DHH23 DIGITAL HUMANITIES HACKATHON HELSINKI

ALIE TACQ
ANNA MOSCHITTA
EFTHYMIOS KOKORDELIS
ELIAS MUCHA
IMAMA ALISHOVA
KOLDO GARAI

LILLI PEURA
MICHELE STEFANI
OLOF FALK
RUXANDRA GÎDEI
TETIANA SHYSHKINA
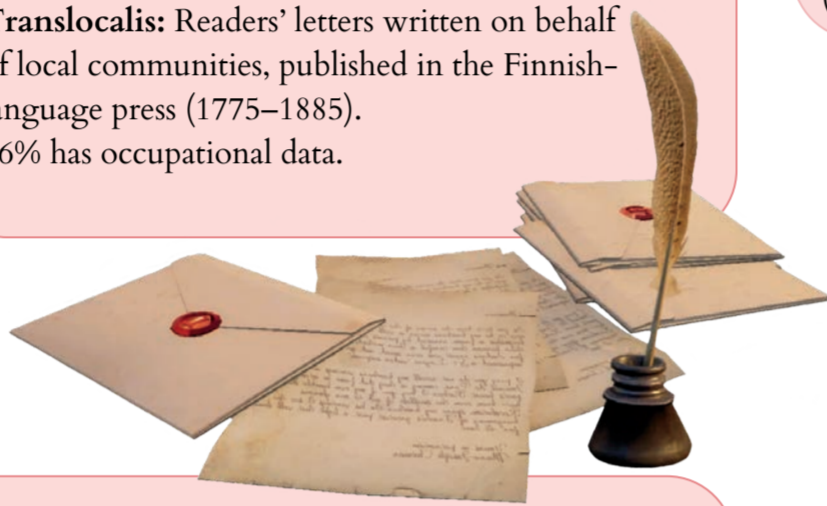VILLE-PEKKA KÄÄRIÄINEN

CAITLIN BURGE
ILONA PIKKANEN
JOUNI TUOMINEN
PETRI LESKINEN

EPISTOLARY PARTICIPATION

LETTER-WRITING AS A HARBINGER OF MODERNITY. EPISTOLARY LITERACY IS A KEY CRITERION FOR SOCIAL INCLUSION AND PARTICIPATION OF THE 'MIDDLING SORT' IN THE MODERNIZING SOCIETY (WHITMAN 2007). WE SET OUT TO ASK IF WE CAN OBSERVE THE BREAKTHROUGH OF THIS PHENOMENON IN EPISTOLARY METADATA ENRICHED WITH OCCUPATIONAL INFORMATION, CONSIDERING THE WAYS IN WHICH 'PARTICIPATION' MIGHT BE DEFINED, OR EVEN RESTRICTED, BY BOTH LETTER-WRITERS AND ARCHIVISTS & ARCHIVAL PROCESSES ALIKE.

RESEARCH QUESTIONS

1. Do occupational groups **cluster** around specific **publications**?
2. How does the advancement of **elementary education** relate to epistolary production?
3. How do different occupational groups **communicate** with each other?
   What occupational groups are most likely to **co-occur**?
4. How is the relationship between occupation and **travelling** patterns reflected in letter metadata?
5. How can metadata provided by **archives** affect our questions?

METHODS

- Utilised **ChatGPT** for data categorisation with further manual data cleaning.
- Examined correlations and connections between entities using Spreadsheets, with **Gephi** for network visualisations.
- **SPARQL query, RStudio and Python** for data extraction and analysis.
- Leveraging **set theory** for comprehensive big data analysis.

