

# Terminology in the age of large language models:

Triumphs, traps and turmoil



# LLMs and terminology (management)

# LLMs as disruptive technology

Large language models (LLMs)

- Cover (a broad range of) natural language processing tasks
- Intuitive
- Based on vast amounts of data
- Quality and speed
- → Huge impact on (language) professions



Source: Image by Lucas from Pixabay: [https://pixabay.com/?utm\\_source=link-attribution&utm\\_medium=referral&utm\\_campaign=image&utm\\_content=5009454](https://pixabay.com/?utm_source=link-attribution&utm_medium=referral&utm_campaign=image&utm_content=5009454)

# Aim of this keynote

Exploring the dual role of LLMs in terminology work

- How does AI depend on terminology?
- How do terminologists benefit from AI?

Hinting at the triumphs, traps and turmoil



# LLMs and terminology (management)

- Hype surrounding large language models (LLMs)
  - LLMs allow for new ways of conducting terminology work and management
  - LLMs and job and role displacement

# Technology and terminology (management)

- Technology was and is crucial in terminology work
  - Corpus analysis tools
  - Terminology extraction tools
  - Terminology management systems
  - Terminology workflow systems
  - Terminology checkers
  - etc.



→ These systems guarantee consistent and predictable outputs  
(Massion 2024)

# The triumphs

# AI4Good

- Inclusivity
- Sustainability
- Ethical considerations



# AI for Good

# Terminology triumphs

- LLMs for terminologists: Term extraction, research, database creation
- Terminology as a resource for LLMs: Impact on automated content generation

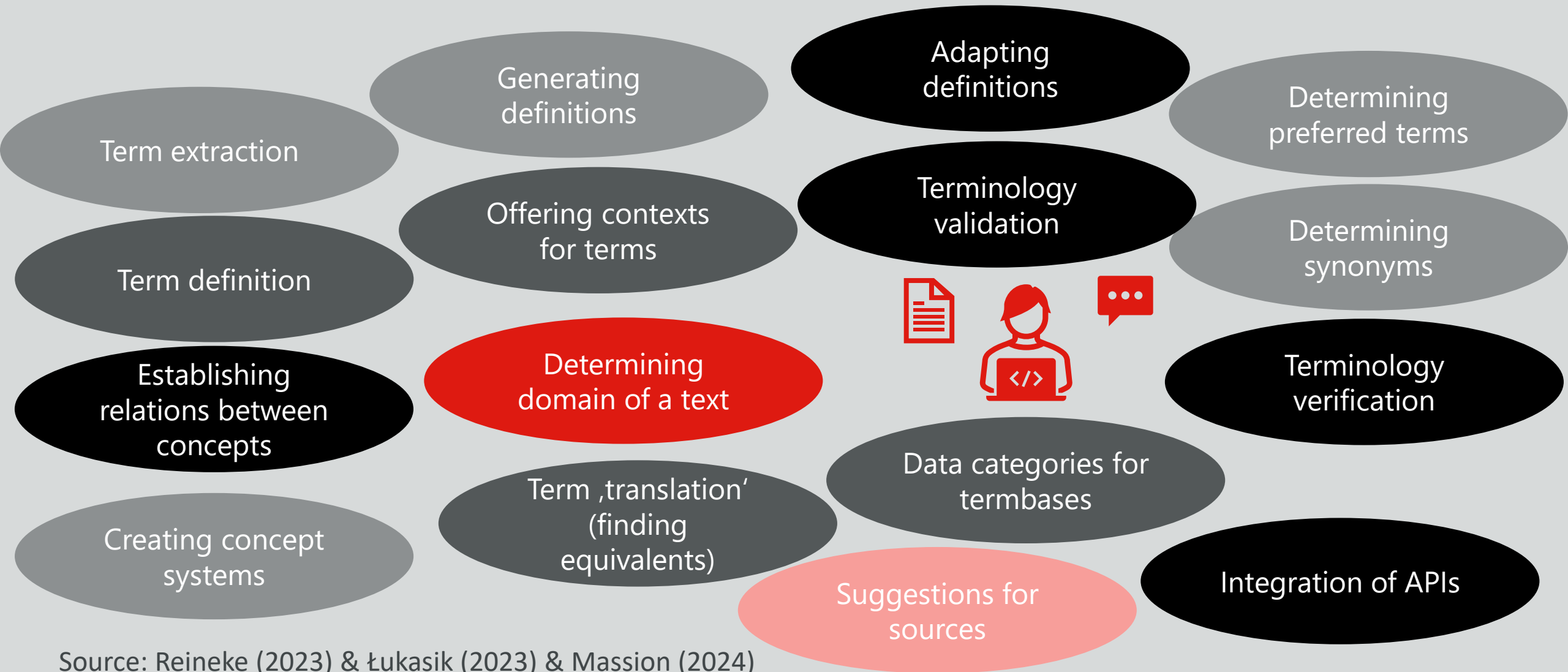


# Terminology triumphs

- LLMs for terminologists: Term extraction, research, database creation
- Terminology as a resource for LLMs: Impact on automated content generation



# LLMs for terminology work



Source: Reineke (2023) & Łukasik (2023) & Massion (2024)

