

Arno Simons, Adrian Wüthrich,
Michael Zichert, Gerd Graßhoff (eds.)

Understanding Science with Large Language Models?

Potentials for the History,
Philosophy, and
Sociology of Science

[transcript] science studies

Arno Simons, Adrian Wüthrich, Michael Zichert, Gerd Graßhoff (eds.)
Understanding Science with Large Language Models?

Arno Simons is senior researcher at the Institute of History and Philosophy of Science, Technology, and Literature at Technische Universität Berlin and a member of the Network Epistemology in Practice (NEPI) project. Trained as an interdisciplinary scholar, he has contributed theoretical and methodological work across policy and environmental science, translational biomedicine, science and technology studies, media studies, and computer science. His expertise includes quantitative and qualitative social science methods, computational linguistics, and natural language processing, which he regularly combines in interdisciplinary, multi-method research.

Adrian Wüthrich is a guest professor for the History and Philosophy of Modern Science at Technische Universität Berlin. He leads the project Network Epistemology in Practice (NEPI), funded by a European Research Council Consolidator Grant, which investigates collective knowledge generation in large collaborations. His earlier work addressed interpretational debates in quantum mechanics, the role of images in science, causal reasoning in discovery processes, and the history of concepts in physics. His research often combines historical, philosophical, and computational approaches.

Michael Zichert is a PhD candidate in the project Network Epistemology in Practice (NEPI) at Technische Universität Berlin, where he studies how scientific knowledge is produced in large-scale collaborations. He holds a BA in History of Science from Technische Universität Berlin and an MA in Digital Humanities from Universität Leipzig. His research combines computational methods with the history of science, focusing on diachronic semantic change, social network analysis, and the application of large language models in the field.

Gerd Graßhoff is a professor for the History of Ancient Science at Humboldt-Universität zu Berlin and former director of the Excellence Cluster Topoi. He pioneered computational approaches to the history of science, applying machine learning and causal reasoning to ancient astronomical texts and scientific discovery processes. A Max Planck Fellow and member of Deutsche Akademie der Naturforscher Leopoldina, his research spans from Babylonian astronomy to digital humanities and AI methods for historical research.

Arno Simons, Adrian Wüthrich, Michael Zichert, Gerd Graßhoff (eds.)

Understanding Science with Large Language Models?

Potentials for the History, Philosophy, and Sociology of Science

[transcript]

This work was supported by the European Union under an ERC Consolidator Grant (Project No. 101044932, “Network Epistemology in Practice (NEPI)”). Views and opinions expressed are however those of the author only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available online at <https://dnb.dnb.de>



This work is licensed under the Creative Commons License BY 4.0. For the full license terms, please visit the URL <https://creativecommons.org/licenses/by/4.0/>. Creative Commons license terms for re-use do not apply to any content (such as graphs, figures, photos, excerpts, etc.) not original to the Open Access publication and further permission may be required from the rights holder. The obligation to research and clear permission lies solely with the party re-using the material.

2026 © Arno Simons, Adrian Wüthrich, Michael Zichert, Gerd Graßhoff (eds.)

transcript Verlag | Hermannstraße 26 | D-33602 Bielefeld | live@transcript-verlag.de

Cover design: Kordula Röckenhaus

Cover illustration: Adrian Wüthrich and Magdalena González (AI-assisted)

Printing: Elanders Waiblingen GmbH, Waiblingen

<https://doi.org/10.14361/9783839447529>

Print-ISBN: 978-3-8376-7994-6 | PDF-ISBN: 978-3-8394-4752-9

ISSN of series: 2703-1543 | eISSN of series: 2703-1551

Printed on permanent acid-free text paper.

