Data-Rich History for 19th Century Literature in Finland

DRDHum2024

Kati Launis , Aino Mäkikalli, Viola Parente-Čapková,
Veli-Matti Pynttäri & Osma Suominen

10.12.2024 (kati.launis@uef.fi)

UEF// University of Eastern Finland

Digital History for Literature in Finland
(Research Council of Finland, 2022-2026)

WP1 Literary history
Literary history in the long 19th century Finland
PI Kati Launis / University of Eastern Finland

WP 2 Data Science
Bibliographic data science
PI Leo Lahti / University of Turku

WP 3 Data and outreach
Digital Resources
PI Osma Suominen/ National Library of Finland

"The Great Unread" (Margaret Cohen)

- What is "Finnish literature in the 19th century"?
- First comprehensive quantitative study
- Fennica - The Finnish National Bibliography
- Focus on first editions of fictive literature in book format written originally in Finnish or Swedish.
- How to construct the dataset?

Data-Rich literary history

- **Katherine Bode**
  - *Reading by Numbers: Recalibrating the Literary Field* (2012): https://directory.doabooks.org/handle/20.500.12854/30197
  - Austlit –database (literature)
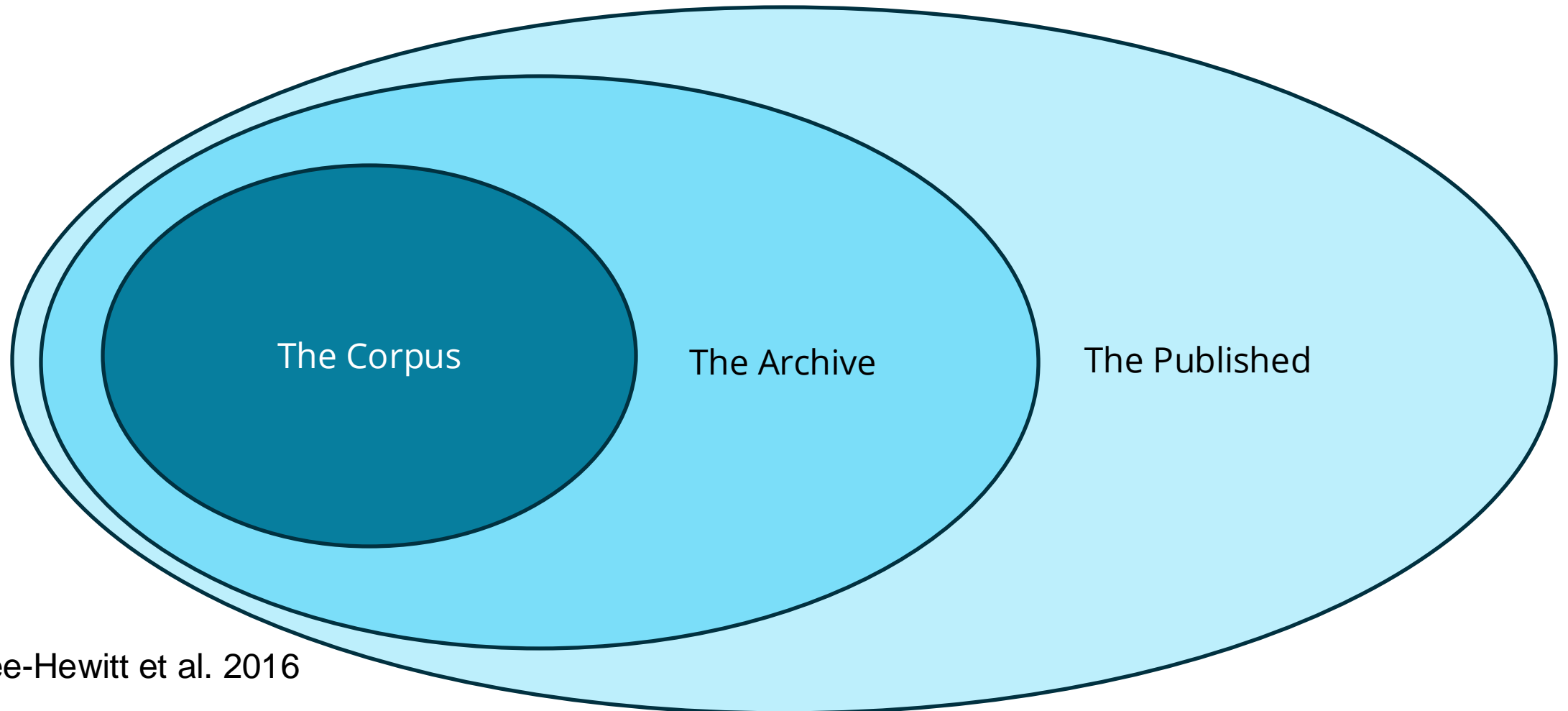 - *A World of Fiction: Digital Collections and the Future of Literary History* (2018) https://www.fulcrum.org/concern/monographs/5q47rp73f#toc
  - Trove –database (newspapers, magazines)
- **Franco Moretti**
  - *Distant reading* (2000/2013)
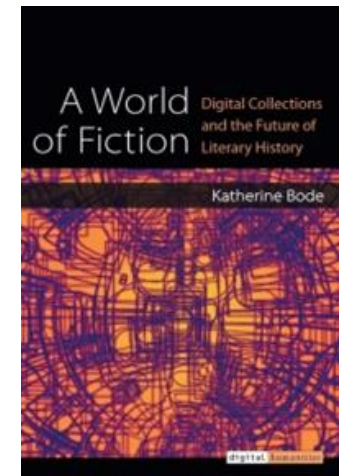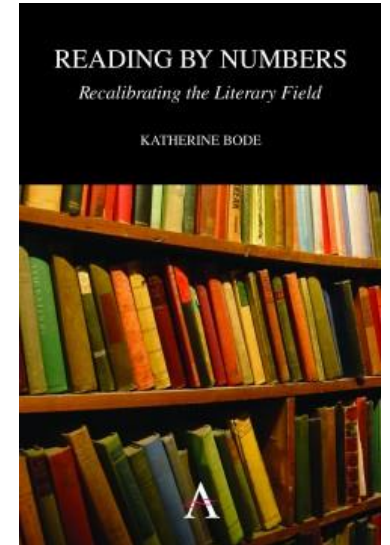
The Published – The Archive – The Corpus