# Example: LLMs for definition generation

## Rules:

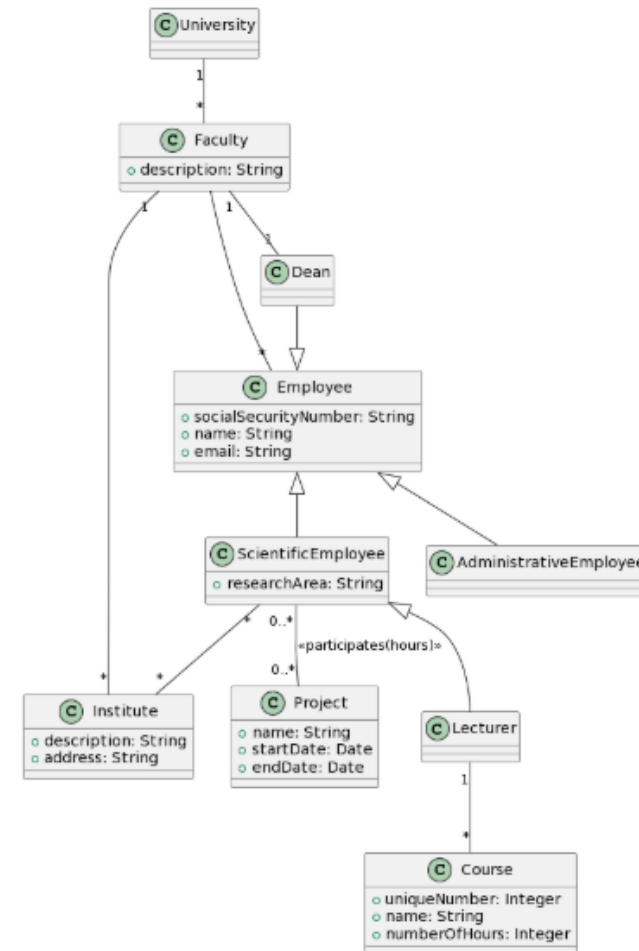
- Try to write the definition in a single sentence
- Always start with a capital letter, preferably an indefinite article and end the definition with a dot (.)
- Never mention the term itself.
- Never mention synonyms of the term.
- Don't add examples to the definition.
- Consider the following subject(s): “nature”
- Consider the following context: “The children played happily along the bank, collecting colorful stones and watching the water flow.”
- Remember, the term to define is “bank” in the language: English

Source: Fleischmann et al. 2025

# Example: LLMs for generating concept systems

## Task Definition 3


A university consists of multiple faculties, which are composed themselves of multiple institutes. Every faculty and each institute have a description. Each institute has an address. Every faculty is led by a dean, who is an employee of the faculty. The total number of employees is known. Each employee has a social security number, a name and an email address. It is distinguished between scientific and administrative employees. Scientific employees are assigned to at least one institute. For each scientific employee, his research area is known. In addition, scientific employees can participate in projects for a particular number of hours. For the projects, the name and the start and end date are known. Some scientific employees give courses. Then they are denoted as lecturers. Courses have a unique number, a name, and a number of hours.



Source: Fill et al. 2023

# Example: Preparation for terminology database

Hier ist eine Tabelle mit deutschen und britischen englischen Benennungen sowie deren Definitionen im Bereich der Studienzulassung:

Deutsche Benennung	Deutsche Definition	Quelle (ac.at)	Englische Benennung	Englische Definition	Quelle (ac.uk)
Studienzulassung	<p>Die Studienzulassung ist das Eingangstor zur Universität Wien. Sie ist zuständig für die Erstzulassung von Bachelor-, Diplom- und Masterstudien. Ihr Aufgabenbereich umfasst alle Angelegenheiten, die mit der Zulassung zum Studium zusammenhängen.</p>	<p>STUDIEN.UNIVIE.AC.AT</p> <p> studieren.univie.ac.at</p> <p>Begriffs-ABC</p>	University admission	The process through which students apply and are accepted into a university program, typically involving the submission of academic qualifications and meeting specific entry requirements.	<p>University of Oxford</p> <p>Admissions</p>

Source: Heinisch in print

## Prompt:

For comparative terminology work in the university sector in German and English, compile a list of the most common terms in the field of university admissions. Add definitions (according to ISO 704 and ISO 1087) from Austrian and British websites and preferably from normative or official sources, such as laws or documents from authorities or organisations. Here is an example: Sammelzeugnis = 'Zeugnis über alle absolvierten Prüfungen eines Studierenden an einer Universität', the English equivalent is: transcript of records = 'official document summarising a university student's academic performance and progress to date'. Add the URL of the website where you found the definitions.