DATASETS

CoCo: Metadata of letters, public and private, kept in archives, libraries, and museums from the period of the Grand Duchy of Finland (1809–1917).
10% of actors have occupational data; however, their letters represent **over 60%** of the collection.
Translocalis: Readers' letters written on behalf of local communities, published in the Finnish-language press (1775–1885).
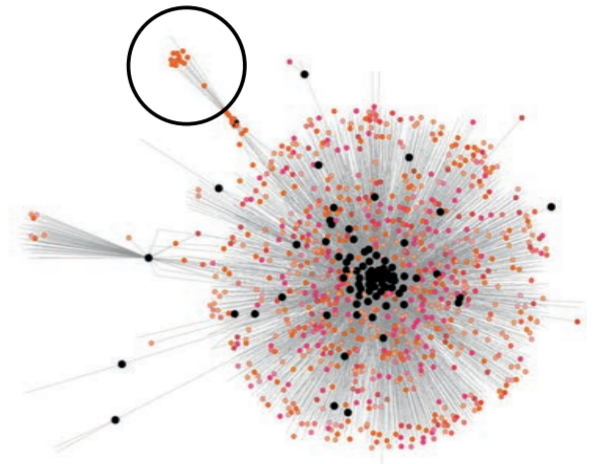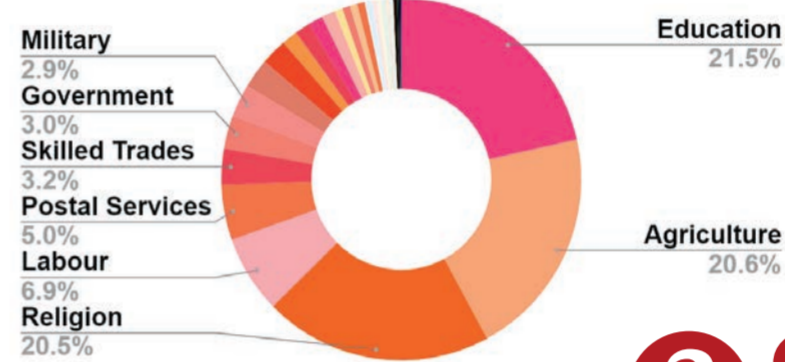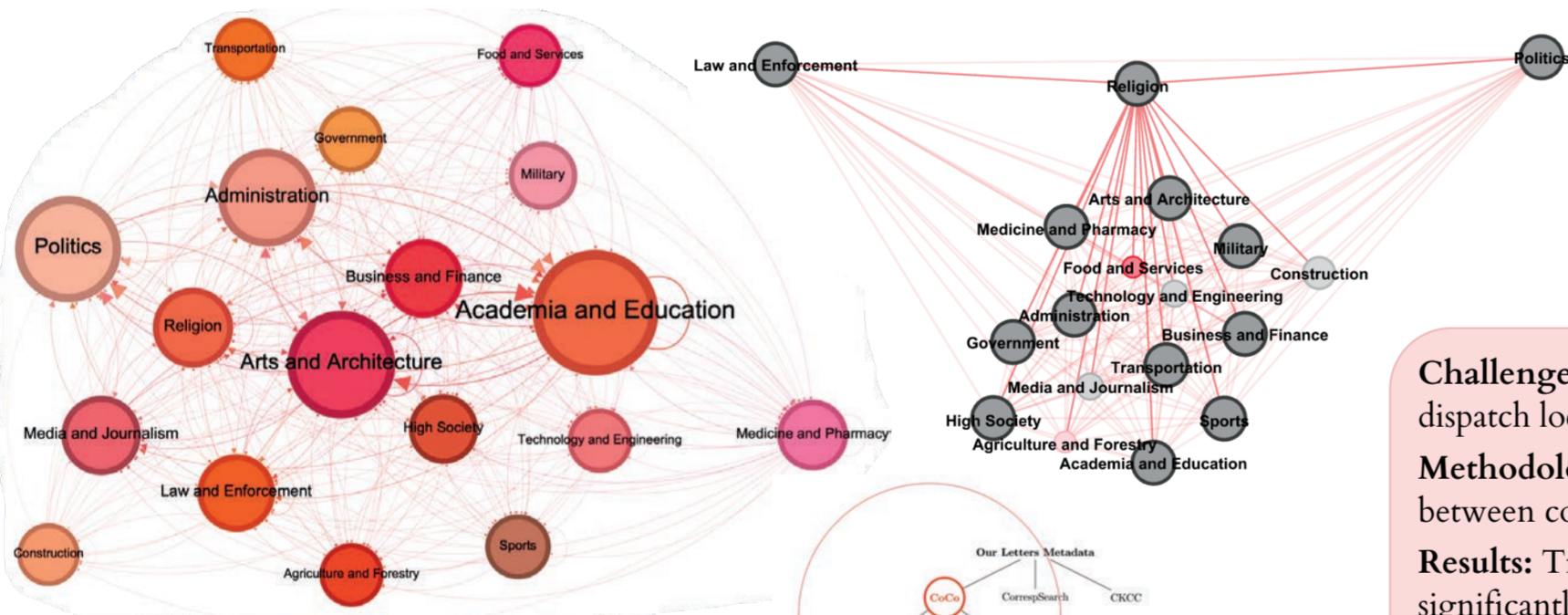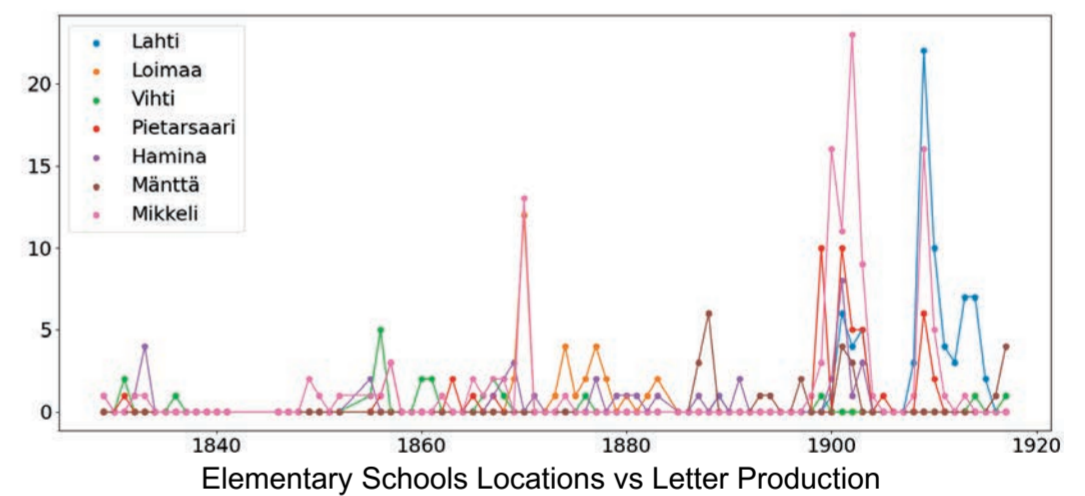16% has occupational data.

