arXiv:2509.02581v1 [cs.DL] 27 Aug 2025

Charting the Future of Scholarly Knowledge with AI: A Community Perspective

Azanzi Jiomekong[1,12]
email: jiomekong@tib.eu

Hande Küçük McGinty[2]
email: hande@ksu.edu

Keith G. Mills[9]
email: Keith.Mills@lsu.edu

Allard Oelen[1]
email: allard.oelen@tib.eu

Enayat Rajabi[3]
email: Enayat_Rajabi@cbu.ca

Harry McElroy[10]

Antrea Christou[7]

Anmol Saini[7]

Janice Anta Zebaze[11]

Hannah Kim[8]
email: hannah.kim0007@temple.edu

Anna M. Jacyszyn[4]
email: Anna.Jacyszyn@fiz-Karlsruhe.de

Sören Auer[1,5]
email: auer@tib.eu

[1]TIB Leibniz Information Centre for Science and Technology, Hannover, Germany, [2]Department of Computer Science Kansas State University, Manhattan KS, [3]Management Science Department, Cape Breton University, Sydney, NS, Canada, [4]FIZ Karlsruhe – Leibniz Institute for Information Infrastructure, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany, [5]L3S Research Center, Leibniz University of Hannover, Germany, [6]Department of Engineering and Computer Science, Wright State University, Dayton, OH, USA, [7]College of Engineering and Computer Science, Dayton, Ohio, USA, [8]Department of Biology, Temple University, Philadelphia, PA, USA, [9]Division of Computer Science and Engineering, Louisiana State University, [10]Department of Development and Research, Performigence Corporation, Los Angeles, CA, USA, [11]Department of Physics, University of Yaounde 1, Cameroon, [12]Department of Computer Science, University of Yaounde 1, Cameroon

Abstract

Despite the growing availability of tools designed to support scholarly knowledge extraction and organization, many researchers still rely on manual methods, sometimes due to unfamiliarity with existing technologies or limited access to domain-adapted solutions. Meanwhile, the rapid

increase in scholarly publications across disciplines has made it increasingly difficult to stay current, further underscoring the need for scalable, AI-enabled approaches to structuring and synthesizing scholarly knowledge. Various research communities have begun addressing this challenge independently, developing tools and frameworks aimed at building reliable, dynamic, and queryable scholarly knowledge bases. However, limited interaction across these communities has hindered the exchange of methods, models, and best practices, slowing progress toward more integrated solutions. This manuscript identifies ways to foster cross-disciplinary dialogue, identify shared challenges, categorize new collaboration and shape future research directions in scholarly knowledge and organization.

1 Introduction

Scholarly work and communication encompass the entire system in which research and creative works are created, evaluated for quality, disseminated to the academic community and beyond, used, and preserved for future use. It includes formal publications, such as journal articles and books, as well as informal sharing through preprints, conference presentations, data sharing, and broader engagement with scholarly works and research outputs. Scholarly knowledge serves as the primary engine of progress, shaping our world and guiding our collective future. It forms the backbone of technological advancement, public health systems, and sustainable environmental practices. Obtained through rigorous methods of observation, experimentation, and validation, it is a reliable resource that helps societies solve complex problems and improve the quality of life by achieving sustainable development goals (SDGs) [6].

To be truly useful, scholarly knowledge must first be systematically extracted and organized. However, the scholarly community of today faces the problem of an overload of scientific papers in their respective domains. There is an increasing number of papers published every year (currently, 3 million), in addition to more than 200 million papers that have already been published . This gives rise to the research question: "How can we provide a reliable and living scholarly knowledge base that empowers researchers to query, synthesize, and analyze the vast body of scholarly knowledge?" [63]. To solve this problem, researchers are increasingly using AI to extract, organize [63, 66, 52, 86, 101], analyze, and synthesize knowledge [8, 78, 50, 42, 36, 20] from scholarly literature.

Researchers have contributed to the development of AI-based systems for scholarly communication. Several recent studies described these studies by considering several dimensions:

- AI systems for literature review [20, 117, 105, 88, 17, 57],
- AI models, datasets, and tools for different aspects of the research life cycle [42, 95, 59, 21, 73, 46],
- The exploration of the capability of Large Language Models (LLMs) in

generating novel research ideas based on information from research papers [42, 20].

The research community working on AI for Scholarly communication (composed of scholarly AI developers and users) has operated more or less independently, with limited exchange of ideas, methodologies, methods, theories, etc. The inaugural edition of the AAAI Bridge on AI for Scholarly Communication (AI4SC) at AAAI 2025 brought together a diverse audience, including students, domain researchers, practitioners, and AI experts, regardless of their current use of AI technologies. The goal was to foster cross-disciplinary dialogue, identify shared challenges, catalyze new collaborations, and shape future research directions in scholarly knowledge extraction, organization, and use. This event spanned two days, and this paper aims to report its findings, results of discussions, and suggested future directions, as well as the challenges discussed.

Through a combination of technical talks, collaborative workshops (such as world café sessions), and thematic discussions, participants explored a range of pressing issues at the intersection of AI and scholarly knowledge production and use. Central themes included the **classification of scholarly knowledge**, where participants examined challenges and advances in categorizing and structuring research across diverse domains. Discussions on **knowledge acquisition and organization** focused on the development of tools and frameworks for extracting, linking, and maintaining scholarly knowledge at scale. The **impacts of Generative AI (GenAI) and Artificial General Intelligence (AGI)** were also a key topic, highlighting both the opportunities and risks associated with these technologies throughout the research lifecycle (see Fig. 1). **Community engagement and participation** emerged as another important theme, emphasizing strategies to foster inclusivity, transparency, and interdisciplinary collaboration. Participants also reflected on **knowledge utilization**, particularly how to empower researchers with actionable insights and decision-support systems derived from structured knowledge. Finally, **ethical considerations** were at the forefront, including concerns about bias, provenance, reproducibility, and the responsible use of AI in scholarly discourse. The following sections present the outcome of these discussions and the opportunities and challenges.

2 Classification of AI Systems in Scholarly Communication

From narrow assistive tools to more autonomous generative systems, AI is reshaping how scholarly knowledge is created, evaluated, disseminated, and consumed [42, 20], and in doing so, playing a transformative role across the entire research lifecycle (see Figure 1). Classifying AI systems by the stages of this lifecycle helps to clarify their capabilities, benefits, and challenges in transforming scholarly knowledge creation, dissemination, and use. This classification will help researchers, publishers, institutions, and policy-makers to select appropriate tools and anticipate their limitations.

Figure 1: Illustration of the research life cycle. For a given problem, scientists first brainstorm potential solutions, incorporating and improving on existing literature in order to propose a new method. They conduct experiments to verify the efficacy of their approach, and analyze the results. Once satisfactory findings have been attained, the research is formalized into a publication and subjected to the peer-review process. This adds to the existing literature body and provides new avenues for brainstorming.

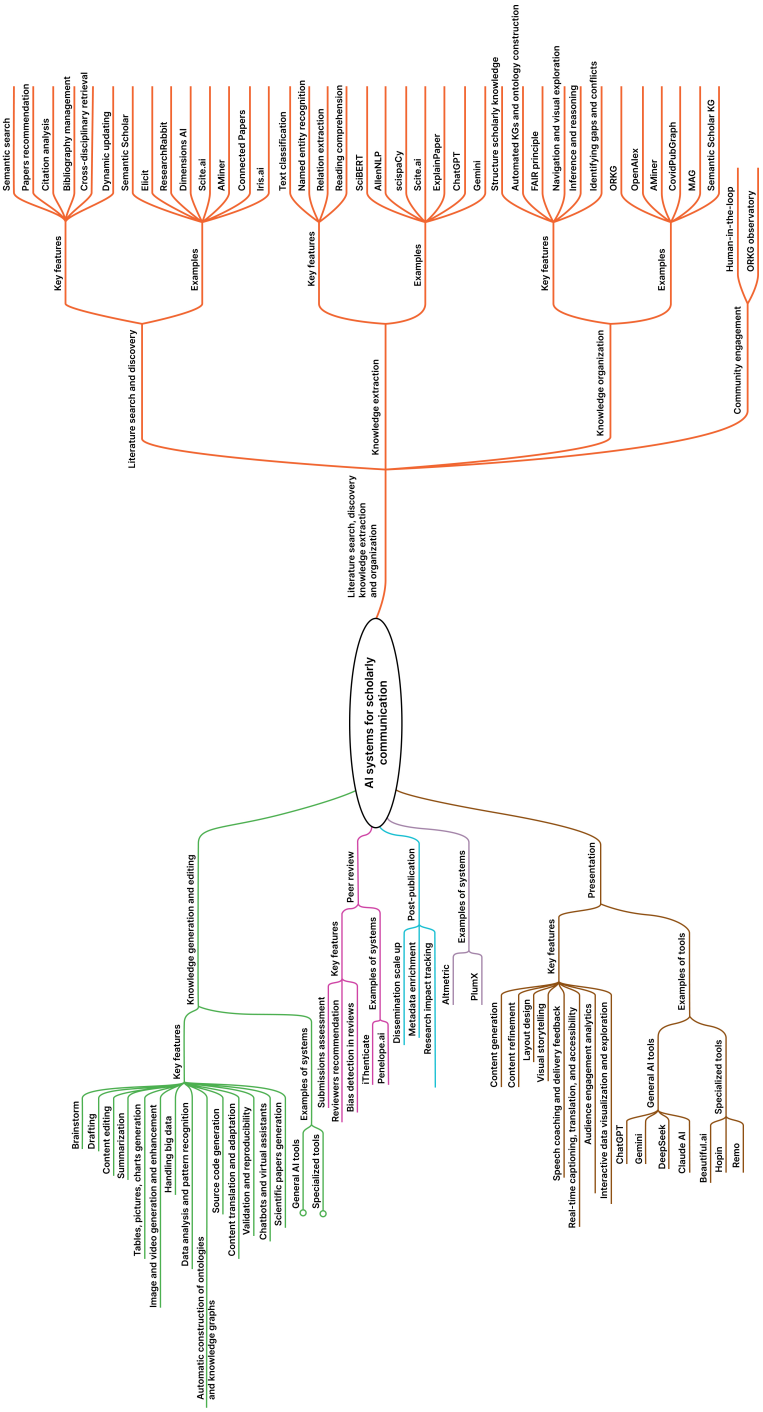

Figure 2: Holistic view of the classification of artificial intelligence for scholarly communication

Generally, AI systems for scholarly communication can be roughly classified as follows:

- Systems for literature discovery, scholarly knowledge extraction and organization;
- Systems for scholarly knowledge generation;
- Systems for peer review, publication, post-publication;
- Systems for slide presentation.

Figure 2 presents a holistic view of AI systems for scholarly communication. The components presented in this figure are described below.

2.1 AI Systems for Literature Search, Discovery, knowledge extraction, and organization

Scholarly communication is recorded, indexed, and disseminated in scholarly publication repositories such as ISI Web of Knowledge, IEE Xplore, Springer, ACM, ScienceDirect, Scopus, Semantic Scholar, arXiv, etc. Based on its structure, knowledge contained in a scholarly paper is broadly classified into two major categories, which are metadata and research contributions [101, 23, 66, 64].

