Niclas Griesshaber

niclasgriesshaber@outlook.com | niclasgriesshaber.com | github.com/niclasgriesshaber
+49 176 4260 4186 | linkedin.com/in/niclasgriesshaber | huggingface.co/niclasgriesshaber

Education

2024	**University College London**, *United Kingdom* MSc Computational Statistics and Machine Learning, Distinction
2023	**University of Oxford**, *United Kingdom* MSc Economic and Social History, Distinction, Dissertation Prize, Top of Class
2022	**National University of Singapore**, *Singapore* Exchange Semester
2022	**University of Tübingen**, *Germany* BSc Cognitive Science, First Class Honours, Top 10%
2021	**University of Tübingen**, *Germany* BSc International Economics, First Class Honours, 3rd Best Bachelor's Degree
2017	**Johannes-Kepler-Gymnasium**, *Germany* Abitur, 4th Best High School Graduate out of 116 Students
2015	**Holy Trinity Catholic Secondary School**, *Canada* Year Abroad during High School

Fellowships & Awards

2023	Joan Thirsk Prize for the Best Dissertation in the MSc Economic and Social History
2021	RWT GmbH Prize for the 3rd Best Bachelor's Degree in International Economics
2019	Scholarship Fellow of the Konrad-Adenauer-Foundation
2017	Winner Baden-Württemberg in the Federal President's History Competition
2017	Special Prize in Economics by Reutlingen's District Savings Bank
2017	History Prize of Reutlingen by the Historical Society of Reutlingen
2017	E-Fellows Online Scholarship for 4th Best High School Graduate

Research Experience

2025 (02–pres.)	Research Fellow, AI Economic History, Prof. Jochen Streb, University of Mannheim
2023 (5 mos.)	Research Assistant, Dr Giulia Caprini, Nuffield College, University of Oxford
2019 (3 mos.)	Research Assistant, Prof. Zhaoping Li, Max Planck Institute, Tübingen, Germany
2017 (1 mo.)	Research Assistant, Archivist Brad Smith, Berks History Center, Reading, PA, USA

Research Presentations

2025	**Upcoming Presentation at the World Economic History Congress, Lund** I will present my multimodal large language model (LLM) pipeline for constructing large-scale datasets from archival image scans at no cost and within minutes.
2024	**Microdata in Economic History: Beyond the Full-Count Census, LMU** My co-author, Gavin Greif (University of Oxford), presented our paper on benchmarking multimodal LLMs for building a dataset from eighteenth- and nineteenth-century German city directories.
2024	**Thirsk Dissertation Prize Lecture, All Souls College, University of Oxford** My lecture was part of the 32nd Annual Graduate Workshop in Economic History. I preceded the Sir John Hicks Memorial Lecture in Economic History, which was delivered by Prof. Joel Mokyr.

Teaching Experience

2022 Summer	**Teaching Assistant in Introduction to Machine Learning** *Chair of Computer Vision, Prof. Andreas Schilling, University of Tübingen*
2020 Winter	**Teaching Assistant in Quantitative Methods** *Chair of Statistics, Prof. Martin Biewen, University of Tübingen*
2019 Winter	**Teaching Assistant in Explorative Data Analysis** *Chair of Statistics, Prof. Martin Biewen, University of Tübingen*

Languages & AI Skills

Languages	English (TOEFL: 111/120), German (Native), Spanish (Native), French (DELF B1)
AI Skills	Python (5 years), PyTorch, LLMs, Hugging Face, Cursor IDE, Git, GitHub, LaTeX

Community Service

2022–2023	Middle Common Room Treasurer at Harris Manchester College, University of Oxford
2016–	Certified DSV Ski Instructor (Trainer C) at the Skiclub Offenburg
2018–2020	Co-founder and HR Director of Enactus Tübingen e.V.

Industry Experience

2019 (2 mos.) **KPMG, Intern in Digital Compliance**, *Stuttgart, Germany*

- Developed automated workflows for SAP audit processes using BluePrism and VBA, enhancing efficiency in audit report generation.
- Collaborated with the IT compliance team to identify automation opportunities, reducing manual input and error rates.

2018 (3 mos.) **Ernst & Young, Intern in Corporate Law**, *San José, Costa Rica*

- Advised clients on selecting legal structures for business incorporation, ensuring compliance with local regulatory requirements.
- Created client-facing presentations and internal reports in PowerPoint, clarifying complex legal concepts for non-legal stakeholders.

