What Do We Mean by GenAI? A Systematic Mapping of The Evolution, Trends, and Techniques Involved in Generative AI

Francisco José García-Peñalvo, Andrea Vázquez-Ingelmo *

GRIAL Research Group, Computer Science Department, Universidad de Salamanca, (https://ror.org/02f40zc51), Salamanca (Spain)

Received 2 July 2023 | Accepted 20 July 2023 | Published 24 July 2023

Abstract

Artificial Intelligence has become a focal point of interest across various sectors due to its ability to generate creative and realistic outputs. A specific subset, generative artificial intelligence, has seen significant growth, particularly in late 2022. Tools like ChatGPT, Dall-E, or Midjourney have democratized access to Large Language Models, enabling the creation of human-like content. However, the concept 'Generative Artificial Intelligence' lacks a universally accepted definition, leading to potential misunderstandings. While a model that produces any output can be technically seen as generative, the Artificial Intelligent research community often reserves the term for complex models that generate high-quality, human-like material. This paper presents a literature mapping of AI-driven content generation, analyzing 631 solutions published over the last five years to better understand and characterize the Generative Artificial Intelligence landscape. Our findings suggest a dichotomy in the understanding and application of the term "Generative AI". While the broader public often interprets "Generative AI" as AI-driven creation of tangible content, the AI research community mainly discusses generative implementations with an emphasis on the models in use, without explicitly categorizing their work under the term "Generative AI".

Keywords

Artificial Intelligence, Content Generation, Generative AI, Generative Models, Machine Learning, Systematic Literature Mapping.

DOI: 10.9781/ijimai.2023.07.006

I. Introduction

Artificial Intelligence (AI) has evolved as an enthralling topic, attracting the attention of researchers, industry experts, and the general public alike. Its growing popularity may be ascribed to its capacity to produce realistic and creative results and its accessibility, which has far-reaching ramifications in fields such as medicine [1-3], education [4], [5], art [6], [7], music [8], [9], marketing [10], [11], software development [12], [13], among several other areas.

While AI has experienced a surge in popularity in recent years, a particular approach within it has undergone explosive growth during the final months of 2022: the field of generative artificial intelligence or GenAI [14].

The introduction of applications such as ChatGPT[1], Dall-E[2], or Midjourney[3], which make Large Language Models (LLMs) [15], [16] accessible to end-users, has set a milestone in the application of artificial intelligence to content generation, enabling wide audiences to effortlessly engage in the creation of human-like texts, realistic images, and even music [17].

But what do we exactly mean when we refer to GenAI? What types of content were being generated prior to the emergence of commercial tools like ChatGPT? And for what purposes?

Before diving into the complexities of generative AI, it is crucial to understand the precise meaning and scope of this term, as well as taking a closer look at the content generation processes that existed prior to putting these approaches in the hands of consumers. By investigating these factors, we can shed light on the underlying objectives and motivations driving the adoption of generative AI solutions.

Taking a closer look at the terminology, the word “generative” is defined as “(being) able to produce or create something”. If we apply this definition to AI, every model can be technically considered as generative, as they always “produce or create something”, whether in the form of numerical predictions or internal rules. However, not every content generation driven by AI is, or has been, considered as Generative AI.

In fact, the term ‘Generative AI’ has been applied more precisely to models that may produce new, previously unseen information dependent on the data on which they were trained. These models are developing fresh, human-like material that can be engaged with and consumed, rather than just numerical forecasts or internal rules.

[1] https://chat.openai.com/

[2] https://openai.com/dall-e-2

[3] https://www.midjourney.com/app/

* Corresponding author.

E-mail addresses: fgarcia@usal.es (F. J. García-Peñalvo), andreavazquez@usal.es (A. Vázquez-Ingelmo).

The lack of a globally agreed definition of 'Generative AI' can result in misunderstanding and misinterpretation. For example, as mentioned before, some may claim that a simple decision tree model that creates rules based on incoming data is a type of Generative AI.

However, the AI research community reserves the term 'generative' for more complex models that can create high-quality, human-like material, unlike discriminative models (such as decision tree models), which are trained to predict probabilities of labels given observations [18]. Some examples of the so-called generative models [19] are Generative Adversarial Networks (GANs) or Variational Autoencoders (VAE), among several others (Fig. 1).

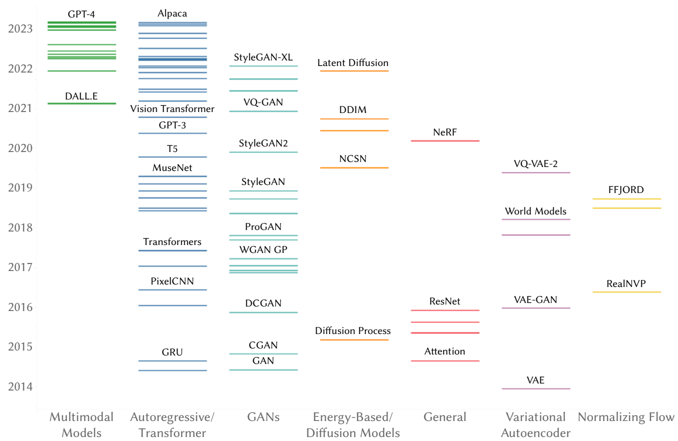


Fig. 1. Timeline of generative models by type. Elaborated by the authors. Source data from [20]. High-resolution version available at https://doi.org/10.5281/zenodo.8165255.

But although generative models have already been differentiated from discriminative models by their internal processes and the probabilities they estimate [19], is Generative AI restricted to the use of these kinds of models? Is the underlying model characteristics crucial in affirming that content has been generated through Generative AI? Or do the nature of the content and the ultimate objective hold greater significance?

Without a clear definition, researchers may use the term 'Generative AI' to refer to various methods to generate content, leading to confusion and misunderstanding. This can stymie research on this subject, making comparing and contrasting various results difficult.

Furthermore, this ambiguity might make it difficult for industry experts and the general public to comprehend what Generative AI is and what it can achieve, leading to unrealistic expectations or fears about its capabilities, which can affect the adoption and acceptance of this technology.

This work aims at providing an analysis of AI-driven solutions for content generation to further define and delimit the meaning and use cases of Generative AI. This analysis has been carried out through a systematic approach, specifically a systematic literature mapping [21]. A systematic mapping provides a structured framework to thoroughly evaluate existing research in the rapidly evolving field of Generative AI, as well as to identify patterns, and spot potential gaps. It focuses on the extent of the subject rather than its depth, which is crucial in emerging and fast-paced domains. This process aids in establishing clear definitions and boundaries within the field, reducing ambiguity, and fostering a consistent discourse.