Metadata provides a brief overview of the scholarly data and contributions that contain valuable information that is beneficial to fellow researchers. AI encompassing Machine Learning (ML), Natural Language Processing (NLP), knowledge representation, and data mining provides transformative tools to automate, accelerate, and enhance the processes of scholarly knowledge search and discovery, acquisition, and organization [42, 20].

2.1.1 Literature Search, and Discovery

AI systems for literature search and discovery aim at accelerating the identification of relevant literature to the research topic from the vast scholarly databases, thereby reducing information overload that traditional keyword searches may cause [20, 117, 105, 88, 17, 57]. These tools go beyond conventional citations by identifying supporting, contrasting, and background references, giving a clearer picture of how a paper has been received in the scholarly community. They use NLP and Knowledge Graph (KG) analytics to provide contextualized search results, helping researchers perform comprehensive literature search, minimizing the risk of missing critical sources, and identifying knowledge gaps [20]. Key features involve:

- **Semantic search:** [53, 44] traditional literature searches are often limited by keywords. Semantic search feature aims to understand the meaning and context of user queries, and retrieve relevant results even if the exact keywords are not present.

- **Papers recommendation:** the recommendation feature consists of analyzing keywords, citations, researchers' reading history, and contextual similarities to suggest related papers, ensuring that researchers have access to the most pertinent literature [9].
- **Citation analysis:** This feature involves analyzing existing references to suggest relevant citations, thereby helping researchers discover key studies they might have overlooked [94].
- **Bibliography management:** the bibliography management consists of finding and managing references, automatic organization of citations, and automatic generation of bibliographies [105].
- **Cross-disciplinary retrieval:** This involves identifying relevant literature across diverse fields by recognizing domain-specific terminologies [88].
- **Dynamic updating:** to be relevant, the system should continuously integrate new publications and data to maintain an up-to-date knowledge base [8].

Several AI systems are developed in this direction: Semantic Scholar [70], Elicit [117, 17], ResearchRabbit [32], Dimensions AI, Scite, AMiner [107], Connected Papers.

2.1.2 Knowledge Extraction

Following the literature discovery, AI systems are used for (semi-)automatic identification and extraction of paper metadata and key insights consisting of authors' research contributions.

Knowledge Acquisition (KA) from scholarly data involves the extraction of structured content in the form of entities, relations, facts, terms, and other types of information that may help researchers to understand the papers and get insights from them [23, 66, 63]. It serves as a pivotal activity for extracting metadata and authors' contributions from scholarly communication. Given the vast amount of papers published yearly, the automatic extraction of scholarly data from scholarly text allows the inclusion of a number of papers related to the research topic in the study.

Scholarly knowledge often resides in unstructured formats (text in papers, figures, tables). AI, particularly NLP and Computer Vision (CV), can extract structured data from these sources. For example, it can extract chemical properties from text, identify protein interactions from biological diagrams, or pull experimental parameters from methods sections. Key features of these systems are:

- **Text classification:** [21] consists of automatic categorization of parts of scholarly texts in order to identify and extract meaningful information.
- **Named entity recognition:** [41] consists of the identification and classification of scholarly entities such as authors, affiliation, methods, datasets, and tools into predefined categories.

- **Relation extraction:** [60] consists of the identification and extraction of semantic relationships between entities mentioned in the text.
- **Reading comprehension:** [15] help users identify and comprehend authors contributions.

Several AI systems such as SciBERT [16], AllenNLP [48], scispaCy [87], Scite.ai, ExplainPaper, ChatGPT, Gemini are used for knowledge extraction.

2.1.3 Knowledge Organization

Acquisition is only one part of the equation. Effective organization of scholarly knowledge is crucial, as it enables its use and reuse by different stakeholders. To this end, KGs play a central role in organizing and representing scholarly knowledge by modeling entities (e.g., authors, institutions, concepts, datasets) and their semantic relationships (e.g., citations, collaborations, method usage), structuring thereby scholarly knowledge around shared vocabularies and ontologies [111, 8, 7, 66, 63, 35, 22, 96].

The growing use of KGs reflects the shift toward structured, machine-readable representations of scholarly knowledge, which are essentially for scalable, AI-driven research discovery and evaluation. Key features of AI systems for knowledge scholarly knowledge organization involves [111, 8, 7, 66, 63, 35, 22, 96]:

- **Structure scholarly knowledge:** entities composed of papers metadata, authors' research contributions, underlying datasets, associated code, and relevant experimental protocols are linked, forming a richer, more interconnected knowledge ecosystem.
- **Automated KGs and ontology construction:** to facilitate the extraction of entities and their relationships from millions of research papers, datasets are used to train ML models.
- **FAIR principle:** during scholarly knowledge organization, the FAIR (Findable, Accessible, Interoperable, and Reusable) principle should be fulfilled.
- **Navigation and visual exploration of scholarly data:** this feature consists of interactive graphs, charts, maps, or a visual interface to help understand, analyze, and navigate scholarly KGs.
- **Inference and reasoning capabilities to derive new knowledge:** the reasoning capability of KGs allows the inference of new connections or the answering of complex multi-hop questions that would be impossible with traditional databases.
- **Identifying gaps and conflicts in knowledge:** This feature involves analyzing the structure and content of scholarly KG to identify areas where knowledge is missing, contradictory, or where new connections can be inferred.

Numerous scholarly knowledge graphs are developed and used: the Open Research Knowledge Graph (ORKG) [8], OpenAlex [96], AMiner [107], CovidPubGraph [92], Microsoft Academic Graph (MAG) [116], Semantic Scholar KG [114].

2.1.4 Community Engagement

Realizing the full potential of AI requires not only technical innovation, but also the active involvement of researchers and scholarly communities to ensure accuracy, inclusivity, and ethical stewardship. Therefore, it is important to keep humans involved in describing articles and validating knowledge (Human-in-the-loop) [43, 115].

To facilitate community engagement during scholarly knowledge extraction and organization, the ORKG observatory has been developed. It allows a whole community to collaboratively build the state-of-the-art in a research domain [66]. To this end, scholarly data such as scientific papers, books, and datasets are collected, compared, and summarized.

2.2 AI systems for Knowledge Generation and Editing

Present AI systems are used in the full scholarly communication workflow for generating scholarly knowledge [42, 20, 36]. During the ideation activity, new ideas are generated, developed, and communicated. From ideas, testable statements that propose a possible explanation for a phenomenon or predict the outcome of an experiment can be stated as a research hypothesis. One or several research questions that define what the researcher wants to investigate may emerge. This often forms the foundation for hypotheses or research objectives. These elements are used to shape the research design, methodologies, methods, tools, and experiments to conduct. The methodology explains which methods are used, why they are chosen, and how they help to answer the research question or test the hypothesis. During the experiment, implementation, or simulation, the method or methodology described is used to translate theoretical concepts into practical steps that can be observed, measured, and evaluated. Models are used to replicate real-world processes or phenomena in a controlled virtual environment. This allows researchers to study complex situations that are difficult, expensive, dangerous, or impossible to observe directly and reduces the need for physical prototypes or large-scale trials. Research results present findings or outcomes that emerge as a result of the experiment, implementation, simulation, or data analysis. The analysis of research results is used to connect the raw results to the broader context of the study. During this activity, the comparison with the previous study allows the identification of similarities, differences, and potential reasons for the results obtained. These elements, together with related work drawn from papers from which knowledge was extracted and organized, form the foundation of the research paper. During these activities, several AI systems may be used for knowledge generation and editing. The key features of these AI systems are:

- **Brainstorm and explore a range of possible directions, questions, ideas, and hypotheses**. This feature consists of exploring millions of scholarly resources to answer research questions, generate novel hypotheses by analyzing complex relationships within data, and identify gaps in current knowledge, emerging areas, and under-explored areas.

- **Drafting:** This feature aims to speed up manuscript preparation by generating initial drafts of papers, books, master's theses, initial responses to reviewer comments, and presentations.

- **Content editing:** content editing feature aims at assisting authors during the improvement of the writing quality and clarity by paraphrasing, improving grammar, style, and coherence, and reference insertion.

- **Summarization:** This feature involves processing thousands or millions of scholarly data, including books, scientific papers, patents, and technical reports, to generate concise summaries, extract key findings, methodologies, and relationships. This enables a quick grasp of complex research without needing to read every detail, drastically reducing the time spent on initial literature surveys.

- **Tables, pictures, charts generation:** This feature involves creating data visualizations, diagrams, or synthetic images to illustrate concepts based on textual descriptions or numerical data.

  Data and results analysis, appropriate comparison tables, pictures, and charts can be generated from numeric data or a textual description of the desired visualization. The best chart type can also be suggested (e.g., bar, line, scatter, heatmap) based on data characteristics and the message to convey. Diagrams or molecular structures can also be generated.

  This feature is beneficial because it reduces manual effort in creating visuals, ensures data accuracy in figures, helps non-designers create professional-looking charts, and allows for rapid integration of visual concepts.

- **Image and video generation and enhancement:** This feature involves enhancing the quality of existing images (e.g., research photos, microscopy images) or creating entirely new illustrative visual assets based on textual prompts. This includes tasks like upscaling resolution, de-noising, enhancing contrast, or even creating synthetic illustrative images or short animations.

- **Handling big data:** research is generating petabytes of data. This feature involves analyzing the data to identify subtle correlations, anomalies, and trends that would be invisible to traditional methods. This aims at accelerating the data analysis phase.

- **Data analysis and pattern recognition:** this feature consists of automatic analysis of scholarly data and pattern recognition. Numerous models show powerful performance on tasks such as image recognition,

analysis of complex chemical spectra, and identification of behaviors in large video datasets.

- **Automatic construction of ontologies and knowledge graphs:** Given the vast amount of scholarly knowledge available, manual construction of ontologies and KGs is not an easy task. This feature consists of automatic identification, extraction, and organization of entities (authors, publications, research contributions, etc.), as well as the relationships between them in a scholarly knowledge graph.

- **Source code generation:** this feature aims at assisting the academic community in coding, debugging, or optimizing source code used for analyses or simulation within their papers.

- **Content translation and adaptation:** This feature aims at reducing language barriers, particularly benefiting non-native English speakers. It supports multilingual communication.

- **Validation and reproducibility:** This feature consists of verifying the correctness of statistical analyses and computational workflows. It allows for enhanced transparency and rigor by automating checks for methodological soundness and enabling the sharing of reproducible research components.

- **Chatbots and virtual assistants:** This feature aims to help authors interact with the AI system, providing instant answers based on existing studies.

- **Scientific papers generation:** this feature consists of the automatic generation of the whole research papers.

AI systems for knowledge generation and editing can be classified into: general AI tools such as ChatGPT, Gemini, DeepSeek, and Claude AI, which cover almost all the features presented above. Specialized tools that involve only several features. For instance, Grammarly, Trinka AI are specialized on content editing; Power BI, plotly, Google Data Studio; Adobe Sensei, DALL-E, Topaz Labs, Midjourney, Synthesia for images and video generation and enhancement; Alphafold for handling big data; StatReviewer, Code Ocean for validation and reproducibility; Sakana AI, TIB assistant.