The Corpus

The Archive

The Published

Algee-Hewitt et al. 2016

From the Published to the Archive

- First collection of Finnish Literature was destroyed in the fire of Turku in 1827.
- The collection was reinstituted in Helsinki by Fredrik Wilhelm Pipping (1783–1868) in 1856-57.
- The only thematically organized collection was Fennica.

Defining Fennica

- Building Fennica entailed definition of what is "Finnish literature"
- Fredrik Wilhelm Pipping: Everything published in Finland and everything published by a Finn
- Sven Gabriel Elmgren (1865):
  - All items published in Finnish regardless of the place of publishing
  - Everything published in Finland regardless of language
  - Every publication by a Finn, or someone who has lived in Finland regardless of language
- Georg Schauman (1913):
  - In addition to Elmgren everything that handles Finland, its citizens, history or languages regardless of place of publication
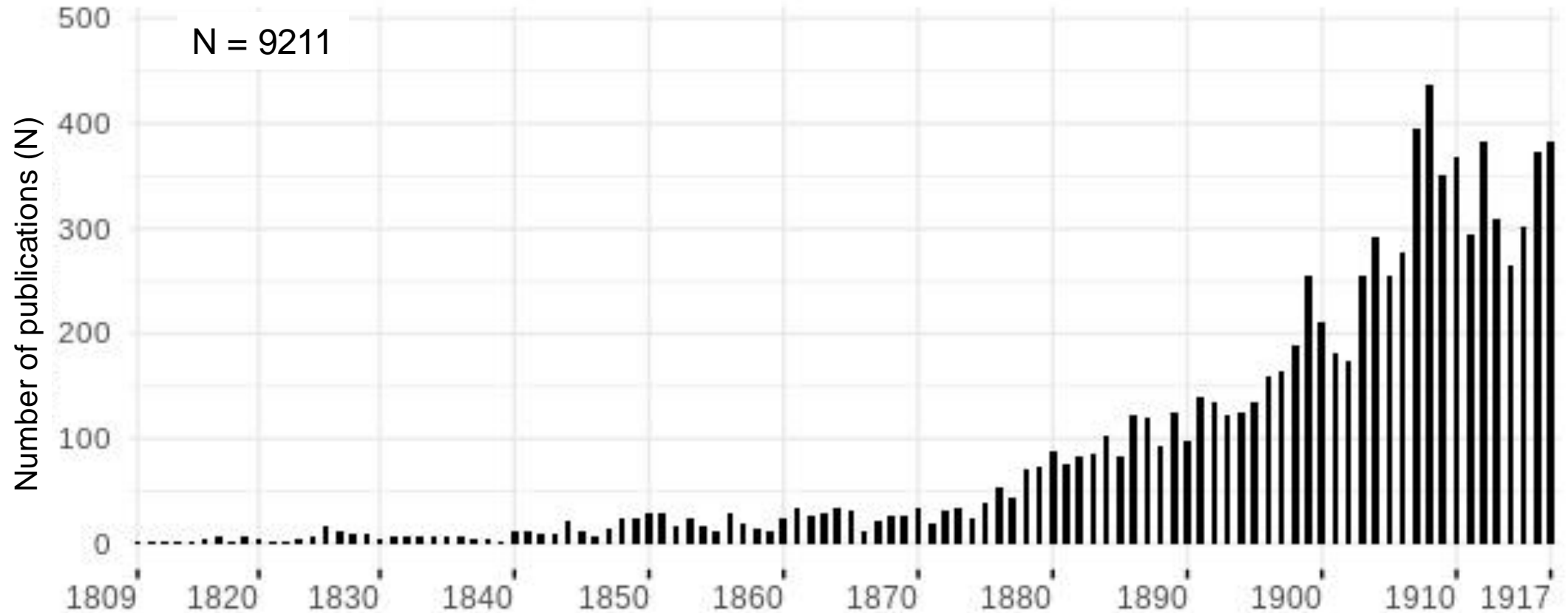
From the Archive to the Corpus