# Prompts for terminology purposes

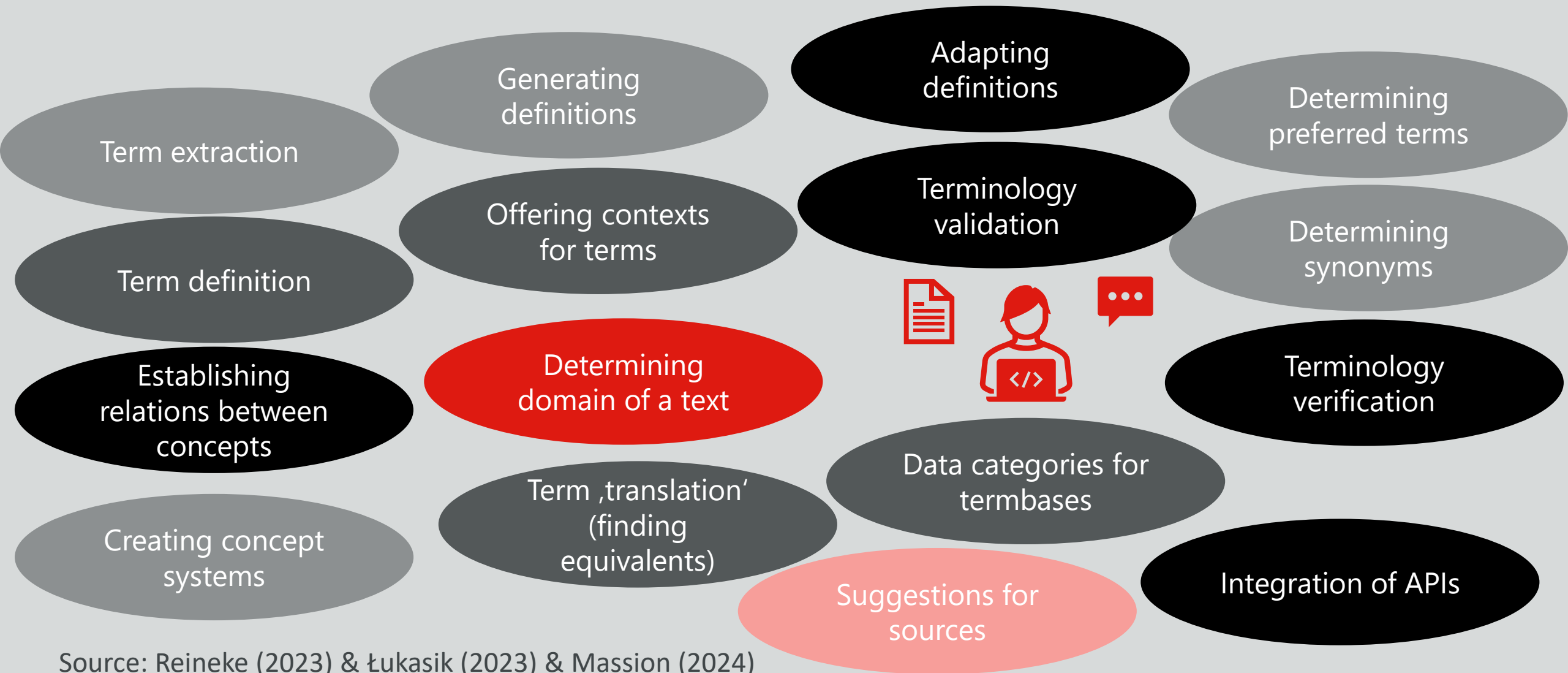
- **Domain**  
Definition of the relevant subject field.
- **System**  
Considering the system boundness of terminology.
- **Context**  
Paying attention to the purpose.



Terminology  
prompts

.

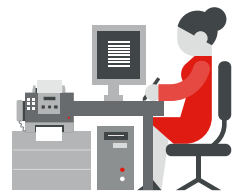
# LLMs for terminology work



Source: Reineke (2023) & Łukasik (2023) & Massion (2024)

# LLMs for terminology work

- LLMs offer opportunities in terminology work
- But:
  - Multilingual terminology work
    - Intransparency regarding training data (which languages)
    - English training data
    - Establishing equivalence
  - Explainable AI vs. deterministic traditional tools
  - Niche domains and emerging domains
- LLM output quality
  - The quality of the prompt determines the quality of the LLM output
- But:
  - Inherent characteristics of the model
  - Language support
  - Biases and hallucinations