2.3 AI systems for Peer Review, Publication, and Post Publication

Peer review happens in two phases: (1) Firstly, scholarly resources are assigned to reviewers, and then, the reviewers review and submit their comments. Given the number of scholarly resources submitted for publication in some conferences (11,239 papers were submitted to ICCV 2025) and the variety of research topics, there is a need for tools that will facilitate the paper attribution to reviewers,

avoiding conflicts of interest and ensuring that papers are assigned to the appropriate reviewer. These systems are used by editorial boards for the following purposes:

- Assess submissions for reproducibility
- Ethics compliance
- Plagiarism
- Methodological flaws
- Reviewer recommendation
- Bias detection in review.

Several AI tools have been developed in this direction: Penelope.ai, iThenticate. (2) Secondly, reviewers assess the relevance of the research, methodology, data, experiments, results, and conclusion of the scholarly work. To this end, several AI tools can be used for knowledge generation and editing, as presented above. In addition, several tools such as SciScore can assist the reviewer during the review writing.

After the publication activity, AI systems may be used to scale up the dissemination, metadata enrichment, and track the broader impact of the research. These systems should be able to:

- Broadening the reach and improving the accessibility of scholarly communication to diverse audiences, beyond just academic peers.
- Automated translation: translating research papers into multiple languages to break down linguistic barriers and facilitate global knowledge sharing.
- Plain language summaries (PLS): automatically generating simplified, accessible summaries of complex scholarly resources for policymakers, journalists, educators, or the general public.
- Adaptive learning: creating personalized educational content from research findings, tailored to a learner's background and progress.
- Multi-format content generation: automatically adapting research content into different formats (e.g., audio summaries, concise social media posts, interactive QA interfaces) for broader dissemination.

Examples of tools currently used are: Altmetric [109] and PlumX [28].

2.4 AI systems for Slide Presentation

After meticulous analysis, scholarly content preparation, and then publication, effective communication of the content to diverse audiences is crucial for impact. This is generally done in conferences, hackathons, etc. AI is increasingly offering powerful tools to enhance this activity, moving beyond simple content generation to sophisticated visual, auditory, and interactive presentation aids in conferences. Key features of existing AI used include:

- **Content generation:** This feature involves creating or refining the core textual and visual components of a presentation. Several types of content, such as slides, posters, voices, and subtitles, can be generated to save time.
- **Content refinement:** This involves grammar correction, optimization of bullet points, suggestions for impactful phrasing, and ensuring consistency in terminology for specific audiences, such as non-experts.
- **Layout design:** This feature assists in the aesthetic and structural organization of the presentation. It suggests or applies aesthetically pleasing layouts, font combinations, colour schemes, and visual hierarchies for slides based on content, brand guidelines, or desired mood. It allows you to transform raw notes into visually coherent slides.
- **Visual storytelling:** this feature consists of the analysis of the presented content (text, data, images). It suggests optimal ways to break it down across slides, arrange elements for maximum impact, grouping related bullets points, recommend moving large blocks to supplementary slides, indicate where an image would enhance a specific point and create a coherent visual narrative.
- **Speech coaching and delivery feedback:** consists of the analysis of the presenter's voice (pitch, pace, volume), body language (via webcam), and verbal content during practice sessions to provide real-time or post-session feedback. It helps presenter refine their delivery, build confidence, and ensure clarity and engagement for the audience, leading to an impactful talk.
- **Real-time captioning, translation, and accessibility:** This feature provides real-time closed captions for spoken content during live presentations, instant translation of spoken or textual content into multiple languages, and the generation of sign language avatars.
- **Audience engagement analytics:** for visual presentation, this feature analyzes audience participation metrics (e.g., chat activity, poll responses, virtual hand raises, even gaze direction if consent is given) to provide presenters with insights into engagement levels, allowing him to adapt their delivery or content for future presentations and assess impact.

- **Interactive data visualization and exploration:** this feature allows presenters to create interactive visualizations where audience members can drill down into data, filter results, or change parameters in real-time, moving beyond static charts

Generative AI tools such as ChatGPT, Gemini, DeepSeek, and Claude AI cover several features presented above while some tools are specialized for several features. For instance, Beautiful.ai, Powerpresent AI[1] for slide generation and refinement; PowerPoint presenter coach for speech coaching and delivery feedback; Hopin and Remo for audience engagement analytics.

3 Using Scholarly Knowledge and Impact

Scholarly communication's utility spans various domains, from deepening our fundamental comprehension to driving practical solutions for real-world problems. These solutions often arise through translational research, which converts basic scientific findings into usable technologies, or through partnerships between academic institutions and industry, which are essential for advancing the SDGs [6]. Achieving the United Nations (UN) Sustainable Development Goals (SDGs) ambitions hinges critically on the effective generation, dissemination, and utilization of scholarly knowledge.

The relationship between the SDGs and scholarly communication is symbiotic. On one hand, scholarly communication is the engine that translates academic work into actionable knowledge, making it a critical enabler for progress on the SDGs. On the other hand, SDGs are profoundly shaping the landscape and priorities of scholarly communication. During this symbiotic relationship, AI offers unprecedented opportunities to real-time understanding and implementation of insights. To ensure that AI enhances rather than distorts the use of scholarly knowledge, it is essential to establish a robust framework for evaluating AI-driven research.

3.1 Use of Scholarly Knowledge and Impact

3.1.1 Use of Scholarly Knowledge in Education

SDG 4 ensures inclusive and equitable quality education and promotes lifelong learning opportunities for all. It is driven by scholarly knowledge and has a direct impact on all the other SDGs. The rapid development of AI systems has significantly transformed the educational field. This increasing trend was accelerated during the COVID-19 pandemic, when remote learning became widespread and reliance on digital technologies reached unprecedented levels [54, 113]. As a result, AI systems have become integral components of educational systems, utilized by administrators, teachers, and students. The widespread availability and increasing accessibility of these systems have created more equitable access to knowledge globally. This spread is particularly significant for countries and

[1] `https://www.powerpresent.ai/`

regions that lack access to top-tier educational resources, helping to bridge the educational gap between developed and developing nations [4]. On one hand, KGs such as the ORKG are used to organize scholarly data to provide teachers and students with synthetic views. Shared using the FAIR principle, this ensures transparency, reproducibility, and enables further analysis by students and teachers.

On the other hand, Generative AI such as ChatGPT, Gemini, DeepSeek, and Claude AI are used to synthesize complex academic work into accessible formats that are tailored for students and teachers. For instance, a set of academic resources (books, research papers, etc.) in English can be uploaded into an AI system, which then provides answers on objective, methodology, results, etc., using the language of the user. AI can also be used to generate plain language summaries, create explanatory visualizations, and provide translation in the user's language. For learners with disability, AI systems can offer support through voice or video for scholarly data illustration [113, 68].

3.1.2 Using Scholarly Knowledge in Food science and Nutrition

According to WHO, every country in the world is affected by one or more forms of malnutrition [110]. Understanding Food information enables people to maintain a healthy diet, enhance productivity and crop quality, and develop systems for food search, recommendation, and question answering. To this end, AI systems are employed through the entire process of food information engineering - from acquisition and processing to the dissemination of up-to-date food information to various stakeholders [65, 65]. Given the large number and diverse type of scholarly resources containing food data, several AI systems are used for automatic extraction and processing [62, 61]. Once acquired, food ontologies such as FoodOn [38], food knowledge graphs such as FoodKG [56] and food linked data such as FOODpedia [71] - AGROVOC [27] are used to provide a shared understanding of food knowledge or tackle data harmonisation problems.

ML models such as neural networks (NNs), are used to learn associations from food data and build models such as food recognition (e.g., VGG-19, AlexNet, GoogLeNet, Resnet-50, DenseNet, MobileNets, ShuffleNets, etc.) [5, 18], food recommendation [104, 24, 31], ingredient substitution [69], automatic fruit classification and counting [122] as well as plant disease detection in large farms [55, 77, 11]. AI/ML models have also been developed and deployed to increase agricultural productivity while reducing adverse impacts on the environment [112, 3], and answer food QA problems [31].

3.1.3 In Physics

Artificial intelligence is revolutionizing physics-based engineering applications across multiple critical domains. The field employs AI technologies ranging from traditional machine learning approaches to cutting-edge deep learning architectures. In structural engineering and seismic control, AI technologies including

reinforcement learning, deep neural networks, and genetic algorithms [47, 97, 51] are enabling model-free intelligent control systems that adapt to earthquake uncertainties without requiring precise system models. Smart materials and vibration control benefit from fuzzy-neural network controllers that overcome the limitations of conventional systems by accurately modeling nonlinear relationships in magneto-rheological elastomer systems [47]. Wind engineering applications leverage machine learning models including Light Gradient Boosting Machines, CNNs, and Generative Adversarial Networks to predict wind loads, crosswind effects, and flutter performance in tall buildings and bridges [34, 74], significantly reducing the need for costly wind tunnel testing.

The integration of AI system in the physic domain have significant positive impacts across multiple SDGs [2]: SDG 9 (Industry, Innovation and Infrastructure) is directly advanced through AI-enhanced structural design, automated construction quality control, and intelligent infrastructure monitoring systems. AI enables the development of more robust infrastructure capable of withstanding natural disasters and adapting to changing environmental conditions; SDG 11 (Sustainable Cities and Communities) benefits from AI applications in urban infrastructure management, including smart building systems that optimize energy consumption, automated monitoring of infrastructures like bridges and tunnels; SDG 13 (Climate Action) is supported through AI-optimized structural designs that minimize material usage and environmental impact, intelligent systems that improve building energy efficiency, and enhanced disaster preparedness capabilities that help communities adapt to climate change impacts.

Despite significant advances, several challenges persist in applications for Physics. Data quality and availability remain critical issues, as many AI models require large, high-quality datasets that are expensive and time-consuming to collect in engineering contexts.

3.1.4 In Environment

The vast corpus of scholarly literature available today, combined with AI's ability to extract and synthesize information from thousands of research papers has enhanced environmental decision-making and policy development [72]. The explainability capacity of Natural Language Processing (NLP) systems has become important in environmental decision making. These tools can process and analyze extensive scholarly databases, tracking environmental data trends and making complex research findings accessible to policymakers and stakeholders, thereby bridging the gap between academic research and practical environmental action.

The increasing global availability of datasets and scholarly research on clean water and sanitation, as well as affordable and clean energy, has created unprecedented opportunities for AI driven environmental solutions. Big data analysis powered by AI systems such as Claude and GPT enables the processing of massive environmental datasets, providing policymakers with timely and accurate information for informed decision making [72]. This capability directly supports the achievement of: SDG 6 (Clean Water and Sanitation): AI systems can an-

alyze water quality data, predict contamination patterns, and optimize water treatment processes [40, 103]; SDG 7 (Affordable and Clean Energy): Machine learning algorithms can optimize renewable energy distribution [100], predict energy demand [10], and improve grid efficiency [121].

