

Proceedings of the

Workshop on Computational Methods in the Humanities 2022 (COMHUM 2022)

University of Lausanne, June 9–10, 2022

The logo for the University of Lausanne (Unil), featuring the word "Unil" in a blue, cursive script font.

UNIL | Université de Lausanne

Section des sciences
du langage
et de l'information

COMHUM
2022

Preface

At the turn of the 2020s, a defining characteristic of digital humanities remains the remarkably wide spectrum of viewpoints they encompass, ranging from a pure engineering perspective applied to humanities data to the use of well established humanities research methods to investigate born-digital artifacts. In this framework, the COMHUM workshop series positions itself as an international forum primarily devoted to the following research questions:

1. Which computational methods are most appropriate for dealing with the particular challenges posed by humanities research, e.g., uncertainty, vagueness, incompleteness, but also with different positions (points of view, values, criteria, perspectives, approaches, readings, etc.)?
2. How can such computational methods be applied to concrete research questions in the humanities?

The second edition of the Workshop on Computational Methods in the Humanities (COMHUM 2022) took place on June 9 and 10, 2022 at the University of Lausanne (UNIL).¹

The first day, introduced by a keynote from Vincent Labatut (Avignon University), was devoted to the specific topic of computational methods for constructing and analyzing character networks. This topic has ramifications in a variety of disciplines, including linguistics, literary analysis, digital humanities, and game studies. It is of particular interest for a number of research initiatives at UNIL and in neighboring institutions. Our goal was to bring together researchers from different communities studying character networks using computational and methodologically explicit approaches, to review the state of the art in this domain and to sketch its future developments.

In the spirit of the first edition of the COMHUM workshop, the second day was open to submissions on any topic pertaining to theoretical or applied research on computational methods for humanities research broadly conceived.

We invited researchers to submit abstracts of 500 to 1000 words. These abstracts were reviewed in a double-blind fashion by members of the program committee; all submissions received several independent reviews. Of the 23 submissions we received, we finally accepted 18. Authors of accepted abstracts presented their research at the workshop as talks; the abstracts are published on the workshop Web site.²

The COMHUM 2022 workshop was organized by members of the Lausanne Lab for Computational and Statistical Text Analysis (LLIST)³: François Bavaud, Guillaume Guex, Coline Métrailler (co-chair), Davide Picca, Michael Piotrowski, Yannick Rochat (chair), and Aris Xanthos, with assistance from Stéphanie Pichot. It was hosted by the Department of Language and Information Sciences⁴ and received support from the Center for Linguistics and the Science of Language,⁵ both part of the Faculty of Arts of the University of Lausanne.

This volume constitutes the proceedings of COMHUM 2022. It contains revised and expanded versions of seven of the papers presented at the workshop. Both the original abstracts and the expanded versions were peer-reviewed by at least two members of the program committee (see below); the seven revised and expanded papers that constitute these proceedings underwent another round of reviews.

We would like to thank everybody who contributed, in one way or another, to making COMHUM 2022 a success.

Yannick Rochat
Coline Métrailler
Michael Piotrowski
(editors)

1. See <https://ceur-ws.org/Vol-2314/> for the proceedings of the first edition, COMHUM 2018.

2. <https://wp.unil.ch/llist/en/event/comhum2022/>

3. <https://wp.unil.ch/llist/>

4. <https://www.unil.ch/sli/>

5. <https://www.unil.ch/cls1/>

Program Committee

- François Bavaud, University of Lausanne (Switzerland)
- Giovanni Colavizza, University of Bologna (Italy)
- Guillaume Gux, University of Lausanne (Switzerland)
- Vincent Labatut, Avignon University (France)
- Coline Métrailler (co-chair), University of Lausanne (Switzerland)
- Fabian Moss, Julius-Maximilians-Universität of Würzburg (Germany)
- Davide Picca, University of Lausanne (Switzerland)
- Michael Piotrowski, University of Lausanne (Switzerland)
- Yannick Rochat (chair), University of Lausanne (Switzerland)
- Elena Spadini, University of Basel (Switzerland)
- Mathieu Triclot, University of Technology of Belfort-Montbéliard (France)
- Aris Xanthos, University of Lausanne (Switzerland)

Please note that some members of the program committee have changed affiliation since the workshop; we have chosen to list them with their current affiliation at the publication date of this volume.

Contents

Modelling Usage Information in a Legacy Dictionary: From TEI Lex-0 to Ontolex-Lemon <i>Bruno Almeida, Rute Costa, Ana Salgado, Margarida Ramos, Laurent Romary, Fahad Khan, Sara Carvalho, Mohamed Khemakhem, Raquel Silva, Toma Tasovac</i>	5
Data Augmentation for Robust Character Detection in Fantasy Novels <i>Arthur Amalvy, Vincent Labatut, Richard Dufour</i>	23
A (Dis)similarity Index for Comparing Two Character Networks Based on the Same Story <i>François Bavaud, Coline Métrailler</i>	33
A Framework for Embedding Entities in a Textual Narrative: a Case Study on Les Misérables <i>Guillaume Guex</i>	43
Introducing VISU: Vagueness, Incompleteness, Subjectivity, and Uncertainty in Art Provenance Data <i>Fabio Mariani</i>	63
Narrative Flow: A Formal Distant Reading Approach for Interactive Narratives <i>Coline Métrailler</i>	85
Exploring Naming Inventories for Architectural Elements for Use in Multi-modal Machine Learning Applications <i>Ronja Utescher, Aaron Pattee, Ferdinand Maiwald, Jonas Bruschke, Stephan Hoppe, Sander Münster, Florian Niebling, Sina Zarriß</i>	95