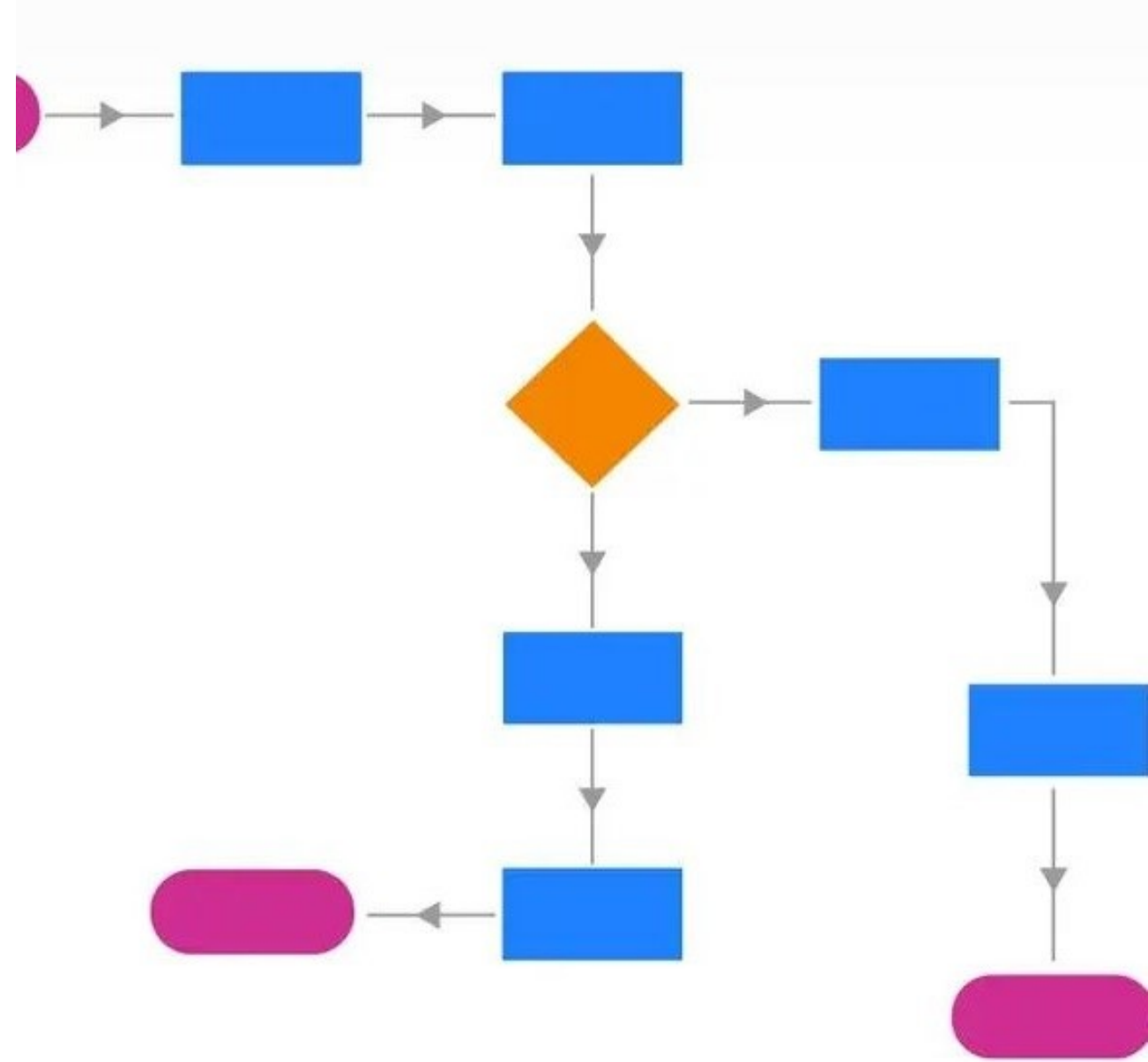
# Terminology triumphs

- LLMs for terminologists: Term extraction, research, database creation
- **Terminology as a resource for LLMs:** Impact on automated content generation

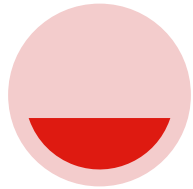


# Terminology as a resource for AI

The renaissance of terminology  
through LLMs

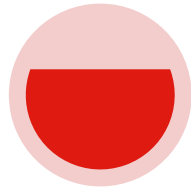


# Terminology as a resource for large language models



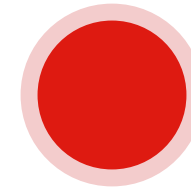
RAG

Retrieval-  
augmented  
generation



TAG

Terminology-  
augmented  
generation



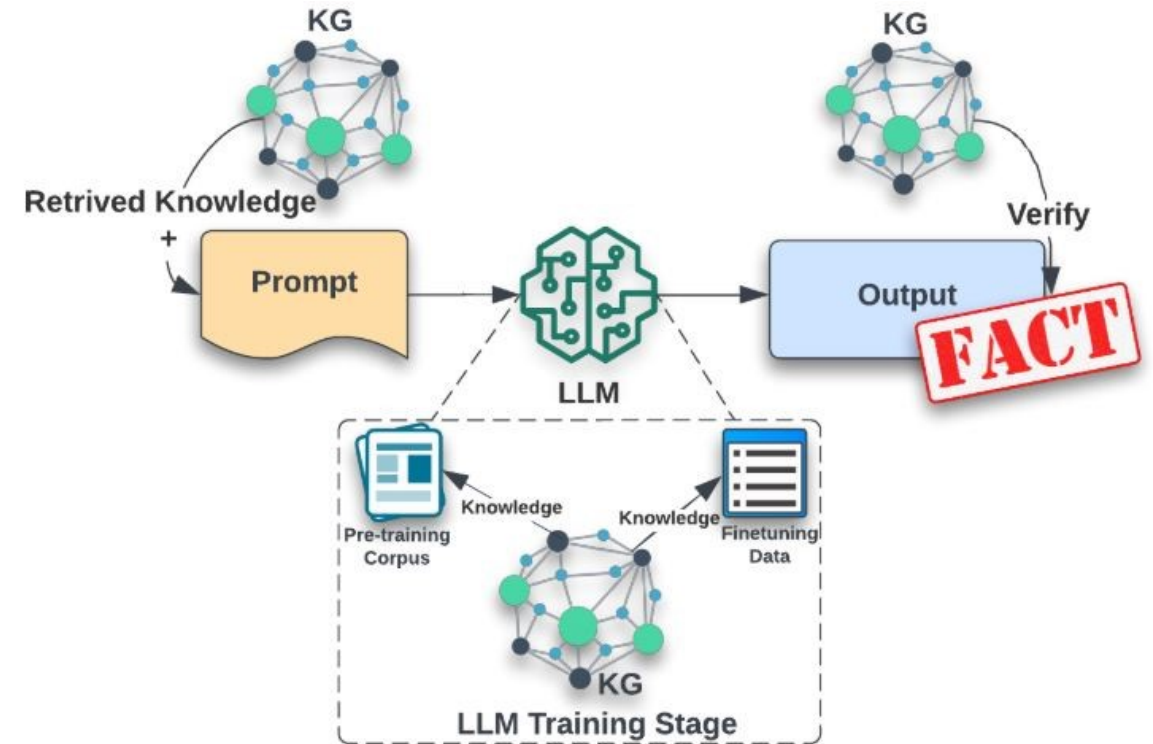
KGAG

Knowledge  
graph-  
augmented  
generation

# Terminology as a resource for AI

## Retrieval-augmented generation

- Terminology-augmented generation (Fleischmann & Lang 2025)
- Knowledge graph-augmented generation / LLMs (Agrawal et al. 2024)