3.1.5 In Economics

AI offers substantial productivity improvements for economists by automating micro-tasks and routine processes. Key applications include research classification, economic forecasting and indicator prediction, data analysis and processing, ideation and feedback for research, background research automation, coding assistance, and mathematical derivations [73, 37]. The field primarily utilizes Large Language Models (LLMs) such as ChatGPT, GPT-3.5, GPT-4, Microsoft Copilot, Google's Gemini and Bard, PaLM-2, Claude 2, and LlaMA 2 [73]. In business related domains like accounting and finance, AI is being used for classification tasks, sentiment analysis, text summarization, and content generation [46]. For economic forecasting specifically, various machine learning techniques are employed including Artificial Neural Networks, adaptive neuro-fuzzy inference systems, genetic programming, support vector regression, and extreme learning machines [90].

The application of AI in economics has significant implications for achieving multiple Sustainable Development Goals (SDGs). Most directly, AI's role in enhancing economic forecasting and strategic planning supports SDG 8: Decent Work and Economic Growth by enabling more informed policy decisions and efficient resource allocation. The automation of routine economic tasks and improved productivity can contribute to sustained economic growth while potentially creating new types of employment opportunities [12]. AI can also support SDG 10: Reduced Inequalities by making sophisticated economic analysis accessible to smaller institutions and developing economies that previously lacked such capabilities [12].

3.1.6 In Law

AI, particularly Large Language Models (LLMs) like GPT, BERT, LEGAL-BERT is being applied across numerous legal tasks including legal judgment prediction, document analysis and drafting, contract review, case prediction, regulatory compliance, legal research, and even providing legal advice [102]. These technologies are also used for document summarization, information extraction, and automated decision-making processes in the field of Law which as other fields, is not exempt of the voluminous amount of data (legal text, proceedings, etc) available. The availability of a voluminous dataset such as: ECHR (European Court of Human Rights, ILDC (Indian Legal Documents Corpus), UNFAIR-ToS (Unfair Terms of Service) is also a plus since it provides enough material for the training of AI models [106].

The integration of AI in law offers several advantages. It can automate routine legal tasks, making legal services more efficient and potentially more

affordable and accessible [80]. AI systems demonstrate high accuracy in statutory reasoning and can effectively manage the complexity of legal language [67]. They also show promise in enhancing legal practice through improved document processing and analysis capabilities.

The application of AI in the legal field has significant impact onc85 SDGs, particularly SDG 16: Peace, Justice and Strong Institutions. By making legal services more accessible and affordable through automation, AI directly contributes to ensuring equal access to justice for all, which is a core target of SDG 16. The technology's ability to process vast amounts of legal information efficiently can strengthen institutional capacity and improve the rule of law. Additionally, AI in law supports SDG 10: Reduced Inequalities by potentially broadening access to legal knowledge and services that are mainly available only to those who can afford expensive legal counsel. The automation of routine legal tasks could bridge the justice gap, particularly benefiting marginalized communities who face barriers to legal representation [83].

However, the bias and fairness concerns identified in AI systems pose risks to these same SDGs. If not properly addressed, AI could perpetuate or even amplify existing inequalities in the justice system, diminishing efforts to achieve fair and inclusive institutions. The privacy and transparency challenges also relate to SDG 16's emphasis on developing effective, accountable, and transparent institutions at all levels [83]. Despite the potential, several major obstacles persist. The inherent "open texture" of law - characterized by vague and ambiguous rules makes automation particularly challenging [67]. AI systems still struggle with the complexity and variability of legal language, and there's a persistent need for large, high-quality annotated datasets [80]. The "black box" nature of many AI systems also creates difficulties in explaining decision-making processes, which is crucial in legal contexts [106].

3.1.7 Improving Collaboration and Communication

SDGs are inherently interdisciplinary. Thus, achieving SDGs demands interdisciplinary, cross-border cooperation that is grounded in scholarly evidence. This cooperation will have a direct impact on SDG 17 (Partnership for the Goals). AI may be used to build scholarly communication networks involving researchers and conferences, thus, connecting researchers across geographical and institutional boundaries and enabling shared efforts to address global challenges. To make research results accessible to different stakeholders, AI may be used personalize content delivery, automate translation, and simplify complex scholarly language into more accessible terms for policymakers, journalists, educators, and the general public. For instance, AI can rapidly gather, filter, and synthesize evidence from multiple studies to answer specific policy questions (e.g., "which foods is appropriate to children with vitamin A deficiency?"). This provides policymakers with comprehensive, evidence-based briefings. AI-powered translation tools may be used to facilitate real-time collaboration between researchers from different linguistic backgrounds.

3.2 Evaluation of the AI-driven use of Research Results

The emergence of AI-driven research necessitates the development of new interdisciplinary methodologies for evaluation that blend traditional scholarly validation with AI-specific assessment techniques.

3.2.1 Evaluation of AI Models

The models used during knowledge discovery, knowledge extraction and generation should be evaluated using context-specific metrics. At the technical level, accuracy, precision, recall, F1 score, MEAN, ROUGE, BLEU, AUC measures should be used. At the operational level, speed, scalability, cost-effectiveness are applicable. At the social level, fairness, accountability, transparency, and ethical acceptability should be applied.

The "black box" nature of many advanced AI models necessitates a deeper look into their internal workings and characteristics. To allow models to be easily evaluated, they should be reproducible and transparent. They should involve open datasets and codebases, with clear documentation of model architectures and parameters, reporting bias and model limitations. These systems should fulfill the following aims:

- Interpretability [99, 84] and explainability (XAI) [82] aims at providing human-understandable explanations for AI decisions, revealing which data features influenced a particular outcome. For instance, which documents were used to generate a content.
- Robustness aims to evaluate the capacity of the model when the data is slightly different or noisy.
- Fairness and bias aims to check for algorithmic bias, concerning demographic groups. In the scholarly communication, the question will be to know if the AI system consider the scholarly data from different geographical region in the world.
- Uncertainty quantification provide the level of confidence in an AI system's predictions. However, currently, AI systems used in scholarly communication do not provide this information. This is problematic because these systems hallucinate at times.
- Computational efficiency and resources aims at assuring that the need in energy consumption, specialized hardware do not limit its broader adoption or sustainability.
- Reproducibility and replicability aims for an AI model to be documented, and ideally made available so as to allow fellow researchers to reproduce the AI's training process and obtain similar results from the same data.
- Validity aims to measure if the AI-generated conclusions, prediction, ... truly accurate and reflective of reality. This means rigorously testing the

AI's outputs against human-verified "ground truth" data or established benchmark dataset.

- Reliability measures the performance consistent over and across different, yet relevant, inputs, or slightly perturbed data. A reliable model should not yield significant different results for minor, imperceptible changes in input data.

3.2.2 Evaluation of Scholarly Communication Supports

Given that AI systems may be used to generate scholarly communication support artifacts such as books, reports, scientific papers, datasets, images, presentation, and graphs, the evaluation of those supports should be considered. The traditional peer-review process, the long-standing cornerstone of scholarly quality control and validation, must evolve significantly to accommodate AI-driven research. New guidelines for reviewing AI-driven research should emerge and peer review should consider AI expertise. Reviewers must possess not only deep domain expertise in the scholarly domain, but also a robust and critical understanding of the use of AI for knowledge extraction, generation and editing to critically assess the appropriateness of the AI generated content, and the contributions of the authors, etc. They should be able to determine what should be accepted for the conference/journal and how to detect the use of AI, etc. This dual expertise is crucial to rigorously assess the ethical considerations and responsible use of AI tools throughout the research.

Journals are increasingly developing and publishing specific guidelines for reviewing AI-driven research while emphasizing AI methodology transparency, ethical considerations, and documented impact. This may necessitate a shift towards more interdisciplinary review panels.

3.2.3 Impact Assessment Based on SDGs

Given the large volume of scholarly communication published yearly, the question on the evaluation of the impact of these data is posed. One of the best ways to evaluate will be to assess the impact of research on SDGs. The ORKG, for instance allows researchers to assign which SDGs a scientific paper impacts. Currently, this task is still manual and many papers indexed by the ORKG do not provide this information. However, AI can help analyze the real-world impact of implemented policies or interventions, evaluate the prediction of future trends or risks such as the spread of infectious diseases, the impact of climate change on specific regions, or the economic effects of technological shifts.

3.2.4 Ethical Evaluation

Beyond technical performance, the broader societal implications of the integration of AI in the research life cycle must be evaluated. The societal impact of the research output will measure the potential for the misuse of the research

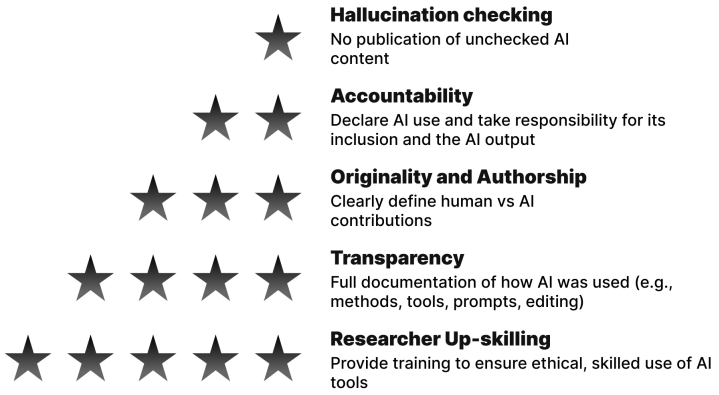

Figure 3: Illustration of the five stars ethics on AI for scholarly communication

results. The accountability aims to determine who is accountable for errors or harmful outcomes when AI contributes significantly to research findings.

The power of AI to synthesize vast amounts of data and identify novel patterns raises fundamental questions about its ethical deployment. AI models learn from the data they are trained on. If this data reflects historical, societal, or sampling biases, the AI will not only perpetuate these biases but can also amplify them in its generated knowledge. For instance, an AI trained predominantly on one demographic group might generate research insights that are less relevant or even harmful to others. If these biases go unchecked, AI can reinforce and legitimate systemic injustices. Addressing this requires inclusive datasets, bias audits and impact assessment.

4 Ethics and Responsibility

As AI becomes a sophisticated tool for knowledge extraction, knowledge generation, content creation, autonomous discovery, etc. critical questions arise regarding reliability, transparency, fairness, accountability, originality, authorship. Responsible use of AI in research demands deliberate ethical reflection, institutional guidelines, and continuous oversight that extend beyond traditional research ethics. Following the five-star model proposed by Tim Berners-Lee for open data, we designed the framework presented in Fig. 3 to illustrate how five-start AI ethics in scholarly communication can be structured.

4.1 Transparency

Transparency is a foundational ethical principle in scientific research and a prerequisite for reproducibility, one of the core pillars of the scientific method. As generative AI systems become increasingly integrated into various stages of the research lifecycle, from data analysis to content generation, it is imperative that

their use be clearly disclosed. This includes providing detailed information on prompts, tool versions, configurations, and any other relevant parameters that could influence the outputs.

Failure to disclose the role of generative AI, or the minimization of its contributions, risks undermining the integrity and reproducibility of the research. Full transparency enables peer reviewers and readers to assess the provenance of research findings and ensures accountability in the use of automated systems.

