Terminology Management Meets AI: The ISO/TC 37/SC 3/WG 6 Initiative

Mohamed Khemakhem[1], Cristina Valentini[2], Natascia Ralli[3], Sérgio Barros[2], Georg Löckinger[5], Federica Vezzani[6], Ana Salgado[9], Zhenling Zhang [7], Sabine Mahr[4], Sara Carvalho[10], Klaus Fleischmann[8], Rute Costa[9]

[1]MandaNetwork, [2]World Intellectual Property Organization,
[3]Eurac Research, Institute for Applied Linguistics, [4]Word b sign Sabine Mahr,
[5]University of Applied Sciences Upper Austria, [6]Università degli Studi di Padova,
[7]Computer Science College, Liaocheng University, Shandong Province, [8]Kaleidoscope,
[9]Centro de Linguística da Universidade NOVA de Lisboa, [10]Universidade de Aveiro

Abstract

The integration of artificial intelligence (AI) with terminology management (TM) has opened new avenues for enhancing efficiency and precision in both fields, necessitating standardized approaches to ensure interoperability and ethical application. The newly formed ISO/TC 37/SC 3/WG 6 represents the first dedicated initiative to study the standardization of the mutual improvements of AI and TM. This group aims to develop standardized frameworks and guidelines that optimize the interaction between AI technologies and terminology resources, benefiting professionals, systems, and practices in both domains. This article presents the state-of-the-art in the mutual relationship between AI and TM, highlighting opportunities for bidirectional advancements. It also addresses limitations and challenges from a standardization perspective. By tackling these issues, ISO/TC 37/SC 3/WG 6 seeks to establish principles that ensure scalability, precision, and ethical considerations, shaping future standards to support global communication and knowledge exchange.

1 Introduction

TM is critical to global communication by supporting the common understanding and exchange of specialized knowledge across languages and cultures. Within ISO's Technical Committee 37 "Language and terminology" (ISO/TC 37) [1], Subcommittee 3 (SC 3) [2] "Management of terminology resources" has played a pivotal role in formulating general principles, recommendations, and best practices for implementing effective, interoperable, and machine-readable TM systems. These systems support various computer language related processes, such as computer-assisted translation (CAT), controlled authoring, search engine optimization, and machine learning.

[1]https://www.iso.org/committee/48104.html
[2]https://www.iso.org/committee/48136.html

Advanced natural language processing (NLP) technologies – especially large language models (LLM) leveraged by generative artificial intelligence (GenAI) since 2022 – today offer new opportunities to enhance and further automate TM tasks which have the potential to make these processes more efficient and terminology resources more interoperable. In addition, the latest AI developments have highlighted how curated terminology can support AI-powered tools and enhance their quality, trustworthiness, and safety.

This bidirectional relationship, where AI functions both as a support technology for TM and as a domain requiring terminological data for improvement, is the focus of ISO/TC 37/SC 3/WG 6 "Terminology Management and Artificial Intelligence". This Working Group (WG) aims to develop standards that facilitate both the successful and ethical integration of AI into TM systems and the effective use of terminology resources to improve AI solutions. This paper discusses the challenges of standardizing this relationship within an unprecedented, rapidly evolving technological landscape.

2 State of the Art

2.1 AI for Terminology Management

The impact of AI on terminology builds upon technological shifts dating back to the early 1990s—when TM transitioned from paper-based to digital systems, as results of the integration of terminology with computational linguistics. Advances in computational power and data availability have subsequently enabled AI systems to perform large-scale data analysis, pattern extraction, and complex decision-making—capacities aligning with the knowledge-driven nature of terminology work.

AI promises to transform TM by enhancing and automating laborious and time-consuming processes, such as term extraction, or Automatic Term

Extraction (ATE), a TM task in which neural networks have been applied in various ways, including the generation of word embeddings as a contextual representation of terms and the development of classifiers for ATE (Lefever and Terryn, 2024; Tran et al., 2023).

Word embeddings are vector representations of words in a high-dimensional space, where words used in similar contexts have similar embeddings and tend to cluster together in that space. This capacity helps in revealing hidden relationships between terms, as seen in specialized domains such as maritime law (Mouratidis et al., 2022) and the broader evaluation of terminology extraction methods (Di Nunzio et al., 2023).

ATE classifiers operate by classifying tokens in sequence, using linguistic and statistical features to determine their status as lexical units. This approach supports a more efficient and context-sensitive TM workflow (Terryn et al., 2022). In enterprise settings [3], AI has been implemented to support terminology harmonization. For example, the Busch Group uses AI to extract terms, detect synonyms, and compare terminology across corporate databases (Beck and Fahlbusch, 2025). In Xu et al. (2025), LLMs are highlighted as a key to the future of terminology extraction tasks.

