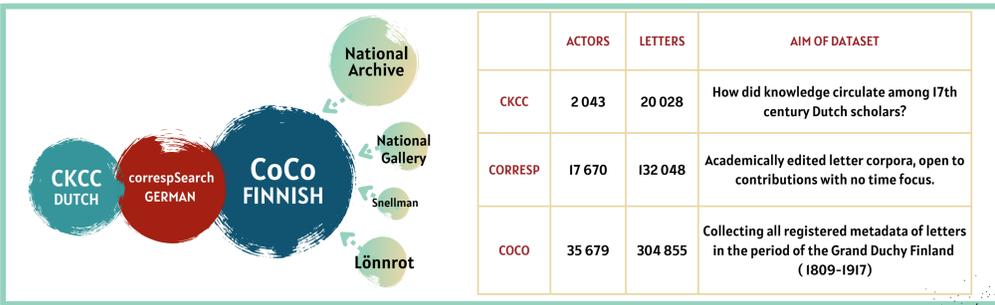




Epistolary Fingerprints

Identity & Representation: Patterns in Historical Letter Metadata

Caitlin Burge, Michael Göggelmann, Milja Ketoja, Mikko Koho, Maelle Le Roux, Petri Leskinen, Anna Sofia Lippolis, Ilona Pikkanen, Heikki Rantala, Voitto Alarik Vuorio, Yikun Xing
<https://dhhackathon.wordpress.com/author/dhh2022letters/>



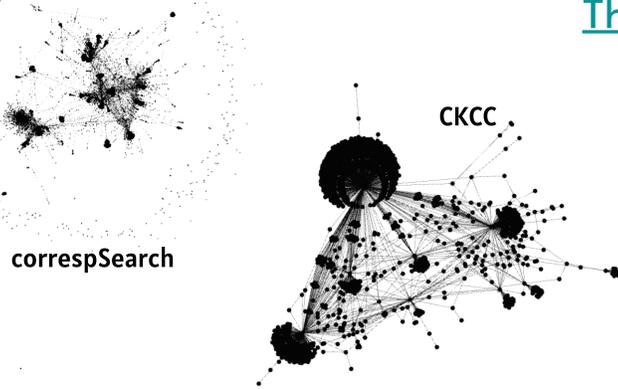
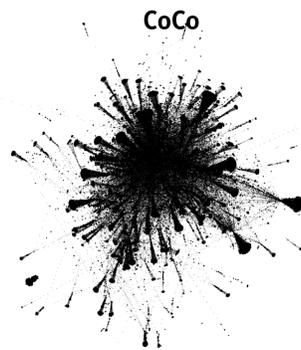
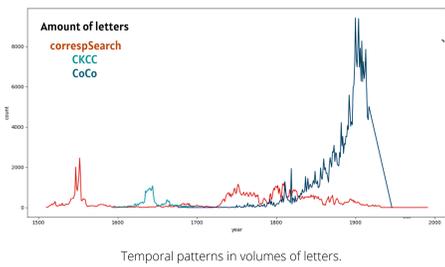
Research Questions

- 1) Epistolary cultures.
- 2) Individuals in correspondence structures.
- 3) Data profiles of collections.

Methodologies

- 1) Python and SPARQL queries.
- 2) Gephi visualisation & network measurements.
- 3) Machine learning.
- 4) Qualitative & historical research.

The Shape of the Archive

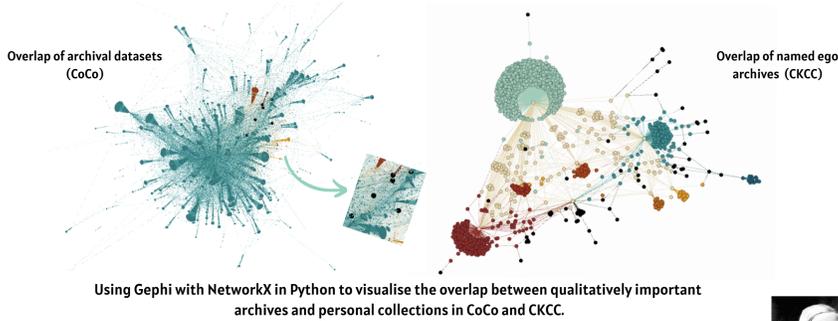


The Median Writer's Profile

JOHANN
Started writing letters at 36 years old
Stopped writing at 50, 14 years before his death
Writer with two other professions

CKCC, CorrespSearch, CoCo

Egos in the Network



CoCo: J.V. Snellman

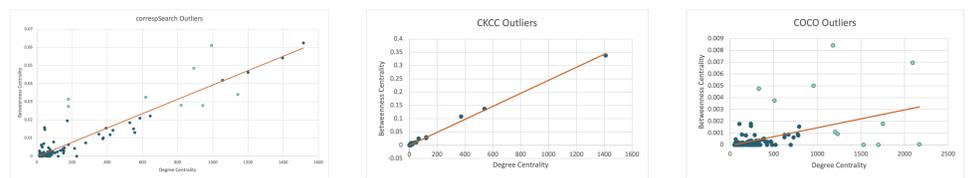
Out of 35,652 people, only six appear in all four datasets contributing to CoCo. Among them was Johan Snellman, one of the most important figures in Finnish nationalism. This is particularly surprising in view of his centrality measures, which did not necessarily suggest that he would be one of the four 'Connectors'. As he was engaged in several occupations though, this should explain his appearance in several archives with different selection criteria.