Current editorial guidelines in many journals recommend that authors explicitly declare the use of generative AI tools when such tools have made substantial contributions to the research process, such as generating content, conducting data analysis, or assisting in the interpretation of results. In contrast, minor uses—such as grammar correction or rephrasing—typically do not require formal disclosure. Nonetheless, as the capabilities of AI tools evolve, clearer and more standardized guidelines may be needed to ensure consistent and ethically sound reporting practices.

At a minimum, transparency in scientific research requires that authors provide a clear and detailed description of their methodology. However, in practice, this requirement is often insufficient—particularly for complex AI projects involving dozens of dependencies, hundreds of files, and thousands of lines of code. Capturing the full intricacy of such systems within the confines of a typical ten-page, double-columned scientific manuscript presents a significant challenge.

To address this limitation, leading AI venues such as AAAI and NeurIPS have introduced reproducibility checklists [19] that prompt authors to report additional details relevant to transparency, such as dataset availability, hyperparameters, code structure, and compute resources. Many conferences also encourage or require the inclusion of source code, which can aid reviewers in assessing the validity and robustness of submitted work. Despite these efforts, reproducibility artifacts typically remain part of the submission's supplementary materials, and their evaluation is often left to the discretion of reviewers [91]. Given the growing volume of submissions to top-tier AI conferences [118], e.g., over 25k submissions for NeurIPS'25 [33], it is not guaranteed that these supplementary components will meaningfully influence the review process or acceptance outcomes.

In contrast, organizations such as the Association for Computational Machinery (ACM) [1] and USENIX take matters a step further by explicitly denoting accepted manuscripts that pass specific transparency and reproducibility thresholds with high-visibly badges on the on the top-right corner of a manuscript's first page in order to foster attention and boost impact. Examples of this process include ANSOR [124] and Optimus [26]. Such assurances provide a greater degree of transparency, for while the peer-review process primarily confirms the novelty/merits of a work, post-acceptance evaluation offers a chance to further confirm implementation correctness and incentivize authors to be more transparent and rigorous in their practices.

4.2 Accountability

AI systems used to assist researchers in scholarly communication often operate as "black boxes," offering little visibility into how decisions are made or outputs are produced. This opacity raises critical questions of accountability. For example, if an AI system autonomously designs experiments, analyzes data, proposes novel theories, or generates hypotheses, who is held responsible if the resulting research is flawed, unethical, or harmful? Similarly, if such a system was trained on copyrighted material without proper attribution, who is liable for any derived content it generates? Ultimately, human researchers, their affiliated institutions, and funding bodies must retain full accountability—not only for their decisions but also for the tools they employ, including AI-generated content. They are responsible for addressing ethical concerns, ensuring the accuracy and fairness of AI-generated contributions, complying with data protection and privacy regulations, and observing copyright and intellectual property norms.

These responsibilities necessitate the development and adoption of clear frameworks for human oversight, continuous review, and transparent reporting in AI-assisted research. Even in highly automated workflows, human-in-the-loop decision-making, validation, and ethical checkpoints remain indispensable. This principle requires that researchers actively interrogate, critique, and verify the outputs of AI tools, rather than passively accepting them.

Efforts such as Verifiable EL Reasoner [58] exemplify how formal verification methods can contribute to this goal. By ensuring that logical inferences made by AI systems are both sound and transparent, such tools provide a foundation for accountable and explainable AI in research contexts. They help bridge the gap between powerful automated reasoning capabilities and the ethical imperatives of human oversight. Current publication practices in leading conferences and journals reinforce this view by placing ultimate responsibility for AI-generated content on the authors. However, as AI becomes increasingly integrated into the research pipeline, the need for verifiable, transparent, and ethically governed AI tools will only grow.

4.3 Originality, Plagiarism, and Authorship

Upholding the integrity of the scientific record requires sustained human vigilance and a deep commitment to quality, transparency, and accountability. Regardless of whether AI systems are used in the research process, it is ultimately the responsibility of human researchers to ensure that all scholarly content they produce is original, properly cited, and intellectually honest. This includes manually verifying the accuracy, appropriateness, and originality of AI-generated content.

As generative AI becomes more capable of producing fluent and coherent text, efficiently synthesizing visual content [125, 85, 45, 30], solving challenging reasoning problems [76, 49] and even full research drafts [50, 78], the peer-review process must evolve accordingly. Reviewers and editors must look beyond surface-level linguistic fluency and assess the conceptual soundness, methodolog-

ical rigor, and genuine novelty of submissions. The potential for AI-generated content to mimic academic style without contributing substantive insight raises serious concerns about research integrity.

When substantial portions of a manuscript are generated by AI without proper citation or disclosure, it may constitute a form of plagiarism or academic misconduct. This is particularly important given that generative AI systems do not create content from first principles; rather, they synthesize outputs based on patterns learned from existing texts, which may include copyrighted or proprietary materials.

The rise of generative AI poses profound ethical questions for scientific communication and authorship. While many leading publishers—including Elsevier, Nature, and Science—have issued guidelines prohibiting the listing of AI systems as authors on the grounds that they cannot assume responsibility for the work, debates continue around the role of AI in the authorship process. Some scholars advocate for the acknowledgment of "AI-assisted authorship" when models contribute significantly to content creation, so long as their involvement is transparently disclosed and human accountability is clearly maintained.

In this rapidly evolving landscape, maintaining the credibility of scientific publishing demands clear ethical frameworks, informed editorial oversight, and a collective commitment to preserving the authenticity of the scholarly record.

4.4 Fabrication and Hallucination Risks

AI models, including large language models (LLMs), learn from the data on which they are trained. When this data contains historical, societal, or sampling biases, these systems not only replicate but often amplify such biases in their generated outputs. This phenomenon contributes to what is commonly referred to as "hallucination" in LLMs—the generation of plausible-sounding but factually incorrect or entirely fabricated content [14, 81, 119, 13, 79].

A prominent example occurred in 2023, when a legal brief drafted using ChatGPT included fictitious case citations, ultimately leading to disciplinary action against the attorneys involved [79]. In scientific contexts, similar hallucinations—such as fabricated studies, inaccurate references, or unsupported claims—pose a serious risk to the integrity of scholarly literature, particularly if undetected during peer review [123, 25, 29].

Recent research highlights the vulnerabilities of LLMs to adversarial prompts and so-called "jailbreaking" techniques, which can elicit harmful or misleading outputs that deviate from expected behavior [120]. These risks underscore the importance of rigorous validation and oversight in the use of generative AI for knowledge generation.

To ensure ethical use of AI in research and publishing, it is essential that all AI-generated content undergo thorough fact-checking and that human oversight is maintained throughout the entire research and writing process. Reliance on AI must never replace the critical judgment, domain expertise, and ethical responsibility of human researchers.

4.5 Human Skill

The growing integration of AI into research workflows has sparked legitimate concerns about its long-term impact on researchers' skills and practices [75]. A central question is whether increasing reliance on AI will lead to a de-skilling of human researchers—eroding their proficiency in core analytical, methodological, or experimental techniques—or, conversely, liberate them to engage more deeply in high-level conceptual thinking and interdisciplinary problem-solving.

AI should be viewed as a tool for augmenting, rather than replacing, human expertise. When used responsibly, it can enable researchers to shift their focus toward strategic reasoning, creative inquiry, and integrative synthesis [75]. However, this requires that researchers maintain critical oversight, actively validate AI-generated outputs, and retain the ability to explain and defend the conclusions drawn from them. Blind reliance on automated systems undermines both scientific rigor and accountability.

Moreover, the significant computational power and technical expertise required for developing and deploying advanced AI systems raise concerns about exacerbating existing global inequities in scientific research. A growing digital divide may emerge, wherein only well-resourced institutions and nations can fully leverage AI capabilities, leaving others behind. Ethical frameworks for the use of AI in research must therefore include considerations of equitable access—to tools, infrastructure, and training—ensuring that the benefits of AI are distributed broadly across the global scientific community.

4.6 Opportunities

As global scientific output continues to expand at an unprecedented rate—surpassing three million publications annually and adding to an archive of over 200 million papers—the research community faces a fundamental challenge: how to meaningfully access, interpret, and build upon this vast and growing body of knowledge. This challenge is not merely logistical but epistemological, raising the question of how collective scientific progress can be sustained in an age of information overload. At the same time, emerging advances in artificial intelligence (AI) offer a historic opportunity to reimagine the entire scholarly knowledge ecosystem.

AI is rapidly becoming a transformative force across all stages of the research lifecycle. From literature discovery to peer review, from knowledge synthesis to post-publication impact analysis, AI-enabled systems promise to augment the capabilities of human researchers, streamline workflows, and expand access to knowledge. When strategically developed and ethically integrated, AI can accelerate discovery, enhance transparency, and democratize scientific participation.

One of the most immediate opportunities lies in leveraging AI for scholarly knowledge extraction and organization. With sophisticated techniques in natural language processing, machine learning, and knowledge representation, AI can distill insights from unstructured scientific texts, map complex relationships between concepts, and build dynamic, queryable knowledge graphs that

evolve with the literature [63, 66]. These systems can help researchers identify relevant works, surface interdisciplinary connections, and synthesize prior findings—functions that are increasingly difficult to perform manually.

The rise of generative AI, particularly large language models (LLMs), presents additional opportunities. These models are capable of summarizing research, generating hypotheses, proposing experimental designs, and even drafting initial research narratives [20]. When used responsibly and transparently, such tools can reduce barriers to participation and allow researchers to focus on conceptual innovation, interdisciplinary collaboration, and socially relevant problem-solving [75].

AI also holds transformative potential for enhancing research integrity and reproducibility. Through automated verification of claims, identification of citation inconsistencies, and replication of computational workflows, AI can support a more trustworthy and transparent scholarly record. Furthermore, AI systems can assist peer reviewers and editors in assessing the factual coherence, methodological rigor, and novelty of submitted manuscripts—thereby strengthening quality control in high-volume publishing environments.

Beyond technical contributions, AI offers strategic pathways toward greater equity in global research. If developed with accessibility and openness in mind, AI-powered tools can help under-resourced institutions, early-career researchers, and non-native English speakers engage more effectively with the scientific literature. Multilingual search, simplified navigation, and context-sensitive recommendations can reduce structural barriers and foster a more inclusive research community. However, realizing this vision requires deliberate investment in equitable infrastructure, governance, and capacity-building.

4.7 Challenges

Despite the transformative potential of artificial intelligence in reshaping scholarly communication, its integration introduces a range of critical challenges that must be addressed with urgency and care. Without proactive intervention, these issues could undermine the integrity, equity, and sustainability of scientific research. Specifically, we highlight the following challenge areas:

Transparency and Reproducibility are a primary concern for AI systems, particularly those relying on large language models and deep learning architectures. These systems often operate as opaque "black boxes," making it difficult to trace how conclusions are derived or to reproduce the same results. In the context of scholarly publishing—where transparency is foundational—this lack of interpretability poses a direct threat to scientific rigor.

One example of the reproducibility challenge is the distinction between whether a contribution can truly be considered 'open-source' or just 'open-weight' [89]. Open-source AI systems provide a sufficiently adequate description of the necessary resources to train a model from experimental hyperparameters, training regime, compute infrastructure and of particular importance, training data. Image generation models [98, 93] are generally closer to open-source in this regard.