The rest of this paper is organized as follows. Section II describes the methodology followed to conduct the systematic mapping, while Section III details the review process. Section IV presents the results obtained from the previous steps. Finally, sections V and VI discuss the results and provide a summary of our main findings.

II. Review Planning

This study adheres to the systematic literature review guidelines established by Kitchenham and Charters [22] and the mapping study guidelines set out by Petersen [23], [24]. Specifically, the process is structured around three core stages: planning, conducting, and reporting the findings.

The initial phase involves establishing the primary goal of the review, followed by its development. The main objective of this review is to collect and analyze the existing studies related to the application of AI in content generation, considering the following dimensions: the generated content, the objective of the content generation, the type of models employed and the application domains.

Once the objective has been defined, it is necessary to complete the next two phases, planning and conducting. In these, we define a set of mapping questions (MQs) that will help characterize Generative AI, the inclusion/exclusion criteria, and the search strategy.

A. Mapping Questions

We defined five mapping questions that characterize the AI-driven content generation landscape.

- **MQ1.** How many studies have been published over the years?
- **MQ2.** Who are the most active authors in the area of AI-driven content generation?
- **MQ3.** Which kinds of algorithms and techniques are employed to develop AI-driven content generation applications?
- **MQ4.** Which domains are applying AI-driven content generation to support their studies?
- **MQ5.** What kind of applications were published before and afterwards the popularization of ChatGPT?

These mapping questions are also set to answer the following research question by analyzing the results from the data extraction: ***RQ. What do researchers understand by Generative Artificial Intelligence?***

B. Inclusion and Exclusion Criteria

To discard irrelevant works (in terms of the scope of this paper) from the search results, a set of inclusion criteria (IC) and a set of exclusion criteria (EC) are defined, being the **inclusion criteria** as follows:

- **IC1.** The paper's main objective is the application of content generation (data, images, text, sound, etc.) through artificial intelligence **AND**
- **IC2.** The artificial intelligence solution technical details are identified and described **AND**
- **IC3.** The field in which the solution was applied is identified and described **AND**
- **IC4.** The paper is not a review, survey, or comparative analysis **AND**
- **IC5.** The paper is written in English **AND**
- **IC6.** The paper is published in peer-reviewed Journals, Books, or Conferences **AND**
- **IC7.** The paper is accessible.

The following items refer to the **exclusion criteria** applied:

- **EC1.** The paper's main objective is not the application of content generation (data, images, text, sound, models, etc.) through artificial intelligence **OR**
- **EC2.** The artificial intelligence solution technical details are not identified nor described **OR**

- **EC3.** The paper is a review, survey, or comparative analysis **OR**
- **EC4.** The field in which the solution was applied is not identified nor described **OR**
- **EC5.** The paper is not written in English **OR**
- **EC6.** The paper is not published in peer-reviewed Journals, Books, or Conferences **OR**
- **EC7.** The paper is not accessible.

These criteria aim at discarding works that are not focusing on generating content through AI. In this sense, we reject studies to benchmark different models, reviews, and works that do not generate tangible content. Following the discriminative and generative models' distinction [18], [19], we want to analyze solutions that generate new data instances in a non-deterministic manner, excluding the outcomes from forecasting, labelling, or classification approaches.

C. Search Strategy

The first step to extracting relevant works for the purpose of this paper is the selection of electronic databases. In this case, two electronic databases are selected: Scopus and Web of Science (WoS). These databases are chosen according to a set of requirements:

- It is a reference database in the research scope.
- It is a relevant database in the research context of this mapping study.
- It allows using similar search strings to the rest of the selected databases and Boolean operators.

An initial search using only the term "generative artificial intelligence" was carried out in these databases. However, this preliminary search yielded a small set of results focused on surveys, editorials, or discussions about the applicability of Generative AI approaches in different domains, such as education.

Given that significant AI-driven content generation applications were not retrieved through this initial search, it was necessary to identify which concepts, approaches or tools are widely associated with Generative AI, to finally collect research literature about these approaches.

Due to the increased accessibility of generative models to consumers after the release of OpenAI's ChatGPT, we analyzed search trends related to "Generative AI" in Google Trends. In this case, we observed that most of the related searches included commercial tools such as "ChatGPT", "Dall-E", or "Midjourney".

Given this trend, we decided to perform another preliminary search, including wildcards to enclose derivations, and the NEAR operator to retrieve works where the terms joined by this operator are separated by an interval of explicitly specified words.

This operator is very handy in this context because we are focused on generative processes driven by AI, so the term related to generation must be near AI-related terms (such as deep learning, machine learning, and so on). This new search was structured as follows:

(("machine learning" OR "deep learning" OR "artificial intelligence" OR "AI" OR "AI-" OR "DL" OR "DL-" OR "ML" OR "ML-") NEAR/0 ("generat")) OR (("ChatGPT" OR "Midjourney" OR "Dall-E" OR "DallE" OR "StableDiffusion" OR "Stable Diffusion") NEAR/1 ("generat*"))*

However, including specific tools would bias the results, as it is nearly impossible to include every released generative AI tool to date. On the other hand, executing a search with the first part of the search string only (*("machine learning" OR "deep learning" OR "artificial intelligence" OR "AI" OR "AI-" OR "DL" OR "DL-" OR "ML" OR "ML-") NEAR/0 ("generat*")*) collected a great set of works, but included an unmanageable quantity of noise, including several articles that were not related to AI-driven content generation.

To overcome these issues, we decided to define further the terminology. As mentioned in section II.B, we want to analyze solutions that generate new data instances in a non-deterministic manner, so we opted to focus on works that explicitly generated content through AI, including images, text, video, audio, sound, etc. We also included terms related to transformations between different types of content, such as text-to-image transformations (e.g., Midjourney and Dall-E).

Once every concept was identified, the specific query strings for each chosen database were specified using their query syntax.