OCCUPATIONAL GROUPS 3

The colourful network represents **inter-sectoral communication** in the CoCo dataset. The grey network illustrates the **likelihood of combining different occupational groups** among individuals in the CoCo dataset. For both men and women, **academics** are most likely to pursue a second career in **Politics**, followed by **Arts and Architecture**.

BIASED ARCHIVES? 5

Set theory with multilevel subsets and intersections was utilised to allow us to navigate our data on different levels (individual letter/collection/database) concurrently and gain valuable insights to a dataset of almost 1 million records and the epistemological implications of archival practices.
Focus on gender; an indicative case of recorded archival bias, since results can be drawn from a binary perspective. In the dataset as a whole, the gender balance of the collections was surprisingly even; however, a more meticulous analysis of the collections shows that this "normal" distribution is artificial since there is a huge divergence on a collection-specific level.
Conclusion: Available data and its curators regulate our access to knowledge and the paths of research: uncovering archival biases is crucial for an accurate and inclusive history, informed decision-making, and social justice.

COCO 2023 — Distribution of Sender Gender
Analysis on the whole dataset — COCO 2022
The Finnish Literature Society · Svenska Literatursällskapet i Fin. · National Archives' Letters Collection · J.V. Snellman Letters
Åbo Akademi · FinNatGa - STYAC · Elias Lönnrot Letters · Albert Edelfelt
Gender distribution differs significantly between collections

1 OCCUPATION & PARTICIPATION IN LOCAL LETTERS

In the decade after 1865, emergent Finnish language publications connected a surprising network of occupations. While our expectation was to find correlations between publications and letter-writers' occupations, this was generally not the case. Instead, we saw **diversity** in the sectors and statuses of occupations with individuals writing to an array of newspapers. An exception is a **cluster of religious professionals** corresponding with a **missionary paper** (circled, upper left).