In contrast, open-weight refers to AI models whose pretrained checkpoints have been released to the open internet. Although one can download the weights and further fine-tune the model on specialized training data, access to the original pretraining data corpus is usually restricted. This is usually the case with LLMs such as Meta's Llama [108, 39] lineup.

Workforce Integration. Equally important, the growing reliance on AI in research workflows raises concerns about the potential de-skilling of human researchers. As automation takes over tasks such as literature reviews, writing assistance, and data interpretation, there is a risk that core competencies in critical thinking, methodological design, and scholarly communication may erode. Ensuring that AI augments rather than replaces human expertise will require robust educational frameworks and a renewed emphasis on human judgment, reflection, and accountability.

Accountability is also increasingly contested in AI-mediated research environments. As AI begins to generate not just summaries or visualizations, but actual hypotheses, analyses, and textual contributions, questions about responsibility and authorship become more pressing. While most major publishers currently prohibit listing AI tools as authors, ambiguities remain around the attribution of AI-assisted contributions. Without clear frameworks, the integrity of authorship and accountability is at risk.

Bias is also a troubling issue. Specifically, AI models are trained on large-scale corpora that often reflect historical inequities, cultural stereotypes, and disciplinary silos. Moreover, large language models are prone to hallucination—producing confident, yet factually incorrect or fabricated information. In a scientific setting, this can result in the introduction of false claims, spurious references, or misleading interpretations into the academic record, especially if such content is not carefully validated by human experts.

Accessibility. The promise of AI is also tempered by its unequal distribution. Access to advanced AI tools, infrastructure, and expertise is highly concentrated in well-funded institutions and nations. Without targeted efforts to address this imbalance, the deployment of AI may widen existing gaps in scientific capacity and global participation. Equitable access to AI technologies—through open-source tools, inclusive design, and infrastructure investment—is essential to prevent a deepening of the digital divide.

Finally, there is a broader structural challenge: the current academic reward system often prioritizes rapid publication, productivity, and novelty. These incentives may encourage misuse of AI tools to accelerate output at the expense of quality, transparency, and ethical responsibility. Without aligning institutional policies and publishing norms with the principles of responsible AI use, the long-term credibility of scholarly communication may be compromised.

Addressing these challenges will require interdisciplinary collaboration, inclusive dialogue, and the development of shared ethical and technical standards. Only through such collective effort can we ensure that AI strengthens rather than destabilizes the global research enterprise.

5 Conclusion

This paper has summarized the outcomes of the AAAI 2025 Bridge Program on AI for Scholarly Communication (AI4SC), which convened a diverse community of students, domain researchers, practitioners and AI experts to examine the rapidly evolving relationship between artificial intelligence and the production, organization, dissemination and use of scientific knowledge.

AI is profoundly reshaping the scholarly landscape. From automating literature reviews and extracting structured knowledge from unstructured texts to assisting in hypothesis generation, experimental design, and data interpretation, AI is accelerating the pace and scale at which research is conducted and consumed. Its influence extends beyond the confines of the research process itself, enabling advanced knowledge synthesis, context-aware dissemination, and evidence-informed policy making—thereby amplifying the societal impact of research and advancing progress toward the Sustainable Development Goals (SDGs).

Yet, alongside these opportunities come deep challenges. The integration of AI into scholarly communication necessitates a rethinking of how we assess research quality, originality, and ethical responsibility. Traditional notions of scientific rigor must now be paired with emerging requirements for transparency, verifiability, fairness, and human accountability in AI-assisted research workflows.

As AI systems become more capable, the human role in research must evolve—not in opposition to automation, but in concert with it. Researchers must maintain critical oversight, validate AI outputs, and ensure that AI augments rather than diminishes scholarly judgment, creativity, and integrity.

References

[1] ACM: Artifact review and badging - current (August 2020), `https://www.acm.org/publications/policies/artifact-review-and-badging-current`, [Online; posted 24-August-2020]

[2] Addagada, L., Ramesh, S.T., Ratha, D.N., Gandhimathi, R., Rout, P.R.: The role of civil engineering in achieving un sustainable development goals. In: Advances in sustainable materials and resilient infrastructure, pp. 373–389. Springer (2022)

[3] Alazzai, W.K., Abood, B.S.Z., Al-Jawahry, H.M., Obaid, M.K.: Precision farming: the power of ai and iot technologies. In: E3S Web of Conferences. vol. 491, p. 04006. EDP Sciences (2024)

[4] AlSagri, H.S., Sohail, S.S.: Evaluating the role of artificial intelligence in sustainable development goals with an emphasis on "quality education". Discover Sustainability **5**(1), 1–26 (2024)

[5] Anthimopoulos, M.M., Gianola, L., Scarnato, L., Diem, P., Mougiakakou, S.G.: A food recognition system for diabetic patients based on an optimized bag-of-features model. IEEE Journal of Biomedical and Health Informatics **18**(4), 1261–1271 (2014). https://doi.org/10.1109/JBHI.2014.2308928

[6] Assembly, U.N.G.: Transforming our world : the 2030 agenda for sustainable development: resolution / adopted by the general assembly. `https://digitallibrary.un.org/record/3923923` (September 2015), a/RES/70/1

[7] Auer, S., D'Souza, J., Farfar, K.E., Jaradeh, M.Y., Jiomekong, A., Karras, O., Oelen, A., Snyder, L., Stocker, M., Vogt, L.: Open research knowledge graph: a large-scale neuro-symbolic knowledge organization system. In: Handbook on Neurosymbolic AI and Knowledge Graphs, pp. 385–420. IOS Press (2025)

[8] Auer, S., Oelen, A., Haris, M., Stocker, M., D'Souza, J., Eddine Farfar, K., Vogt, L., Prinz, M., Wiens, V., Jaradeh, M.Y.: Improving access to scientific literature with knowledge graphs. BIBLIOTHEK – Forschung und Praxis (2020). https://doi.org/http://dx.doi.org/10.18452/22049

[9] Bai, X., Wang, M., Lee, I., Yang, Z., Kong, X., Xia, F.: Scientific paper recommendation: A survey. Ieee Access **7**, 9324–9339 (2019)

[10] Bai, Y., Camal, S., Michiorri, A.: News and load: A quantitative exploration of natural language processing applications for forecasting day-ahead electricity system demand. IEEE Transactions on Power Systems **39**(5), 6222–6234 (2024)

[11] Bakr, M., Abdel-Gaber, S., Nasr, M., Hazman, M.: Densenet based model for plant diseases diagnosis. European Journal of Electrical Engineering and Computer Science **6**(5), 1–9 (2022)

[12] Bali Swain, R., Yang-Wallentin, F.: Achieving sustainable development goals: predicaments and strategies. International Journal of Sustainable Development & World Ecology **27**(2), 96–106 (2020)

[13] Banerjee, S., Agarwal, A., Singla, S.: Llms will always hallucinate, and we need to live with this. arXiv preprint arXiv:2409.05746 (2024)

[14] Bang, Y., Cahyawijaya, S., Lee, N., Dai, W., Su, D., Wilie, B., Lovenia, H., Ji, Z., Yu, T., Chung, W., et al.: A multitask, multilingual, multimodal evaluation of chatgpt on reasoning, hallucination, and interactivity. arXiv preprint arXiv:2302.04023 (2023)

[15] Baradaran, R., Ghiasi, R., Amirkhani, H.: A survey on machine reading comprehension systems. Natural Language Engineering **28**(6), 683–732 (2022)

[16] Beltagy, I., Lo, K., Cohan, A.: Scibert: A pretrained language model for scientific text. arXiv preprint arXiv:1903.10676 (2019)

[17] Bernard, N., Sagawa Jr, Y., Bier, N., Lihoreau, T., Pazart, L., Tannou, T.: Using artificial intelligence for systematic review: the example of elicit. BMC medical research methodology **25**(1), 75 (2025)

[18] Bettadapura, V., Thomaz, E., Parnami, A., Abowd, G.D., Essa, I.: Leveraging context to support automated food recognition in restaurants. In: 2015 IEEE Winter Conference on Applications of Computer Vision. pp. 580–587. IEEE (2015)

[19] Beygelzimer, A., Dauphin, Y., Liang, P., Vaughn, J.W.: Introducing the neurips 2021 paper checklist (March 2021), `https://neuripsconf.medium.com/introducing-the-neurips-2021-paper-checklist-3220d6df500b`, [Online; posted 26-March-2021]

[20] Bolanos, F., Salatino, A., Osborne, F., Motta, E.: Artificial intelligence for literature reviews: Opportunities and challenges (2024), `https://arxiv.org/abs/2402.08565`

[21] Brack, A., Hoppe, A., Buschermöhle, P., Ewerth, R.: Cross-domain multitask learning for sequential sentence classification in research papers. In: Proceedings of the 22nd ACM/IEEE Joint Conference on Digital Libraries. pp. 1–13 (2022)

[22] Brack, A., Hoppe, A., Stocker, M., Auer, S., Ewerth, R.: Requirements analysis for an open research knowledge graph. In: International Conference on Theory and Practice of Digital Libraries. pp. 3–18. Springer (2020)

[23] Brack, A., Hoppe, A., Stocker, M., Auer, S., Ewerth, R.: Analysing the requirements for an open research knowledge graph: Use cases, quality requirements, and construction strategies. Int. J. Digit. Libr. **23**(1), 33–55 (mar 2022). https://doi.org/10.1007/s00799-021-00306-x

[24] Brintha, N.C., Nagaraj, P., Tejasri, A., Durga, B.V., Teja, M.T., Venkata Pavan Kumar, M.N.: A food recommendation system for predictive diabetic patients using ann and cnn. In: 2022 7th International Conference on Communication and Electronics Systems (ICCES). pp. 1364–1371 (2022). https://doi.org/10.1109/ICCES54183.2022.9835808

[25] Buchanan, J., Hill, S., Shapoval, O.: Chatgpt hallucinates non-existent citations: Evidence from economics. The American Economist **69**(1), 80–87 (2024). https://doi.org/10.1177/05694345231218454

[26] Cai, X., Wang, Y., Zhang, L.: Optimus: towards optimal layer-fusion on deep learning processors. In: Proceedings of the 22nd ACM SIGPLAN/SIGBED International Conference on Languages, Compilers, and Tools for Embedded Systems. pp. 67–79 (2021)

[27] Caracciolo, C., Stellato, A., Morshed, A., Johannsen, G., Rajbhandari, S., Jaques, Y., Keizer, J.: The agrovoc linked dataset. Semantic Web **4**(3), 341–348 (2013)

[28] Champieux, R.: Plumx. Journal of the Medical Library Association: JMLA **103**(1), 63 (2015)

[29] Chelli, M., Descamps, J., Lavoué, V., Trojani, C., Azar, M., Deckert, M., Raynier, J.L., Clowez, G., Boileau, P., Ruetsch-Chelli, C.: Hallucination rates and reference accuracy of chatgpt and bard for systematic reviews: Comparative analysis. J Med Internet Res **26**, e53164 (May 2024). https://doi.org/10.2196/53164