Recent advances in corpus alignment and domain-specific language modeling (Ye et al., 2024; Gururangan et al., 2020) indicate potential for LLMs trained on terminology-rich corpora to improve machine translation and domain understanding. Multilingual term alignment has also advanced, with embeddings facilitating the matching of equivalents across languages. Recent work has shown their success in aligning Arabic–French terminology (Setha and Aliane, 2023).

AI-backed chatbot tools have further enabled semi-automatic validation through document retrieval and concept checking. While these tools are not yet fully autonomous and require human oversight, they enhance TM (Bezobrazova et al., 2024). AI systems can be trained to assist conceptual validation by recognizing criteria such as term frequency, source authority, and definitional quality. AI also facilitates the automatic discovery of less intuitive concept relations, allowing for the scalable construction of domain-specific knowledge graphs. LLMs have recently gained attention for their ability to carry out concept system extraction (ISO 5394:2024; Gromann et al., 2022), hypernym detection (Cai et al., 2025), and term extraction from knowledge graphs (Pan et al., 2023; Cao et al., 2021).

AI can also help identify authoritative documents, which enhances corpus quality for terminology work (Nagendra and Chandra, 2022). Moreover, AI expedites terminology mining at scale—minimizing manual tasks for terminologists and improving efficiency in large organizations (Hamm, 2025).

2.2 Terminology Management for AI

While ontologies have historically guided the structuring of domain knowledge for AI systems, curated terminology resources in many cases remain underutilized. Terminology data—complete with designations, definitions, contexts and other data categories—provides valuable added structure for language models operating in specialized communication.

When LLMs are guided by curated terminology resources, either through training, fine-tuning, prompting, or retrieval, they exhibit improved factual grounding, contextual relevance, and reduced hallucination. Hallucination—wherein LLMs produce plausible but incorrect outputs—is especially problematic in critical fields. Terminologies act as semantic anchors, reinforcing verified content and supporting better error mitigation (Warburton, 2025).

Structured terminological data also expand reasoning in LLMs and support neuro-symbolic methods by integrating inferential pathways such as hierarchical (e.g., generic-specific) and non-hierarchical (e.g., associative) concept relationships. This aligns with the principles outlined in ISO 704:2022, which standardizes the representation of concepts, their relations, and definitions in terminology work, providing a foundational framework for knowledge organization in AI systems (Iantosca, 2022).

Conceptual Model-Augmented Generative AI (CMAG) leverages conceptual schemas to impose structure on Human-LLM interactions (Fill et al., 2024). CMAG encourages prompt generation within conceptual templates, facilitating cross-domain application in software engineering, heritage studies, and knowledge organization. Retrieval augmented generation (RAG), a prominent paradigm for augmenting LLMs with external knowledge - usually unstructured in text docu-

[3] https://kaleidoscope.at/en/blog/ai-and-terminology/

ments - to improve contextual relevance. Ontology-grounded RAG (OG-RAG) expands this idea: documents are modeled as hypergraphs using ontology-based clustering to enhance contextual understanding and decrease inference costs (Sharma et al.). Similarly, CLEAR (Clinical Entity Augmented Retrieval) utilizes entity recognition and ontology links to maximize retrieval precision in clinical contexts (Lopez et al., 2025). Such ontology-based approaches can inspire the leverage of conceptual knowledge in terminology resources to augment, improve, or evaluate LLM capabilities. Terminology Augmented Generation (TAG) (Fleischmann and Lang, 2025) is a recently proposed paradigm that integrates domain-specific terminologies into text retrieval processes, aiming at achieving high precision and computational efficiency.

Graph-structured retrieval (GraphRAG) offers further optimization by using Knowledge Graphs for query expansion, retrieval, and generative consistency (Han et al., 2025).

Terminology also plays a central role in adapting general-purpose LLMs to specialized domains. The VEGAD approach (Liu et al., 2024) automatically identifies and integrates optimal domain vocabularies into LLMs while mitigating catastrophic forgetting.

Terminology-based integration has been pivotal in neural machine translation tasks as well. Legacy efforts include structured incorporation of glossaries into machine translation (MT) models (Michon et al., 2020), while newer methods like trie-based extraction have led to efficient LLM training using limited specialized data (Kim et al., 2024). Instruction-based fine-tuning further enhances cross-linguistic term consistency, as evidenced by MT+G-Align's success in tasks like WMT 2023 (Zhao et al., 2024).

Collectively, these methods underscore terminology's central value in improving contextual relevance, factuality, and domain alignment of AI systems.