1. Web of Science

TS=(("image generation" OR "text generation" OR "video generation" OR "audio generation" OR "sound generation" OR "3D generation" OR "content generation" OR "code generation" OR "dataset generation" OR "data generation" OR "text to text" OR "text-to-text" OR "text to image" OR "text-to-image" OR "text to audio" OR "text-to-audio" OR "text to video" OR "text-to-video" OR "text to code" OR "text-to-code" OR "text to 3D" OR "text-to-3D" OR "audio to text" OR "audio-to-text") AND ("artificial intelligence" OR AI OR "deep learning" OR "machine learning"))

2. Scopus

TITLE-ABS-KEY(("image generation" OR "text generation" OR "video generation" OR "audio generation" OR "sound generation" OR "3D generation" OR "content generation" OR "code generation" OR "dataset generation" OR "data generation" OR "texttotext" OR "texttoimage" OR "texttoaudio" OR "texttovideo" OR "texttocode" OR "textto3D" OR "audiototext") AND ("artificial intelligence" OR AI OR "deep learning" OR "machine learning"))

Additionally, we limited the search to **journal articles published over the last 5 years** to analyze recent, established, and complete research works. Using this approach, we collected the final set of articles to analyze and outline the landscape of AI-driven generative solutions.

III. Review Process

The data-gathering process to conduct the present Systematic Literature Mapping has been divided into different phases in which various activities are carried out. The PRISMA 2020 [25], [26] guidelines were followed for data extraction.

Once the search was performed (on May 17th, 2023), the paper selection process was carried out through the following steps:

1. The raw results (i.e., the records obtained from each selected database) were gathered in a GIT repository[4] and arranged into a spreadsheet. A total of 3295 papers were retrieved: 1835 from Scopus and 1460 from Web of Science.
2. After organizing the records, duplicate works were removed. Specifically, 1332 records were removed, retaining 1963 works (59.58% of the raw records) for the next phase.
3. The maintained papers were analyzed by reading their titles, abstracts, and keywords and by applying the inclusion and exclusion criteria. A total of 1332 papers were discarded as they did not meet the criteria, retaining 631 papers (32.14% of the unique papers retrieved) for the next phase.
4. The selected 631 papers were finally characterized following the mapping questions. For each paper, the following information was collected:
   a. Content being generated by the AI technique (text, images, code, etc.)
   b. AI model type employed (transformers, generative adversarial networks, etc.)

[4] https://github.com/AndVazquez/slm-gen-ai

c. Objective of the AI content generation (data augmentation, image enhancement, text summarization, text translation, style transfer, etc.)
d. Domain of application of the AI-driven content generation

Fig. 2 shows the PRISMA 2020 [25], [26] flow diagram detailing the data extraction process. The dataset containing the works collected in every phase, along with the 631 selected and characterized works, is available at https://doi.org/10.5281/zenodo.8162484 [27].

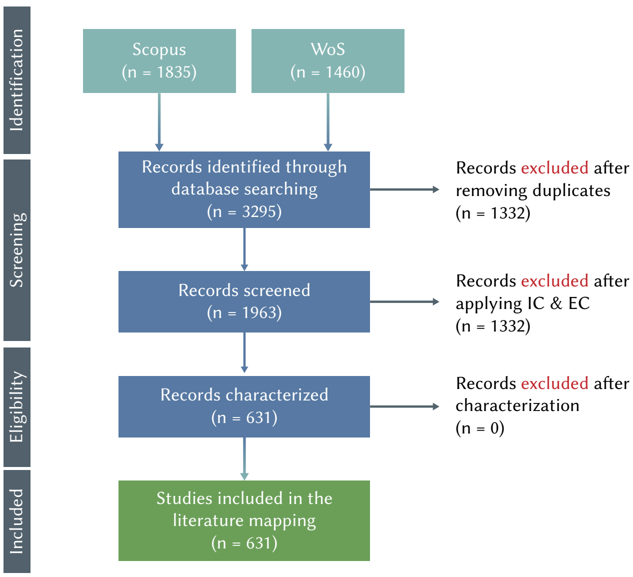


Fig. 2. PRISMA 2020 flow diagram of the literature mapping. High-resolution version available at https://doi.org/10.5281/zenodo.8167557.

IV. Results

The following results have been obtained from the analysis of the obtained records. For a comprehensive review of the records, including title, authors, abstract, and characterization, please refer to the "Characterization" sheet of the provided dataset: https://doi.org/10.5281/zenodo.8162484 [27].

A. How Many Studies Have Been Published Over the Years?

The first mapping question aims at inspecting the temporal landscape of AI-driven content generation. Over the last five years, we can clearly see an increase in the number of works published.

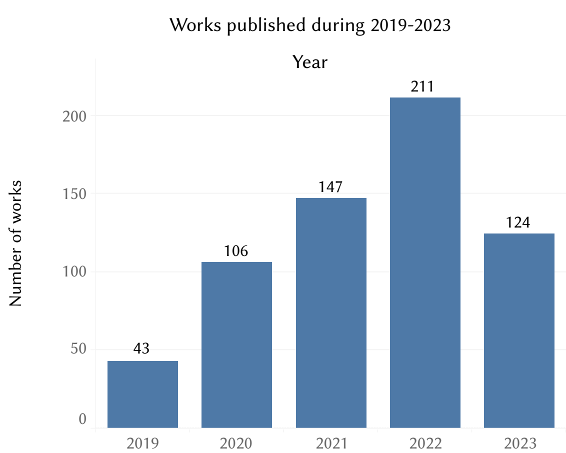


Fig. 3. Number of works published over the last five years. High-resolution version available at https://doi.org/10.5281/zenodo.8167574.

It seems that this trend will continue throughout 2023, with 124 articles already published in the first 5 months of the year (Fig. 3).

B. Who Are the Most Active Authors in the Area of AI-driven Content Generation?

We performed an analysis and normalization of the authors involved in the 631 articles retrieved. This analysis allows us to identify influential authors that likely guide the research direction.

TABLE I. Most Prolific Authors

Articles	Authors
3	Yang, Yang; Chen, Peng; Li, Yibin; Yoon, Hyunsoo; Togo, Ren; Pang, Zhiqi; Ogawa, Takahiro; Haseyama, Miki; Li, Wei; Fujita, Hiroshi; Schlaefer, Alexander; Scarselli, Franco; Fabelo, Himar; Andreini, Paolo; Bianchini, Monica
4	Byun, Yung-Cheol

Most authors (2493) have only published one article in the context of this literature mapping, while 169 have published more than one article. Table I displays the most prolific authors from the records retrieved during the data extraction process.

