

Automatic Classification of Subjects and Sustainable Development Goals (SDGs) in Documents with Generative AI

An Experience from the Unicamp Library System

Francisco Tadeu Gonçalves de Oliveira Foz, Márcio Souza Martins, Alessandra Karyne Clemente de Souza Neves, Erica Cristina de Carvalho Mansur, and Oscar Eliel

ABSTRACT

This study evaluates the effectiveness of the Artificial Intelligence for Theme Generation tool (original Portuguese acronym name: IAGeraTemas), developed with generative artificial intelligence (AI; Google Gemini), for automating thematic classification and the assignment of Sustainable Development Goals (SDGs) in documents. The methodology combined quantitative analyses (metrics of precision, recall, and accuracy) on 50 articles published by authors from the State University of Campinas (Unicamp), using classification from the SciVal database and qualitative analyses (analysis of the relevance of terms indexed by librarians from the Unicamp Library System in 40 articles available in the Unicamp Institutional Repository), comparing them with manual indexing performed by librarians. The quantitative results in SDG classification showed a recall of 0.785, while the “precision” and “accuracy” metrics were moderate. The qualitative analysis deepened the evaluation of term coherence and relevance suggested by the AI versus human indexing. It revealed the tool’s potential for suggesting relevant terms and expanding concepts, but it also exposed limitations in addressing complex topics. The research, conducted as an experiment at Unicamp Library System, concludes that IAGeraTemas is a valuable auxiliary tool, complementing but not replacing manual indexing, reinforcing the importance of human expertise in validating and refining results, and emphasizing the synergistic potential between AI and information professionals.

INTRODUCTION

Artificial intelligence (AI) has rapidly evolved, significantly impacting various sectors of society. Originating in the 1950s and first implemented in 1956, its history is closely tied to the development of computers.¹ AI was conceived as a set of computing systems designed to perform cognitive tasks similar to those carried out by the human mind, encompassing machine learning (ML), natural language processing, data mining, neural networks, and problem-solving algorithms.²

About the Authors

Francisco Tadeu Gonçalves de Oliveira Foz (fhoz@unicamp.br; corresponding author) is Project Manager, State University of Campinas (UNICAMP). **Márcio Souza Martins** (marciosm@unicamp.br) is Deputy Director, State University of Campinas (UNICAMP). **Alessandra Karyne Clemente de Souza Neves** (akaryne@unicamp.br) is Supervisor of the Information Processing Division, State University of Campinas (UNICAMP). **Erica Cristina de Carvalho Mansur** (ericacc@unicamp.br) is Coordinator of the Information Processing Division, State University of Campinas (UNICAMP). **Oscar Eliel** (oeliel@unicamp.br) is Director, State University of Campinas (UNICAMP). © 2026.

Submitted: 24 July 2025. Accepted for Publication: 22 October 2025. Published: 16 March 2026.

AI in academic libraries is an emerging field of research and development. Recent studies explore AI's potential to enhance multiple areas within library and information science, including user services, translation of materials into other languages, automation of cataloging and classification processes, information retrieval technologies, and knowledge management.³ AI enables libraries to offer high-value services, improving user experience, simplifying library workflows, and enabling fast and efficient access to information.⁴ Its impact is evident both in internal processes, such as the automation of classification, cataloging, and content description tasks, and in external processes, such as enhancing user services, supporting research, and promoting information literacy.

The growing production and digitization of documents, driven by information technologies, have posed significant challenges for organizing and retrieving information, especially in contexts such as libraries and documentation centers.⁵ In such contexts, thematic classification and the assignment of Sustainable Development Goals (SDGs) are essential tasks to ensure that documents promote access to strategic information supporting the development of global and local solutions to social, economic, and environmental challenges. The SDGs comprise a set of 17 global goals established by the United Nations in 2015 as part of the 2030 Agenda for Sustainable Development. This global framework aims to guide public policies, institutional actions, and research practices toward a development model that integrates economic growth, social inclusion, and environmental preservation. Linking scientific production to the SDGs strengthens the role of knowledge as a driver for addressing issues such as poverty, inequality, climate change, and equitable access to education and information.

Therefore, document classification and SDG mapping are not merely technical activities; they play an essential role in fulfilling the social mission of libraries as agents of transformation and promoters of development. However, these processes, traditionally carried out manually by librarians, require time, expertise, and a careful analysis of content. This challenge is even greater for book chapters, which, unlike full books, whose catalog records include the subjects of the entire work and serve as guides for classification and cataloging, usually lack specific indexing terms. This gap in cataloging can hinder both thematic identification and alignment with global guidelines such as the SDGs. Given the relevance of books and chapters in the academic output of schools and research institutes, the Library System of the University of Campinas (SBU) has implemented procedures to catalog these materials more efficiently.

The use of AI for this type of task has proven to be highly efficient. AI-based classifiers demonstrate the ability to process large volumes of textual data, identify semantic patterns, and perform automatic classification with high accuracy.⁶ This approach optimizes the time and resources required for cataloging and ensures greater consistency in indexing and subject assignment, directly contributing to the alignment of documents with global SDGs.

Clearly defining objectives and success criteria is critical to the effectiveness of any classification system, especially when using AI. The complexity inherent in thematic classification lies not only in selecting the most appropriate technology but, more importantly, in precisely defining the task and what constitutes a "good enough" outcome.⁷ No matter how sophisticated an AI model is, it cannot compensate for a poorly defined problem. In the context of classifying documents for SDG alignment, this clarity involves specifying which SDGs will be considered and how themes will be defined and mapped. Prioritizing these aspects enhances the tool's effectiveness and ensures that the results are relevant and aligned with the intended objectives. Conversely, ambiguity in task definition can lead to inaccurate outcomes and the inefficient use of resources, compromising the system's utility.

In this context, this article explores the SBU's experience in applying generative AI to the automatic classification of subjects and SDGs in documents. The comparison presented in this study focuses on two main aspects: SDG assignment and thematic classification. For SDG assignment, AI performance is compared with the automatic classification process used by SciVal (Scopus), which applies search strategies and ML algorithms.⁸ For thematic classification, AI performance is assessed against manual indexing performed by librarians, considering the consistency, comprehensiveness, and relevance of the AI-suggested terms versus those assigned manually. This initiative seeks to assess the effectiveness of AI in automating these complex processes, investigating its potential to improve document organization and expand access to strategic information related to the SDGs. The ultimate goal is to reinforce the role of academic libraries as agents of social transformation and technological innovation, demonstrating how AI can be a valuable ally in optimizing library services and enhancing access to information.

Artificial Intelligence in University Libraries

University libraries must explore the opportunities offered by advances in AI to enhance the quality of their products and services, meeting user needs more efficiently. Many libraries already recognize the benefits of AI, whether for improving operational efficiency, enriching user services, or supporting research activities.⁹

In this context, the adoption of AI in university libraries is growing, particularly in developed countries. This trend is mainly driven by AI's ability to streamline operations, increase staff productivity, and consequently deliver high-quality services to a new generation of users.¹⁰

Implementing AI allows libraries to integrate technologies that optimize repetitive processes, freeing librarians for more strategic tasks.¹¹ It also provides a more personalized and innovative experience for users. Thus, AI not only improves the operational efficiency of libraries but also contributes to the evolution of library services, aligning them with the expectations of an increasingly digital and demanding audience.¹² Recent case studies have illustrated this potential in concrete operational contexts. For example, initiatives based on generative AI and large language models have been applied to the automatic tagging of graduate theses, demonstrating significant gains in scalability and consistency in subject description. Although primarily focused on improving efficiency in descriptive metadata creation, such approaches highlight the growing role of AI as a supportive tool for librarians, particularly in handling large and homogeneous document corpora.¹³

The 5th Strategic Plan of the Spanish University Library Network (2024–2027) highlights AI as a valuable opportunity to optimize library operations. According to the report, AI can automate tasks, enhance user experience, improve data analysis, and offer more efficient information services.¹⁴

The "IFLA Statement on Libraries and Artificial Intelligence" from the International Federation of Library Associations (IFLA) also emphasizes several AI and ML applications with the potential to introduce new services and functionalities in libraries.¹⁵ For example, AI and ML technologies can improve optical character recognition (OCR) in texts, enabling better digitization and information retrieval, and allow new uses for machine-readable library collections, such as content categorization and discovery. Moreover, AI and ML can add new dimensions to knowledge management processes in libraries, particularly in aspects like information organization, storage, and integration. These innovations not only increase the efficiency of library services but also

expand the potential for use and reuse of stored knowledge, making it more accessible and useful to the academic community.