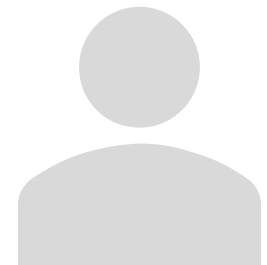
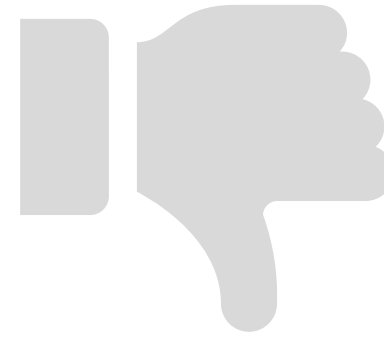


Source: Agrawal et al. 2024

# The traps

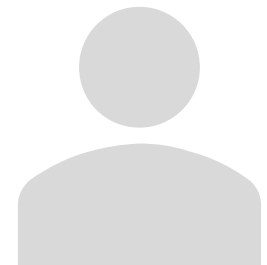
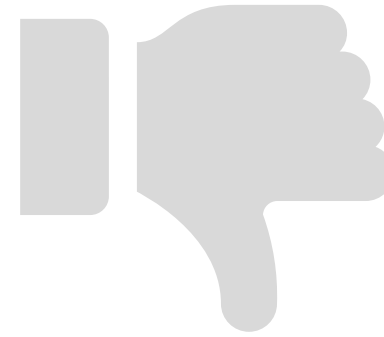
# The traps

- **Training costs and sustainability:** Environmental and socio-economic impact of LLMs
- **Language change and homogenisation risks:** Effects on linguistic diversity within and across languages



# The traps

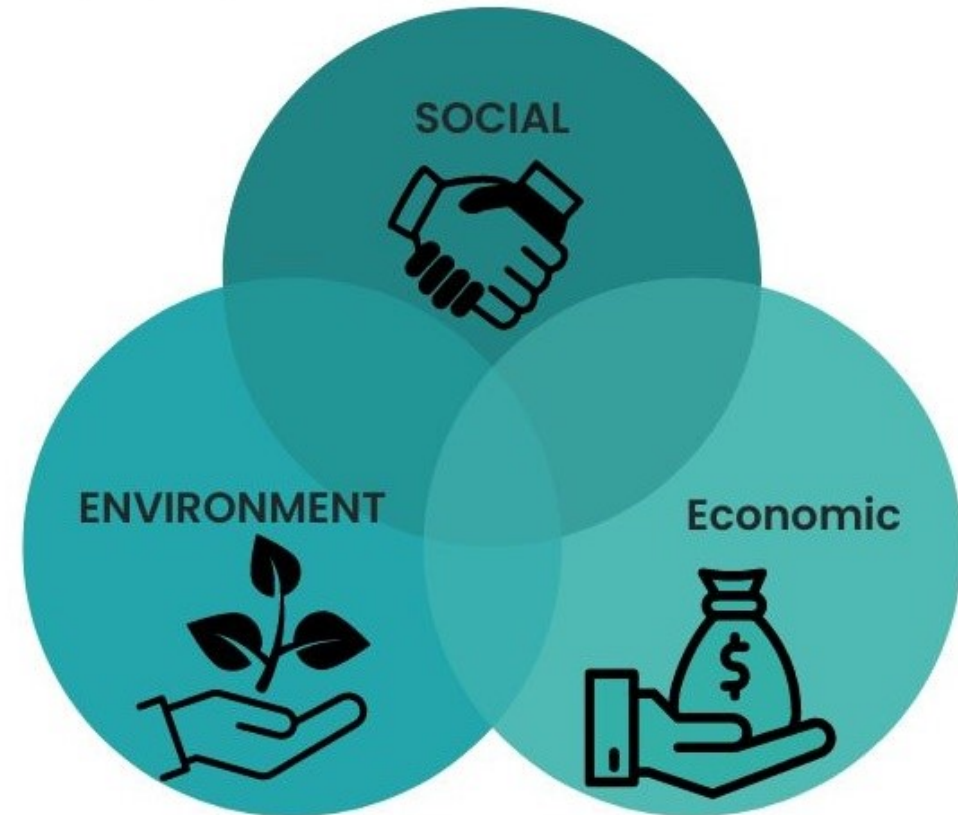
- **Training costs and sustainability:** Environmental and socio-economic impact of LLMs
- **Language change and homogenisation risks:** Effects on linguistic diversity within and across languages