Proportions of occupations in Translocalis by sector (for letter senders with occupation data only)

Military 2.9%
Government 3.0%
Skilled Trades 3.2%
Postal Services 5.0%
Labour 6.9%
Religion 20.5%
Education 21.5%
Agriculture 20.6%

2 EPISTOLARY LITERACY

In 1866, Finland passed a statute for the establishment of **elementary schools**. By linking the locations of these schools to place of sender information in the CoCo dataset, we sought to explore the relationship between the passing of the act and the increase of letter production. Although there was spike in letters in **Mikkeli** and **Loimaa** around 1870, on the whole there is no drastic increase in letter-writing until after the turn of the century. Further investigation is needed to explore the causes of this letter boom.
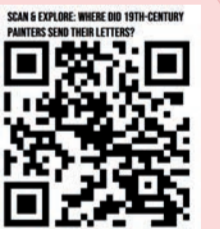

Elementary Schools Locations vs Letter Production
(Legend: Lahti, Loimaa, Vihti, Pietarsaari, Hamina, Mänttä, Mikkeli)
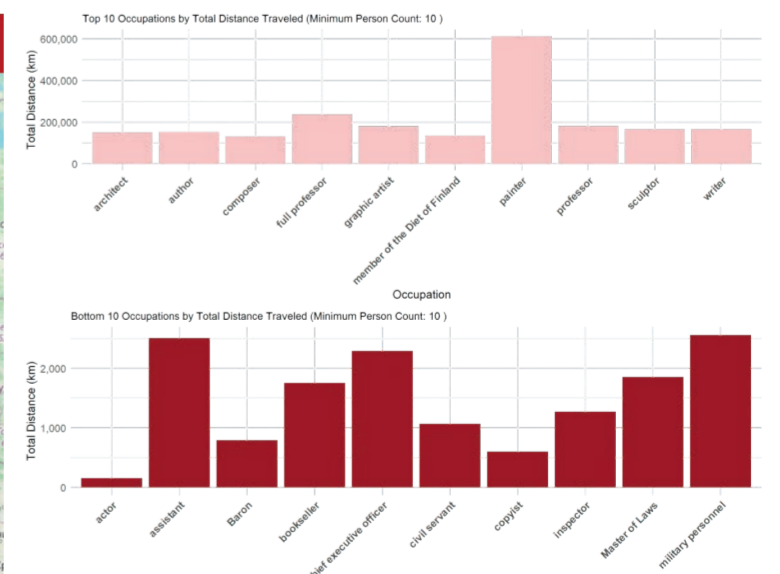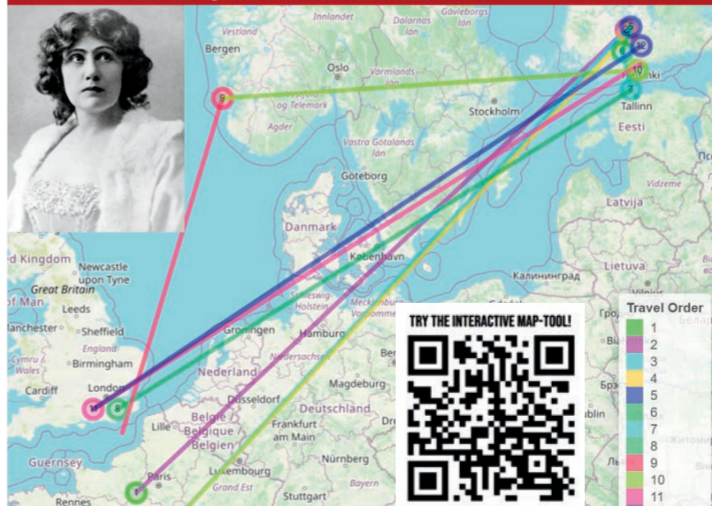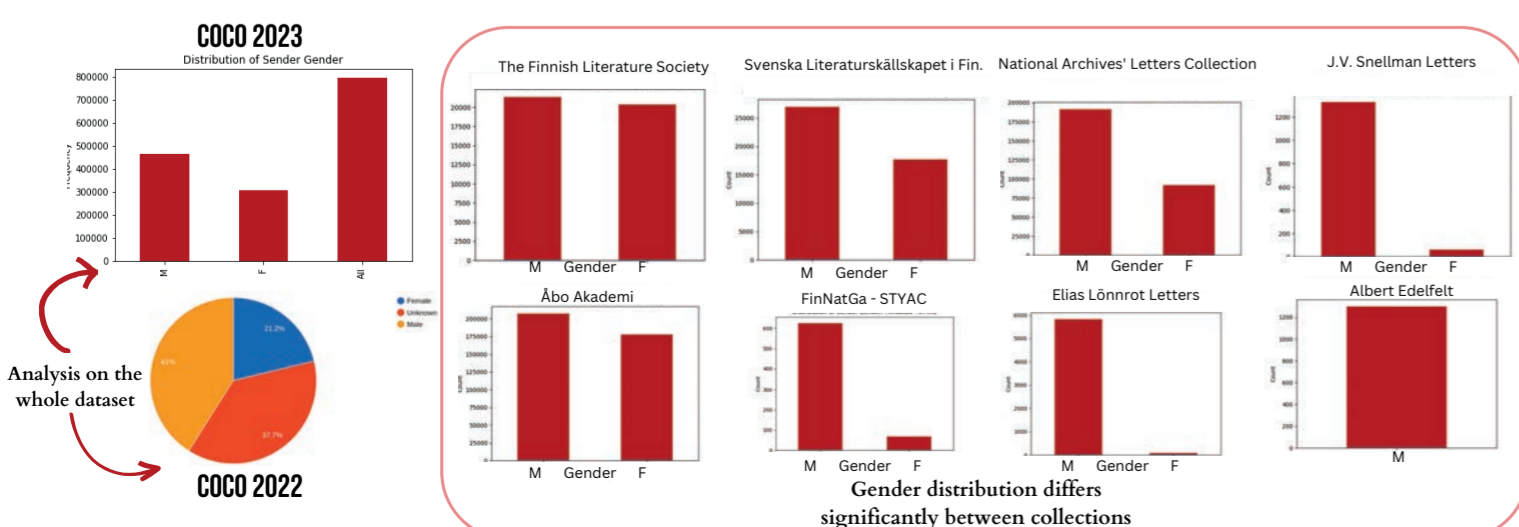
4 OCCUPATION AND TRAVELLING PATTERNS