C. Which Kinds of Algorithms and Techniques Are Employed to Develop AI-driven Content Generation Applications?

As introduced, one of the main concepts of Generative AI is using certain models. But are these models limited to generative models? Or are discriminative models being employed to generate content?

Each article's primary AI model or technique was identified to answer this question. Fig. 4 shows a clear tendency to use GANs for content generation, followed by encoder-decoder networks (such as Autoencoders) and other types of neural networks. We also found several solutions based on Transformers [28], [29].

Finally, the remaining methods have been grouped under the category "others," which include hidden Markov methods, evolutionary algorithms, and Bayesian models.

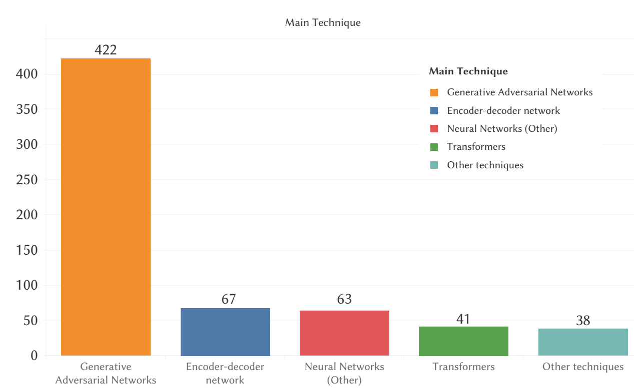


Fig. 4. Number of works grouped by AI technique employed. High-resolution version available at https://doi.org/10.5281/zenodo.8167582.

It is important to clarify that although Transformers are encoder-decoder networks, we have decided to include them in a separate category. This separation allows us to analyze the impact of the release of ChatGPT, which is a transformer-based solution, on the usage of this particular model type.

Fig. 5 illustrates the evolution of the number of works utilizing these models. There is a growing trend in adopting Transformer-based solutions, which could be attributed to the recent popularity of GPT models.

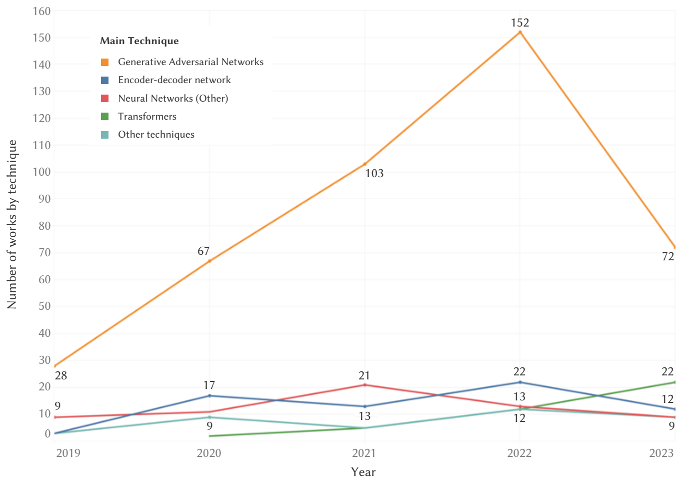


Fig. 5. Number of works grouped by AI technique employed. High-resolution version available at https://doi.org/10.5281/zenodo.8167600.

D. Which Domains Are Applying AI-driven Content Generation to Support their Studies?

Another interesting insight regarding generative AI solutions is identifying which domains or fields of study benefit from generative models' outputs.

The main domain of application was extracted from each article. Fig. 6 shows that the principal domains in which generative AI is being applied are medicine and computer vision. Other domains include natural language processing (NLP), machine learning, remote sensing, art, software/videogames development, and cybersecurity.

But for what purposes is AI-driven content generation being used in each domain? This analysis allows us to better understand how and with what objectives generative artificial intelligence is being used in each field of study.

Fig. 7 breaks down the main content generation tasks by domain. It is possible to see that generative AI (and, specifically, Generative Adversarial Networks) is supporting data augmentation in most domains, but especially in medicine. Sample generation provides a minimally intrusive, fast, and effective method to augment or balance

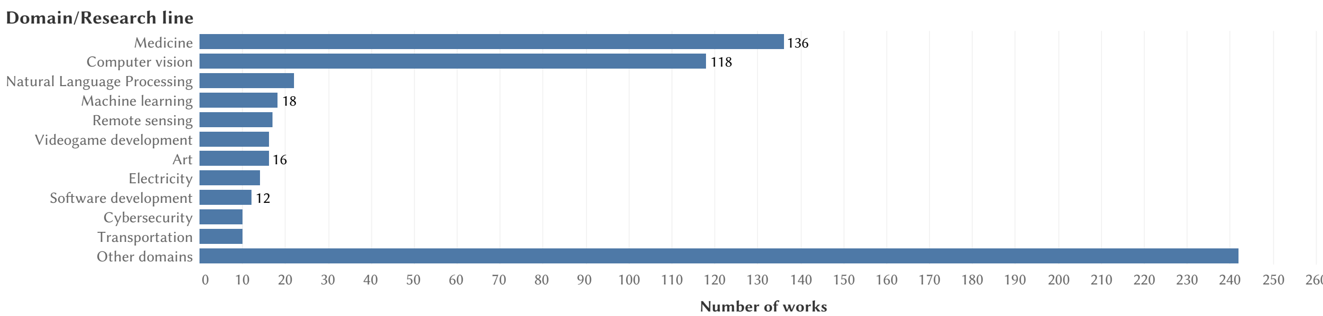


Fig. 6. Number of works grouped by domain of application. High-resolution version available at https://doi.org/10.5281/zenodo.8167627.

Fig. 7. Number of works grouped by objective, domain and AI technique employed. High-resolution version available at https://doi.org/10.5281/zenodo.8167632.