[30] Chen, R., Mills, K.G., Niu, D.: Fp4dit: Towards effective floating point quantization for diffusion transformers. arXiv preprint arXiv:2503.15465 (2025)

[31] Chen, Y., Subburathinam, A., Chen, C.H., Zaki, M.J.: Personalized food recommendation as constrained question answering over a large-scale food knowledge graph. In: Proceedings of the 14th ACM international conference on web search and data mining. pp. 544–552 (2021)

[32] Cole, V., Boutet, M.: Researchrabbit. The Journal of the Canadian Health Libraries Association **44**(2), 43 (2023)

[33] CSPaper: Icml 2025 position paper: Why ai peer review is crumbling and what we can do about it (May 2025), https://cspaper.org/topic/75/icml-2025-position-paper-why-ai-peer-review-is-crumbling-and-what-we-can-do-about-it/2

[34] Deng, J., Lu, Y., Lee, V.C.S.: Concrete crack detection with handwriting script interferences using faster region-based convolutional neural network. Computer-Aided Civil and Infrastructure Engineering **35**(4), 373–388 (2020)

[35] Dessí, D., Osborne, F., Reforgiato Recupero, D., Buscaldi, D., Motta, E.: Cs-kg: A large-scale knowledge graph of research entities and claims in computer science. In: The Semantic Web – ISWC 2022: 21st International Semantic Web Conference, Virtual Event, October 23–27, 2022, Proceedings. p. 678–696. Springer-Verlag, Berlin, Heidelberg (2022). https://doi.org/10.1007/978-3-031-19433-7_39

[36] van Dijk, S.H., Brusse-Keizer, M.G., Bucsán, C.C., van der Palen, J., Doggen, C.J., Lenferink, A.: Artificial intelligence in systematic reviews: promising when appropriately used. BMJ open **13**(7), e072254 (2023)

[37] Dong, M.M., Stratopoulos, T.C., Wang, V.X.: A scoping review of chatgpt research in accounting and finance. International Journal of Accounting Information Systems **55**, 100715 (2024)

[38] Dooley, D.M., Griffiths, E.J., Gosal, G.S., Buttigieg, P.L., Hoehndorf, R., Lange, M.C., Schriml, L.M., Brinkman, F.S., Hsiao, W.W.: Foodon: a harmonized food ontology to increase global food traceability, quality control and data integration. npj Science of Food **2**(1), 23 (2018)

[39] Dubey, A., Jauhri, A., Pandey, A., Kadian, A., Al-Dahle, A., Letman, A., Mathur, A., Schelten, A., Yang, A., Fan, A., et al.: The llama 3 herd of models. arXiv preprint arXiv:2407.21783 (2024)

[40] Durgun, Y.: Real-time water quality monitoring using ai-enabled sensors: Detection of contaminants and uv disinfection analysis in smart urban water systems. Journal of King Saud University-Science **36**(9), 103409 (2024)

[41] D'Souza, J., Auer, S.: Computer science named entity recognition in the open research knowledge graph. In: International Conference on Asian Digital Libraries. pp. 35–45. Springer (2022)

[42] Eger, S., Cao, Y., D'Souza, J., Geiger, A., Greisinger, C., Gross, S., Hou, Y., Krenn, B., Lauscher, A., Li, Y., Lin, C., Moosavi, N.S., Zhao, W., Miller, T.: Transforming science with large language models: A survey on ai-assisted scientific discovery, experimentation, content generation, and evaluation (2025), `https://arxiv.org/abs/2502.05151`

[43] Enqvist, L.: 'human oversight' in the eu artificial intelligence act: what, when and by whom? Law, Innovation and Technology **15**(2), 508–535 (2023). https://doi.org/10.1080/17579961.2023.2245683

[44] Fazzinga, B., Lukasiewicz, T.: Semantic search on the web. Semantic web **1**(1-2), 89–96 (2010)

[45] Feng, W., Yang, C., Qin, H., Li, X., Wang, Y., An, Z., Huang, L., Diao, B., Zhao, Z., Xu, Y., et al.: Q-vdit: Towards accurate quantization and distillation of video-generation diffusion transformers. arXiv preprint arXiv:2505.22167 (2025)

[46] Fisher, I.E., Garnsey, M.R., Hughes, M.E.: Natural language processing in accounting, auditing and finance: A synthesis of the literature with a roadmap for future research. Intelligent Systems in Accounting, Finance and Management **23**(3), 157–214 (2016)

[47] Fu, J., Lai, J., Yang, Z., Bai, J., Yu, M.: Fuzzy-neural network control for a magnetorheological elastomer vibration isolation system. Smart Materials and Structures **29**(7), 074001 (2020)

[48] Gardner, M., Grus, J., Neumann, M., Tafjord, O., Dasigi, P., Liu, N., Peters, M., Schmitz, M., Zettlemoyer, L.: Allennlp: A deep semantic natural language processing platform. arXiv preprint arXiv:1803.07640 (2018)

[49] Ghasemabadi, A., Mills, K.G., Li, B., Niu, D.: Guided by gut: Efficient test-time scaling with reinforced intrinsic confidence. arXiv preprint arXiv:2505.20325 (2025)

[50] Gottweis, J., Weng, W.H., Daryin, A., Tu, T., Palepu, A., Sirkovic, P., Myaskovsky, A., Weissenberger, F., Rong, K., Tanno, R., et al.: Towards an ai co-scientist. arXiv preprint arXiv:2502.18864 (2025)

[51] Greco, A., Cannizzaro, F., Pluchino, A.: Seismic collapse prediction of frame structures by means of genetic algorithms. Engineering Structures **143**, 152–168 (2017)

[52] Greiner-Petter, A.: Semantification of Mathematical LaTeX, pp. 57–93. Springer Fachmedien Wiesbaden, Wiesbaden (2023). https://doi.org/10.1007/978-3-658-40473-4_3

[53] Guha, R., McCool, R., Miller, E.: Semantic search. In: Proceedings of the 12th International Conference on World Wide Web. p. 700–709. WWW '03, Association for Computing Machinery (2003). https://doi.org/10.1145/775152.775250

[54] Han, J.W., Park, J., Lee, H.: Analysis of the effect of an artificial intelligence chatbot educational program on non-face-to-face classes: a quasi-experimental study. BMC Medical Education **22**(1), 830 (2022)

[55] Hassena, M., Alemu, D., Dey, B.: Quality declared seed mechanism in ethiopia. A Feed the Future Global Supporting Seed Systems for Development activity report (2023)

[56] Haussmann, S., Seneviratne, O., Chen, Y., Ne'eman, Y., Codella, J., Chen, C.H., McGuinness, D.L., Zaki, M.J.: Foodkg: a semantics-driven knowledge graph for food recommendation. In: International Semantic Web Conference. pp. 146–162. Springer (2019)

[57] Heikkilä, J.T.: Human intelligence versus artificial intelligence in classifying economics research articles: exploratory evidence. Journal of Documentation **81**(7), 18–30 (2024)

[58] Ileri, A.M., Rangarajan, N., Cannell, J., McGinty, H.: Vel: A formally verified reasoner for el++ description logic (2024), `https://arxiv.org/abs/2412.08739`

[59] Jhajj, K.S., Jindal, P., Kaur, K.: Use of artificial intelligence tools for research by medical students: a narrative review. Cureus **16**(3) (2024)

[60] Jiang, M., D'Souza, J., Auer, S., Downie, J.S.: Targeting precision: A hybrid scientific relation extraction pipeline for improved scholarly knowledge organization. Proceedings of the Association for Information Science and Technology **57**(1), e303 (2020)

[61] Jiomekong, A.: Food information processing. https://orkg.org/review/R640411 (2022), [Online; accessed 2023-09-16]

[62] Jiomekong, A.: Collecting food information. https://orkg.org/review/R609854 (2023). https://doi.org/10.48366/r609854, [Online; accessed 2023-09-16]

[63] Jiomekong, A., Auer, S., Oelen, A.: Linked open literature review using the neuro-symbolic open research knowledge graph. In: Companion Proceedings of the ACM on Web Conference 2024. p. 1015–1018. Association for Computing Machinery, New York, NY, USA (2024). https://doi.org/10.1145/3589335.3651238

[64] Jiomekong, A., Auer, S., Oelen, A.: Linked open literature review using the neuro-symbolic open research knowledge graph. In: Companion Proceedings of the ACM Web Conference 2024. p. 1015–1018. WWW '24, Association for Computing Machinery (2024). https://doi.org/10.1145/3589335.3651238

[65] Jiomekong, A., Oelen, A., Auer, S., Anna-Lena, L., Lars, V.: Food information engineering. AI Mag. **45**(3), 338–353 (Sep 2024). https://doi.org/10.1002/aaai.12185

[66] Jiomekong, A., Tiwari, S.: An approach based on open research knowledge graph for knowledge acquisition from scientific papers. The Electronic Library (2024). https://doi.org/10.1108/EL-06-2023-0154

[67] Katz, D.M., Hartung, D., Gerlach, L., Jana, A., Bommarito II, M.J.: Natural language processing in the legal domain. arXiv preprint arXiv:2302.12039 (2023)

[68] Kaur, A., Singh, S., Chandan, J.S., Robbins, T., Patel, V.: Qualitative exploration of digital chatbot use in medical education: A pilot study. Digital Health **7**, 20552076211038151 (2021)

[69] Kim, H., Venkataramanan, R., Sheth, A.: A survey on food ingredient substitutions. arXiv preprint arXiv:2501.01958 (2024)

[70] Kinney, R., Anastasiades, C., Authur, R., Beltagy, I., Bragg, J., Buraczynski, A., Cachola, I., Candra, S., Chandrasekhar, Y., Cohan, A., et al.: The semantic scholar open data platform (2025), `https://arxiv.org/abs/2301.10140`

[71] Kolchin, M., Chistyakov, A., Lapaev, M., Khaydarova, R.: Foodpedia: Russian food products as a linked data dataset. In: European Semantic Web Conference. pp. 87–90. Springer (2015)

[72] Konya, A., Nematzadeh, P.: Recent applications of ai to environmental disciplines: A review. Science of The Total Environment **906**, 167705 (2024)

[73] Korinek, A.: Generative ai for economic research: Use cases and implications for economists. Journal of Economic Literature **61**(4), 1281–1317 (2023)

[74] Lin, P., Ding, F., Hu, G., Li, C., Xiao, Y., Tse, K.T., Kwok, K., Kareem, A.: Machine learning-enabled estimation of crosswind load effect on tall buildings. Journal of Wind Engineering and Industrial Aerodynamics **220**, 104860 (2022)

[75] Lira, B., Rogers, T., Goldstein, D.G., Ungar, L., Duckworth, A.L.: Learning not cheating: Ai assistance can enhance rather than hinder skill development (2025), https://arxiv.org/abs/2502.02880

[76] Liu, A., Feng, B., Xue, B., Wang, B., Wu, B., Lu, C., Zhao, C., Deng, C., Zhang, C., Ruan, C., et al.: Deepseek-v3 technical report. arXiv preprint arXiv:2412.19437 (2024)