Contents

Foreword	9
-----------------------	---

Doing science studies with large language models

An introduction

<i>Arno Simons, Adrian Wüthrich, and Michael Zichert</i>	13
--	----

Part 1: General challenges and limitations

Critical concerns for using LLMs in the (computational) humanities and beyond

<i>Sarah Lang</i>	35
-------------------------	----

AI and the scientist

On the fracture of epistemic authority

<i>Kristina Khutsishvili</i>	51
------------------------------------	----

Grounding AI in humanistic inquiry

Interdisciplinary challenges for evaluation and interpretability

<i>Oliver Eberle</i>	57
----------------------------	----

On the use and limitations of large language models in historical scholarship

<i>Holle Meding and Aurel Daugs</i>	71
---	----

Encoded humanities, or: not everything has to be generative

A dialogue on AI tasks and roles

<i>Andreas Wagner and Jürgen Hermes</i>	101
---	-----

LLMs as philosophers

What can they do, and why aren't they better?

<i>William D'Alessandro</i>	115
-----------------------------------	-----

Part 2: Historicizing LLMs

Why pursue temporally-grounded AI for historical disciplines, and what makes it so challenging?

Jochen Büttner 127

Generative LLMs and history research

Limitations for languages, periods, and tasks

Fernanda Olival, Helena Freire Cameron, António Branco, and Renata Vieira 143

Zero-shot generation of synthetic historical data with LLMs

Vera Danilova, Julia Reed, Andrew Burchell, Gijs Aangenendt, and Ylva Söderfeldt 151

LLMs and multilingual historical corpora in a digital history project

Reflections from the Berlin workshop

Jeffrey C. Wolf 179

Part 3: Linguistic change, markers, and differences

Epistemic framings in science

Charting scientific knowledge with embeddings and LLMs

Christophe Malaterre and Francis Lareau 191

Computational conceptual history of scientific concepts

From early digital methods to LLMs

Michael Zichert and Arno Simons 197

Meaning at the Planck scale?

Contextualized word embeddings for doing history, philosophy, and sociology of science

Arno Simons 217

Exploring disciplinary differences in semantic uniformity

A computational approach to codification

Elaheh Sadat Ahmadi 243

Use of large language models in the classification of scientific texts into disciplines

Diego Alves, Sergei Bagdasarov, Badr Mohammed Abdullah, and Stefania Degaetano-Ortlieb 257

Discursive parallels of the chemical revolution

Topic modelling and distributional analysis

Sofía Aguilar-Valdez, Bách Phan-Tăt, Dirk Speelman, Dirk Geeraerts, and Stefania Degaetano-Ortlieb 271

Part 4: Knowledge graphs and topic modeling

The potential of LLMs for constructing a socio-legal knowledge graph

Christian Boulanger 293

From source to structure

Extracting knowledge graphs with LLMs

Raphael Schlattmann, Aleksandra Kaye, and Malte Vogl 307

Large language models in interdisciplinary research settings

A reflection

Malte Vogl, Alexander von Schwerin, and Sabrina Kirschke 335

When less is more?

Topic model performance across title, abstract, and full-text corpora

Francis Lareau and Christophe Malaterre 343

Part 5: Retrieval-augmented generation (RAG)

From RAGs to rich responses

Enhancing LLM reliability through retrieval-augmented generation

Holle Meding and Aurel Daus 353

The data interview

Reflexive integration of large language models in qualitative content analysis

Miira Hill 369

Co-creation of AI technology, empowering curators of cultural heritage information and guarding research commons

Andrea Scharnhorst, Han Yang, Jetze Toubert, Kim Ferguson, Philipp Mayr,

and Vyacheslav Tykhonov 383

AI-Reporter: a path to a new genre of scientific communication

From presentation to publication through agentic LLMs

Gerd GraBhoff 413

Recursive reflections

Using LLMs to discuss LLMs in HPSS research

Jeffrey C. Wolf 439

Part 6: Citation context analysis

Supporting citation context analysis with large language models raises questions that should have been asked 40 years ago

Lydia Liesegang and Jochen Gläser 453

Reconstructive citation context analysis using large language models

A roadmap
Arno Simons, Hiba Arnaout, and Iryna Gurevych 471

Scaling In, Not Up?

Testing Thick Citation Context Analysis with GPT-5 and Fragile Prompts
Arno Simons..... 493