- **Initially 9211 titles**
  - All the titles classified as fiction based on call numbers (signum), not UDC-classification

Fiction in Fennica – Number of publications in 1809-1917

Data: Satu Niininen (National Library) Chart: Julia Matveeva (Univ. of Turku)
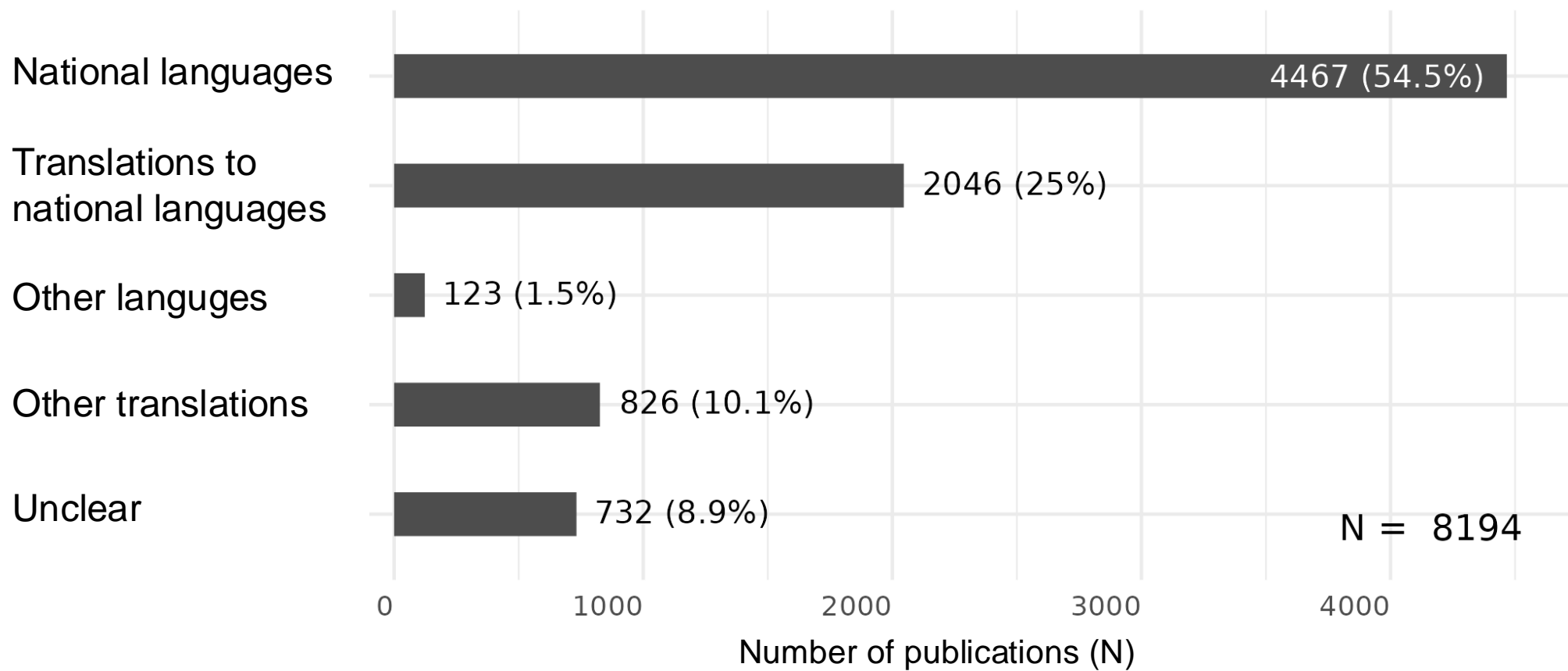
From the Archive to the Corpus

