The Journey to Making a Termbase AI-Ready: The Philips Use Case

Maria Bruno[1,*,†], Daniela Fleck[1,†]

[1]*Philips Medizin Systeme, Hewlett-Packard Straße 2, 71034, Böblingen*

Abstract

In the age of artificial intelligence (AI), terminology management has evolved from a tool for ensuring linguistic accuracy to a critical enabler of AI system performance. This paper explores the concept of AI readiness in terminology, focusing on the transformation of the Philips termbase into a resource optimized for both human and machine use. The study highlights the challenges posed by AI limitations, such as reasoning, bias generation, hallucinations, multilingual inconsistency, and domain inaccuracy, and outlines strategies to address these issues. Key steps include reducing ambiguity through clear definitions, avoiding abbreviations, and modeling concept and term relations. The paper also emphasizes the importance of labeling terms to mitigate bias and ensure consistent usage, as well as the role of continuous cleanup in maintaining data accuracy and relevance. Collaboration and change management are central to these efforts, fostering shared ownership of data quality and aligning terminology practices with organizational needs. By integrating these strategies, the Philips termbase has become a dynamic and reliable resource, supporting the evolving demands of both AI systems and human users. The paper concludes by advocating for ongoing research, experimentation, and collaboration to advance the field of AI-ready terminology management.

Keywords

Terminology, Termbase, AI readiness, Bias-aware language, Standardized formats

1. Introduction

A terminology database (termbase) serves as a multilingual repository of structured knowledge, and it can fuel diverse AI use cases. Historically, terminology management was primarily associated with translators and technical writers, focusing on ensuring linguistic accuracy and consistency. However, its role has evolved significantly in the age of artificial intelligence (AI), where terminology now plays a pivotal role in enhancing the reliability and consistency of AI outputs.

The relationship between terminology and AI can be understood through two perspectives: (1) Terminology in AI, where termbases are used to support AI applications, and (2) AI in terminology, where AI technologies assist in managing and optimizing terminology workflows. While both perspectives are relevant, this paper focuses on the former—how termbases can serve as high-quality inputs for AI systems.

The principle of "garbage in, garbage out" underscores the importance of data quality in AI systems. Terminology, as structured and accurate language data, directly impacts the correctness and consistency of AI outputs. Inconsistent or ambiguous terms lead to flawed AI-generated results, whereas clear, structured, and connected terminology data enhances the quality of AI outputs. An AI-ready termbase is one that is designed and populated to be seamlessly integrated into AI systems, ensuring efficiency and precision in applications such as documentation generation, chatbots, translation, multimedia content creation, and information retrieval.

This paper argues that terminology resources must evolve toward AI readiness, "the state of an organization in terms of willingness and ability of its stakeholders, and suitability of its environment, processes,

5th International Conference on "Multilingual digital terminology today. Design, representation formats and management systems" (MDTT) 2026, June 25-26, 2026, Zadar, Croatia.

*Corresponding author.

†These authors contributed equally.

maria.bruno@philips.com (M. Bruno); daniela.fleck@philips.com (D. Fleck)

0000-0003-1746-3469 (M. Bruno)

data and resources for adopting and operating AI" [1]. It situates AI readiness within terminology theory and emerging AI research, and illustrates its practical implementation through the Philips termbase.

2. AI Readiness in Terminology

Although the notion of "AI readiness" is increasingly discussed, it has not yet been defined in the context of terminology databases. This paper proposes the following definition of an AI-ready terminology database: concept oriented, semantically structured terminological knowledge base that transforms implicit domain expertise into explicit, machine interpretable data, including multilingual equivalences, conceptual relations, term level distinctions, bias aware signals, and metadata, such as definitions and usage labels, encoded in standardized formats to support semantic interoperability, reduce ambiguity, and provide AI systems with an explicit, authoritative conceptual framework. AI readiness in terminology refers to the process of preparing a termbase to meet the specific requirements of AI systems. A helpful framework for understanding AI readiness includes four key dimensions [2]: data readiness, cultural readiness, skills readiness, and infrastructure readiness.

Data readiness is foundational, as high-quality data is essential for AI systems to function effectively. In terminology work, this involves organizing terminology into machine-readable formats, ensuring metadata is complete, and addressing inconsistencies. This highlights the importance of creating a solid foundation for both human and machine use.

Cultural readiness emphasizes the human aspect of AI readiness. It requires fostering a collaborative environment that promotes trust, curiosity, and openness among stakeholders. This is particularly important in terminology work, as it often involves cross-functional teams, including linguists, domain experts, and AI developers.