Personal Information

Citizenships	Ecuadorean and German

References

Sheilagh Ogilvie, FBA
Chichele Professor of Economic History
All Souls College, University of Oxford

Pontus Stenetorp
Professor of Natural Language Processing
University College London

Andreas Schilling
Professor of Computer Vision
University of Tübingen

Giulia Caprini
Prize Postdoctoral Fellow
Nuffield College, University of Oxford

Papers under Review

Transplanting Craft Guilds to Colonial Latin America: A Large Language Model Analysis
with Sheilagh Ogilvie; currently under review at The Journal of Economic History

What can we learn about institutional transplantation by analyzing craft guilds in colonial Latin America? We use large language models (LLMs) to investigate colonial guild ordinances, addressing two major bottlenecks in assessing institutions: digitizing qualitative sources efficiently and analyzing them quantitatively. Our newly designed methodology reveals both long-term continuities and striking differences between craft guilds in colonial Mexico and Peru, particularly with regard to human capital and product quality. The LLM-based approach identifies patterns that were previously not discernible using standard methods in economic history, its results are reproducible, and it can easily be extended to other historical settings.

Benchmarking Multimodal LLMs for OCR, OCR-Post-Correction, and Named-Entity Recognition in Historical Documents
with Gavin Greif; currently under review at ICDAR 2025

Archives hold a wealth of data, yet historical researchers continue to face transcription and classification bottlenecks to transform these sources into analyzable datasets. This paper presents a novel approach that uses multimodal large language models (LLMs) to accurately transcribe and parse historical printed documents. We demonstrate and benchmark the performance of our methodology using a corpus of pages from German-language city directories published between 1754 and 1870 in various historical fonts. The corpus was carefully curated to exhibit the typical challenges faced in automating the transcription of historical documents. We make three main contributions. First, we benchmark the optical character recognition (OCR) capabilities of multimodal LLMs, demonstrating that Gemini-2.0-Flash outperforms fine-tuned OCR algorithms and other LLMs while being much faster, accurate and at no cost. Second, we demonstrate that multimodal LLMs can further improve Transkribus OCR text, yielding the highest accuracy across all our experiments. Finally, we test how accurately LLMs can perform named-entity recognition on our historical documents, both from the transcribed ground truth and directly from the archival image scans. We find initial evidence that, under specific circumstances, multimodal LLMs are accurate enough to substantially accelerate the process of constructing datasets from historical structured documents.

Papers in Progress

Multimodal Large Language Models for Dataset Creation from Archival Image Scans: The Complete Patent Registers of the German Empire
with Jochen Streb; in preparation for the NeurIPS 2025 Datasets and Benchmarks Track

We constructed the complete dataset of German Empire patent registers (1871 to 1914), comprising approximately 700,000 patent entries, including patent descriptions and many other variables. We use Gemini-2.0-Flash to construct the dataset directly from archival image scans, rapidly, accurately, at zero cost, and with full replicability. Our multimodal LLM dataset construction pipeline effectively manages older, type-faced fonts without relying on traditional optical character recognition (OCR) algorithms or manual annotations, overcoming longstanding barriers in dataset construction from archival image scans. By utilizing prompt engineering, researchers can select and precisely define the variables they want extracted before dataset construction, thus integrating domain-specific historical expertise into the dataset construction process. Crucially, our pipeline significantly outperforms human research assistants in accuracy, speed, and cost efficiency. Additionally, it can be readily extended to other archival materials, including more legible handwritten sources. Therefore, our multimodal LLM dataset construction pipeline represents an unprecedented advance for the digital humanities, substantially enhancing access to vast historical data previously locked in archival image scans.

A New Historical Dataset for Machine Translation between Early German and English: Cross-Lingual Transfer, Many-Shot In-Context Learning, and Fine-Tuning
with Sheilagh Ogilvie, Jiayi Wang, and Yao Lu; in preparation for submission to ACL

We present a new historical text dataset for machine translation consisting of 3,873 paragraph-level pairs between Early Modern Bohemian German and English. The dataset contains more than 750,000 words and is based on handwritten legal court records dated from 1582 to 1787. The primary sources were manually transcribed and translated across several years by the Chichele Professor of Economic History at the University of Oxford. We evaluate the translation accuracy of open-source large language models (LLMs) on this low-resource historical language pair. These models struggle to generalize their abilities through cross-lingual transfer, as evident from their low translation scores (e.g., for Early Modern Bohemian German to English: BLEU 11.50, COMET 0.57; for the reverse direction: BLEU 2.19, COMET 0.48). To address these challenges, we successfully replicated a recent many-shot in-context learning paradigm for low-resource languages, demonstrating continuous improvements from 0 to 128 shots. Using LLaMA-3.1-8B-Instruct, BLEU scores increased to 24.32 (COMET 0.66) for Early Modern Bohemian German to English and to 7.23 (COMET 0.54) for the reverse direction. Fine-tuning remains the performance ceiling across all experiments, with BLEU scores of 36.70 (COMET 0.69) and 9.55 (COMET 0.57) respectively, highlighting that in-context learning may not reach the same level of effectiveness in handling out-of-distribution data. These findings reinforce the need for more interdisciplinary research, as adapting LLMs to historical languages not only advances our understanding of transformer architectures but also sharpens new analytical tools to advance the field of history.