- **Initially 9211 titles**
  - All the titles classified as fiction based on call numbers not UDC-classification.
- **Then about 4500 titles**
  - Automatic filtering out of children's literature and translations.

Fiction in Fennica (excluding children's literature)

Data: Satu Niininen (National Library) Chart: Julia Matveeva (Univ. of Turku)

National languages — 4467 (54.5%)

Translations to national languages — 2046 (25%)

Other languges — 123 (1.5%)

Other translations — 826 (10.1%)

Unclear — 732 (8.9%)

N = 8194

Number of publications (N)

Translations

- Titles translated from English (579), German (378), French (309) and Russia (285).
- Titles translated from Swedish to Finnish (430) and from Finnish to Swedish (95)
- Most translated authors: **Leo Tolstoi** (87) and **Arthur Conan Doyle** (85)

First translation of Doyle is *Salapoliisin seikkailuja* published in Kuopio in 1894 and translated by Ida Wickstedt.

In 1907 altogether 38 Sherlock Holmes -titles

From the Archive to the Corpus

- **Initially 9211 titles**
  - All the titles classified as fiction based on call numbers not UDC-classification
- **Then about 4500 titles**
  - Automatic filtering out of children's literature, translations and reprints
- **Finally about 2800 titles**
  - Manual exclusion of reprints, anthologies, collected works, translations, broadsheet literature and other
  - Time consuming work with metadata
  - Shortcomings in metadata, consideration of historical circumstances