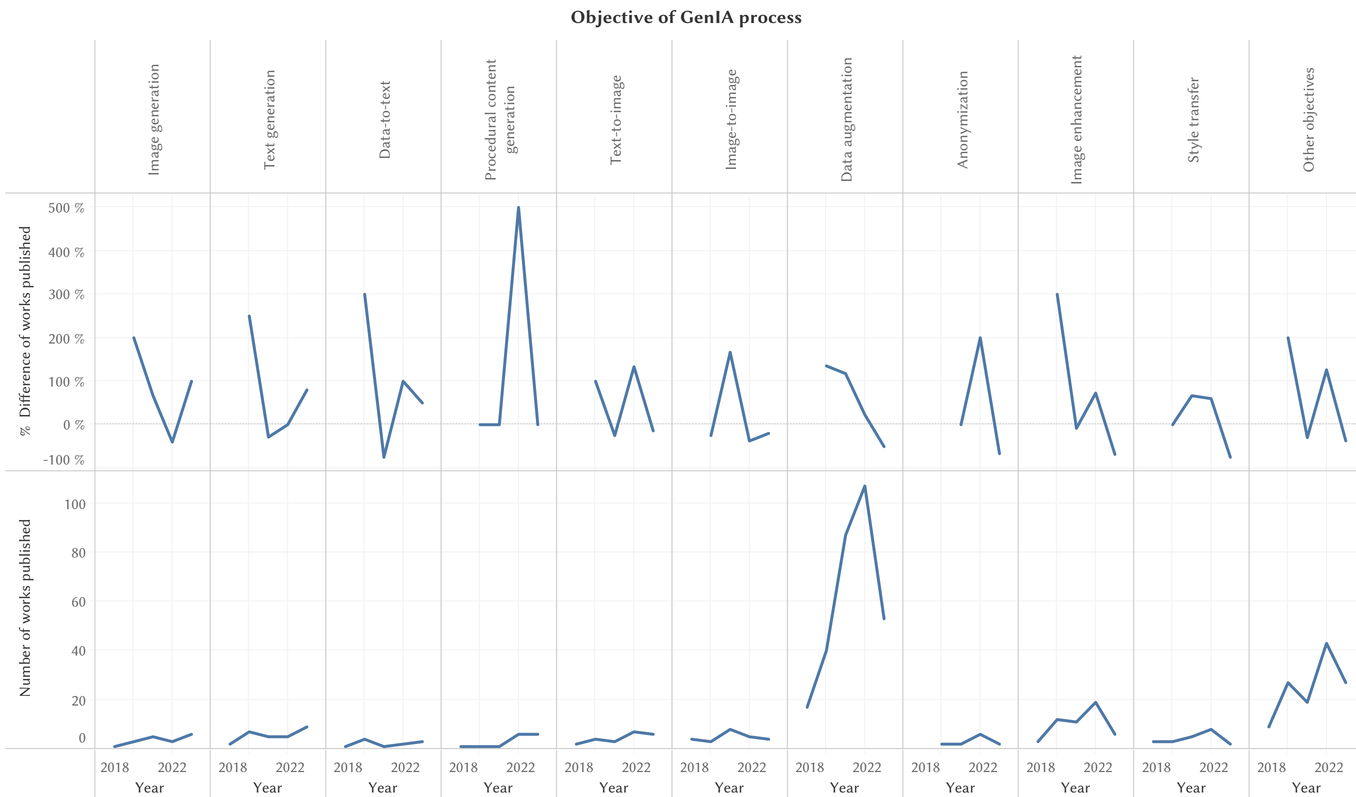


Fig. 8. Difference and number of works published over the years grouped by the objective of the generative process. High-resolution version available at https://doi.org/10.5281/zenodo.8167647.

datasets that will subsequently be used to train or fine-tune complex models in each area (e.g., with the aim of diagnosing or segmenting images in the case of medicine).

Another interesting objective of AI-driven content generation is anonymization. By generating new samples, AI can preserve both crucial information and privacy. On the other hand, we also found several objectives related to image processing, including image enhancement (increasing the image resolution, completing missing or damaged parts, or even colorization), style transfer, and text-to-image translations.

Finally, we can also find specific objectives such as procedural content generation in the field of videogame development, and text generation (mostly related to the NLP area).

E. What Kind of Applications Were Published Before and Afterwards the Popularization of ChatGPT?

This question is focused on analyzing the trends in generative AI after the release of ChatGPT[5], which has significantly reshaped the landscape of AI-driven communication [30], [31]. We have computed the trend of the top tasks supported by AI and compared it over the last five years (Fig. 8).

Although our analysis has covered only the first five months of 2023, we observe a growing trend in text and image generation tasks. This may be influenced by the release of commercial tools for these tasks (ChatGPT, Bing chat[6], Midjourney, Dall-E, etc.), although it is too early to draw robust conclusions about this.

Additionally, we examined the number of works published annually, segmented by the type of content being generated (such as images, data, text, etc.). It can be observed that images and data (referring to tabular, geospatial, time-series, or network data) are the types of resources most frequently generated (see Fig. 9).

However, the same trend observed in Fig. 8 is evident for text content. The number of works focused on generating text (spanning activities like human-like text generation, summarization, translation, etc.) has been increasing over the past two years, unlike most other types of generated content analyzed.

Fig. 10 summarizes the main findings of this literature mapping. We can see that Generative Adversarial Networks (GANs) are the preferred generative model in several domains for a wide range of tasks, especially for data augmentation and image-related tasks.

V. Discussion

A. What Do Researchers Understand by Generative Artificial Intelligence (GenAI)?

The concept of "Generative AI" has been in use for several years, but it wasn't until the mid to late 2010s that it gained widespread recognition. This surge in popularity was in sync with the rise and acceptance of generative models like GANs in the AI research community.

These models, pioneered by Ian Goodfellow and his team in 2014 [32], were crucial in bringing the term "generative AI" into the spotlight.

However, the term gained broader recognition beyond the research community recently. By inspecting Google Trends, we can see that users became interested in this concept around November and December 2022 (Fig. 11), which aligns with the release of ChatGPT and other commercial tools.

Thus, although generative models have existed for several years, the term "generative AI" has only gained popularity with the widespread availability of tools aimed at the general public.

[5] Launched on November 30, 2022.

