

Distant Reading for European Literary History. A COST Action

Christof Schöch (schoech@uni-trier.de), University of Trier, Germany

Maciej Eder (maciej.eder@ijp.pan.pl), Institute of Polish Language, Polish Academy of Sciences, Kraków, Poland

Carolin Odebrecht (carolin.odebrecht@hu-berlin.de), Humboldt-Universität Berlin, Germany

Mike Kestemont (mike.kestemont@gmail.com), University of Antwerp, Belgium

Antonija Primorac (antonija.primorac@uniri.hr), University of Rijeka, Croatia

Justin Tonra (justin.tonra@nuigalway.ie), National University of Ireland Galway, Ireland

Katja Mihurko Poniz (katja.mihurko-poniz@guest.arnes.si), University of Nova Gorica, Slovenia

Catherine Kanellopoulou (catherine.kanellopoulou@gmail.com), Ionian University, Kérkyra, Greece

Introduction

This poster aims to stimulate awareness of the existence of the newly-established COST Action on “Distant Reading for European Literary History” (2017-2021). In the context of this networking project, “distant reading” is understood as an umbrella term for recent computational, and particularly quantitative, approaches to the study of large collections of texts. This paradigm is here applied to the multilingual literary traditions of Europe in the long nineteenth century.

What is a COST Action?

COST (www.cost.eu) stands for ‘European Cooperation in Science and Technology’: COST Actions are essentially networking initiatives focused on a particular, timely and innovative research topic, aiming to bring together a critical mass of researchers from Europe and beyond. COST Actions coordinate their activities through working group meetings and offer Training Schools and opportunities for scientific exchange, for example so-called Short Term Scientific Missions. Examples of previous COST Actions in Digital Humanities include Interedition (<http://www.interedition.eu/>, 2008-2012) and e-Lexicography (<http://www.elexicography.eu/>, 2013-2017).

Aims of the “Distant Reading” Action

The contribution of the Distant Reading paradigm to Literary Studies continues to be a matter of intense debate. (See, for example, the contributions to a recent issue of PMLA: Goldstone 2017; Piper 2017; So 2017 and others, as well as Underwood 2017. The term itself has been introduced by Moretti 2005.) In our view, recent, quantitative approaches clearly provide an important methodological perspective that usefully complements, and at times challenges, more established approaches to literary history and theory in areas like authorship attribution, genre analysis, periodization, canonization and intertextuality.

We aim to create a vibrant and diverse network of researchers jointly developing the resources and methods necessary to change the way European literary history is written. Fostering insight into cross-national, large-scale patterns and evolutions across European literary traditions, we will facilitate the creation of a broader, more inclusive and better-grounded account of European literary history and cultural identity. We will foster

distributed research, the systematic exchange of expertise, and the visibility of all participants, activities and resources.

In terms of scientific objectives, we will coordinate the creation of a multilingual European Literary Text Collection (ELTeC). We will use the ELTeC to establish best practices and develop innovative methods of Distant Reading for the multiple European literary traditions. Furthermore, we will engage in an investigation into the theoretical consequences of Distant Reading approaches for literary history and literary theory. We also aim to foster the acquisition of state-of-the-art methods related to data curation, standards, best practices and quantitative analysis in workshops and training schools. Last but not least, we aim to address the current gender imbalance among practitioners of Distant Reading research.

The network

Our network of members is currently comprised of researchers in Corpus Linguistics, Computational Linguistics, (Digital) Literary History and Literary Theory from 26 different countries and more than 40 cities across Europe and beyond (Figure 1).