Skills readiness involves upskilling people to work effectively with AI technologies. This includes, in the context of terminology, training in data structuring, metadata management, and understanding AI capabilities and limitations. As AI continues to evolve, ongoing education and skill development are critical.

Infrastructure and budget readiness focus on organizational-level requirements, such as investing in the right tools, computational resources, and scalable systems. These elements ensure that AI integration is both sustainable and adaptable to future needs.

The process of achieving AI readiness is not a one-time effort but a continuous evolution. It involves transforming unstructured data—akin to a pile of loose bricks—into a well-built wall that serves as a solid foundation for both humans and machines. That said, the question is: why should a termbase be AI-ready? It should be because it can be used by AI systems, in addition to people. The end goal is for this structured foundation to aid AI systems to perform tasks that fall under text generation and information retrieval with greater reliability and efficiency. It is also important to build termbases that are resilient to new AI uses that might arise in the future.

The growing importance of AI readiness in terminology is reflected in initiatives such as the ISO/TC 37/SC 3/WG 6 [3] , a new working group of the ISO's Technical Committee 37 "Language and terminology" that aims to develop standardized frameworks for integrating AI and terminology resources. These efforts focus on ensuring interoperability, scalability, and ethical application, addressing the challenges of aligning traditional terminology practices with rapidly evolving AI technologies . Another notable initiative is that described by [4], who have tested how to feed a termbase to LLMs to customize it to specific topics and terms with a technique they named Terminology Augmented Generation (TAG).

3. Key Steps to Make a Termbase AI-Ready

The first thing to consider for making a termbase AI-ready is that high-quality input is crucial to produce reliable outputs. Secondly, it is necessary to think differently about how information should be stored and represented. Humans can fill in the blanks; we possess what is called "commonsense knowledge" [5]. This is the ability to understand things about the world by living in it. However, AI systems struggle

with implicit or assumed information. Therefore, part of the work of making a termbase AI-ready is transforming implicit knowledge into explicit knowledge. These improvements do not help only AI – they help human users too. The same clarity and consistency that make a termbase machine-usable also make it easier to navigate for writers, translators, reviewers, and any other content creator.

Another fundamental concern is that of interoperability. For a termbase to integrate with other systems – like AI pipelines , translation tools, or content management systems – and exchange data, it needs to follow standards; i.e. standardized formats and tags. Formats such as TBX (TermBase eXchange) [6], an ISO standard for terminology data exchange, are particularly effective in ensuring compatibility with other tools.

While the most recent and used AI models, Large Language Models (LLMs), excel at tasks such as fluent text generation, pattern recognition, and instant translation, they struggle with certain areas that are essential for terminology work: reasoning, bias generation, hallucinations, multilingualism, and maintaining domain accuracy (especially in healthcare), among others [7]. Those are the areas on which to focus to make a termbase AI-ready:

- As a way to aid reasoning, ambiguity can be reduced by adding clear definitions, avoiding abbreviations inside data fields and creating relations among entries.
- To address potential bias generation, inclusive and non-inclusive language can be labeled.
- Multilingual consistency can be strengthened by adding terms and metadata in different languages so that meaning stays aligned across languages.
- Lastly, to contrast hallucinations and improve accuracy in the healthcare domain, a termbase should be kept current and precise through continuous cleanup.

4. Turning a Termbase AI-ready: The Use Case of the Philips termbase

The Philips termbase serves as a corporate repository of terminology, encompassing a wide range of subject fields related to medical technology. Accessible exclusively to Philips employees, the termbase was initially developed in the 1990s as a multilingual glossary intended for use by the Localization Team. Over the years, its scope and utility have expanded significantly. Today, the Philips termbase is a critical resource for various stakeholders, including technical writers, user interface (UI) designers, internal and external translators, sales specialists, and other content creators. This evolution reflects the growing importance of terminology management in supporting diverse organizational functions and ensuring consistency across content.

As Philips embarked on the journey to make its termbase AI-ready, the focus shifted to addressing the limitations of AI systems, such as reasoning, bias generation, hallucinations, multilingualism, and domain accuracy. The following sections outline specific steps taken to mitigate these challenges.

4.1. Avoiding Abbreviations

Abbreviations are ubiquitous in corporate environments, and Philips is no exception, with over 7,000 abbreviations in use across the organization. They are ambiguous by nature and can lead to misinterpretations, particularly when AI systems are tasked with generating or translating content.