[6] https://www.bing.com/

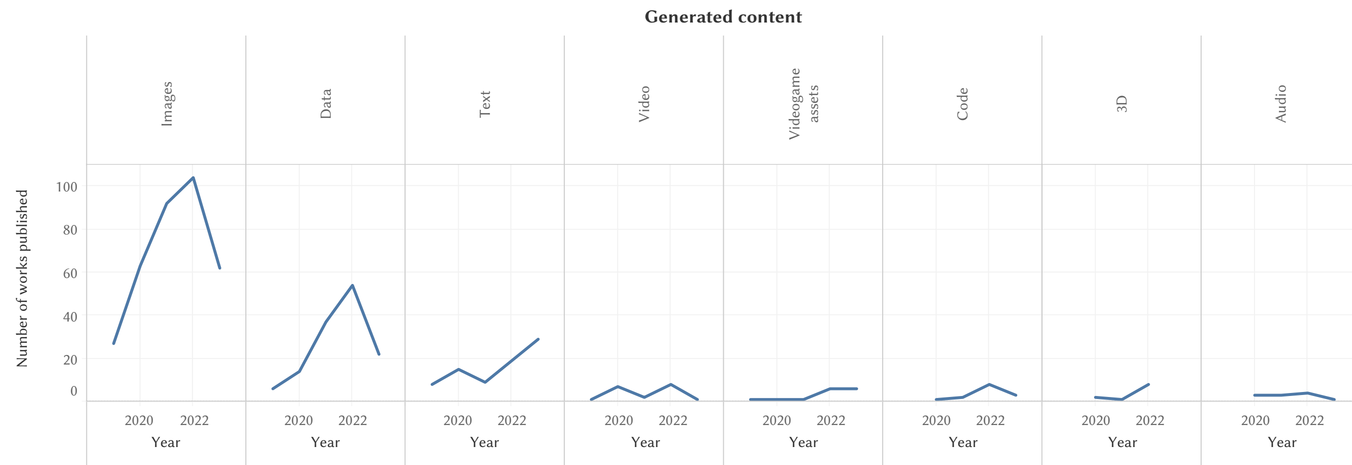


Fig. 9. Number of works published over the years grouped by generated content type. High-resolution version available at https://doi.org/10.5281/zenodo.8167655.

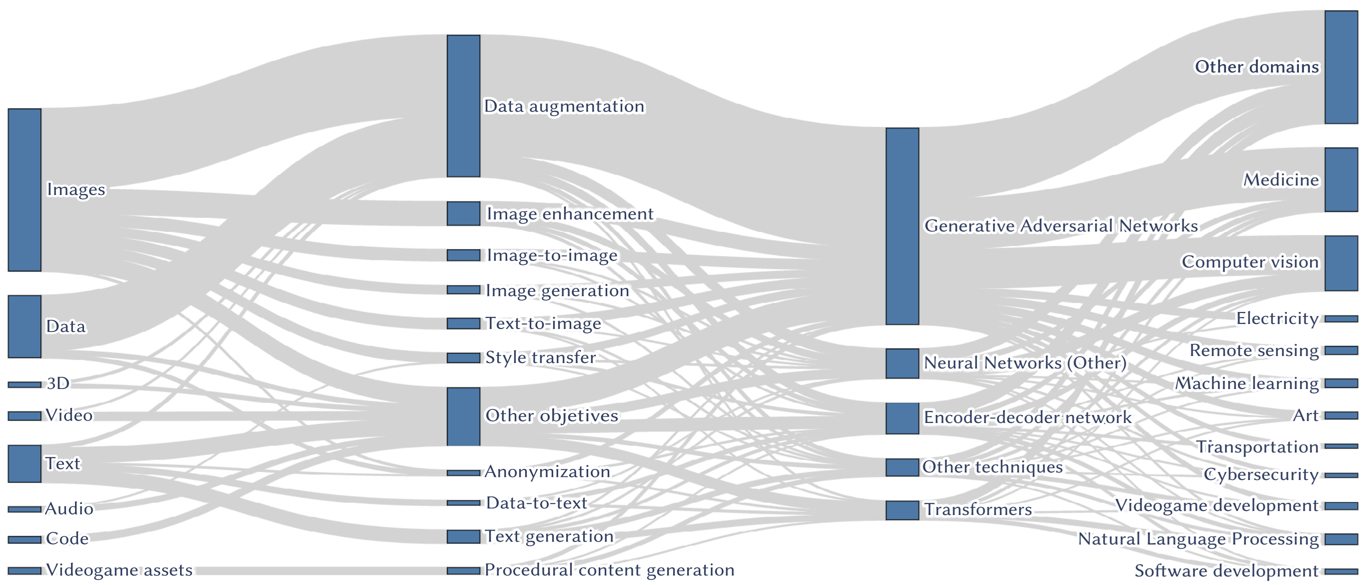


Fig. 10. Relationships among the generated content type, task, AI technique, and application domain in the retrieved works. High-resolution version available at https://doi.org/10.5281/zenodo.8167662.

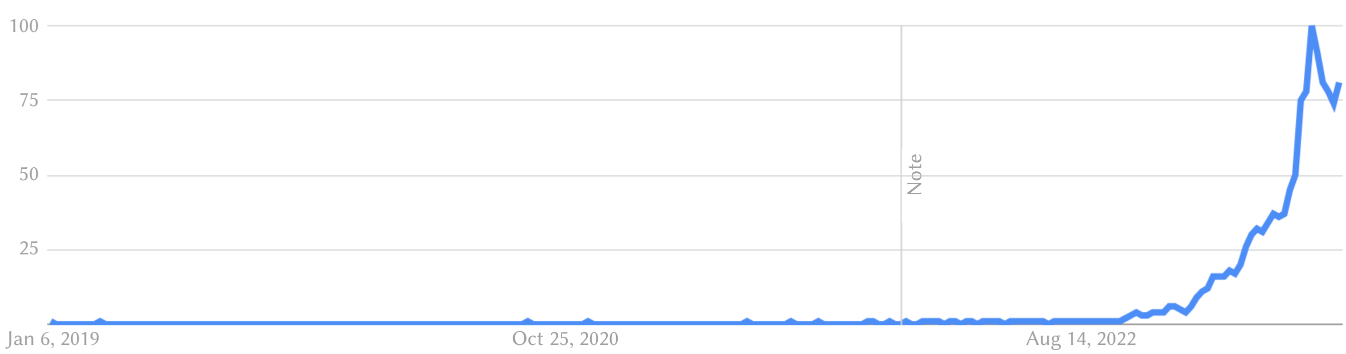


Fig. 11. Worldwide interest over time in the term "Generative AI". Source: Google Trends, https://bit.ly/44IzLmI.

But what does the research community understand by Generative AI? After retrieving and analyzing 631 articles published between January 2019 to May 2023, we obtained a curated set of real-world applications for AI-driven content generation. These solutions included generating a wide variety of resources (images, tabular data, 3D models, videogame assets, etc.) to support different tasks in several domains. What they do have in common is that every solution employed **generative**, not discriminative models, which could drive the definition of the term Generative AI.

If we further analyze the keywords employed by the researchers to refer to their solutions, only 1 work from 2023 includes the term "Generative AI" (record no. 615 from the "Characterization" sheet of the mapping dataset, https://doi.org/10.5281/zenodo.8162484 [27]).