However, despite the clear benefits of AI for university library operations, it also brings challenges that cannot be overlooked. Among these are issues related to privacy, security, and ethical considerations.¹⁶ IFLA recommends that AI usage in libraries be aligned with clear ethical standards, such as those outlined in the IFLA Code of Ethics for Librarians. Additionally, other obstacles have been identified, such as financial uncertainty regarding the adoption of AI tools, emerging skills gaps, and the lack of adequate infrastructure, which may hinder the full implementation of AI in many university libraries.¹⁷

Once these challenges are addressed, it is possible to identify various successful examples of AI use in university libraries around the world. One of the most common cases is the use of AI-based chatbots to provide real-time support to users, answering frequently asked questions and assisting in navigating digital resources and virtual reference services.¹⁸ The implementation of chatbots often results in reduced wait times for users and increased availability of support, particularly outside of regular library hours. However, challenges may arise in the chatbot's ability to handle complex issues or nuances in natural language, requiring human librarian intervention in some cases.

Other AI applications in library services have been highlighted, such as the analysis of scientific publication data.¹⁹ This application shows promise in identifying emerging research trends through alternative methods to traditional approaches and in supporting academic collaboration mapping and generating strategic insights for collection and service management. However, the quality of input data and the interpretation of results are critical for validating such analyses.

Halaychik adds another compelling example through the initiatives undertaken by the University of Texas Libraries in Austin.²⁰ According to the author, to reduce repetitive tasks, an AI-driven initiative was launched to catalog a collection of approximately 150,000 sound recordings on vinyl and compact disc (CD). Halaychik notes that only about 6,000 items had been cataloged over the past 20 years, which, at the current pace, would take another 480 years to complete. The use of AI, leveraging information from the album covers to create basic catalog records, accelerated the process and made the collection more accessible and usable for library users. This initiative demonstrates the potential of AI to handle large volumes of data and expedite lengthy processes, although the need for human review to ensure record quality remains essential.

In Brazil, some university libraries are beginning to explore AI's potential, though in a more modest way. One example is the initiative by the Pontifical Catholic University of Rio de Janeiro (PUC-Rio), which developed a pilot project integrating AI into knowledge management. The project used advanced technologies such as embedding generation, vector databases, semantic search, and large language models to create a question-and-answer interface to support library staff. This support is primarily operational, enabling staff to quickly retrieve validated internal knowledge, such as procedures, services, and documented routines, to respond consistently and accurately to users' inquiries, rather than focus on formal onboarding or training activities.²¹

The opportunities presented by AI to optimize processes such as cataloging and user support, alongside the challenges inherent to knowledge management in a context of exponential document production, as discussed in the article's introduction, underscore the possibilities for innovative and efficient solutions with AI-based tools.

Building on the landscape of AI applications in university libraries and recognizing both the potential and limitations of existing solutions, the next section details the development and functionality of the *IAGeraTemas* tool, designed specifically to address the challenge of automatic subject and SDG classification within the SBU context.

Development of the IAGeraTemas Tool

The *IAGeraTemas* tool was developed to automate the thematic classification of subjects and the assignment of SDGs to documents using generative AI, thereby serving as a guiding instrument.²² Traditionally performed manually by librarians, this process requires considerable time and expertise. The automation provided by *IAGeraTemas* aims to optimize this workflow, reducing the human effort involved.

The tool's architecture is based on the Python programming language, integrating the "[PyPDF2](#)", "[streamlit](#)", and "[google.generativeai](#)" libraries. The PyPDF2 library handles PDF documents, enabling the extraction of the full text regardless of formatting and converting it into a TXT file stored in temporary memory. This extracted text serves as input for the generative AI model. Integration with the Gemini application programming interface (API), through the "[google.generativeai](#)" library allows access to advanced language models capable of understanding and interpreting textual content. In this context, *IAGeraTemas* then analyzes documents from the State University of Campinas (Unicamp), using classification from the SciVal database and qualitative analyses of the relevance of terms indexed by librarians from 40 articles available in the Unicamp Institutional Repository (IR).

Specifically, interaction with the Gemini API is achieved by sending carefully crafted prompts that instruct the model on the tasks of subject classification and SDG identification. The full set of prompts employed in the implementation and evaluation of the tool is provided in the Appendix. It is important to highlight that the *IAGeraTemas* tool was designed to process documents in a structured and guided manner, not only from the Unicamp IR, the official platform for preserving and disseminating the university's academic output, but also from any documents submitted to the tool. The tool's internal instructions, embedded in the prompts sent to the AI model, are carefully formulated to ensure that the content is interpreted according to specific indexing standards (machine-readable cataloging [MARC 21]) and to identify the relevant SDGs. This design choice aligns with recent studies that emphasize that the effectiveness of generative AI applications in libraries depends less on the choice of the language model itself and more on the structuring of instructions, workflows, and contextual constraints imposed on the system. Case studies on the use of large language models for automatic subject tagging in academic libraries have shown that carefully designed prompt strategies and controlled processing flows are essential to achieve consistent and usable descriptive metadata, distinguishing task-oriented tools from generic interactions with large language models.²³

This structured approach, combined with text extraction and preprocessing, ensures that the AI-generated results are aligned with the context and requirements of document classification, distinguishing the use of *IAGeraTemas* from a generic use of the Gemini API, which would not guarantee consistency or contextual relevance. From an operational and economic perspective, the feasibility of *IAGeraTemas* is directly related to its scale of use. The Google Gemini platform provides a daily free processing allowance, which enables the routine use of the tool for regular activities, occasional applications, and pilot studies without direct costs. However, when applied to large document collections or large-scale automation processes, the intensive use of large language models entails higher API processing costs. Evidence reported in the literature indicates

that large-scale thematic tagging and SDG assignment initiatives may involve significant computational costs. In this context, *IAGeraTemas* was designed as a support tool for librarians' work, suitable for controlled use scenarios and indexing support, rather than as a fully automated solution for mass collection processing.²⁴ The model's parameters are configured to prioritize accuracy and relevance in responses, maintaining a balance with creativity in generating indexing terms. For error handling, the tool implements basic input validations, such as API key verification and confirmation that the uploaded file is an OCR-enabled PDF. In case of communication failure with the Gemini API, informative messages are displayed to the user, guiding them on possible causes and solutions. The "[streamlit](#)" library provides the graphical user interface, simplifying interaction with the tool and visualization of results.

The choice of Google's "gemini-2.0-flash" model is justified by its capability to handle complex natural language processing tasks, such as context comprehension and coherent text generation. Security measures are configured in the model to ensure that no content is blocked, enhancing the tool's ability to analyze a wide range of documents.

Processing in *IAGeraTemas* occurs sequentially. After the user inputs their API key and uploads an OCR-enabled PDF file, the PyPDF2 library extracts the document's text, which is temporarily stored. This text file is then processed by the Google generative AI platform. Two specific queries are made to the model: one for subject cataloging, following the MARC 21 standard, field 653\$a, and another for identifying relevant SDGs.

The instructions, formulated as prompts, ensure precision and consistency in results. For subject cataloging, the model returns up to five indexing terms in Portuguese, representative of the document's content. For SDG identification, the system provides one to three goals, listed with their corresponding numbers and descriptions. The results, including subjects and SDGs, are displayed in a structured manner on the tool's graphical interface, developed with "streamlit." The *IAGeraTemas* interface is designed to be intuitive and user-friendly, as shown in Figure 1.

Figure 1. Interface of the *IAGeraTemas* tool, showing the fields for API key entry and PDF file upload.

IA GeraTemas

Gerador de Assuntos e ODS

Bem-vindo ao IA GeraTemas, uma aplicação que utiliza inteligência artificial generativa (Gemini - Google) para identificar e gerar os principais assuntos e Objetivos de Desenvolvimento Sustentável (ODS) de documentos. A ferramenta é projetada para facilitar a análise de textos, proporcionando uma compreensão rápida e precisa dos temas abordados.

