Life and Death, War and Peace,
*c.*1550-*c.*1700

Voices of Women in Early Modern Ireland

Newsletter December, 2024

# *Welcome to the first VOICES Newsletter!*

Just a year into the VOICES Project, we can already see the impact of combining pioneering digital approaches with historical scholarship to recover the marginalised voices of women hidden in the ever-growing 'digital windfall' of historical documents from c.1550-c.1700.

We began our collaboration one year ago with a trip to Oxford to meet with the team from the Civil War petitions project. This was followed in November 2023, by the launch by Taoiseach Simon Harris of Making Empire: Ireland, imperialism, & the early modern world (Oxford University Press). Chapter 3 (on women and assimilation) is the starting point for VOICES.



In the meantime, our research has been progressing at pace. In April, we held a workshop which brought together leading historians, literary scholars, data analysts, and computer scientists to discuss our primary sources and how we can use innovative machine learning technologies such as Transkribus and generative AI for historical analysis.

In the meantime, our research has been progressing at pace. In April, we held a workshop which brought together leading historians, literary scholars, data analysts, and computer scientists to discuss our primary sources and how we can use innovative machine learning technologies such as Transkribus and generative AI for historical analysis.

Much of our focus has been on how best to use existing digitised data, especially the 1641 Depositions, the records of the Statute Staple, and other sources available thanks to the Virtual Record Treasury of Ireland. Research fellow Dr. Bronagh McShane is focusing on the wills, testamentary materials and funeral entries and Dr. Daniel Patterson is working on the Court of Chancery records to capture and digitise cases where women were plaintiffs or defendants. We were joined most recently by Dr. Madina Kurmangaliyeva who is

implementing statistical analysis across the data. Our Computer Science team led by Professor Declan O'Sullivan has welcomed Dr. Diego Rincon-Yanez who is using Natural Language Processing to extract the data for upload to the Knowledge Graph by Dr. Lucy McKenna.

The project has seen some attention with both academic and mainstream media coverage, and the project team is reporting early results at conferences such as the recent Tudor Ireland Conference and ENCHOS in Cambridge. Most recently, we offered an MPhil module in the History Department on 'The Lived Experiences of Women in Early Modern Ireland'.

The VOICES Team has also been collaborating closely with Dr. Peter Crooks and the VRTI team as we both follow similar paths in the use of AI, Natural Language Processing (NLP) and Knowledge Graph technologies to analyse and present our source materials.

And most recently, I visited India for the official Indian launch of *Making Empire: Ireland, Imperialism and the Early Modern World* and to speak about the VOICES project.
We are excited for the year ahead and will continue to deliver groundbreaking advances in using AI, together with NLP, to transform how we can access and interrogate previously inaccessible information on these non-elite women in early modern Ireland and recover their voices and experiences during this transformative period.


Professor Jane Ohlmeyer
VOICES Principal Investigator

*Sign up for future issues of the Newsletter here*
*Unsubscribe here*

View our VOICES video!

# *Wills and Testamentary material - Capturing and annotating AI test bed*

Dr Bronagh McShane

Wills proved in the Prerogative Court of Canterbury (PCC) in London include a small proportion (c. 57) made by women (mostly of English background) who lived in Ireland during the seventeenth century but who had some property interests in England. As the single largest surviving collection of *original* wills made by early modern women resident in Ireland, the PCC testaments are of intrinsic value. This corpus was used as a case study to perform some initial experiments with Transkribus and generative AI (in this case ChatGPT-4o) in order to produce and proof automatic transcriptions that could then be subject to further in-depth analysis (for example, using Natural Language Processing techniques).

The results were a mixed bag. In the case of Transkribus, some of the major limitations identified were not in fact limitations of the technology but rather of the quality of the digital images themselves. The experiment was conducted using digitised black and white microfilms of the original Probate 11 documents– which are noisy, poor-quality images and come with several inherent problems, even for manual transcription purposes.



These limitations were ultimately reflected in the accuracy of the transcriptions produced by the machine. Chat GPT-4o was then trialled for the purposes of proofing and correcting the Transkribus transcriptions. The model was deployed in a controlled manner that involved manual verification of the output. The importance of careful 'prompt engineering' in achieving the desired output cannot be overstated, demonstrating the inherent value of the 'human' in this human-machine collaboration.

Further exploratory work into the generation of customised and source specific prompts, in addition to the construction of 'walled gardens', offer a way forward that may, in time, allow us to harnesses the potential of generative AI without compromising empirical rigor and transparency in a large-scale project such as VOICES.

# Irish Chancery Pleadings: Breaking new archival ground

*By Dr Daniel Patterson*

Alongside its effort to harness the "digital windfall" of previously digitised sources for early modern Irish history, the VOICES project is also working to break new archival ground with work on under-studied manuscripts. Our primary focus in this regard has been work on the recently conserved Irish Chancery Pleadings, a corpus of just over 5,000 documents rescued from the fire in 1922.

Dating predominantly from 1570 to 1630, the Irish Chancery Pleadings represent the largest single body of early modern legal sources in Ireland. Despite their significance, they have been largely overlooked by historians (with some notable exceptions), leaving their rich contents underexplored.