In fact, during the preliminary search that we carried out with only words related to Generative AI, we noted that the works referring to this term were mainly editorials or discussions about the implications of commercial tools such as ChatGPT in different domains, but no actual nor applications on generative models producing actual content as we collected in this literature mapping.

So how did the authors refer to the collected solutions? Fig. 12 presents the most common terms found within the abstracts of the 631 retrieved works. We can observe that generation-related terms (generative, generate, generated, etc.) and the generated content (images or data) are very common.

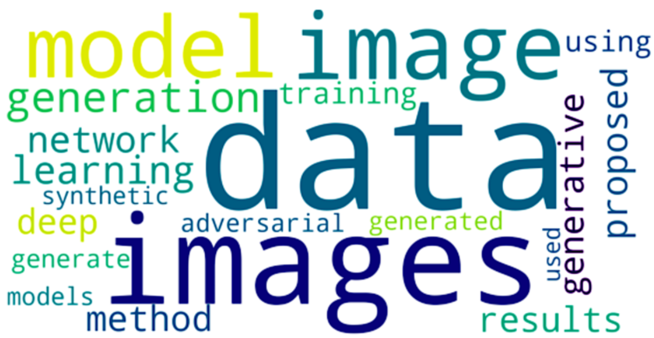


Fig. 12. Most common terms found within the abstracts. High-resolution version available at https://doi.org/10.5281/zenodo.8167676.

We also analyzed the most common bigrams within the abstracts to gain more insights into this terminology. Fig. 13 shows how the term "generative" (one of the most common terms found in Fig. 12) was mostly used to refer to generative adversarial networks, the model employed in most works (Fig. 4).

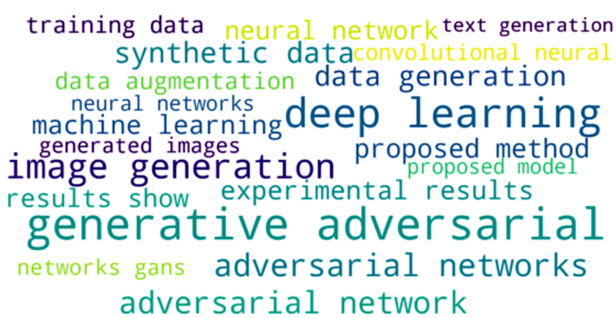


Fig. 13. Most common bigrams found within the abstracts. High-resolution version available at https://doi.org/10.5281/zenodo.8167693.

Based on these findings, we can conclude that the general public commonly uses the term "Generative AI" to refer to the creation of tangible content (such as images, text, code, models, audio, etc.) via AI-powered tools. However, the AI research community primarily discusses generative applications focusing on the models used, without explicitly categorizing their work under the term "Generative AI".

To sum up, following the insights reached through this literature mapping, we can define "Generative AI" as the **production of previously unseen synthetic content, in any form and to support any task, through generative modeling**[7].

VI. Conclusions

This work presents the results of a literature mapping about AI-driven content generation. A total of 1963 unique works related to this topic were analyzed, obtaining 631 categorized articles to better understand the landscape of generative AI solutions. The entire process and final characterization can be reviewed in the provided dataset at https://doi.org/10.5281/zenodo.8162484 [27].

We found a clear trend in using specific models, such as GANs or encoder-decoder networks, to generate various resources, especially images and tabular data. These solutions have been mostly applied to augment datasets and enhance subsequent models' predictions.

Although preliminary, it is possible to see how the release of commercial solutions is shifting the landscape of generative solutions, with a slight increase of solutions focused on text generation in the first months of 2023, but also with the advent of new ethical issues and dilemmas, as the widespread accessibility of AI-driven content generation tools has triggered a deep polarization of society regarding Generative AI.

Some individuals are optimistic, envisaging a plethora of opportunities, while others predict dystopian ramifications.

Considering, for example, the domain of education, we can see through the obtained results that Generative AI was marginally applied within this field compared to other areas, such as medicine. However, introducing these tools in education is triggering several concerns among educators, parents, and policymakers [33].

The potential of AI to transform pedagogical methods, assessment systems, and learning experiences opens a new frontier for education. On the one hand, the scalability and personalization offered by AI can improve educational processes by providing a more differentiated and inclusive learning environment.

But just as opportunities and great potential benefits have emerged, there have also grown significant concerns. Some ethical concerns educators raise include assessment, academic integrity [34], and data privacy [35], among others.

The software development field presents a similar narrative. While AI may speed up development processes, automate routine tasks, and significantly reduce debugging time, critics express apprehension about potential job losses and the ethical implications associated with accountability and transparency in AI-generated code [36]. In fact, these powerful applications of AI-driven code generation are also influencing computer science education.

Traditional programming approaches, which frequently rely on manual code writing, debugging, and learning the complexities of programming languages, might be replaced by teaching students how to interface with and manage AI-driven development tools. This change might result in a more efficient learning process, allowing students to handle more complex problems early in their education [37].

However, all these concerns and acceptance issues of Generative AI could be alleviated through a more comprehensive understanding of what precisely Generative AI entails. As introduced, we can technically refer to "Generative AI" as any process of producing any content by any AI technique. But the nuances obtained through this literature mapping offer a deeper view into this term.

By including the technique employed (generative modelling) in the definition of Generative AI, we focus on the process of generating new content from existing resources rather than on the generated content.

It is crucial to understand that Generative AI is not some form of arcane magic but the procedure of training a model with input data. Its capacity to generate original content is based on learning patterns within the available data and then creating outputs that represent these patterns in new ways [18].

By demystifying Generative AI, it is possible to tackle its acceptance issues and address its potential challenges more pragmatically and

[7] Understood as modeling the joint distribution of inputs and outputs.

effectively. Understanding Generative AI as a data-driven tool rather than an omnipotent solution helps set realistic expectations of what it can accomplish. This viewpoint can facilitate us to successfully integrate AI into different domains without expecting utopian results, minimizing the disappointment that may occur when AI does not perform as anticipated.

Acknowledgement

This research was partially funded by the Ministry of Science and Innovation through the AvisSA project grant number (PID2020-118345RB-I00).

References

[1] C. Zhang, B. Vinodhini, and B. A. Muthu, "Deep Learning Assisted Medical Insurance Data Analytics With Multimedia System," *International Journal of Interactive Multimedia and Artificial Intelligence,* vol. 8, no. 2, pp. 69-80, 2023.