## LLMs and sustainability

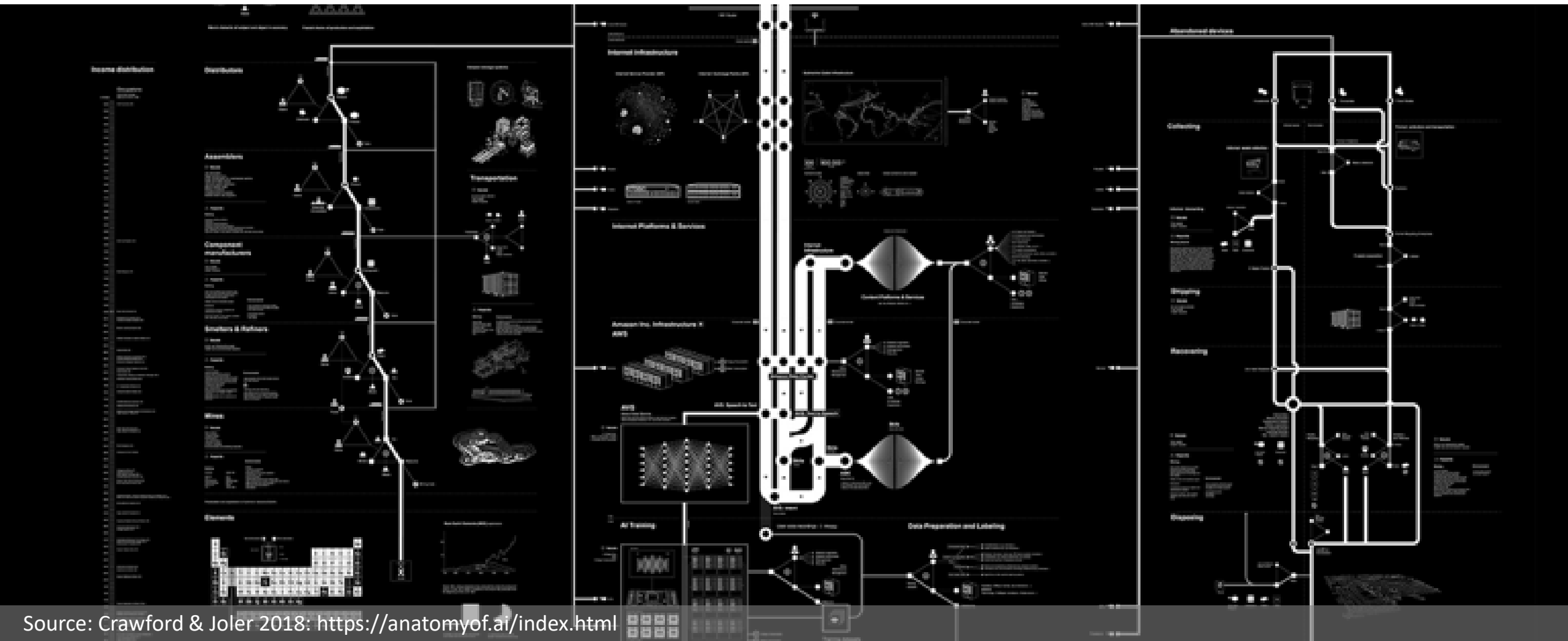
- Sustainability encompasses:
  - Social
  - Environmental
  - Economic