Como usar:

1. Insira a API Key: No campo designado, insira sua chave de API para ativar a funcionalidade de geração de assuntos e ODS.
2. Faça o upload do arquivo: Carregue o documento desejado usando o botão de upload. O documento deve ser um PDF OCR.
3. Visualização dos Resultados: Após o processamento, os assuntos e ODS identificados serão exibidos na tela, proporcionando uma visão clara dos tópicos principais e como eles se relacionam com os Objetivos de Desenvolvimento Sustentável.

A aplicação é projetada para ser intuitiva e balizadora, permitindo que você obtenha informações norteadoras para a análise descritiva de seus documentos com apenas alguns cliques.

Insira sua API Key do Gemini

A1ps50AfgGsd3P3g_dfrv0sdL78L7u8LH_md-cg

Faça o upload do arquivo:

Drag and drop file here
 Limit 20MB per file • PDF

Browse files

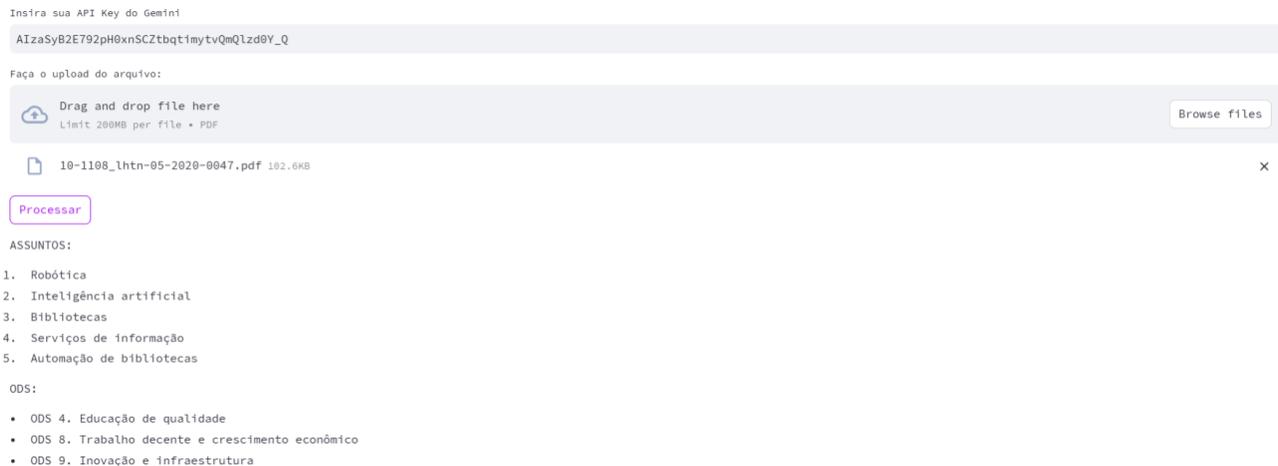
Processar

Elaborado por Francisco Tadeu Gonçalves de Oliveira Foz.

GitHub: [FranciscoFoz](#)

Upon accessing the tool, the user is prompted to enter their Google Gemini API key. Subsequently, a button allows for the upload of the PDF file to be analyzed. After processing, the results are clearly displayed, presenting the suggested indexing terms along with the identified SDGs, their numbers, and corresponding descriptions, as illustrated in Figure 2. This structured presentation of results facilitates understanding and utilization of information for document cataloging and analysis.

Figure 2. Demonstration of the *IAGeraTemas* functionality, showing the list of subjects and Sustainable Development Goals automatically assigned to the processed document.



It is important to note that the tool has certain limitations. First, it depends on the quality of the PDF's OCR; if the OCR is inadequate, text extraction and, consequently, classification may be compromised. Second, the quality of results can vary depending on the document's complexity and clarity.

MATERIALS AND METHODS

This study employs both quantitative and qualitative methodologies to evaluate the performance of the *IAGeraTemas* tool in two interrelated tasks: the automatic classification of SDGs and the subject classification in documents. This dual approach is justified by the need to objectively measure the tool's effectiveness as a support mechanism for cataloging and indexing processes in academic libraries, especially given the increasing complexity and volume of information.

Automatic Classification of the Sustainable Development Goals

The first phase of the research focuses on evaluating *IAGeraTemas*' ability to identify SDGs. For this purpose, a dataset was constructed consisting of 50 open-access scientific articles authored by researchers from Unicamp, selected from the Scopus database via the SciVal tool. The choice of open-access articles authored by Unicamp researchers ensures reproducibility of results and methodological transparency, while also enabling control over variables related to institutional academic output. The SDG attribution made by the Scopus platform serves as the standard reference against which the tool's performance will be compared. It is important to emphasize, however, that this reference should not be interpreted as an absolute measure of accuracy. The Scopus classification provides a widely recognized and systematically curated set of SDG attributions, offering a consistent baseline for evaluation. Nevertheless, the inherent limitations of any bibliometric database imply that discrepancies between the tool and Scopus may reflect

differences in indexing approaches rather than errors in the tool itself. Article selection was guided by criteria aimed at mitigating sampling bias, prioritizing thematic diversity and open access within Unicamp's output, in order to reflect different research contexts.

The evaluation of *IAGeraTemas'* performance in SDG classification employs three statistical metrics: precision, recall, and accuracy. The choice of these metrics is grounded in previous studies that highlight their relevance in assessing automatic text classification systems, particularly in the SDG context. Authors have used these metrics in their research on the automatic classification of sustainability reports, research outputs, and publications in relation to SDGs, demonstrating their utility in measuring performance in similar scenarios.²⁵

Precision measures the proportion of correctly identified SDGs among all SDGs suggested by the tool. Accuracy represents the proportion of correct identifications relative to the total number of identifications (correct and incorrect). Recall measures the proportion of relevant SDGs effectively identified by the tool relative to the total number of SDGs in the standard reference.

Although accuracy is relevant, this study prioritizes recall, based on the rationale that emphasizes the importance of minimizing false negatives (relevant SDGs not identified), even at the expense of a potential increase in false positives (incorrectly attributed SDGs).²⁶ This is because the tool is intended to assist, not replace, professional cataloging activities. A high recall indicates the tool's effectiveness in capturing the most relevant information, thereby supporting human work.

The analysis of results goes beyond numerical comparisons between the metrics and the standard reference. For each article, the correct identification of at least one SDG by the tool will be checked. The absence of correctly identified SDGs will trigger a more detailed qualitative analysis, examining the tool's suggestions individually and investigating the reasons for discrepancies with the standard reference classification. This approach acknowledges the tool's auxiliary nature: it is designed to support, not supplant, the cataloging process carried out by professionals.

Subject Classification (Thematic Indexing)

The second phase of the research, focused on subject classification, complements the evaluation of *IAGeraTemas* by expanding the analysis beyond SDG identification to investigate its ability to extract relevant indexing terms for bibliographic outputs. For this purpose, a corpus of 40 open-access scientific articles was compiled from the Unicamp IR. The choice to use articles from the IR, in contrast to the Scopus database used in the first phase, is justified by the availability of indexing terms manually assigned by librarians, which serve as the standard reference for evaluating the tool. As in the previous phase, the article selection aimed to reflect the thematic diversity present in the repository, ensuring representation across different fields of knowledge and enabling the evaluation of *IAGeraTemas'* performance in varied contexts. After selection, the articles were submitted to the tool for the automatic extraction of indexing terms.

The evaluation of subject classification, unlike the quantitative approach used in the SDG analysis, adopted a predominantly qualitative perspective. This choice stemmed from the inherent complexity of indexing, an activity that involves interpretative nuances and subjective judgments that resist measurement through isolated quantitative metrics. The qualitative analysis, conducted by librarians experienced in indexing, considered the relevance of the terms suggested by *IAGeraTemas* in relation to the content of each article, comparing them with the terms manually assigned in the IR. Four criteria guided this evaluation: concordance between the suggested terms and the manual indexing (taking into account synonyms and related terms), specificity (assessing the conceptual granularity and precision of the terms), exhaustiveness (analyzing the tool's ability

to capture all relevant themes), and coherence (examining the semantic consistency among the suggested terms).

The expertise of the librarians involved in the qualitative analysis is crucial for contextualizing the terms extracted by *IAGeraTemas* and assessing their appropriateness in relation to the subject area and indexing principles. This approach allows not only the identification of specific successes and failures but also a broader understanding of the tool's performance patterns and the underlying reasons for its limitations. In this way, the qualitative analysis offers valuable insights for the refinement of *IAGeraTemas*, going beyond what purely quantitative evaluation can provide. The results of this analysis will be described in detail, highlighting the strengths and shortcomings of the tool in subject classification and informing the development of strategies to improve automated indexing by combining the potential of AI with the expertise of information professionals.

