

# Niclas Griesshaber

niclasgriesshaber@outlook.com  
+49 \*\*\*\*\*

niclasgriesshaber.com  
linkedin.com/in/niclasgriesshaber

github.com/niclasgriesshaber  
huggingface.co/niclasgriesshaber

## Education

---

2028/10	<b>University of Oxford, <i>United Kingdom</i></b> DPhil in Economic and Social History, In Progress
2024	<b>University College London, <i>United Kingdom</i></b> MSc Computational Statistics and Machine Learning, Distinction
2023	<b>University of Oxford, <i>United Kingdom</i></b> MSc Economic and Social History, Distinction, Dissertation Prize, Top of Class
2022	<b>National University of Singapore, <i>Singapore</i></b> Exchange Semester
2022	<b>University of Tübingen, <i>Germany</i></b> BSc Cognitive Science, First Class Honours, Top 10%
2021	<b>University of Tübingen, <i>Germany</i></b> BSc International Economics, First Class Honours, 3rd Best Bachelor's Degree
2017	<b>Johannes-Kepler-Gymnasium, <i>Germany</i></b> Abitur, 4th Best High School Graduate out of 116 Students
2015	<b>Holy Trinity Catholic Secondary School, <i>Canada</i></b> Year Abroad during High School

## Grants & Awards

---

2026	NFDI4Memory Incubator Funds for building historical datasets with AI, €65,000
2025–2028	ESRC Advanced Quantitative Methods Award, £180,000+ DPhil Scholarship
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
2019–2024	Fellow of the Konrad-Adenauer-Foundation, €40,000+ Scholarship
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

## Papers

---

2025	Niclas Griesshaber and Jochen Streb, "Multimodal LLMs for Historical Dataset Construction from Archival Image Scans: German Patents (1877–1918)", <i>arXiv preprint arXiv:2512.19675</i> , <a href="https://arxiv.org/abs/2512.19675">https://arxiv.org/abs/2512.19675</a> .
2025	Gavin Greif, Niclas Griesshaber, and Robin Greif, "Multimodal LLMs for OCR, OCR Post-Correction, and Named Entity Recognition in Historical Documents", <i>arXiv preprint arXiv:2504.00414</i> , <a href="https://arxiv.org/abs/2504.00414">https://arxiv.org/abs/2504.00414</a> .
2025	Niclas Griesshaber and Sheilagh Ogilvie, "Transplanting Craft Guilds to Colonial Latin America: A Large Language Model Analysis", <i>CEPR Discussion Paper No. 20556</i> , CEPR Press, Paris & London. <a href="https://cepr.org/publications/dp20556">https://cepr.org/publications/dp20556</a> .

## Research Fellowships & Assistantships

---

2025 (8 mos.)	Prof. Jochen Streb, University of Mannheim <ul style="list-style-type: none"><li>• Led 4 research assistants in document scanning and ground truth production</li><li>• Built an AI pipeline to construct historical datasets from archival image scans</li></ul>
2023 (5 mos.)	Dr. Giulia Caprini, Nuffield College, University of Oxford
2019 (3 mos.)	Prof. Zhaoping Li, Max Planck Institute for Biological Cybernetics, Tübingen
2017 (1 mo.)	Bradley Smith, Berks History Center in Reading, Pennsylvania, USA

## Invited Seminars & Conferences

---

2026	Computational History Workshop, CAGE, University of Warwick
2025	The Oxford, Warwick and LSE (OWL) Workshop in Economic History
2025	Economic and Social History Graduate Seminar, University of Oxford
2025	Economic History Colloquium, University of Mannheim
2025	World Economic History Congress, Lund University
2024	Microdata in Economic History: Beyond the Full-Count Census, LMU
2024	Joan Thirsk Dissertation Prize Lecture, All Souls College, University of Oxford

## Teaching

---

2022 Summer	Teaching Assistant in Introduction to Machine Learning, University of Tübingen
2020 Winter	Teaching Assistant in Quantitative Methods in Economics, University of Tübingen
2019 Winter	Teaching Assistant in Explorative Data Analysis, University of Tübingen

## Machine Learning Projects

---

2025	<b>Multimodal LLMs for Historical Dataset Construction from Archival Image Scans: German Patents (1877–1918)</b> <a href="https://github.com/niclasgriesshaber/llm_patent_pipeline">github.com/niclasgriesshaber/llm_patent_pipeline</a> <ul style="list-style-type: none"><li>• Built an LLM-based data pipeline to extract 300,000+ German patent entries (1877–1918) from 9,000+ archival image scans with \$1,000 API cost</li><li>• Constructed a benchmarking dataset to evaluate the accuracy of multimodal LLMs on historical dataset construction</li></ul>
2025	<b>Multimodal LLMs for Historical OCR and NER Benchmarking</b> <a href="https://github.com/niclasgriesshaber/llm_historical_dataset_benchmarking">github.com/niclasgriesshaber/llm_historical_dataset_benchmarking</a> <ul style="list-style-type: none"><li>• Built OCR and NER benchmark comparing GPT-4o, Gemini-2.0-Flash, and standard OCR algorithms on German 18th- and 19th-century city directories</li></ul>
2024	<b>Many-Shot Learning and LoRA Fine-tuning for Historical Translations</b> <a href="https://github.com/niclasgriesshaber/llm_historical_machine_translation">github.com/niclasgriesshaber/llm_historical_machine_translation</a> <ul style="list-style-type: none"><li>• Conducted experiments on Llama-3.1-8b-instruct, Gemma-2-9b-instruct, and other small open-source models using UCL’s GPU cluster</li><li>• Experiments on English-to-early-modern-German machine translation using privately held legal dataset from the 16th-18th centuries (1.4M tokens)</li></ul>
2023	<b>Historical Text Classification with GPT-4</b> <a href="https://github.com/niclasgriesshaber/guilds_llm">github.com/niclasgriesshaber/guilds_llm</a> <ul style="list-style-type: none"><li>• Developed LLM-based sentence classification pipeline for historical text analysis</li></ul>
2022	<b>Learning Feature Point Descriptor with Triplets of Image Patches</b> <a href="https://github.com/niclasgriesshaber/computer_vision_bachelor_thesis">github.com/niclasgriesshaber/computer_vision_bachelor_thesis</a> <ul style="list-style-type: none"><li>• Trained a feature point descriptor using residual neural networks and triplet loss to identify corresponding 3D points across different views of the same scene</li><li>• Applied epipolar geometry constraints on the MegaDepth Dataset to match pixels</li></ul>

## 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

## 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.

## Languages & Software Skills

---

- Languages English (Professional Proficiency), German (Native), Spanish (Native), French (B1)
- Skills Python, Git, SQL, Docker, Gemini and OpenAI API, Agentic IDEs

## Personal Information

---

- Citizenship Ecuadorean and German
- Interests Table Football, Video Games, Skiing, and Gym

## 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

**Stephen Broadberry, FBA**  
Professor of Economics  
Nuffield College, University of Oxford

**Andreas Schilling**  
Professor of Computer Vision  
University of Tübingen

**Jochen Streb**  
Professor of Economic History  
University of Mannheim