- **First editions of fiction in book format published originally in Finnish or Swedish**
  - The idea of "originality" is visible in the titels, e.g. *Murgrönan. Finsk original* (1840), Mikko Kukkonen: *Aallotar. Alkuperäisiä koelmia runollisuuden alalla* (1869)
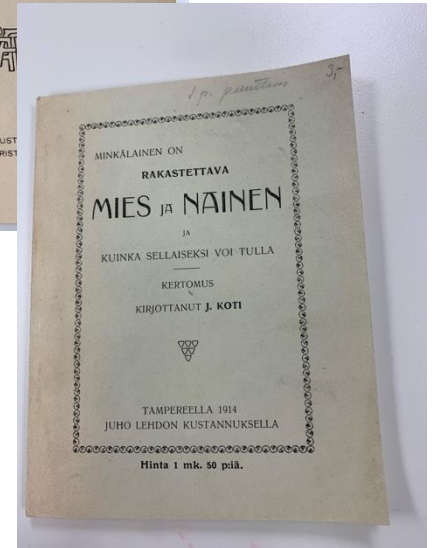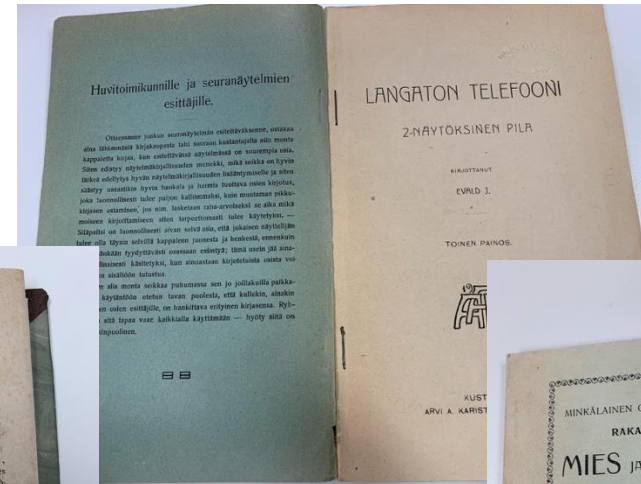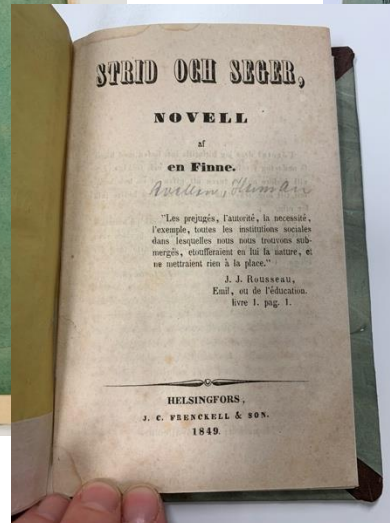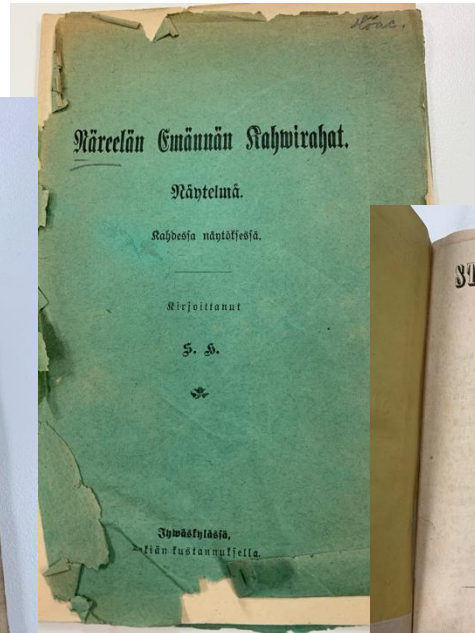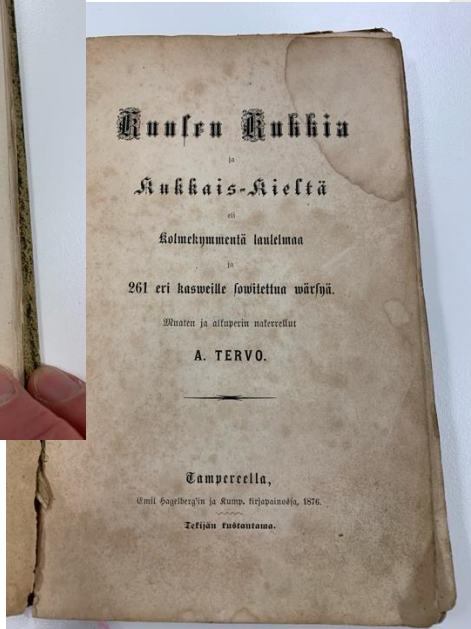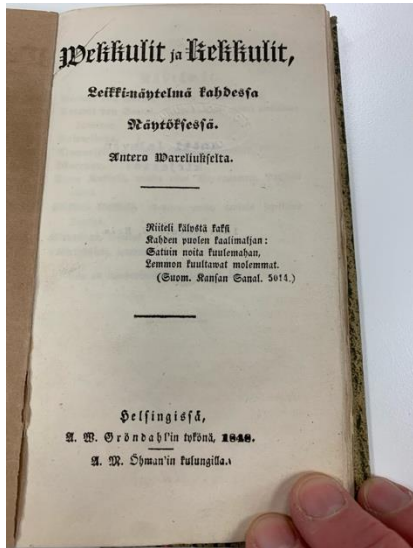
Language of the first editions