RESULTS

Evaluation of the Automatic Classification of the Sustainable Development Goals

The results of the evaluation of the *IAGeraTemas* tool in the automatic classification of SDGs reveal a promising performance, although there is room for improvement, as shown in Table 1.

Table 1. Comparison between Sustainable Development Goal classification (Scopus/SciVal) and automatic classification by *IAGeraTemas*.

DOI	SDG SciVal (Scopus)	SDG <i>IAGeraTemas</i>
10.2166/wh.2024.205	3, 6, 9, and 11	3, 6, and 11
10.5212/PraxEduc.v.19.22596.022	4	4, 10, and 17
10.1038/s41598-024-70305-w	15	15, 12, and 17
10.26848/rbgf.v17.1.p213-228	15	15
10.3390/plants13213075	2, 8, and 12	2, 12, and 15
10.47197/retos.v52.101121	4	3
10.1016/j.fertnstert.2024.06.013	3	3
10.1186/s13023-024-03273-z	3 and 10	3 and 10
10.5935/0004-2749.20230047	3	3
10.1590/S0034-759020240203	4, 8, and 15	8, 12, and 13
10.1073/pnas.2215677121	2, 8, and 13	2, 8, and 13
10.1186/s12934-024-02578-9	7	7, 9, and 10
10.1016/j.slasd.2024.100158	3	3
10.20396/BJOS.V23I00.8672997	7, 9, and 12	3
10.1186/s13717-024-00486-9	13	13 and 15
10.1163/22116257-bja10076	16	16
10.3390/eng5020060	7	7, 12, and 13
10.1080/02692171.2024.2382099	8	8, 9, and 10
10.3390/su16135730	7 and 13	7, 9, and 12
10.1097/RHU.0000000000001882	1, 3, and 10	3, 10, and 17
10.1590/1413-7054202448019724	9	2, 12, and 13
10.1590/1516-3180.2022.0647.R1.260723	3	3, 5, and 10
10.1021/acsomega.4c06126	9	9 and 12

DOI	SDG SciVal (Scopus)	SDG <i>IAGeraTemas</i>
10.1038/s41598-024-78763-y	3	3
10.1590/1516-3180.2022.0682.R1.180723	5 and 16	3, 5, and 10
10.3390/buildings14092762	12	9, 11, and 12
10.1007/s40487-024-00282-7	3	3
10.1186/s40793-024-00627-9	7, 12, and 13	7, 12, and 13
10.3389/frsus.2024.1409418	11, 12, 14, and 16	11, 12, and 17
10.3390/plants13202907	3	2, 3, and 12
10.1016/j.cie.2024.110181	7	11, 7, and 12
10.1007/s11356-024-33304-y	7, 13, and 15	6, 7, and 12
10.4317/medoral.26518	3	3
10.33361/RPQ.2024.v.12.n.32.659	6	3, 10, and 11
10.3390/su16188065	3 and 11	3, 11, and 13
10.1016/j.ekir.2024.09.005	3	3 and 10
10.3390/resources13020021	7, 9, and 12	7, 12, and 13
10.1590/0101-31572024-3614	8, 13, 15	7, 13, and 17
10.1590/0001-3765202420231344	15	12, 13, and 15
10.1007/s40336-024-00621-8	3	3
10.3390/antiox13030255	3 and 7	3
10.61622/RBGO/2024RBGO75	3 and 4	3, 5, and 17
10.1177/10732748241303425	3	3
10.3390/urbansci8030092	7	7, 11, and 13
10.5935/1518-0557.20230063	5	3
10.3897/BDJ.12.E133976	15	15
10.1038/s41598-024-61624-z	13 and 15	13 and 15
10.4102/JAMBA.V16I1.1649	6	6, 11, and 13
10.1590/1678-49442024V30N2E2024017.PT	2 and 8	2, 8, and 9
10.1016/j.foodcont.2024.110559	2	2, 8, and 12

It is important to acknowledge that the classification of SDGs in Scopus, although widely used as a reference, is carried out through search strategies and ML algorithms based on textual patterns and keywords. This approach, inherently tied to a specific sociocultural framework, may reflect biases and interpretations that do not fully align with the complex and multifaceted reality of the SDGs. Therefore, the comparison with Scopus in this study serves as an important evaluation benchmark but should not be interpreted as an absolute measure of accuracy. The definition of what is “good enough” in classification tasks is complex and context-dependent, with no single universally correct classification.²⁷ The goal here is to evaluate the performance of *IAGeraTemas* against an established standard, while recognizing its limitations and the possibility of interpretive divergences.

The quantitative analysis is based on a set of 50 open-access scientific articles authored by researchers from Unicamp, indexed in the Scopus database, and selected using the SciVal tool. For each article, recall was calculated individually by comparing the SDGs assigned by the

IAGeraTemas tool with those assigned by Scopus, which served as the reference standard. The reported average recall of 0.785 corresponds to the mean recall across all 50 articles, indicating that, on average, the tool correctly identified 78.5% of the SDGs associated with each article. This is a positive result, considering the complexity of SDG classification and the supportive role of the tool. Using the same per-article evaluation approach adopted for recall, precision and accuracy were calculated individually for each article and then averaged across the dataset. Following the same per-article evaluation approach adopted for recall, precision and accuracy were also calculated individually for each article and then aggregated using the arithmetic mean across the dataset. The average precision obtained was 0.576, indicating that approximately 57.6% of the SDGs suggested by the tool were correct. The average accuracy, which measures the exact match between the tool's classifications and the standard reference, was 0.3. This lower value reflects the difficulty in achieving a perfect match, especially considering that an article may relate to multiple SDGs.

The qualitative analysis of the results complemented the quantitative evaluation, providing insights into the limitations and potential of *IAGeraTemas*. Of the 50 articles analyzed, five had no SDGs correctly identified by the tool. A detailed investigation of these cases, presented in Table 2, revealed that the tool, in some instances, assigned SDGs that were tangentially related to the article's topic but not those considered central according to Scopus classification. For example, the article "Perfil de Composición Corporal de los Estudiantes de la Escuela Nacional de Circo de Brasil" ["Body Composition Profile of Brazilian National Circus School Students"], classified by Scopus as related to SDG 4 (Quality Education), was classified by the tool as SDG 3 (Good Health and Well-being). The qualitative analysis shows that although SDG 4 is relevant given the school context, the main focus of the study is on the physical health of the students, justifying the tool's classification. Similar cases were observed in articles addressing topics such as oral health, waste reuse, healthcare territorialization, and medical complications, highlighting the need for refinement in the classification criteria used by Scopus. Considering that Scopus's methodology for SDG mapping²⁸ is based on keyword queries and ML algorithms, while *IAGeraTemas* employs the large language model Gemini, version 2.0 Flash, which is capable of a deeper semantic understanding of full-text documents, the qualitative analysis suggests that in some cases, *IAGeraTemas'* classification may be more accurate and aligned with the true objectives of the articles, thus exposing the limitations of the keyword-based approach used by Scopus.

The results indicate that *IAGeraTemas* performs well in identifying SDGs, with a satisfactory recall of 0.785. The moderate precision and accuracy scores (0.576 and 0.3, respectively) highlight the need for refinement in the tool's ability to distinguish between relevant and tangential SDGs. However, the qualitative analysis presented in Table 2 suggests that part of the discrepancies may be attributed to inconsistencies in Scopus's classification. Considering the analyzed cases in which *IAGeraTemas'* classification proved to be more aligned with the article's subject matter, the tool's recall would increase to 0.885, reinforcing its potential.

The qualitative analysis constitutes an essential component of this study; it demonstrates the unquestionable importance of critical analysis and the specialized knowledge of the librarian professional. Nuances of interpretation, thematic contextualization, and the distinction between superficial relationships and deep connections with the SDGs are aspects that, as shown in Table 2, require human expertise. Therefore, *IAGeraTemas* does not aim to replace but to assist and enhance the librarian's work. The tool provides a starting point, a set of suggestions that, after analysis and validation by the professional, result in a more accurate and comprehensive classification.