One of the key challenges lies in the use of abbreviations within data fields, such as definition or note fields. While subject matter experts and experienced employees can often interpret these abbreviations based on context, they pose significant challenges for AI systems, which rely on explicit and unambiguous data. To address this issue, Philips implemented a policy to minimize the use of abbreviations within the termbase. Wherever possible, full terms are used instead of abbreviations. For example:

- "Brand Desk Question" is used instead of "BDQ."
- "Query Log" replaces "QL."
- "Instructions for Use" is written out instead of "IFU."

By explicitly writing out terms, Philips reduces ambiguity and ensures that the termbase is both human- and machine-readable. This approach not only aids AI systems in processing terminology but also improves the usability of the termbase for employees across different roles and levels of expertise.

4.2. Clear Definitions

A critical aspect of ensuring the usability of a termbase, both for humans and AI systems, is the inclusion of clear and precise definitions for each entry. The Philips termbase, originally developed as a simple multilingual glossary, initially lacked definitions, containing only equivalent terms in different languages. Over time, this limitation became evident, as entries without definitions led to ambiguity and misinterpretation.

The absence of definitions poses significant challenges. For instance, external translators unfamiliar with specific entries may struggle to discern the intended concept, increasing the risk of mistranslations. Similarly, AI systems, might be unable to correctly infer the meaning of undefined terms, leading to errors in content generation or translation.

Recognizing these challenges, Philips has adopted a policy to ensure that every new entry in the termbase is accompanied by a clear and comprehensive definition. This practice not only reduces ambiguity but also enhances the termbase's utility across various applications, including technical writing, translation, and AI integration. By systematically adding definitions to older entries and mandating their inclusion for new ones, Philips has improved the clarity and reliability of its terminology resources.

4.3. Concept Relations

Concepts and terms within a termbase are inherently interconnected, and making these relationships explicit is essential for improving both human and machine understanding. At Philips, efforts are underway to model these relationships systematically, using standard relation types such as generic (IS-A), partitive (PART-OF), and associative (non-hierarchical) relations [8].

By explicitly modeling relationships between concepts, the Philips termbase achieves several key benefits:

- Improved comprehension: definitions become easier to understand when contextualized within a network of related concepts.
- Consistency across teams: explicit relationships support consistent terminology use across different departments and teams.
- Enhanced AI capabilities: AI systems can leverage these relationships to improve search functionality, reasoning, and content generation.
- Data reuse and duplication reduction: explicit relationships help identify and eliminate redundant entries, streamlining the termbase.
- Alignment with taxonomies: the termbase aligns more effectively with internal taxonomies, ensuring consistency across organizational resources.

Despite these advancements, challenges remain in determining how to represent the directionality of relationships within the termbase. For example, the logistics of displaying and managing these relationships in entry structures are still under development. Addressing these challenges will further enhance the termbase's functionality and its integration with AI systems.

4.4. Term Relations

In addition to concept-level relationships, Philips is developing a new data field to manage term-level relations, specifically to indicate terms that should not be confused with one another. This initiative addresses a common issue in terminology management: the potential for confusion between terms that are similar in form (near-homonyms) or meaning but represent distinct concepts.

For example, the Philips termbase includes entries for "outpatient" and "ambulant patient." While these terms are related, they are not interchangeable. To prevent confusion, the terminology team plans to introduce a "Not to be confused with" data field, which will link each term to its corresponding entry. This standardized approach will replace the current practice of indicating such relationships in various non-standardized fields, such as "Differentiator", "Note," and "See also." The new data field will serve several purposes:

- Avoiding misinterpretation: content creators will have clear guidance to distinguish between similar terms.
- Standardization: a dedicated field ensures consistency in how term-level relationships are represented across the termbase.
- Improved usability: by explicitly linking related terms, the termbase becomes easier to navigate for both human users and AI systems.

4.5. Labeling Terms

Bias is a persistent challenge in AI systems, arising not only from human perspectives but also from the data used to train these systems [9]. AI learns from human-created content, which inherently reflects societal values, stereotypes, and cultural norms. Consequently, if the training data contains biases, AI systems are likely to replicate and amplify them in their outputs. This underscores the critical need for strategies to mitigate bias in AI applications.