# THE THREE COMPONENTS OF SUSTAINABILITY



Source: Sustainability-Components CC BY-SA 4.0 Mcnultym

# LLMs, ethics and sustainability



Source: Crawford & Joler 2018: <https://anatomyof.ai/index.html>

# The turmoil

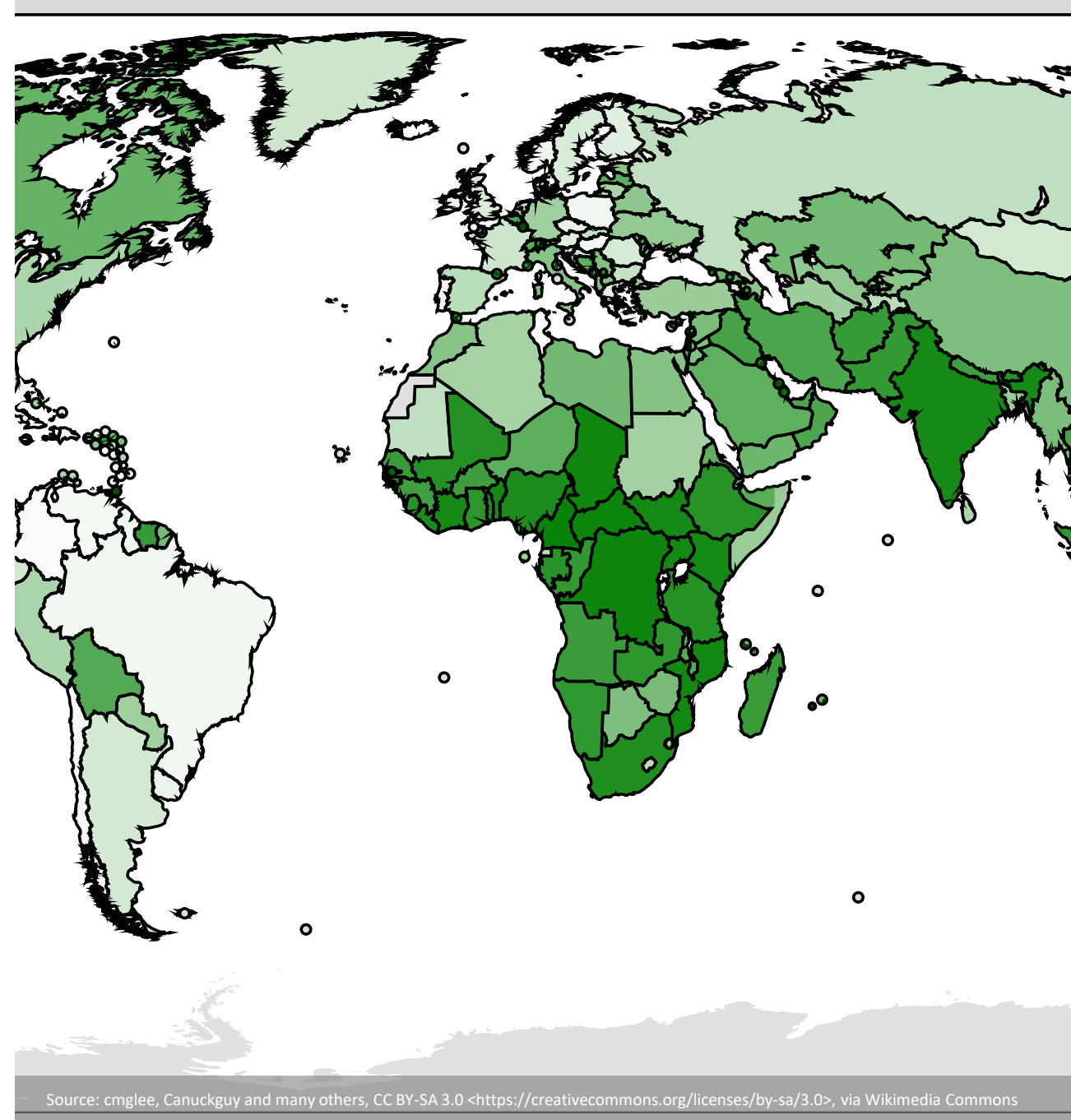
# The turmoil

- LLM's role in shaping (specialised) language, and thus terminology:  
Impact on language(s) and its implications



# LLMs and language (change)

- How do LLMs affect...?
  - Language change
  - Language diversity
  - Linguistic creativity



Source: cmglee, Canuckguy and many others, CC BY-SA 3.0 <<https://creativecommons.org/licenses/by-sa/3.0/>>, via Wikimedia Commons

# LLMs and language homogenisation

Over-reliance on AI-generated content may result in loss of “originality, diversity and authenticity in language and communication” (Massion 2024)

# Language modelling bias

Language modelling bias (Bella et al. 2024)

- LLMs focus on the world's 2-3% most widely spoken languages
- “multilingual language processing systems often exhibit a hardwired, yet usually involuntary and hidden representational preference towards certain languages” (Bella et al. 2024)

LLMs do not give access to the underlying data on which its output is based (Rundell 2024)

# LLMs and low-resource languages

LLMs show...

- degradation of performance for low-resource languages (Dwivedi et al. 2024)
- resistance to low-resource language varieties (Heinisch in print)
- Despite prompting according to system boundness, context dependence and language variety

Terminology characteristics	Aspect	Considered in the prompt
Domain specificity	University terminology >> studies >> admission	„in the university context“ „in the field of university admission“
System boundness (system)	(Austrian) university system	„Austrian higher education system“ „universities in Austria“
System boundness (language variety)	Austrian	„Austrian German“ „Austrian variety of the German language“
Context dependence	University vs. University of applied sciences; terminological variation, individual universities	By including source hierarchy, domains (ac.at) from which terms and definitions should be extracted and vary the terminology used in the prompt, e.g. ‘university’ or ‘higher education system’; ‘Benennungen’ or ‘Termini’ (in German)

# LLMs impact (specialised) language

- LLMs contributing to language change
  - Some phrases becoming more popular (Matsui 2024)
- ‘Synthetic’ data: LLMs re-using their own output
  - Can we (or do we have) to create ‘authentic’ corpora in future?
- Support of languages (and language varieties)
- Hallucinations especially relevant in new fields of knowledge or emerging domains
- Biases (linguistic representation, schools of thought, etc.)

# Addressing the traps and the turmoil

# Language support

- EuroLLM
  - Multilingual AI model
  - Supporting all 24 official EU languages
- Small and very small language models

**EURO\*LLM**

**Meet EuroLLM**  
**Large language model**  
**made in Europe**  
**built to support all**  
**official 24 EU languages**



Source: <https://euollm.io/>

# Digital Humanism

- argues for technologies that are shaped
- according to human needs and values
- "instead of allowing technologies to shape humans" (Werthner et al. 2019)



# VIENNA MANIFESTO ON DIGITAL HUMANISM

*VIENNA, MAY 2019*

»The system is failing« – stated by the founder of the Web, Tim Berners-Lee – emphasizes that while digitalization opens unprecedented opportunities, it also raises serious concerns: the monopolization of the Web, the rise of extremist opinions and behavior orchestrated by social media, the formation of filter bubbles and echo chambers as islands of disjoint truths, the loss of privacy, and the spread of digital surveillance. Digital technologies are disrupting societies and questioning our understanding of what it means to be human. The stakes are high and the challenge of building a just and democratic society with humans at the center of technological progress needs to be addressed with determination as well as scientific ingenuity. Technological innovation demands social innovation, and social innovation requires broad societal engagement.

**This manifesto is a call to deliberate and to act on current and future technological development.** We encourage our academic communities, as well as industrial leaders, politicians, policy makers, and professional societies all around the globe, to actively participate in policy formation. Our demands are the result of an emerging process that unites scientists and practitioners across fields and topics, brought together by concerns and hopes for the future. We are aware of our joint responsibility for the current situation and the future – both as professionals and citizens.