Table 2. Qualitative analysis of articles with discrepancies in Sustainable Development Goal (SDG) classification between Scopus/SciVal and *IAGeraTemas*.

Title	DOI	SDG SciVal (Scopus)	SDG <i>IAGeraTemas</i>	Analysis
“Perfil de Composición Corporal de los Estudiantes de la Escuela Nacional de Circo de Brasil” (“Body Composition Profile of Brazilian National Circus School Students”)	10.47197/retos.v52.101121	4	3	The study aligns with SDG 3, as it focuses on the analysis of physical health and body composition of the students. Indicators such as body fat, fat-free mass, and body mass index (BMI) are directly related to health and well-being. Despite the tangential link to SDG 4 (Quality Education) due to the school context, the main objective of the study is to optimize physical training and prevent health problems, reinforcing its connection with SDG 3.
“Evaluation of the Physicochemical Properties of Intracanal Medications used in Traumatized Teeth”	10.20396/BJOS.V23I00.8672997	7, 9, and 12	3	The paper addresses the physicochemical properties of intracanal medications for treating traumatized teeth, being directly related to health, not to energy, infrastructure, or consumption patterns (SDGs 7, 9, and 12). The misclassification may have occurred due to a model based on generic keywords, without considering the actual context of the research. Such models may mistakenly associate terms like <i>physicochemical properties</i> with innovation and consumption. The correct classification is SDG 3, because the study seeks to improve the treatment of traumatized teeth, contributing to oral health and well-being. The comparative analysis of medications aims to improve patients’ quality of life.
“Impact of the Extraction Temperature on the Antifungal Efficiency and Valorization Potential of Anise (<i>Pimpinella anisum</i> L.) Essential Oil Residues: A Sustainable Strategy”	10.1590/1413-7054202448019724	9	2, 12, and 13	The study on the use of residues from the extraction of anise essential oil aligns with SDG 2 by exploring the production of flour with nutritional potential, contributing to food security and agricultural sustainability. It also aligns with SDG 12 by promoting the valorization of residues, fostering responsible consumption and production. Finally, by reducing waste and using resources more efficiently, the study indirectly impacts SDG 13, related to climate change mitigation. SDG 9 (Industry, Innovation, and

Title	DOI	SDG SciVal (Scopus)	SDG <i>IAGeraTemas</i>	Analysis
				Infrastructure), which was initially assigned, does not fit well. Although there is innovation in the extraction process, the main focus of the work is not on infrastructure development or large-scale industrialization.
"Diagnóstico Situacional e Processos de Territorialização em Saúde em Alfenas-Mg: Discursos da Gestão" ("Situational Diagnosis and Territorialization Processes in Health in Alfenas-MG: Management Discourses")	10.33361/RPQ.2024.v.12.n.32 .659	6	3, 10, and 11	Although tangentially related to SDG 6 due to the sanitation aspect, the research focuses on the analysis of health indicators, socioeconomic inequalities, and health territorialization, aligning more appropriately with SDG 3 (Health and Well-being), SDG 10 (Reduced Inequalities), and SDG 11 (Sustainable Cities and Communities). These aspects more accurately capture the study's objectives and scope.
"Ureterovaginal Fistula: A Case Report and Review of an Unexpected Contemporary Complication Secondary to Transvaginal Ultrasound-guided Oocyte Retrieval"	10.5935/1518-0557.20230063	5	3	The study on complications arising from the transvaginal oocyte retrieval focuses on the analysis and treatment of a specific medical condition, the ureterovaginal fistula. Despite the tangential connection to SDG 5 due to its association with female reproductive health, the article's main contribution is to promote health and well-being, aligning directly with SDG 3.

Evaluation of Thematic Classification (Subjects)

The evaluation of thematic classification aimed to analyze the capacity of *IAGeraTemas* to generate subject terms that are relevant and consistent with the content of the articles. For this purpose, four evaluation criteria were applied to each article: agreement, specificity, exhaustiveness, and coherence. The results of this analysis, detailed in Table 3, demonstrate the tool's potential to assist the indexing process, although they also reveal areas that require improvement.

Table 3. Analysis of *IAGeraTemas*' automatic subject indexing.

Link	Subjects – Unicamp Institutional Repository	Subjects – <i>IAGeraTemas</i>
https://repositorio.unicamp.br/acervo/detalhe/1347274	Lasers; Vidros Teluritos; Artigo Original	Laser aleatório; Vidro telurito; Neodímio; Emissão estimulada; Espalhamento da luz
https://repositorio.unicamp.br/acervo/detalhe/1212506	Capitalismo; Resenha	Capitalismo tardio; Cultura de massa; Mídias; Sono; Tecnologia
https://repositorio.unicamp.br/acervo/detalhe/1404021	Cosmologia; Geografia Física; Espaço; Natureza; Teleologia; Artigo Original	Cosmologia; Geografia física; Espaço; Natureza; Teleologia
https://repositorio.unicamp.br/acervo/detalhe/1402114	Solos – Formação; Terra Preta; Artigo Original	Solos urbanos; Antropossolos; Pedogênese; Terra preta de índio; Biomas brasileiros
https://repositorio.unicamp.br/acervo/detalhe/1397396	História Das Idéias Linguísticas; Análise Do Discurso; Artigo Original	Análise do discurso; História das ideias linguísticas; Gramatização brasileira do português; Política linguística; Institucionalização da linguística
https://repositorio.unicamp.br/acervo/detalhe/1214204	Religiões; Cristianismo; Islamismo; Judaísmo; Resenha	Cristianismo; Islamismo; Judaísmo; Monoteísmo; Abraão (Personagem bíblico)
https://repositorio.unicamp.br/acervo/detalhe/1374334	Ciência E Tecnologia – Brasil; Ciência E Tecnologia – Política; Capitalismo; Ciência E Tecnologia – Aspectos Sociais; Artigo Original	Política de ciência e tecnologia; Capitalismo periférico; Elaboração de políticas; Neutralidade da tecnociência; Determinismo tecnológico
https://repositorio.unicamp.br/acervo/detalhe/1175166	Jogos E Brinquedos; Crianças; Assistência Ambulatorial; Relato De Experiência	Criança com câncer; Tratamento ambulatorial; Brinquedoteca hospitalar; Cuidado infantil; Oncologia pediátrica
https://repositorio.unicamp.br/acervo/detalhe/1211812	Biodiesel; Políticas Públicas; Estrutura Produtiva; Artigo Original	Biodiesel; Cadeia produtiva agroindustrial; Programa Nacional de Produção e Uso de Biodiesel (PNPB); Soja; Óleo de algodão
https://repositorio.unicamp.br/acervo/detalhe/1396839	Expressionismo Abstrato; Arte Minimalista; Arte E Literatura; Artigo Original	Pintura; Linguagem; Arte conceitual; Minimalismo; Teatro
https://repositorio.unicamp.br/acervo/detalhe/1182960	Memória; Século Xix; Dossiê	Arquitetura escolar; Século XIX; Campinas (SP); Memória Cidade