2.3 Limits

Despite the synergy between AI and TM, challenges persist. One concern is bias in training data, which can lead to culturally insensitive or non-inclusive terminological usage. Studies on the alignment of LLMs highlight its ethical criticality, inherent complexity, and the numerous challenges involved (Ji et al., 2024; Wang et al., 2024; Shen et al., 2023). Studies in gender-inclusive translation point toward unresolved ethical issues in AI integration (Gromann et al., 2023; Měchura, 2022; Corral and Saralegi, 2022; Touileb et al., 2022; Costa-jussà and de Jorge, 2020).

Another concern is the dependence of machine learning models on training data whose features and characteristics are unknown to its end-users, which makes it impossible to understand clearly how and why a given AI system makes certain decisions. This is a lack of transparency that could be addressed in the future, e.g. by building specialized LLMs (Ling et al., 2024), which are already being tested across several domains (e.g. Calamo et al. (2023) in e-justice, Chen et al. (2025) in biomedicine), as well as by investing on prompt engineering that could teach an AI system to make better decisions (Xu et al., 2025; Heinisch, 2024).

Another area of concern is the automatic extraction of terms and definitions. Bezobrazova et al. (2024) observed that while ChatGPT is not specifically designed for terminology extraction, it can perform the task effectively when properly prompted. However, compiling a thorough set of terms requires considerable time. Di Nunzio et al. (2024) corroborate this finding, noting that leveraging LLMs for term extraction, assessing term difficulty, and generating definitions for complex concepts is still in its early stages but holds potential for the future of ATE.

AI tools also face limitations in the automatic recognition of terminological variation, language varieties (e.g. American English, British English, Austrian German, Swiss German) and low-resource languages (e.g.. Ladin, South Tyrolean German) (Heinisch, 2024; Frontull and Moser, 2024; De Camillis et al., 2023).

Automated deep learning techniques for bias detection and inclusive language identification remain an emerging area (Savoldi et al., 2023; Piergentili et al., 2023). Ethical consideration must thus accompany technical innovation in deploying AI-enhanced terminology systems (Kumar et al., 2024).

3 Standardization Challenges in the Bidirectional Integration of Terminology Management and AI

3.1 Standardization Challenges in AI-Enhanced Terminology Workflows

AI is increasingly used to automate TM tasks. These capabilities promise greater efficiency and

scalability—but also reveal substantial gaps in traceability, quality assurance, and semantic precision. These weaknesses are particularly problematic in domains where terminological accuracy is crucial for regulation, safety, or legal compliance.

Several challenges must be addressed through standardization efforts:

- **Validation Protocols:** There is a pressing need for standardized validation mechanisms to ensure that AI-generated terminological units are accurate, contextually appropriate, and domain-compliant. Human-in-the-loop validation workflows, quality thresholds, and expert review loops should be formally defined.
- **Metadata Schemas:** Current metadata standards do not accommodate the unique characteristics of AI-generated content. New schemas are needed to record confidence scores, sources of training data, contextual prompts, versioning information, and post-generation modifications.
- **Tool Interoperability:** AI-assisted TM tools must integrate smoothly with existing terminological ecosystems, such as termbases, content management systems (CMS), and translation tools. Standardized APIs and data exchange formats are essential to maintain consistency and usability across platforms.
- **Licensing and Attribution:** AI outputs frequently draw on heterogeneous training data, much of which may include third-party resources. Without appropriate licensing metadata, the reuse of such outputs in professional contexts raises intellectual property and compliance concerns. Machine-readable, standardized descriptors for licensing and use rights must be incorporated into the data pipeline.
- **Explainability and Traceability:** The opaque nature of neural networks and generative models hinders the auditing and validation of terminology-related outputs. Standards are needed to enable traceability of model behavior—potentially through techniques such as RAG, hybrid symbolic methods, or prompt-driven audit trails.

Collectively, these challenges call for the evolution of terminology standards to address emerging AI-specific requirements. Without rigorous validation frameworks, interoperable infrastructures, and transparent metadata, the integration of AI into TM risks undermining the precision and credibility traditionally associated with terminology work.

3.2 Standardization Challenges for the Integration of Terminologies into AI Systems

Integrating standardized terminologies into AI systems presents a complex set of challenges that span linguistic, technical, and ethical domains. AI systems—ranging from knowledge graphs to LLMs—often rely on representational formats (such as vector embeddings or ontology languages) that differ significantly from those used in traditional TM environments such as ISO 12620-1:2022, ISO 12620-2:2022, ISO 30042:2019, and ISO16642:2017. As a result, bridging the gap between terminology resources and AI's inference or generation processes requires addressing several foundational standardization issues.

