte der Frühen Neuzeit*, edited by M. Hochedlinger and T. Winkelbauer (Wien) pp. 243–269.
- Härter, K. (2010), “Statut und Policeyordnung: Entwicklung und Verhältnis des Statutarrechts zur Policeygesetzgebung zwischen spätem Mittelalter und Fruher Neuzeit in mitteleuropäischen Reichs-

und Landstädten.” In: Von der Ordnung zur Norm: Statuten in Mittelalter und Früher Neuzeit, edited by G. Drossbach,, (Paderborn) pp. 127–152.

- Härter, K. (2012), ‘Konfliktregulierung im Umfeld frühneuzeitlicher Strafgerichte: Das Konzept der Infrajustiz in der historischen Kriminalitätsforschung’, in: *KritV* 2 (2012), pp. 130-144.
- Kotkas, T. (2014) *Royal police ordinances in early modern Sweden: the emergence of voluntaristic understanding of law* (Leiden).
- Pihlajamäki, H. T. (2017), *Conquest and the Law in Swedish Livonia (ca. 1630–1710): A Case of Legal Pluralism in Early Modern Europe* (Leiden).
- Romein, C.A. (2016/2017): *Rubicon application*. Online: [http://www.caromein.nl/wp-content/uploads/2017/04/Rubicon\\_application\\_CARomein-website.pdf](http://www.caromein.nl/wp-content/uploads/2017/04/Rubicon_application_CARomein-website.pdf)
- Simon, T. (2004). “Gute Policey”: *Ordnungsleitbilder und Zielvorstellungen politischen Handels in der Frühen Neuzeit* (Frankfurt a/Main).
- Stolleis, M. (1988). *Geschichte Des öffentlichen Rechts in Deutschland. Volume 1, Reichspublizistik und Policeywissenschaft 1600–1800*. (Munich).
- Vermeir, R. (2005), ‘Comblent les lacunes. L’édition des ordonnances de Philippe IV (1621- 1665)’ in: *Commission Royale Pour La Publication Des Anciennes Lois et Ordonnances De Belgique, Bulletin XLVI* (Brussels).
- Von Friedeburg, R.C.F. (2016), *Luther’s Legacy. The Thirty Years War and the Modern Notion of ‘State’ in the Empire, 1530s to 1790s* (Cambridge).