Link	Subjects – Unicamp Institutional Repository	Subjects – IAGeraTemas
https://repositorio.unicamp.br/acervo/de talhe/1387286	Espectroscopia De Infravermelho Com Transformada De Fourier; Nir; Espectroscopia No Ultravioleta Visível; Artigo Original	Espectroscopia de infravermelho com transformada de Fourier; Espectroscopia de infravermelho; Espectroscopia de ultravioleta-visível; Espectroscopia de ressonância magnética nuclear de próton; Análise de variância – análise de componentes simultânea (ASCA)
https://repositorio.unicamp.br/acervo/de talhe/1202844	Modernismo (Arte); Artigas, João Batista Vilanova, 1915–1985; Artigo Original	Arquitetura moderna; Arte concreta; Modernismo; Pós-colonialismo; Vilanova Artigas, João Batista, 1915–1985
https://repositorio.unicamp.br/acervo/de talhe/1222114	Geologia; Livros Didáticos; Artigo Original	Ensino de geologia; Geociências; Livro didático; Alfabetização científica; Representações visuais
https://repositorio.unicamp.br/acervo/de talhe/1174942	Heterogeneidade; Artigo Original	Avaliação educacional; Educação médica; Portfólio; Currículo; Aprendizagem baseada em competências
https://repositorio.unicamp.br/acervo/de talhe/1189564	Cidades E Vilas; Editorial	Cidades e patrimônio cultural; Memória urbana; Planejamento urbano; Historiografia; Imaginário urbano
https://repositorio.unicamp.br/acervo/de talhe/1187370	Etanol; Artigo Original	Etanol; Logística; Transporte intermodal; Simulação; Equilíbrio espacial
https://repositorio.unicamp.br/acervo/de talhe/1404227	Logística; Competitividade; Informação; Território Nacional	Logística; Competitividade; Circuitos espaciais de produção; Território brasileiro; Globalização
https://repositorio.unicamp.br/acervo/de talhe/1175424	Hipertensão; Carta	Hipertensão resistente; Hipertensão; Prevalência; Adesão à medicação; Tratamento da hipertensão
https://repositorio.unicamp.br/acervo/de talhe/1213108	Dependência; Dossiê	Alain Touraine; Sociologia brasileira; Dependência (Ciências sociais); Atores sociais; Modernização
https://repositorio.unicamp.br/acervo/de talhe/1212188	Redes Neurais (Computação); Mercado De Opções; Modelo De Black-Scholes; Artigo Original	Apreçamento de opções; Taxa de câmbio; Redes neurais artificiais; Modelo de Black; Mercado de opções brasileiro
https://repositorio.unicamp.br/acervo/de talhe/1202864	Arte E Antropologia; Artigo Original	Fotografia; Memória; Tempo; Georges Didi-Huberman; Antropologia visual
https://repositorio.unicamp.br/acervo/de talhe/1350894	Artigo Original; Crise Financeira Global, 2008–2009; Forças Armadas; Capitalismo; Democracia	Crise mundial; Forças armadas dos Estados Unidos; Capitalismo; Democracia; Governabilidade internacional

Link	Subjects – Unicamp Institutional Repository	Subjects – IAGeraTemas
https://repositorio.unicamp.br/acervo/detalhe/1274388	Música – Instrução E Estudo; Parte De Evento	Educação musical; Músicos; Formação musical; Musicalidade; Talento musical
https://repositorio.unicamp.br/acervo/detalhe/1175994	Traumatismos Da Medula Espinal; Ressonância Magnética; Artigo Original	Lesão medular; Ombro; Ressonância magnética; Dor no ombro; Traumatismos da medula espinal
https://repositorio.unicamp.br/acervo/detalhe/1403706	Imagem Tridimensional; Artigo Original	Mapeamento geomorfológico; Imagens tridimensionais; Fotografias aéreas; StereoPhoto; Maker; Anáglifo
https://repositorio.unicamp.br/acervo/detalhe/1269229	Arte; Memória; Arte E Fotografia; Parte De Evento	Memória; Fotografia; Primeira Guerra Mundial; Arte; Narrativas visuais
https://repositorio.unicamp.br/acervo/detalhe/1341346	Vigas De Concreto; Fibras De Carbono; Concreto Armado; Artigo Original	Reforço estrutural; Fibra de carbono; Vigas de concreto armado; Capacidade resistente; Modos de ruptura
https://repositorio.unicamp.br/acervo/detalhe/1350893	Artigo Original; Mercantilismo; Capitalismo	Capital comercial; Acumulação primitiva; Economia-mundo; Antigo Sistema Colonial; Segunda servidão
https://repositorio.unicamp.br/acervo/detalhe/1175718	Fibrose Cística; Pneumopatias; Espirometria; Artigo Original	Fibrose cística; Capnografia volumétrica; Espirometria; Doença pulmonar; Função pulmonar
https://repositorio.unicamp.br/acervo/detalhe/1406448	Território Nacional; Urbanização; Artigo Original	Círculos de informação; Urbanização; Uso do território; Geografia do Brasil; Redes sociais na Internet
https://repositorio.unicamp.br/acervo/detalhe/1213154	Capitalismo; Dossiê	Capitalismo; Globalização; Políticas públicas; Estado de bem-estar social; Democracia representativa
https://repositorio.unicamp.br/acervo/detalhe/1194912	Pressão Arterial – Medição; Hipertensão; Pressão Arterial; Artigo De Pesquisa	Pressão arterial; Determinação da pressão arterial; Hipertensão; Esfigmomanômetros; Sons de Korotkoff
https://repositorio.unicamp.br/acervo/detalhe/1211310	Capitalismo Financeiro; Mercado Imobiliário; Artigo Original	Mercado imobiliário; Financeirização; Capital portador de juros; Mercadoria fictícia; Capital fictício
https://repositorio.unicamp.br/acervo/detalhe/1207982	Vírus De Rna; Citoquímica; Túbulos De Malpighi; Triatominae; Corpos De Inclusão Viral; Nota	Vírus RNA; Triatomíneos; Microscopia eletrônica; Citoquímica; Tubos de Malpighi

Link	Subjects – Unicamp Institutional Repository	Subjects – <i>IAGeraTemas</i>
https://repositorio.unicamp.br/acervo/detail/1211792	Agricultura Familiar; Inovações Tecnológicas; Agropecuária; Artigo Original	Agricultura familiar; Inovações tecnológicas; Produtividade agropecuária; Censo agropecuário; Brasil
https://repositorio.unicamp.br/acervo/detail/1176952	Anemia; Artigo Técnico	Anemia; Doença renal crônica; Diagnóstico; Hemoglobina; Ferritina;
https://repositorio.unicamp.br/acervo/detail/1362981	Educação Inclusiva; Acessibilidade; Pessoas Com Deficiência – Direitos Fundamentais; Artigo Original	Acessibilidade arquitetônica; Acessibilidade pedagógica; Educação inclusiva; Pessoa com deficiência física; Análise do discurso francesa
https://repositorio.unicamp.br/acervo/detail/1404043	Filmes Finos; Células Solares De Perovskita; Semicondutores; Artigo De Pesquisa	Filmes finos de iodeto de chumbo; Deposição por sputtering; Propriedades óticas; Propriedades estruturais; Segregação de chumbo
https://repositorio.unicamp.br/acervo/detail/1228086	Nanopartículas De Prata	Nanopartículas de prata; Nanotoxicologia; Toxicidade in vitro; Toxicidade in vivo; Biocida

In general, *IAGeraTemas* demonstrated satisfactory performance in thematic classification, with high agreement between the suggested terms and the original subjects manually assigned in most cases analyzed. The tool was particularly effective in expanding general terms into more specific ones, as observed in the article about lasers and tellurite glasses (first row of Table 3), where terms such as “random laser,” “Tellurite glass,” “neodymium,” “stimulated emission,” and “light scattering” were added, enriching the theme description. Similarly, in the indexing of works on the biodiesel production chain and land financialization, the tool generated specific and relevant terms, showing high specificity and exhaustiveness.

In some cases, *IAGeraTemas* replicated the original terms without adding relevant information, such as in the article about cosmology and physical geography (third row of Table 3). In these cases, specificity and exhaustiveness were considered low, indicating the need for the tool’s improvement to expand concepts and provide more granular terms. Semantic coherence, on the other hand, was predominantly high, indicating that the suggested terms, even when not specific, maintained adequate semantic relationships among themselves.

Despite the promising performance, some limitations were observed. In articles with more complex and multidisciplinary themes, such as the article about the various dimensions of the global crisis, the tool presented relevant terms, but the relationship between them was not fully explored, resulting in medium coherence. Specificity and exhaustiveness were also medium in these cases, indicating the need for improvements in the tool’s capacity to handle thematic complexity.

As with the evaluation of the SDGs, the qualitative analysis of thematic classification reinforces the importance of the librarian’s role in validating and refining the terms suggested by *IAGeraTemas*.

Although effective in generating relevant terms, the tool does not replace human expertise in interpreting and contextualizing themes, serving instead to assist the information professional in indexing by providing a starting point for a more precise and comprehensive classification. The combination of AI with human expertise thus appears to be the most promising approach for optimizing indexing processes.

DISCUSSION

This study explored the development and application of the *IAGeraTemas* tool²⁹ within the SBU, aiming to automate thematic classification of subjects and the assignment of SDGs to documents. The research demonstrated the potential of generative AI to optimize traditionally manual processes in university libraries, contributing to greater efficiency in cataloging and indexing, especially in the face of the growing volume of information.