Key challenges include:

- **Fragmentation Across Industries and Languages:** Terminology resources are often created in isolated institutional or national settings. In combination with a lack of adoption of standards, this results in a rather low level of interoperability. Cross-domain, multilingual frameworks are underdeveloped.
- **Under-Resourced Languages:** Most terminology resources exist only in high-resource languages. This skews AI inclusivity and limits access in culturally diverse contexts.
- **Access Limitations:** High-quality terminologies in medicine, law, and engineering are often proprietary. Licensing issues restrict open usage, especially in academic/non-commercial AI applications.
- **Complex and Costly Terminology Development:** Building terminology resources requires linguistic and domain expertise and is not easily scalable. Formats are often not compatible with contemporary AI pipelines.

These challenges point to the need for a new generation of interoperability and governance standards capable of:

- Mapping between traditional terminological formats and AI-ready knowledge structures (e.g., ontologies, embeddings) (Roche et al., 2009),
- Supporting modular, scalable, and semi-automated methods for terminology validation and enrichment,
- Embedding ethical and provenance metadata within terminology resources,
- Encouraging the creation and standardization of multilingual terminologies, especially in under-resourced language settings,
- Defining lifecycle-aware structures that enable alignment between terminological data and AI workflows.

3.3 Cross-Cutting Standardization and Governance Challenges

Efforts to integrate TM and AI systems face systemic challenges that go beyond technical interoperability.

A fundamental issue stems from the contrasting development dynamics of the fields: AI is characterized by rapid, iterative innovation cycles, often driven by data-centric experimentation and agile tooling. Conversely, TM evolves through slower, institutionalized processes grounded in expert validation, domain specificity, and long-term conceptual stability. These contrasting tempos result in mismatches between the needs of AI systems and the availability or granularity of terminology resources, complicating real-time integration, update synchronization, and cross-domain scalability.

Additionally, several cross-cutting challenges persist:

- **Fragmented Standardization Ecosystem:** ISO/TC 37 (language & terminology), W3C [4] (semantic web), and ISO/IEC JTC 1/SC 42 [5] (AI) are separate organizational frameworks. The lack of cross-committee collaboration creates isolated representations and challenges interoperability.
- **Legal and Licensing Barriers:** Terminology resources frequently lack standardized, machine-readable licensing information. This creates uncertainty around reuse, especially in AI workflows that involve automatic ingestion, transformation, and content generation. Initiatives to integrate licensing metadata—such as RDF-based descriptors or Creative Commons tags—in standardized terminology formats are still nascent and require broader institutional adoption.
- **Discipline-Specific Governance Models:** Governance structures in TM, AI, data ethics, and digital infrastructure are organized under different frameworks. As a result, collaboration between terminologists, AI developers, data stewards, legal experts, public-sector institutions, and domain knowledge authorities remains limited. Effective integration of TM into AI ecosystems demands multi-stakeholder governance models that support cross-disciplinary decision-making, accountability mechanisms, and shared compliance frameworks.

Addressing these challenges will require proactive coordination between standardization bodies, the development of bridging standards across communities, and the rethinking of governance models to ensure that integration strategies reflect both the conceptual rigor of terminology and the dynamic realities of AI innovation.

4 ISO/TC 37/SC 3/WG 6: A Strategic Player in Bridging Terminology Management and AI

4.1 Positioning the New WG in the Global Standardization and Regulatory Landscape

ISO/TC 37 develops international standards that support the language industry, emphasizing interoperability, consistency, and quality in multilingual content management. Within this structure, SC 3 focuses on standards for computer-assisted processes for managing terminological data. It is responsible for key infrastructural standards such as ISO 5078:2025, ISO 12620-1:2022, ISO 12620-2:2022, ISO 30042:2019, and ISO16642:2017.

In early 2024, ISO/TC 37/SC 3 initiated Ad Hoc Group 1 (AHG 1) to explore the growing synergies of TM and AI, ultimately leading to the formal establishment of WG 6 "Terminology Management and Artificial Intelligence" in 2025.

WG 6 occupies a unique niche within the broader standards ecosystem. While important AI and NLP

[4] https://www.w3.org/
[5] https://www.iso.org/committee/6794475.html

developments are underway in ISO/IEC JTC 1/SC 42—especially through its Joint Working Group 5 (JWG 5) on NLP—these initiatives often treat linguistic resources at a high level of abstraction and do not sufficiently address the specificities of controlled vocabularies and terminological data. In contrast, WG 6 focuses specifically on integrating TM with AI systems, with attention to multilingual and domain-specific precision.