Challenges: Determining methods to calculate travel distances from letter dispatch locations and addressing potential bias in archival data.
Methodology: Organised the data rigorously and calculated beeline distances between correspondence locations.
Results: Travel increased as the century progressed, but the increase was not significantly high in the context of the quantity of letters sent. The occupation of the letter sender influenced travel patterns. People in cultural and business roles were more likely to travel abroad, while those in administrative, legal, and military occupations were more likely to correspond domestically.

SCAN & EXPLORE: WHERE DID 19TH-CENTURY PAINTERS SEND THEIR LETTERS?

Geographically Known Correspondence of Aino Ackté (Achté), a renowned opera singer (b. 1876, d. 1944), from the years 1900 to 1917.
TRY THE INTERACTIVE MAP-TOOL!
Travel Order: 1 2 3 4 5 6 7 8 9 10 11

Top 10 Occupations by Total Distance Traveled (Minimum Person Count: 10)
Bottom 10 Occupations by Total Distance Traveled (Minimum Person Count: 10)

FUTURE CHALLENGES

- Mapping shifts in **newspaper themes** against **local letter content.**
- Comparing compositions of occupation across datasets by aligning **ontologies** of occupational categories.
- Enriching the database with **more occupational data** to produce more accurate and representative results.
- Utilising network analysis further to identify **hubs of participatory importance** and influence.