Terminology management provides a solution to address certain types of bias in AI systems. By leveraging labeled data within a termbase, organizations can guide AI systems toward more inclusive and bias-aware language use. For example, the Philips termbase incorporates specific guidelines to replace potentially biased terms with neutral alternatives. Terms such as "whitelist" are substituted with "allowlist" or "include list," depending on the context, while "master monitor" and "slave monitor" are replaced with "primary monitor" and "secondary monitor." These modifications not only promote inclusivity but also align with ethical standards for language use.

In addition to mitigating bias, labeling terms enhances the consistency of terminology use across languages and applications. For instance, the "Usage" data field in the Philips termbase has been streamlined to include only four values: Preferred, Preferred Abbreviation, Admitted, and Do Not Use. These usage labels are now mandatory for all terms in both English (the primary language of the termbase) and target languages. This standardization ensures that terms are used consistently across different contexts, reducing ambiguity and improving the reliability of AI outputs.

Another significant improvement is the introduction of a new data field called "Concept Type", which categorizes concepts based on their nature, such as chemical compounds, Philips products, symbols, or trademarks. This field replaces the older, free-text "Differentiator" field with a controlled picklist, ensuring consistent categorization and eliminating free-text variations that machines cannot reliably interpret. Aligning this categorization with other company resources, such as the product names list from the Brand Team, ensures that all teams and systems within the organization use a unified terminology framework.

By implementing these labeling strategies, Philips has enhanced the AI readiness of its termbase, ensuring that it is not only inclusive and bias-aware but also structured and consistent for both human and machine use.

4.6. Multilingual Consistency

The composition of training data for LLMs reveals a significant imbalance in language representation: English dominates these datasets [10, 11]. This dominance means that AI systems are primarily trained to "think" in English, often prioritizing English-language structures, expressions, and cultural nuances in their outputs [12]. In contrast, languages with fewer digital resources, such as Thai, Indonesian, and Vietnamese, are significantly underrepresented. Their linguistic and cultural nuances are less frequently

captured in training data, which can lead to biased multilingual outputs and a lack of inclusivity in AI-generated content. This underrepresentation reinforces systemic biases, as AI systems may fail to adequately process or generate content in these languages, further marginalizing them in digital and AI-driven contexts.

To address this imbalance, multilingual termbases play a crucial role. By populating entries in underrepresented languages with equivalent terms, usage labels, and comprehensive metadata, these termbases ensure that such languages are included in the conversation.

4.7. Continuous Cleanup

Maintaining the accuracy and relevance of a termbase is a complex and ongoing challenge, particularly in domains where terminology evolves rapidly. Without proper management, outdated or irrelevant entries can accumulate, increasing the risk of hallucinations and biases in AI systems. These issues may arise when AI systems rely on terminology that is no longer accurate or reflective of current standards, leading to errors in reasoning, content generation, and decision-making. To mitigate these risks, regular reviews and updates are essential to ensure that the termbase remains a reliable resource for both human users and AI systems. At Philips, the termbase has a long history and, over time, it has accumulated entries that no longer meet internal current standards. These include entries that refer to outdated concepts, lack clear definitions, or duplicate other entries. Such entries can mislead users and compromise the quality of AI outputs. To address this, Philips has implemented a continuous cleanup process aimed at maintaining the termbase's accuracy and trustworthiness.

Entries that are outdated or cannot be clearly defined are either updated or removed. However, deleting data from a shared resource is not a decision taken lightly, as the termbase is a collective asset managed by terminologists but used by a wide range of stakeholders. To ensure that valuable information is not lost, outdated entries that may still hold reference value are moved to a dedicated archive termbase. This approach preserves historical data while maintaining the currentness of the active termbase.

The cleanup process is designed to be collaborative, involving input from linguists, domain experts, and other stakeholders. Discussion tasks are used to facilitate decision-making, allowing internal linguists to vote on whether to retain or remove specific entries. This participatory approach ensures that decisions are informed by diverse perspectives and that the termbase continues to meet the needs of its users.

The cleanup process is not merely a technical activity but also a change management initiative. Recognizing that many entries were created with significant time, care, and expertise, the process emphasizes transparency and collaboration. Information sessions are conducted to explain the rationale behind the cleanup, and feedback is collected to refine guidelines and processes. This approach fosters a sense of shared ownership for data quality, reinforcing the idea that maintaining clarity and accuracy is a collective responsibility.

By involving users in the cleanup process and emphasizing the "why" behind each decision, Philips has strengthened both the quality of its terminology and the organizational culture surrounding it. This collaborative approach ensures that the termbase remains a dynamic and reliable resource, capable of supporting the evolving needs of both humans and AI systems.