WG 6's relevance extends beyond ISO's internal architecture. The adoption of the European Union's *Artificial Intelligence Act (AI Act)* (Parliament and Council, 2024) provides an external regulatory catalyst for WG 6's mission. As the AI Act introduces legally binding requirements for transparency, traceability, and documented safety in AI systems—especially in high-risk domains such as medicine, law, and finance—terminological precision emerges as a crucial safeguard. Misuse or absence of validated domain-specific terminology in such contexts can result in legal liability, safety hazards, and public mistrust.

Against this backdrop, WG 6 not only contributes to ISO's standardization efforts but also functions as a strategic actor helping stakeholders operationalize the AI Act's compliance requirements. By aligning semantic resources with ethical, legal, and technical expectations, WG 6 enables AI systems to meet the regulatory demand for explainability, contextual accuracy, and trustworthy outputs. For example, the output created by AI systems needs to reflect best practices of data modelling and conform to relevant data categories (ISO 12620-1:2022 and ISO 12620-2:2022).

4.2 Strategic Objectives and Operational Approach

WG 6 was established with a dual mandate:

- To develop standards and technical guidance for the integration of AI technologies into terminology workflows—covering tasks such as term extraction, clustering, multilingual alignment, and automated definition generation.
- To support the integration of standardized terminologies into AI systems, enabling them to benefit from linguistic precision, multilingual equivalence, and concept-level clarity in interpretation and generation tasks.

Addressing this dual mission requires WG 6 to operate in new, adaptable ways that diverge from legacy standardization procedures. Traditional standardization timelines are outpaced by AI's rapid, weekly or monthly cycles. To address this, WG 6 introduced methods adapted to AI's fast and volatile developments.

One notable methodological adjustment occurred during the exploratory phase. AHG 1 was organized around four agile task forces, each assigned to focus on key dimensions stemming from the different stages of terminology work. This distributed structure greatly accelerated production timelines and was key to transforming AHG 1 into a fully operational WG, culminating in its first Technical Report (TR) project (ISO/AWI TR 25896) [6]. TR had been chosen as the appropriate type of deliverable for its considerably shortened production time.

This approach has proven useful: instead of relying solely on static deliverables designed for long-term generality, WG 6 embraces a portfolio of standardization formats—including TRs, Technical Specifications, and eventually International Standards—that are scalable, iterative, and adaptable to fast-moving technological ecosystems. More broadly, this flexibility reflects a growing consensus across standardization bodies that AI not only uses standards, but is itself a target for standardization—thus requiring a rethinking of the pace, process, and governance of standards development.

5 Conclusion and Outlook

In conclusion, the integration of AI with TM marks a transformative step toward enhancing efficiency, precision, and global communication in both fields. The establishment of ISO/TC 37/SC 3/WG 6 represents a pioneering effort to standardize the symbiotic advancements of AI and TM, fostering interoperable, scalable, and ethically sound frameworks. By addressing the challenges of aligning traditional standardization processes with the rapid evolution of AI technologies, WG 6 introduces innovative methodologies that balance stability with adaptability. This strategic approach ensures that emerging standards remain rigorous yet agile, enabling public and private stakeholders to responsibly leverage AI in complex, multilingual knowledge environments. Through these efforts, WG 6 is poised to shape future standards that support seamless knowledge exchange and drive ethical, effective applications of AI and TM worldwide.

[6] https://www.iso.org/standard/91875.html

Acknowledgments

We would like to express our sincere thanks to the International Organization for Standardization (ISO) for providing a collaborative framework to advance terminology and language resource standards. In particular, we are grateful to the members of ISO/TC 37/SC 3/WG 6 for the valuable discussions, insights, and collective expertise that have informed and enriched this work.

We also gratefully acknowledge the support of the authors' respective institutions and national standardization bodies, which have provided the academic and organizational environment necessary for the completion of this work.

This work has been supported, in part, by the *StandICT.eu* initiative through financial assistance provided to some of the authors. This support has facilitated their participation in standardization-related activities and contributed to the development of certain aspects of the research. We appreciate the initiative's continued efforts to promote European engagement in international ICT standardization. The views expressed in this paper are those of the authors and do not necessarily reflect those of the *StandICT.eu* initiative.

References

Katharina Beck and Fabian Fahlbusch. 2025. TerminoKI. Terminologiearbeit 4.0 mit BuschGPT. In *Terminologie in der KI - KI in der Terminologie. Akten des Symposions, Worms, 27-29 März 2025*, pages 39–50, München, Deutschland. Deutscher Terminologie-Tag e.V.

Anastasiia Bezobrazova, Miriam Seghiri, and Constantin Orasan. 2024. Benchmarking terminology building capabilities of ChatGPT on an English-Russian Fashion Corpus. *Computing Research Repository*, arXiv:2412.03242.

Guohui Cai, Jiangchuan Gong, Junliang Du, Hao Liu, and Anda Kai. 2025. Investigating Hierarchical Term Relationships in Large Language Models. *Journal of Computer Science and Software Applications*, 5(4).