CoCo: Elias Lönnrot

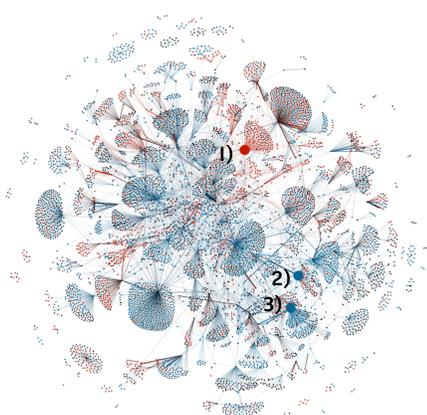
Elias Lönnrot, a major figure in Finnish history, surprisingly has not turned out as a main contributor within the overall network, even though the Finnish dataset contains his ego archive (appr. 6,300 letters and 1,200 correspondents). Although he is generally regarded as a reserved and humbled person, he nevertheless occupies the status of an "information gatherer" within the network. This is also being reflected in the genesis of his life's work, the national epic 'Kalevala', for which he gathered Finnish folk poetry and literature passed down orally during several expeditions in Finland, Russian Karelia, the Kola Peninsula and the Baltic countries.

'Information Gatherers' vs 'Information Redundant'



'Information Gatherers' are above the orange trendline in each graph, representing writers who act as intermediary more often for a limited number of correspondents; 'Information Redundant' writers, below the trendline have fewer opportunities to act as intermediary and operate in tightly knit communities

Genders in the Network

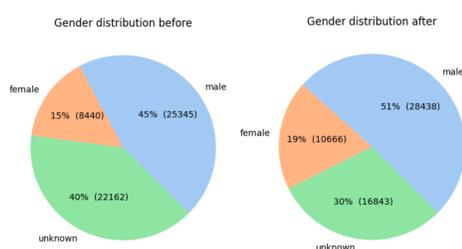


Network of correspondences in the CoCo dataset with nodes coloured by the gender (red for female; blue for male; and black for unknown).

The marked nodes are 1) Maikki Friberg; 2) J.V. Snellman; 3) Elias Lönnrot.

By observing the clusters around them it can be noticed that Friberg has been in correspondence with many other females, while the proportion of female correspondences for Lönnrot is lower.

Gender distribution among actors



Machine learning algorithms used to determine the likely gender from the known first name of the actor across datasets.

Changing places



Datasets with place information interrogated to study the geography of Lönnrot's and Melanchthon's networks, and the influence of modernisation on local and global correspondence networks.

Formalising Interestingness



CKCC: Christina, Queen of Sweden

Christina (1626-1689), was anything but an ordinary queen: she refused to marry and was known for her androgynous personality. She is considered one of the most learned women of the 17th century with a broad interest in many scientific disciplines. Although she is not a top contributor in qualitative terms, she still plays an important quantitative role in connecting scientists in the network, which also fits her stated mission to make Stockholm the "Athens of the North" by inviting them for a residence.

CoCo: Maikki Friberg

Maikki Friberg founded the 'International Union of Women for Peace and Freedom' and was editor of the magazine 'Women's Voice' - and there are no words that could better describe her role within the network. She formed the interface between important political actors and women activists of her time. This is also reflected in her network measures; she ranks tenth in all three centralities surveyed.



correspSearch: Philipp Melanchthon

It was to be expected that Philipp Melanchthon, polymath and intellectual leader of the Reformation - second only to Martin Luther himself - would play an important role in the German dataset. Whilst Melanchthon ranks first as an intermediary and in the number of connections in the dataset, he does not rank high in 'eigenvector', suggesting that he was connected to lots of people who held limited influence in the wider network. This is not exclusively due to the "Shape of the Archive", as Melanchthon holds similar network measures in the State Papers of England (1523-1540), which implies similar scholarly activity in different network structures/countries.

Challenges & Future Research

- Disambiguation (CoCo) - the amount of actors will become more accurate/ the unknowns.
- Using complete ego networks or archival collections as a 'prototype' to model 'complete' networks and estimate missing data and correspondence connections.
- Building more network profiles (like 'Gatherer' vs 'Redundant') to consider similar roles in network structures, both in these datasets and beyond.
- Thinking critically about the role and development of letter metadata in cultural heritage study and preservation.