[77] Liu, X., Min, W., Mei, S., Wang, L., Jiang, S.: Plant disease recognition: A large-scale benchmark dataset and a visual region and loss reweighting approach. IEEE Transactions on Image Processing **30**, 2003–2015 (2021)

[78] Lu, C., Lu, C., Lange, R.T., Foerster, J., Clune, J., Ha, D.: The ai scientist: Towards fully automated open-ended scientific discovery. arXiv preprint arXiv:2408.06292 (2024)

[79] Lyon, C.F.: Fake cases, real consequences: misuse of chatgpt leads to sanctions. NY Litigator **28**(2), 8–12 (2023)

[80] Magnusson Sjöberg, C.: Legal automation: Ai and law revisited. Legal Tech, Smart Contracts and Blockchain pp. 173–187 (2019)

[81] Maleki, N., Padmanabhan, B., Dutta, K.: Ai hallucinations: a misnomer worth clarifying. In: 2024 IEEE conference on artificial intelligence (CAI). pp. 133–138. IEEE (2024)

[82] Mersha, M., Lam, K., Wood, J., Alshami, A.K., Kalita, J.: Explainable artificial intelligence: A survey of needs, techniques, applications, and future direction. Neurocomputing **599**, 128111 (2024)

[83] Meskic, Z., Albakjaji, M., Jevremovic, N., Omerovic, E., Adams, J.: Digitalization and innovation in achieving sdgs–impacts on legislation and practice. In: IOP Conference Series: Earth and Environmental Science. vol. 1026, p. 012061. IOP Publishing (2022)

[84] Mills, K.G., Han, F.X., Salameh, M., Lu, S., Zhou, C., He, J., Sun, F., Niu, D.: Building optimal neural architectures using interpretable knowledge. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 5726–5735 (06 2024)

[85] Mills, K.G., Salameh, M., Chen, R., Hassanpour, N., Lu, W., Niu, D.: Qua2sedimo: Quantifiable quantization sensitivity of diffusion models. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 39, pp. 6153–6163 (2025)

[86] Nasar, Z., Jaffry, S.W., Malik, M.K.: Information extraction from scientific articles: a survey. Scientometrics **117**(3), 1931–1990 (2018). https://doi.org/10.1007/s11192-018-2921-5, `https://doi.org/10.1007/s11192-018-2921-5`

[87] Neumann, M., King, D., Beltagy, I., Ammar, W.: Scispacy: fast and robust models for biomedical natural language processing. arXiv preprint arXiv:1902.07669 (2019)

[88] O'Mara-Eves, A., Brunton, G., McDaid, D., Kavanagh, J., Oliver, S., Thomas, J.: Techniques for identifying cross-disciplinary and 'hard-to-detect'evidence for systematic review. Research Synthesis Methods **5**(1), 50–59 (2014)

[89] OSI: Open weights: Not quite what you've been told, `https://opensource.org/ai/open-weights`

[90] Pakholchuk, V., Horiacheva, K.: Methodological concepts of appling ai into military and economic capabilities data analysis. Challenges to National Defence in Contemporary Geopolitical Situation **1**(1) (2024)

[91] PCs, N..: Neurips call for papers 2025 (April 2025), `https://neurips.cc/Conferences/2025/CallForPapers`, [Online; posted 3-April-2025]

[92] Pestryakova, S., Vollmers, D., Sherif, M.A., Heindorf, S., Saleem, M., Moussallem, D., Ngomo, A.C.N.: Covidpubgraph: A fair knowledge graph of covid-19 publications. Scientific Data **9**(1), 389 (2022)

[93] Podell, D., English, Z., Lacey, K., Blattmann, A., Dockhorn, T., Müller, J., Penna, J., Rombach, R.: Sdxl: Improving latent diffusion models for high-resolution image synthesis. arXiv preprint arXiv:2307.01952 (2023)

[94] Pornprasit, C., Liu, X., Kiattipadungkul, P., Kertkeidkachorn, N., Kim, K.S., Noraset, T., Hassan, S.U., Tuarob, S.: Enhancing citation recommendation using citation network embedding. Scientometrics pp. 1–32 (2022)

[95] Pride, D., Cancellieri, M., Knoth, P.: Core-gpt: Combining open access research and large language models for credible, trustworthy question answering. In: Linking Theory and Practice of Digital Libraries: 27th International Conference on Theory and Practice of Digital Libraries, TPDL

2023, Zadar, Croatia, September 26–29, 2023, Proceedings. p. 146–159. Springer-Verlag, Berlin, Heidelberg (2023). https://doi.org/10.1007/978-3-031-43849-3_13

[96] Priem, J., Piwowar, H., Orr, R.: Openalex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. arXiv preprint arXiv:2205.01833 (2022)

[97] Rahmani, H.R., Chase, G., Wiering, M., Könke, C.: A framework for brain learning-based control of smart structures. Advanced Engineering Informatics **42**, 100986 (2019)

[98] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B.: High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 10684–10695 (06 2022)

[99] Ru, B., Wan, X., Dong, X., Osborne, M.A.: Interpretable neural architecture search via bayesian optimisation with weisfeiler-lehman kernels. In: ICLR (2021)

[100] Saxena, N., Kumar, R., Rao, Y.K., Mondloe, D.S., Dhapekar, N.K., Sharma, A., Yadav, A.S.: Hybrid knn-svm machine learning approach for solar power forecasting. Environmental Challenges **14**, 100838 (2024)

[101] Shahid, A., Afzal, M.T., Abdar, M., Basiri, M.E., Zhou, X., Yen, N.Y., Chang, J.: Insights into relevant knowledge extraction techniques: a comprehensive review. J. Supercomput. **76**(3), 1695–1733 (2020). https://doi.org/10.1007/s11227-019-03009-y, `https://doi.org/10.1007/s11227-019-03009-y`

[102] Siino, M., Falco, M., Croce, D., Rosso, P.: Exploring llms applications in law: A literature review on current legal nlp approaches. IEEE Access (2025)

[103] Singh, P., Mahor, V., Lakshmaiya, N., Shanker, K., Kaliappan, S., Muthukannan, M., Rajendran, G.: Prediction of groundwater contamination in an open landfill area using a novel hybrid clustering-based ai model. Environment Protection Engineering **50**(1) (2024)

[104] Stefanidis, K., Tsatsou, D., Konstantinidis, D., Gymnopoulos, L., Daras, P., Wilson-Barnes, S., Hart, K., Cornelissen, V., Decorte, E., Lalama, E., et al.: Protein ai advisor: a knowledge-based recommendation framework using expert-validated meals for healthy diets. Nutrients **14**(20), 4435 (2022)

[105] Stephen McMinn, H.: Library support of bibliographic management tools: a review. Reference services review **39**(2), 278–302 (2011)

[106] Sun, Z.: A short survey of viewing large language models in legal aspect. arXiv preprint arXiv:2303.09136 (2023)

[107] Tang, J.: Aminer: Toward understanding big scholar data. In: Proceedings of the ninth ACM international conference on web search and data mining. pp. 467–467 (2016)

[108] Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., Bashlykov, N., Batra, S., Bhargava, P., Bhosale, S., et al.: Llama 2: Open foundation and fine-tuned chat models. arXiv preprint arXiv:2307.09288 (2023)

[109] Trueger, N.S., Thoma, B., Hsu, C.H., Sullivan, D., Peters, L., Lin, M.: The altmetric score: a new measure for article-level dissemination and impact. Annals of emergency medicine **66**(5), 549–553 (2015)

[110] UNICEF, et al.: Who. unicef-who-the world bank: Joint child malnutrition estimates (jme)—levels and trends–2023 edition (2023)

[111] Verma, S., Bhatia, R., Harit, S., Batish, S.: Scholarly knowledge graphs through structuring scholarly communication: a review. Complex & intelligent systems **9**(1), 1059–1095 (2023)

[112] Visentin, F., Cremasco, S., Sozzi, M., Signorini, L., Signorini, M., Marinello, F., Muradore, R.: A mixed-autonomous robotic platform for intra-row and inter-row weed removal for precision agriculture. Computers and Electronics in Agriculture **214**, 108270 (2023)

[113] VIZCONDE, E., CAGUIAT, M.R., et al.: Generative artificial intelligence in education: Evaluating students' self-efficacy and utilization in their homework. In: International Conference on Computers in Education (2024)

[114] Wade, A.D.: The semantic scholar academic graph (s2ag). Companion Proceedings of the Web Conference 2022 (2022), `https://api.semanticscholar.org/CorpusID:251597885`

[115] Walter, Johannes: Human oversight done right : the AI Act should use humans to monitor AI only when effective. https://hdl.handle.net/10419/271285 (1 2023), `https://hdl.handle.net/10419/271285`

[116] Wang, K., Shen, Z., Huang, C., Wu, C.H., Dong, Y., Kanakia, A.: Microsoft Academic Graph: When experts are not enough. Quantitative Science Studies **1**(1), 396–413 (02 2020). https://doi.org/10.1162/qss_a_00021

[117] Whitfield, S., Hofmann, M.A.: Elicit: Ai literature review research assistant. Public Services Quarterly **19**(3), 201–207 (2023)

[118] Xin, L.: Conference-acceptance-rate. https://github.com/lixin4ever/Conference-Acceptance-Rate (2025)

[119] Yao, J.Y., Ning, K.P., Liu, Z.H., Ning, M.N., Liu, Y.Y., Yuan, L.: Llm lies: Hallucinations are not bugs, but features as adversarial examples. arXiv preprint arXiv:2310.01469 (2023)

[120] Yi, S., Liu, Y., Sun, Z., Cong, T., He, X., Song, J., Xu, K., Li, Q.: Jailbreak attacks and defenses against large language models: A survey (2024), https://arxiv.org/abs/2407.04295

[121] Yu, S., Vautard, R.: A transfer method to estimate hub-height wind speed from 10 meters wind speed based on machine learning. Renewable and Sustainable Energy Reviews **169**, 112897 (2022)

[122] Zhang, Y., Zhang, W., Yu, J., He, L., Chen, J., He, Y.: Complete and accurate holly fruits counting using yolox object detection. Computers and Electronics in Agriculture **198**, 107062 (2022)

[123] Zhang, Z., Wang, C., Wang, Y., Shi, E., Ma, Y., Zhong, W., Chen, J., Mao, M., Zheng, Z.: Llm hallucinations in practical code generation: Phenomena, mechanism, and mitigation. Proc. ACM Softw. Eng. **2**(ISSTA) (Jun 2025). https://doi.org/10.1145/3728894

[124] Zheng, L., Jia, C., Sun, M., Wu, Z., Yu, C.H., Haj-Ali, A., Wang, Y., Yang, J., Zhuo, D., Sen, K., et al.: Ansor: Generating {High-Performance} tensor programs for deep learning. In: 14th USENIX symposium on operating systems design and implementation (OSDI 20). pp. 863–879 (2020)

[125] Zheng, Z., Peng, X., Yang, T., Shen, C., Li, S., Liu, H., Zhou, Y., Li, T., You, Y.: Open-sora: Democratizing efficient video production for all. arXiv preprint arXiv:2412.20404 (2024)