Marco Calamo, Francesca De Luzi, Mattia Macrì, Tommaso Mencattini, and Massimo Mecella. 2023. CICERO: a GPT2-based writing assistant to investigate the effectiveness of specialized LLMs' applications in e-justice. In *ECAI 2023*, pages 3196–3203. IOS Press.

Boxi Cao, Hongyu Lin, Xianpei Han, Le Sun, Lingyong Yan, Meng Liao, Tong Xue, and Jin Xu. 2021. Knowledgeable or Educated Guess? Revisiting Language Models as Knowledge Bases. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 1860–1874, Online. Association for Computational Linguistics.

Linqing Chen, Weilei Wang, Yubin Xia, Wentao Wu, Peng Xu, Zilong Bai, Jie Fang, Chaobo Xu, Ran Hu, Licong Xu, Haoran Hua, Jing Sun, Hanmeng Zhong, Jin Liu, Tian Qiu, Haowen Liu, Meng Hu, Xiuwen Li, Fei Gao, and 14 others. 2025. Streamlining Biomedical Research with Specialized LLMs. *Computing Research Repository*, arXiv:2504.12341.

Ander Corral and Xabier Saralegi. 2022. Gender bias mitigation for NMT involving genderless languages. In *Proceedings of the Seventh Conference on Machine Translation (WMT)*, pages 165–176.

Marta R Costa-jussà and Adrià de Jorge. 2020. Fine-tuning neural machine translation on gender-balanced datasets.

Flavia De Camillis, Egon Waldemar Stemle, Elena Chiocchetti, and Francesco Fernicola. 2023. The MT@ BZ corpus: machine translation & legal language. In *Proceedings of the 24th Annual Conference of the European Association for Machine Translation 12–15 June 2023, Tampere, Finland*, pages 171–180.

Giorgio Maria Di Nunzio, Stefano Marchesin, and Gianmaria Silvello. 2023. A systematic review of Automatic Term Extraction: What happened in 2022? *Digital Scholarship in the Humanities*, 38(Supplement_1):i41–i47.

Giorgio Maria Di Nunzio, Federica Vezzani, Vanessa Bonato, Hosein Azarbonyad, Jaap Kamps, Liana Ermakova, and 1 others. 2024. Overview of the CLEF 2024 SimpleText task 2: identify and explain difficult concepts. *Working Notes of CLEF*.

HG Fill, F Härer, I Vasic, D Borcard, B Reitemeyer, F Muff, S Curty, and M Bühlmann. 2024. CMAG: A Framework for Conceptual Model Augmented Generative Artificial Intelligence, ER2024: Companion. In *Proceedings of the 43rd International Conference on Conceptual Modeling: ER Forum, Special Topics, Posters and Demos, Pittsburgh, PA, USA*, pages 28–31.

Klaus Fleischmann and Christian Lang. 2025. Terminologie in der KI: Wie mit Terminologie der Output von LLMs und GenAI optimiert werden kann. In *Terminologie in der KI - KI in der Terminologie. Akten des Symposions, Worms, 27-29 März 2025*, pages 83–96, München, Deutschland. Deutscher Terminologie-Tag e.V.

Samuele Frontull and Georg Moser. 2024. Traduzione automatica per il ladino della Val Badia. *Ladinia XLVIII/2024*.

Dagmar Gromann, Manuel Lardelli, Katta Spiel, Sabrina Burtscher, Lukas Daniel Klausner, Arthur Mettinger, Igor Miladinovic, Sigrid Schefer-Wenzl,

Daniela Duh, and Katharina Bühn. 2023. Participatory research as a path to community-informed, gender-fair machine translation. *Computing Research Repository*, arXiv:2306.08906.

Dagmar Gromann, Lennart Wachowiak, Christian Lang, and Barbara Heinisch. 2022. Extracting Terminological Concept Systems from Natural Language Text. In *European Language Grid: A Language Technology Platform for Multilingual Europe*, pages 289–294. Springer International Publishing Cham.

Suchin Gururangan, Ana Marasović, Swabha Swayamdipta, Kyle Lo, Iz Beltagy, Doug Downey, and Noah A. Smith. 2020. Don't stop pretraining: Adapt language models to domains and tasks. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8342–8360, Online. Association for Computational Linguistics.

Julian Hamm. 2025. Terminologische Konsistenz und generative KI – ein Perfect Match? Produktiver Einsatz von Sprachmodellen im Terminologiemanagement und beim Post-Editing. In *Terminologie in der KI - KI in der Terminologie. Akten des Symposions, Wörms, 27-29 März 2025.*, pages 151–164, München, Deutschland. Deutscher Terminologie-Tag e.V.