The evaluation results of *IAGeraTemas*, both in SDG classification and subject assignment, revealed a promising performance. In SDG identification, the tool achieved an average recall of 0.785, calculated as the arithmetic mean of recall values obtained individually for each of the 50 analyzed articles, demonstrating its ability to retrieve most of the relevant SDGs. The qualitative analysis also suggested that part of the divergences relative to the standard reference (Scopus) may be attributed to inconsistencies in Scopus's own classification, highlighting the complexity of the task and the need for human validation. In thematic classification, the tool exhibited high agreement with manual indexing in most cases, showing its effectiveness in suggesting relevant terms and expanding general terms into more specific ones. However, limitations were observed in articles with complex and multidisciplinary themes, indicating the need for improvements in the tool's ability to handle such nuances.

The research reinforces the librarian's importance as a central figure in the process, even with the automation provided by AI. *IAGeraTemas* presents itself as an auxiliary tool that supports but does not replace human expertise in interpreting, contextualizing, and validating classifications. The combination of artificial intelligence with the specialized knowledge of information professionals appears as the most effective approach to improving cataloging and indexing processes, ensuring the quality and coherence of information.

Tests carried out by the SBU and plans for the tool's implementation in 2025, including training for librarians and application to theses, dissertations, and other materials in the IR, demonstrate SBU's commitment to innovation and the search for efficient solutions to information management challenges. The application of *IAGeraTemas* in the IR will not only optimize librarians' work but also enrich document metadata, facilitating information retrieval and contributing to knowledge dissemination. Future research could explore the application of *IAGeraTemas* to other document types and contexts, as well as investigate the integration of the tool with other AI technologies, aiming to further enhance automation and efficiency in material cataloging processes. The investigation of new evaluation metrics that consider the particularities of indexing and the inherent subjectivity of textual interpretation also presents a relevant perspective for future research.

ACKNOWLEDGMENTS

Generative AI tools (ChatGPT 4o and Gemini 2.0 flash) were utilized to ensure the orthographic accuracy of the text.

APPENDIX: PROMPTS USED IN THE IAGERATEMAS TOOL

This appendix presents the prompts employed in the *IAGeraTemas* tool for subject classification and Sustainable Development Goals (SDG) identification. The original prompts were written in Brazilian Portuguese (PT-BR) and are presented here with their respective English translations in order to ensure transparency and reproducibility of the experimental procedures.

A.1 Prompt for Subject Classification (MARC 21 – Field 653\$a)

Original prompt (PT-BR)

Catalogue os assuntos desse documento conforme o campo 653\$a do MARC 21.

- Analise o documento e insira apenas 5 assuntos/termos indexadores, em PT-BR, mais relevantes de acordo com o conteúdo.
- Responda apenas os assuntos.
- Liste os assuntos.

English translation

Catalogue the subjects of this document according to MARC 21 field 653\$a.

- Analyze the document and provide only 5 subject headings/indexing terms, in Brazilian Portuguese (PT-BR), that are most relevant to the content.
- Respond only with the subject terms.
- List the subjects.

A.2 Prompt for SDG Identification

Original prompt (PT-BR)

Catalogue a ODS (Objetivos do Desenvolvimento Sustentável) que esse documento representa.

- Analise o documento e imprima de 1 a 3 Objetivos do Desenvolvimento Sustentável (ODS) pertinentes, de acordo com o conteúdo.

Seguem ODSs:

ODS 1. Erradicação da pobreza:

Acabar com a pobreza em todas as suas formas, em todos os lugares.

ODS 2. Erradicação da fome:

Acabar com a fome, alcançar a segurança alimentar e melhoria da nutrição e promover a agricultura sustentável.

ODS 3. Saúde e bem-estar:

Assegurar uma vida saudável e promover o bem-estar para todas e todos, em todas as idades.

ODS 4. Educação de qualidade:

Assegurar a educação inclusiva, equitativa e de qualidade, e promover oportunidades de aprendizagem ao longo da vida para todas e todos.

ODS 5. Igualdade de gênero:

Alcançar a igualdade de gênero e empoderar todas as mulheres e meninas.

ODS 6. Água potável e saneamento:

Assegurar a disponibilidade e gestão sustentável da água e saneamento para todas e todos.

ODS 7. Energia acessível e limpa:

Assegurar o acesso confiável, sustentável, moderno e a preço acessível à energia para todas e todos.

ODS 8. Trabalho decente e crescimento econômico:

Promover o crescimento econômico sustentado, inclusivo e sustentável, emprego pleno e produtivo e trabalho decente para todas e todos.

ODS 9. Inovação e infraestrutura:

Construir infraestruturas resilientes, promover a industrialização inclusiva e sustentável e fomentar a inovação.

ODS 10. Redução das desigualdades:

Reduzir a desigualdade dentro dos países e entre eles.

ODS 11. Cidades e comunidades sustentáveis:

Tornar as cidades e os assentamentos humanos inclusivos, seguros, resilientes e sustentáveis.

ODS 12. Consumo e produção responsáveis:

Assegurar padrões de produção e consumo sustentáveis.

ODS 13. Ação contra a mudança global do clima:

Tomar medidas urgentes para combater a mudança climática e seus impactos.

ODS 14. Vida na água:

Conservação e uso sustentável dos oceanos, dos mares e dos recursos marinhos para o desenvolvimento sustentável.

ODS 15. Vida terrestre:

Proteger, recuperar e promover o uso sustentável dos ecossistemas terrestres, gerir de forma sustentável as florestas, combater a desertificação, deter e reverter a degradação da terra e deter a perda de biodiversidade.

ODS 16. Paz, justiça e instituições eficazes:

Promover sociedades pacíficas e inclusivas para o desenvolvimento sustentável, proporcionar o acesso à justiça para todos e construir instituições eficazes, responsáveis e inclusivas em todos os níveis.

ODS 17. Parcerias e meios de implementação:

Fortalecer os meios de implementação e revitalizar a parceria global para o desenvolvimento sustentável.

– Responda apenas o número e o nome da ODS.

– Caso haja mais de uma, responda em formato de lista, conforme o exemplo:

- ODS 10. Redução das desigualdades
- ODS 8. Trabalho decente e crescimento econômico

English translation

Identify the Sustainable Development Goal(s) (SDGs) represented by this document.

– Analyze the document and output between 1 and 3 relevant Sustainable Development Goals (SDGs), according to its content.

The SDGs are listed below:

SDG 1. No Poverty:

End poverty in all its forms everywhere.

SDG 2. Zero Hunger:

End hunger, achieve food security and improved nutrition, and promote sustainable agriculture.

SDG 3. Good Health and Well-being:

Ensure healthy lives and promote well-being for all at all ages.

SDG 4. Quality Education:

Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all.

SDG 5. Gender Equality:

Achieve gender equality and empower all women and girls.

SDG 6. Clean Water and Sanitation:

Ensure availability and sustainable management of water and sanitation for all.

SDG 7. Affordable and Clean Energy:

Ensure access to affordable, reliable, sustainable and modern energy for all.

SDG 8. Decent Work and Economic Growth:

Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all.

SDG 9. Industry, Innovation and Infrastructure:

Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation.

SDG 10. Reduced Inequalities:

Reduce inequality within and among countries.

SDG 11. Sustainable Cities and Communities:

Make cities and human settlements inclusive, safe, resilient and sustainable.

SDG 12. Responsible Consumption and Production:

Ensure sustainable consumption and production patterns.

SDG 13. Climate Action:

Take urgent action to combat climate change and its impacts.

SDG 14. Life Below Water:

Conserve and sustainably use the oceans, seas and marine resources for sustainable development.

SDG 15. Life on Land:

Protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, halt and reverse land degradation and halt biodiversity loss.

SDG 16. Peace, Justice and Strong Institutions:

Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels.

SDG 17. Partnerships for the Goals:

Strengthen the means of implementation and revitalize the global partnership for sustainable development.