In late 2023 and early 2024, our team undertook an intensive "scoping" process of the collection to uncover the voices and perspectives of early modern women. This effort identified 1,212 target documents involving female litigants, either as plaintiffs or defendants.
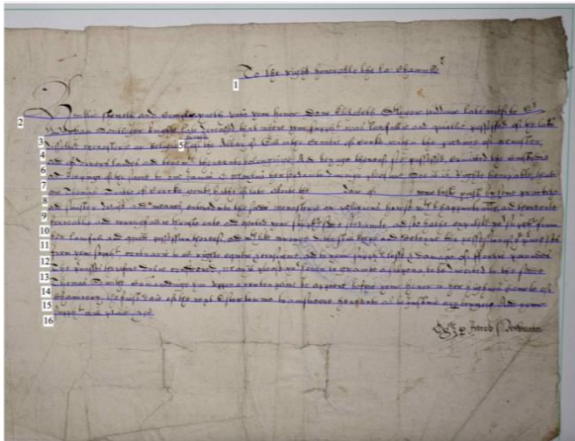


Over the summer of 2024, we conducted weekly visits to the National Archives of Ireland, where we meticulously photographed over 1,000 documents from this target list. These invaluable sources have since been edited and uploaded to Transkribus, a cutting-edge AI-based transcription tool, enabling us to process the material with unprecedented efficiency.

By blending traditional archival methods with innovative digital techniques, the VOICES project is unlocking insights into a collection of unparalleled historical significance in a timeframe that would have been unimaginable in the past. In the coming years, this work will allow us to share data and stories of hundreds of Irish women who appeared before the Court of Chancery in early modern Ireland.

# *Potential for analysis of quantitative data from early modern Ireland*

By Dr Madina Kurmangaliyeva

In the VOICES project, we're using new technology to uncover stories from the past.

Using traditional Deep Learning approaches and Generative AI, we're processing vast amounts of archival data from different archives into data that is searchable within the knowledge graph as part of the Virtual Record Treasury of Ireland, but also available for quantitative analysis and visualization. This will allow us to understand the lives -- and tell the stories -- of women in Ireland in the 16th and 17th century.

The VOICES project is an interdisciplinary group of researchers, from history, computer science and economics. I contribute as an economic historian and am fascinated by how our different knowledge and backgrounds reinforce this project. Economic history is a rich subfield within economics, that has advanced our understanding of the incredible growth in economic output and human flourishing that has occurred in the course of human history, with a pivotal moment in the Industrial Revolution. More recently, it has studied the long-term effects of colonial practices on the livelihoods of the colonized.  In my previous research at Université Libre de Bruxelles, I studied the impact of colonisation by crop-cultivating Russian settlers on the property rights and institutions of Kazakh nomad pastoralists at the turn of the 20th century.

Importantly for the VOICES project, economic history has developed a rich toolkit to work with historical data, being mindful of the biases they may contain. Unlike experimental sciences where new data is generated by running experiments, historical social sciences deal with "observational" data, which has been collected in the past, and not always for analytical purposes.  In VOICES, we use qualitative material - court records, wills, funeral entries - and quantitative information extracted from the 1641 Depositions, the Statute Staple records, the Books of Survey and distribution, poll taxes and hearth money roll taxes. These sources often have patchy geographical coverage and are usually biased towards documenting the experiences of wealthier men. The challenge is to account for the historical context and understanding the selection process that leads to data being available or not. For example, victims of violence who lived far away from Cork or Dublin were much less likely to testify in the 1641 Depositions, which makes comparisons of the extent of violence across different geographical areas difficult. Any quantitative analysis of the Depositions will have to account for this.

# *Using AI to unearth hidden VOICES*
**By Dr Diego Rincon-Yanez and Anne English**

At the core of the VOICES project is leveraging the power of Artificial Intelligence AI to harvest and curate the vast amount of data on women hidden in the many manuscripts from the 16th and 17th centuries. This windfall, while a treasure trove of data, would likely require years of historian time to sift through and catalogue manually. AI is the key here to analysing and extracting this data at scale and speed.

VOICES has implemented a series of structured workflows (pipelines) to transform these archives into knowledge that can be understood and utilised by both historians, economists and knowledge engineers.



We are using Natural Language Processing (NLP) techniques, and specifically Named Entity Recognition (NER) to identify and extract "entities," which, in this case, identify women's roles and experiences in C16 and C17. With the application of these technologies to source data ranging from Chancery pleadings to Wills and Testamentary materials as well as the 1641 Depositions, we can establish the occupations, land ownership or indeed familial relationships and experiences of violence of large numbers of women across the island of Ireland that would otherwise be extremely difficult to do from 'messy' or 'dirty' data at this scale and at this speed.

Named Entity Recognition is only the first step however—once the individuals are identified, the project matches findings across different contemporary documents to extract and infer more comprehensive relationships and knowledge. With the help of Knowledge Graph Technology, Semantic Web and Linked Open Data, we will use visualization tools and novel ontologies specifically for women and violence, which heretofore did not exist, to display the historical data in a searchable, user-friendly interface.

We hope that the use of these technologies in VOICES will offer a new paradigm of historical research, and enable us to recover, at scale and in a holistic and integrated way, the marginalised voices and lived experiences of women in early modern Ireland.

https://voicesproject.ie/

https://x.com/VOICES_TCD

https://x.com/janeohlmeyer

https://twitter.com/BA_McShane

https://bsky.app/profile/voicestcd.bsky.social