Haoyu Han, Yu Wang, Harry Shomer, Kai Guo, Jiayuan Ding, Yongjia Lei, Mahantesh Halappanavar, Ryan A Rossi, Subhabrata Mukherjee, Xianfeng Tang, and 1 others. 2025. Retrieval-augmented generation with graphs (graphrag). *Computing Research Repository*, arXiv:2501.00309.

Barbara Heinisch. 2024. Large language model prompting in a university course on terminology. *Terminology Science & Research/Terminologie: Science et Recherche*, (27):28–51.

Michael J. Iantosca. 2022. From microcontent to neurons A practical guide for building a cognitive AI content supply chain for highly personalized user assistance. Online. Online article, accessed on 2025-05-29.

ISO 12620-1:2022. Management of terminology resources — Data categories. Part 1: Specifications. International Organization for Standardization. https://www.iso.org/standard/79078.html.

ISO 12620-2:2022. Management of terminology resources — Data categories. Part 1: Repositories. International Organization for Standardization. https://www.iso.org/standard/79018.html.

ISO 30042:2019. Management of terminology resources – TermBase eXchange (TBX). International Organization for Standardization. https://www.iso.org/standard/62510.html.

ISO 5078:2025. Management of terminology resources — Terminology extraction. International Organization for Standardization. https://www.iso.org/standard/81917.html.

ISO 5394:2024. Information technology — Criteria for concept systems. International Organization for Standardization. https://www.iso.org/standard/81656.html.

ISO 704:2022. Terminology work — Principles and methods. International Organization for Standardization. https://www.iso.org/standard/79077.html.

ISO16642:2017. Computer applications in terminology — Terminological markup framework. International Organization for Standardization. https://www.iso.org/standard/56063.html.

ISO/AWI TR 25896. Management of terminology resources - Artificial intelligence for terminology management. International Organization for Standardization. https://www.iso.org/standard/91875.html.

Jiaming Ji, Boyuan Chen, Hantao Lou, Donghai Hong, Borong Zhang, Xuehai Pan, Tianyi Alex Qiu, Juntao Dai, and Yaodong Yang. 2024. Aligner: Efficient alignment by learning to correct. *Advances in Neural Information Processing Systems*, 37:90853–90890.

Sejoon Kim, Mingi Sung, Jeonghwan Lee, Hyunkuk Lim, and Jorge Froilan Gimenez Perez. 2024. Efficient Terminology Integration for LLM-based Translation in Specialized Domains. *Computing Research Repository*, arXiv:2410.15690.

Naveen Kumar, Xiahua Wei, and Han Zhang. 2024. Addressing Bias in Generative Ai: Challenges and Research Opportunities in Information Management. *Available at SSRN 4976889*.

Els Lefever and Ayla Rigouts Terryn. 2024. Computational Terminology. In *New Advances in Translation Technology: Applications and Pedagogy*, pages 141–159. Springer.

Chen Ling, Xujiang Zhao, Jiaying Lu, Chengyuan Deng, Can Zheng, Junxiang Wang, Tanmoy Chowdhury, Yun Li, Hejie Cui, Xuchao Zhang, Tianjiao Zhao, Amit Panalkar, Dhagash Mehta, Stefano Pasquali, Wei Cheng, Haoyu Wang, Yanchi Liu, Zhengzhang Chen, Haifeng Chen, and 5 others. 2024. Domain Specialization as the Key to Make Large Language Models Disruptive: A Comprehensive Survey. *Computing Research Repository*, arXiv:2305.18703.

Chengyuan Liu, Shihang Wang, Lizhi Qing, Kun Kuang, Yangyang Kang, Changlong Sun, and Fei Wu. 2024. Gold Panning in Vocabulary: An Adaptive Method for Vocabulary Expansion of Domain-Specific LLMs. *Computing Research Repository*, arXiv:2410.01188.

Ivan Lopez, Akshay Swaminathan, Karthik Vedula, Sanjana Narayanan, Fateme Nateghi Haredasht, Stephen P Ma, April S Liang, Steven Tate, Manoj Maddali, Robert Joseph Gallo, and 1 others. 2025. Clinical entity augmented retrieval for clinical information extraction. *npj Digital Medicine*, 8(1):45.

Michal Měchura. 2022. A taxonomy of bias-causing ambiguities in machine translation. In *Proceedings of the 4th Workshop on Gender Bias in Natural Language Processing (GeBNLP)*, pages 168–173.

Elise Michon, Josep M Crego, and Jean Senellart. 2020. Integrating domain terminology into neural machine translation. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 3925–3937.