- Respond only with the SDG number and name.
- If more than one SDG applies, present the answer as a list, for example:
 - SDG 10. Reduced Inequalities
 - SDG 8. Decent Work and Economic Growth

ENDNOTES

- ¹ Óscar López-Regalado, Norka Nuñez-Roja, Óscar R. Lopez-Gil, and José Sanchez-Rodriguez, "Análisis del Uso de la Inteligencia Artificial en la Educación Universitaria: Una Revisión Sistemática," *Pixel-Bit. Revista de Medios y Educación* 70 (2024): 97–122, <https://doi.org/10.12795/pixelbit.106336>.
- ² López-Regalado, Nuñez-Roja, Lopez-Gil, and Sanchez-Rodriguez, "Análisis del Uso de la Inteligencia Artificial."
- ³ Muhammad Asim, Muhammad Arif, Muhammad Rafiq, and Rafiq Ahmad, "Investigating Applications of Artificial Intelligence in University Libraries of Pakistan: An Empirical Study," *The Journal of Academic Librarianship* 49, no. 6 (2023): 102803, <https://doi.org/10.1016/j.acalib.2023.102803>; Caio Saraiva Coneglian et al., "Inteligência Artificial Generativa e Recuperação da Informação: Tendências e Oportunidades de Pesquisa," in *Anais do 23. Encontro Nacional de Pesquisa em Ciência da Informação*, 1–16 (ANCIB, 2023), <https://ancib.org/enancib/index.php/enancib/xxxiiienancib/paper/view/1944>; Giuliano Ferreira, "O Impacto da Inteligência Artificial na Gestão do Conhecimento em Bibliotecas Universitárias: Desafios e Soluções," in *Anais do 22. Seminário Nacional de Bibliotecas Universitárias*, 1–19 (SNBU, 2023), <https://portal.febab.org.br/snbu2023/article/view/2934/2832>; J. Liu, J. Liu, and Y. Chen, "Application of Artificial Intelligence Technology in Information Retrieval of University Library," in *Innovative Computing*, edited by J. C. Hung, J. W. Chang, Y. Pei, and W. C. Wu (Singapore: Springer, 2022), 221–28, https://doi.org/10.1007/978-981-16-4258-6_28.
- ⁴ Rosario Gil García, "La Inteligencia Artificial en las Bibliotecas Universitarias Españolas," *Boletín de la Asociación Andaluza de Bibliotecarios* 35, no. 138 (2024): 200–12, <https://aab.es/wp-content/uploads/2024/12/128-BAAB-200-212.pdf>.
- ⁵ A. C. S. Santos et al., "A Literacia Digital nas Práticas Arquivísticas: Uma Revisão da Literatura," *Cuadernos de Educación y Desarrollo* 15, no. 10 (2023): 12522–45, <https://doi.org/10.55905/cuadv15n10-133>; Sarah Rúbia de Oliveira Santos and Célia da Consolação Dias, "Desafios para a Organização do Conhecimento no Âmbito das Humanidades Digitais e do Patrimônio Cultural: Um Olhar a Partir da Área de Informação e Ciência da Computação na Dimensões," *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação* 29 (2024): 1–21, <https://doi.org/10.5007/1518-2924.2024.e98058>; Maycon Soave, and Daniela Lucas da Silva Lemos, "Curadoria Digital em Acervos do Patrimônio Cultural Digital: Aspectos Teóricos e Práticos no Âmbito da Ciência da Informação," *Brazilian Journal of Information Science: Research Trends* 16 (2022): e02152, <https://doi.org/10.36311/1981-1640.2022.v16.e02152>.
- ⁶ Cristina Arhiliuc, Raf Guns, Walter Daelemans, and Tim C. E. Engels, "Journal Article Classification Using Abstracts: A Comparison of Classical and Transformer-based Machine Learning Methods," *Scientometrics* 130, no. 1 (2024): 313–42, <https://doi.org/10.1007/s11192-024-05217-7>.
- ⁷ Arhiliuc, Guns, Daelemans, and Engels, "Journal Article Classification Using Abstracts."
- ⁸ Alexander Bedard-Vallee, Chris James, and Guillaume Roberge, "Sustainable Development Goals (SDGs) Mapping," Elsevier Data Repository, v. 1, July 13, 2023, <https://doi.org/10.17632/y2zzy9vwzy.1>.
- ⁹ A. O. Isiaka et al., "Perceived Awareness and Usefulness of Artificial Intelligence Technology for Efficient Library Operations in University Libraries in Kwara State, Nigeria," *Journal of Library Services and Technologies* 6, no. 1 (2024): 120–34, <https://credence-publishing.com/journal/uploads/archive/202417141644679010325466.pdf>; A. Tella, "Robots Are Coming to the Libraries: Are Librarians Ready to Accommodate Them?" *Library Hi Tech News* 9 (2020): 13–17, <http://doi.org/10.47524/jlst.v6i1.1217>.
- ¹⁰ Isiaka et al., "Perceived Awareness and Usefulness"; S. Olusegun Oyetola et al. "Artificial Intelligence in the Library: Gauging the Potential Application and Implications for Contemporary Library Services in Nigeria," *Data and Metadata* 2 (2023): 36, <https://doi.org/10.56294/dm202336>.
- ¹¹ C. Halaychik. "Embrace AI in Libraries: Freeing Staff for Meaningful Work While Preserving Human Touch," *IFLA Academic and Research Libraries Section Blog*, July 24, 2024, <https://blogs.ifla.org/arl/2024/07/24/embrace-ai-in-libraries-freeing-staff-for-meaningful-work-while-preserving-human-touch/>.
- ¹² Isiaka et al., "Perceived Awareness and Usefulness"; Olusegun Oyetola et al. "Artificial Intelligence in the Library."

- ¹³ Kyle Morgan, "Using AI to Auto-Tag Graduate Theses," *Information Technology and Libraries* 44, no. 4 (2025), <https://doi.org/10.5860/ital.v44i4.17381>.
- ¹⁴ García, "La Inteligencia Artificial en las Bibliotecas Universitarias Españolas."
- ¹⁵ "IFLA Statement on Libraries and Artificial Intelligence," IFLA, 2020, <https://repository.ifla.org/handle/20.500.14598/1646>.
- ¹⁶ Isiaka et al., "Perceived Awareness and Usefulness."
- ¹⁷ T. I. Yusuf, O. A. Adebayo, L. A. Bello, and J. O. Kayode, "Adoption of Artificial Intelligence for Effective Library Service Delivery in Academic Libraries in Nigeria," *Library Philosophy and Practice*, no. 6804 (2022), <https://digitalcommons.unl.edu/libphilprac/6804>.
- ¹⁸ D. Arora, A. Bansal, N. Kumar, and A. Suri, "Invigorating Libraries with Application of Artificial Intelligence," *Library Philosophy and Practice*, no. 3630 (2020), <https://digitalcommons.unl.edu/libphilprac/3630>; S. Hervieux and A. Wheatley, "Perceptions of Artificial Intelligence: A Survey of Academic Librarians in Canada and the United States," *The Journal of Academic Librarianship* 47, no. 1 (2021): 102270, <https://doi.org/10.1016/j.acalib.2020.102270>.
- ¹⁹ Yusuf, Adebayo, Bello, and Kayode, "Adoption of Artificial Intelligence."
- ²⁰ Halaychik, "Embrace AI in Libraries."
- ²¹ Ferreira, "O Impacto da Inteligência Artificial na Gestão do Conhecimento em Bibliotecas Universitárias."
- ²² FranciscoFoz, "Classificador de Assuntos e ODS (generative-ia)," Zenodo, June 12, 2024, <https://doi.org/10.5281/zenodo.11592704>.
- ²³ Morgan, "Using AI to Auto-Tag Graduate Theses."
- ²⁴ Morgan, "Using AI to Auto-Tag Graduate Theses."
- ²⁵ Merih Angin et al., "A RoBERTa Approach for Automated Processing of Sustainability Reports," *Sustainability* 14, no. 23 (2022): 16139, <https://doi.org/10.3390/su142316139>; M. H. L. F. da S. Barros et al., "AI and Data Visualization for Enhanced Policy-Making: Aligning Research Initiatives with Sustainable Development Goals," *Sustainability* 16, no. 24 (2024): 11050, <https://doi.org/10.3390/su162411050>; H. Yin et al., "Leveraging Artificial Intelligence Technology for Mapping Research to Sustainable Development Goals: A Case Study," arXiv, November 9, 2023, <https://arxiv.org/abs/2311.16162>.
- ²⁶ Angin et al., "A RoBERTa Approach"; da S. Barros et al., "AI and Data Visualization."
- ²⁷ Arhiliuc, Guns, Daelemans, and Engels, "Journal Article Classification Using Abstracts."
- ²⁸ Bedard-Vallee and Roberge, "Sustainable Development Goals (SDGs) Mapping."
- ²⁹ FranciscoFoz, "Classificador de Assuntos e ODS (generative-ia)."