- Swedish and Finnish in 19th century literature in Finland
- Consequences of national awakening and economic growth in the end of the century
- Newspapers and periodicals?
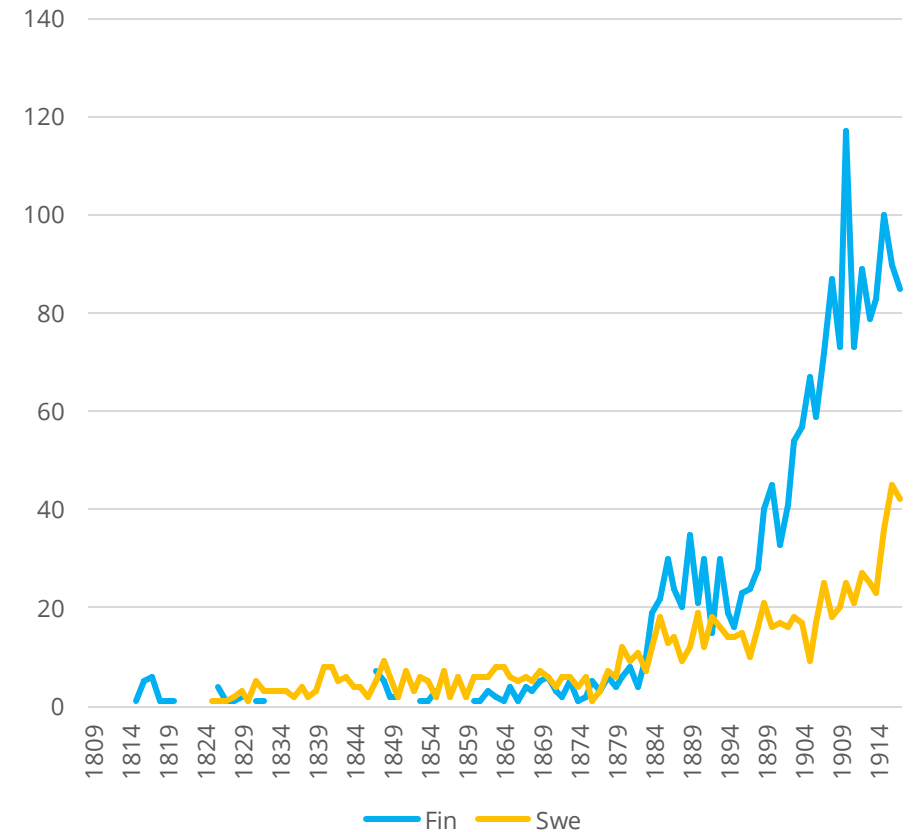
Wekkulit ja kekkulit
and other first editions

- Familiar authors but also new acquaintances, especially plays

Challenges today

- **Nature of metadata**: inadequte data especially in relation to source language, number of editions, pseudonyms or mistaken identities

- **Nature of collections**: uncatalogued items, description is subjective process that has taken place and continues to do so in historical circumstances

- **Nature of history**: bilingual culture, transnational reality, undefined publishing environment

- **Nature of subjective component**: how to be consistent?

Reading with Bibliographic Metadata – What changes?

- With bibliographic metadata the information on literary history is more systematic and comprehensive: enables more diverse understanding of the 19th century literature in Finland.

- Quantitative analysis of metadata can highlight forgotten or unnoticed works, authors, or literary phenomena, and motivates further close reading.

- Method changes: focus on the "mesospace" represented by metadata
    - Describes the literature as an object of close reading
    - Construes the literary system as an object of quantitative analysis
    - New object for data-rich literary history (Bode 2018)?

Thank you!

(veli-matti.pynttari@uef.fi)

https://sites.utu.fi/digital-history-literature-finland/

References

- Bode, K. (2018). *World of Fiction. Digital Collections and the Future of Literary History.* University of Michigan Press.
- Bode, K. & Dixon, B. (2010). Resourceful Reading: A New Empiricism in the Digital Age. In K. Bode & R. Dixon (Eds.), *Resourceful Reading: The New Empiricism, ERe-search and Australian Literary Culture*, 1–27. Sydney University Press.
- Genette, G. (1987/1997). *Paratexts. Tresholds of Interpretation*. Transl. Jane E. Lewin. Cambridge University Press.
- Lahti L., Marjanen J., Roivainen H. & Tolonen, M. (2019). Bibliographic data science and the history of the book (c. 1500–1800). *Cataloging & Classification Quarterly* 57(1), 5–23. http://dx.doi.org/10.1080/01639374.2018.1543747
- Moretti, F. (2000/2013). *Distant Reading*. Verso.
- Parente-Čapková, V ., Launis, K. & Westerlund, J. (2023). Digitaaliset metadata-arkistot ja mesoanalyysi kulttuurisen vaihdon kartoittamisessa. *Avain - Kirjallisuudentutkimuksen aikakauslehti*, vol 20, nro 1, 100-111. https://journal.fi/avain/article/view/127486/77833.
- Saint-Amour, P. (2019). The Medial Humanities: Toward a Manifesto for Meso-Analy-sis. *M/m,* vol. 3, cycle 4. https://doi.org/10.26597/mod.0092.