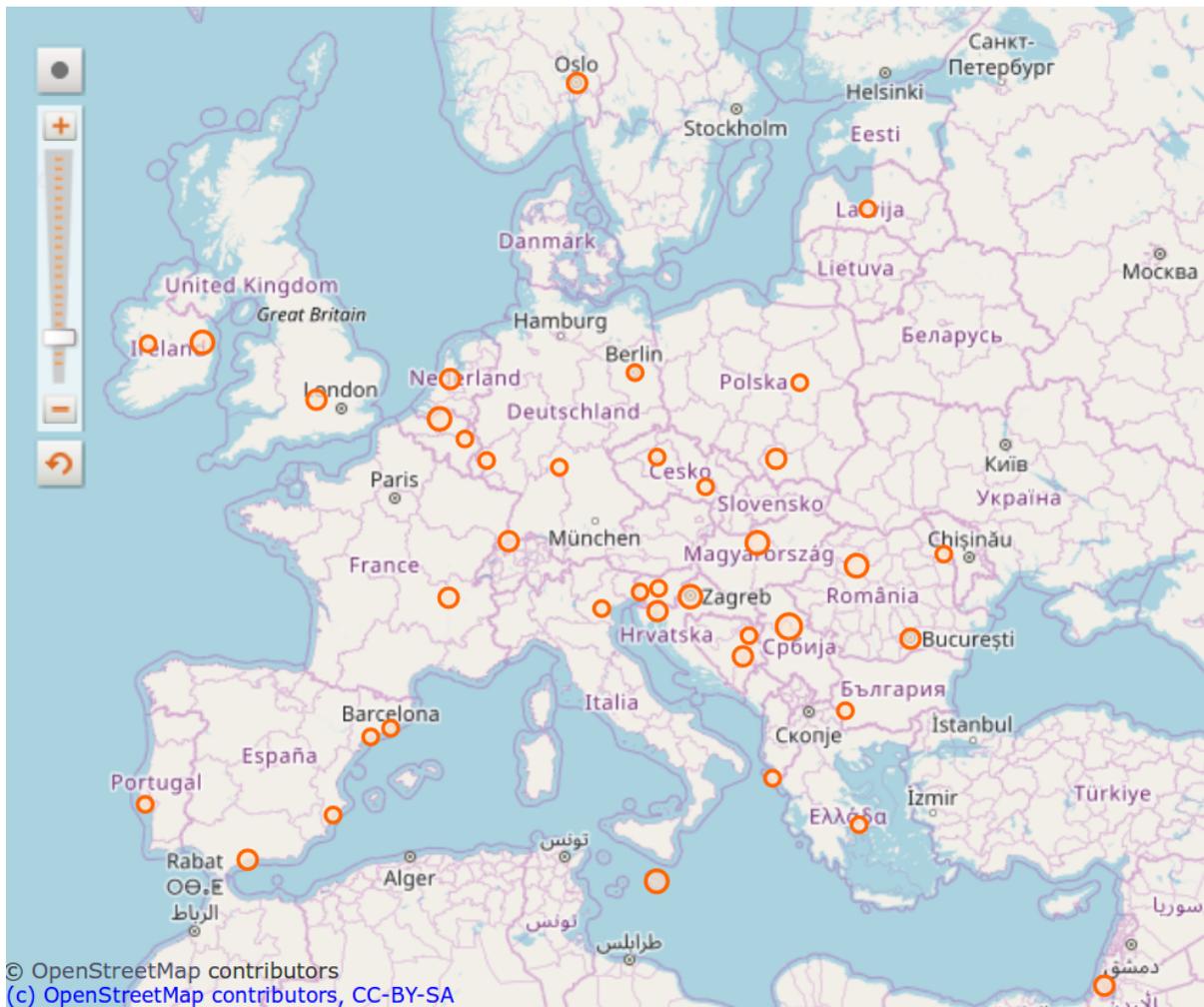


Figure 1: Map of Europe with the locations of Action members; for more details, see: <https://www.distant-reading.net/network/>.

Our key deliverable: the ELTeC

Our key deliverable is the European Literary Text Collection (ELTeC) that brings together comparable sets of nineteenth-century novels from at least 10 different European languages. Each set will comprise 100 different novels, from both inside and outside the canon, published in the late nineteenth century, with extensions covering the early nineteenth century or adding additional novels from the late nineteenth century. The purpose of the ELTeC is to serve as a benchmark corpus for the evaluation and development of annotation tools and distant reading methods across languages and as the basis for investigations into patterns and trends in literary history in multiple literary traditions.

Research strands

Our activities are divided into three main research strands (organized in working groups):

- “Scholarly Resources”, focused on structuring, annotating and publishing the ELTeC;
- “Methods and Tools”, concerned with using, evaluating and developing methods for distant reading analysis;
- “Literary History and Theory”, dedicated to the theoretical consequences of distant reading methods for literary history and theory.

In addition, a working group on “Dissemination” provides infrastructure services, enables communication within the Action and gives visibility to the Action’s activities and results.

Learn more, learn how to join

To learn more, see the Action’s website at <http://www.distant-reading.net>, the Action’s profile page at http://www.cost.eu/COST_Actions/ca/CA16204 and the full proposal linked there (“Memorandum of Understanding”). Researchers from Computational Linguistics, (Digital) Literary Studies as well as Computer Scientists and Librarians are welcome to get involved!

Acknowledgements

This poster describes the COST Action "Distant Reading for European Literary History" (CA16204 - "Distant-Reading"). Find out more at: <http://www.distant-reading.net>. COST is funded by the Horizon 2020 Framework Programme of the EU.

References

- Goldstone, Andrew (2017). The Doxa of Reading. *PMLA*, 132(3): 636–42.
- Moretti, Franco (2005). *Graphs, Maps, Trees: Abstract Models for a Literary History*. New York: Verso.
- Piper, Andrew (2017). Think Small: On Literary Modeling. *PMLA*, 132(3): 651–58.
- So, Richard Jean (2017). ‘All Models Are Wrong’. *PMLA*, 132(3): 668–73.
- Underwood, Ted (2017). A Genealogy of Distant Reading. *Digital Humanities Quarterly*, 11(2). <http://www.digitalhumanities.org/dhq/vol/11/2/000317/000317.html>.