Source: Werthner et al. 2019

# AI ethics

The Principles of the **Recommendation on the Ethics of Artificial Intelligence**:

- proportionality and do no harm
- safety and security
- fairness and non-discrimination
- sustainability
- right to privacy and data protection
- human oversight and determination
- transparency and explainability
- responsibility and accountability
- awareness and literacy
- multi-stakeholder and adaptive governance and collaboration



# Conclusion: LLMs and (language) change

- LLMs raise serious ethical questions → corporate social responsibility
- LLMs offer opportunities and threats with regard to sustainability
  - AI4Good
  - Negative environmental, social and economic impact
- LLMs affect language

# Conclusion:

## LLMs for terminology and terminology for LLMs

- LLMs for terminology
  - LLMs as another tool in the toolbox of terminologists (and those who are not)
  - Require (expert) knowledge and scrutiny
- Terminology for LLMs:
  - Terminology as a valuable resource for content generation by means of LLMs



# Acknowledgement

The project *University terminology in German in the age of Chat GPT (UniTermGPT)* has received funding from the Autonomous Province of Bozen/Bolzano – Department for Innovation, Research and University in the frame of the EC-MCSA Seal of Excellence Programme.

AUTONOME PROVINZ BOZEN - SÜDTIROL



PROVINCIA AUTONOMA DI BOLZANO - ALTO ADIGE

PROVINZIA AUTONOMA DE BULSAN - SÜDTIROL

# Contact

**Barbara Heinisch**

barbara.heinisch@eurac.edu

**www.eurac.edu**

# Literature

Agrawal, G., Kumarage, T., Alghamdi, Z., & Liu, H. (2024). Can Knowledge Graphs Reduce Hallucinations in LLMs? : A Survey. In K. Duh, H. Gomez, & S. Bethard (Eds.), Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers) (pp. 3947–3960). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.naacl-long.219>

Bella, G., Helm, P., Koch, G., & Giunchiglia, F. (2024). Tackling Language Modelling Bias in Support of Linguistic Diversity. Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency. [https://hal.science/hal-04564896/file/FACCT\\_PREPRINT\\_Tackling\\_Language\\_Modelling\\_Bias\\_to\\_Support\\_Linguistic\\_Diversity-1.pdf](https://hal.science/hal-04564896/file/FACCT_PREPRINT_Tackling_Language_Modelling_Bias_to_Support_Linguistic_Diversity-1.pdf)

Crawford, K., & Joler, V. (2018). Anatomy of an AI system: The Amazon Echo as an anatomical map of human labor, data and planetary resources. <https://anatomyof.ai/index.html>

Dwivedi, S., Ghosh, S., & Dwivedi, S. (2024). Navigating Linguistic Diversity: In-Context Learning and Prompt Engineering for Subjectivity Analysis in Low-Resource Languages. SN Computer Science, 5(4). <https://doi.org/10.1007/s42979-024-02770-z>

Fill, H.-G., Fettke, P., & Köpke, J. (2023). Conceptual Modeling and Large Language Models: Impressions From First Experiments With ChatGPT. Advance online publication. <https://doi.org/10.18417/EMISA.18.3>

Fleischmann, K., & Lang, C. (2025). Terminologie in der KI: Wie mit Terminologie der Output von LLMs und GenAI optimiert werden kann. In P. Drewer, F. Mayer, & D. Pulitano (Eds.), Terminologie in der KI - KI in der Terminologie: Akten des Symposions Worms, 27.-29. März 2025 (pp. 83–95). Deutscher Terminologie-Tag e.V.

Fleischmann, K., Lang, C., & Lackner, A. (May 2025). Terminology Augmented Generation - How to Teach LLMs Your Corporate Language, online. <https://kaleidoscope.at/en/webinars/terminology-augmented-generation/>

He, Z., Liang, T., Jiao, W., Zhang, Z., Yang, Y., Wang, R., Tu, Z., Shi, S., & Wang, X. (2024). Exploring Human-Like Translation Strategy with Large Language Models. Transactions of the Association for Computational Linguistics, 12, 229–246. [https://doi.org/10.1162/tacl\\_a\\_00642](https://doi.org/10.1162/tacl_a_00642)

Heinisch, B. (in print). Large language models for terminology work: A question of the right prompt?

Łukasik, M. (2023). Terminology Work in the Era of AI-generated Content. A Case of ChatGPT. Presentation at 5th International Conference on Terminology: Scientific, Administrative and Educational Dimensions of Terminology. 19-20 Oct. 2023.

Massion, F. (2024). Terminology in the Age of AI: The Transformation of Terminology Theory and Practice. Journal of Translation Studies, 1, 67–94. <https://doi.org/10.3726/JTS012024.04>

Matsui, K. (2024). Delving into PubMed Records: Some Terms in Medical Writing Have Drastically Changed after the Arrival of ChatGPT. Advance online publication. <https://doi.org/10.1101/2024.05.14.24307373>

Reineke, D. (2023). Terminologearbeit mit ChatGPT & Co. Fachzeitschrift Für Terminologie, 19(1), 25–28. Retrieved from <http://dttev.org/images/edition/ausgaben/edition-2023-1-e-version.pdf>

Rundell, M. (2024). Automating the Creation of Dictionaries: Are We Nearly There? Humanising Language Teaching, 26(1).

UNESCO (2022). Recommendation on the Ethics of Artificial Intelligence. Edited by United Nations Educational, Scientific and Cultural Organization. Paris. Available online at <https://unesdoc.unesco.org/ark:/48223/pf0000381137/PDF/381137eng.pdf.multi>

Werthner, H., Lee, E. A., Akkermans, H., Vardi, M., & et al. (2019). Vienna Manifesto on Digital Humanism. Retrieved from [https://caiml.dbai.tuwien.ac.at/dighum/dighum-manifesto/Vienna\\_Manifesto\\_on\\_Digital\\_Humanism\\_EN.pdf](https://caiml.dbai.tuwien.ac.at/dighum/dighum-manifesto/Vienna_Manifesto_on_Digital_Humanism_EN.pdf)