Despoina Mouratidis, Eirini Mathe, Yorghos Voutos, Klio Stamou, Katia Lida Kermanidis, Phivos Mylonas, and Andreas Kanavos. 2022. Domain-specific term extraction: a case study on Greek Maritime legal texts. In *Proceedings of the 12th Hellenic Conference on Artificial Intelligence*, pages 1–6.

N Nagendra and J Chandra. 2022. A Systematic Review on Features Extraction Techniques for Aspect Based Text Classification Using Artificial Intelligence. *ECS Transactions*, 107(1):2503.

Jeff Z Pan, Simon Razniewski, Jan-Christoph Kalo, Sneha Singhania, Jiaoyan Chen, Stefan Dietze, Hajira Jabeen, Janna Omeliyanenko, Wen Zhang, Matteo Lissandrini, and 1 others. 2023. Large language models and knowledge graphs: Opportunities and challenges. *Computing Research Repository*, arXiv:2308.06374.

European Parliamant and Council. 2024. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). Official Journal of the European Union.

Andrea Piergentili, Dennis Fucci, Beatrice Savoldi, Luisa Bentivogli, and Matteo Negri. 2023. Gender neutralization for an inclusive machine translation: from theoretical foundations to open challenges. *Computing Research Repository*, arXiv:2301.10075.

Christophe Roche, Marie Calberg-Challot, Luc Damas, and Philippe Rouard. 2009. Ontoterminology: A new paradigm for terminology. In *International Conference on Knowledge Engineering and Ontology Development*, pages 321–326, Madeira, Portugal.

Beatrice Savoldi, Marco Gaido, Matteo Negri, and Luisa Bentivogli. 2023. Test suites task: Evaluation of gender fairness in MT with MuST-SHE and INES. *Computing Research Repository*, arXiv:2310.19345.

Imene Setha and Hassina Aliane. 2023. Bilingual Terminology Alignment Using Contextualized Embeddings. In *Proceedings of the Workshop on Computational Terminology in NLP and Translation Studies (ConTeNTS) Incorporating the 16th Workshop on Building and Using Comparable Corpora (BUCC)*, pages 1–8, Varna, Bulgaria. INCOMA Ltd., Shoumen, Bulgaria.

Kartik Sharma, Peeyush Kumar, and Yunqing Li. OG-RAG: Ontology-Grounded Retrieval-Augmented Generation For Large Language Models. *Computing Research Repository*, arXiv:2412.15235.

Tianhao Shen, Renren Jin, Yufei Huang, Chuang Liu, Weilong Dong, Zishan Guo, Xinwei Wu, Yan Liu, and Deyi Xiong. 2023. Large Language Model Alignment: A Survey. *Computing Research Repository*, arXiv:2309.15025.

Ayla Rigouts Terryn, Veronique Hoste, and Els Lefever. 2022. Tagging terms in text: A supervised sequential labelling approach to automatic term extraction. *Terminology: international journal of theoretical and applied issues in specialized communication*, 28(1):157–189.

Samia Touileb, Lilja Øvrelid, and Erik Velldal. 2022. Occupational biases in Norwegian and multilingual language models. In *Proceedings of the 4th Workshop on Gender Bias in Natural Language Processing (GeBNLP)*, pages 200–211.

Hanh Thi Hong Tran, Matej Martinc, Jaya Caporusso, Antoine Doucet, and Senja Pollak. 2023. The recent advances in automatic term extraction: A survey. *Computing Research Repository*, arXiv:2301.06767.

Han Wang, An Zhang, Nguyen Duy Tai, Jun Sun, Tat-Seng Chua, and 1 others. 2024. ALI-Agent: Assessing LLMs' Alignment with Human Values via Agent-based Evaluation. *Advances in Neural Information Processing Systems*, 37:99040–99088.

Kara Warburton. 2025. Terminology in the Age of AI. `https://multilingual.com/magazine/march-2025/terminology-in-the-age-of-ai/`. Online article, accessed on 2025-05-29.

Kang Xu, Yifan Feng, Qiandi Li, Zhenjiang Dong, and Jianxiang Wei. 2025. Survey on terminology extraction from texts. *Journal of Big Data*, 12(1):1–40.

Qi Ye, Zicheng Yao, Peihong Hu, Xiang Ji, Tong Ruan, and Ruihui Hou. 2024. Alignment of chinese-english medical terminology in small-sample scenarios: A two-stage approach. In *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 3908–3911. IEEE.

Xuan Zhao, Chong Feng, Shuanghong Huang, Jiangyu Wang, and Haojie Xu. 2024. Incorporating Terminology Knowledge into Large Language Model for Domain-Specific Machine Translation. In *China Conference on Machine Translation*, pages 82–96. Springer.