5. Conclusion: Future Directions for AI-Ready Terminology

This paper has explored the critical role of AI-ready termbases in enhancing the quality and consistency of AI outputs. By emphasizing the importance of structured, accurate, and explicit terminology, we have argued how these resources can address challenges such as ambiguity, bias, and multilingual inconsistencies. The integration of clear definitions, metadata, and concept relationships has been highlighted as essential for improving both human and AI usability. Furthermore, the discussion underscored the need for collaboration between terminologists, domain experts, and AI developers to ensure that terminology work aligns with the evolving demands of AI systems.

Looking ahead, our objectives are clear. We must continue to refine termbases by incorporating metrics to measure their impact on user satisfaction and AI performance, such as reduced errors in automated content generation. Experimentation and iteration will remain central to our approach, as will fostering a culture of curiosity, collaboration, and openness. Beyond data quality, we aim to focus on the cultural, skill-based, and infrastructural readiness required for successful AI integration. By staying committed to these principles, we can ensure that terminology work not only keeps pace with AI advancements but also contributes to making language consistent, inclusive, and fair for all stakeholders.

Declaration on Generative AI

During the preparation of this work, the authors used Philips Enterprise AI Chat for: paraphrasing and rewording, and abstract drafting. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] A. N. Tehrani, S. Ray, S. K. Roy, R. L. Gruner, F. P. Appio, Decoding ai readiness: An in-depth analysis of key dimensions in multinational corporations, Technovation 131 (2024) 102948. URL: https://www.sciencedirect.com/science/article/pii/S0166497223002596. doi:https://doi.org/10.1016/j.technovation.2023.102948.

[2] K. Ganesan, The Business Case for AI: A Leader's Guide to AI Strategies, Best Practices & Real-World Applications, Opinosis Analytics Publishing, 2022.

[3] M. Khemakhem, C. Valentini, N. Ralli, S. Barros, G. Löckinger, et al., Terminology management meets AI: The ISO/TC 37/SC 3/WG 6 initiative, in: TermTrends 2025 Workshop at LDK, Naples, Italy, 2025. Hal-05162487v1.

[4] A. Lackner, A. Vega-Wilson, C. Lang, Terminology augmented generation: A systematic review of terminology formats for in-context learning in LLMs, in: MDTT, 2025. URL: https://api.semanticscholar.org/CorpusID:280408311.

[5] J. Scott, G. Marshal, A Dictionary of Sociology, 3 ed., Oxford University Press, 2009. doi:10.1093/acref/9780199533008.001.0001.

[6] International Organization for Standardization, ISO 30042:2019(en) Management of terminology resources — TermBase eXchange (TBX), Standard 30042, 2019. URL: https://www.iso.org/obp/ui/en/#iso:std:iso:30042:ed-2:v1:en.

[7] A. Kostikova, Z. Wang, D. Bajri, O. Pütz, B. Paaßen, S. Eger, LLLMS: A data-driven survey of evolving research on limitations of large language models, arXiv preprint arXiv:2505.19240 (2025). URL: https://arxiv.org/abs/2505.19240.

[8] International Organization for Standardization, ISO 1087:2019(en) Terminology work and terminology science — Vocabulary, Standard 1087, 2019. URL: https://www.iso.org/obp/ui/en/#iso:std:iso:1087:ed-2:v1:en.

[9] I. O. Gallegos, R. A. Rossi, J. Barrow, M. M. Tanjim, S. Kim, F. Dernoncourt, T. Yu, R. Zhang, N. K. Ahmed, Bias and fairness in large language models: A survey, Computational Linguistics 50 (2024) 1097–1179. URL: https://aclanthology.org/2024.cl-3.8/. doi:10.1162/coli_a_00524.

[10] Y. Xu, L. Hu, J. Zhao, Z. Qiu, Y. Ye, H. Gu, A survey on multilingual large language models: corpora, alignment, and bias, Frontiers of Computer Science 19 (2024). URL: https://api.semanticscholar.org/CorpusID:268819377.

[11] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, D. Amodei, Language models are few-shot learners, CoRR abs/2005.14165 (2020). URL: https://arxiv.org/abs/2005.14165. arXiv:2005.14165.

[12] W. Wang, W. Jiao, J. Huang, R. Dai, J.-T. Huang, Z. Tu, M. R. Lyu, Not all countries celebrate thanksgiving: On the cultural dominance in large language models, ArXiv abs/2310.12481 (2023). URL: https://api.semanticscholar.org/CorpusID:264305810.