[2] F. García-Peñalvo *et al.*, "KoopaML: a graphical platform for building machine learning pipelines adapted to health professionals," *International Journal of Interactive Multimedia and Artificial Intelligence,* vol. In Press, 2023.

[3] J. Zhou, T. Li, S. J. Fong, N. Dey, and R. González-Crespo, "Exploring ChatGPT's Potential for Consultation, Recommendations and Report Diagnosis: Gastric Cancer and Gastroscopy Reports' Case," *International Journal of Interactive Multimedia and Artificial Intelligence,* vol. 8, no. 2, pp. 7-13, 2023.

[4] J.-M. Flores-Vivar and F.-J. García-Peñalvo, "Reflections on the ethics, potential, and challenges of artificial intelligence in the framework of quality education (SDG4)," *Comunicar,* vol. 31, no. 74, pp. 37-47, 2023.

[5] H. Khosravi *et al.*, "Explainable artificial intelligence in education," *Computers and Education: Artificial Intelligence,* vol. 3, p. 100074, 2022.

[6] A. Chatterjee, "Art in an age of artificial intelligence," *Frontiers in Psychology,* vol. 13, p. 1024449, 2022.

[7] B. Agüera y Arcas, "Art in the age of machine intelligence," in *Arts*, 2017, vol. 6, no. 4, p. 18: MDPI.

[8] P. Álvarez, J. García de Quirós, and S. Baldassarri, "RIADA: A Machine-Learning Based Infrastructure for Recognising the Emotions of Spotify Songs," *International Journal of Interactive Multimedia and Artificial Intelligence,* vol. 8, no. 2, pp. 168-181, 2023.

[9] M. Civit, J. Civit-Masot, F. Cuadrado, and M. J. Escalona, "A systematic review of artificial intelligence-based music generation: Scope, applications, and future trends," *Expert Systems with Applications,* p. 118190, 2022.

[10] J. Lies, "Marketing Intelligence: Boom or Bust of Service Marketing?," *International Journal of Interactive Multimedia and Artificial Intelligence,* vol. 7, no. 7, pp. 115-124, 2022.

[11] P. Mikalef, N. Islam, V. Parida, H. Singh, and N. Altwaijry, "Artificial intelligence (AI) competencies for organizational performance: A B2B marketing capabilities perspective," *Journal of Business Research,* vol. 164, p. 113998, 2023.

[12] R. H. Kulkarni and P. Padmanabham, "Integration of artificial intelligence activities in software development processes and measuring effectiveness of integration," *IET Software,* vol. 11, no. 1, pp. 18-26, 2017.

[13] A. Mashkoor, T. Menzies, A. Egyed, and R. Ramler, "Artificial intelligence and software engineering: Are we ready?," *Computer,* vol. 55, no. 3, pp. 24-28, 2022.

[14] T. van der Zant, M. Kouw, and L. Schomaker, "Generative Artificial Intelligence," in *Philosophy and Theory of Artificial Intelligence*, V. C. Müller, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 107-120.

[15] J. Yang *et al.*, "Harnessing the power of llms in practice: A survey on chatgpt and beyond," *arXiv preprint arXiv:2304.13712,* 2023.

[16] W. X. Zhao *et al.*, "A survey of large language models," *arXiv preprint arXiv:2303.18223,* 2023.

[17] F. J. García-Peñalvo, "The perception of Artificial Intelligence in educational contexts after the launch of ChatGPT: Disruption or Panic?," *Education in the Knowledge Society,* vol. 24, art. e31279, 2023.

[18] D. Foster, "Generative deep learning: teaching machines to paint," *Write, Compose, and Play (Japanese Version) O'Reilly Media Incorporated,* pp. 139-140, 2019.

[19] H. Gm, M. K. Gourisaria, M. Pandey, and S. S. Rautaray, "A comprehensive survey and analysis of generative models in machine learning," *Computer Science Review,* vol. 38, p. 100285, 2020/11/01/ 2020.

[20] D. Foster. (2023). *Generative Deep Learning - 2nd Edition Codebase.* Available: https://github.com/davidADSP/Generative_Deep_Learning_2nd_Edition/tree/main

[21] F. J. García-Peñalvo, "Developing robust state-of-the-art reports: Systematic Literature Reviews," *Education in the Knowledge Society,* vol. 23, p. e28600, 2022.

[22] B. Kitchenham and S. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering. Version 2.3," School of Computer Science and Mathematics, Keele University and Department of Computer Science, University of Durham, Technical Report EBSE-2007-01, 2007, Available: https://goo.gl/L1VHcw.

[23] K. Petersen, S. Vakkalanka, and L. Kuzniarz, "Guidelines for conducting systematic mapping studies in software engineering: An update," *Information and software technology,* vol. 64, pp. 1-18, 2015.

[24] K. Petersen, R. Feldt, S. Mujtaba, and M. Mattsson, "Systematic mapping studies in software engineering," in *12th International Conference on Evaluation and Assessment in Software Engineering (EASE) 12*, 2008, pp. 1-10.

[25] M. J. Page *et al.*, "PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews," *BMJ,* vol. 372, p. n160, 2021.

[26] M. J. Page *et al.*, "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," *BMJ,* vol. 372, p. n71, 2021.

[27] A. Vázquez-Ingelmo and F. J. García-Peñalvo, "Dataset for the mapping study 'What do we mean by GenAI?'," (1.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.8162484, 2023.

[28] R. Gruetzemacher and D. Paradice, "Deep transfer learning & beyond: Transformer language models in information systems research," *ACM Computing Surveys (CSUR),* vol. 54, no. 10s, pp. 1-35, 2022.

[29] A. Vaswani *et al.*, "Attention is all you need," in *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems*, Long Beach, CA, USA, 2017, vol. 30, pp. 5998-6008.

[30] A. S. George and A. H. George, "A review of ChatGPT AI's impact on several business sectors," *Partners Universal International Innovation Journal,* vol. 1, no. 1, pp. 9-23, 2023.

[31] A. Bozkurt *et al.*, "Speculative Futures on ChatGPT and Generative Artificial Intelligence (AI): A collective reflection from the educational landscape," *Asian Journal of Distance Education,* p. (Published), 2023.

[32] I. Goodfellow *et al.*, "Generative adversarial nets," *Advances in neural information processing systems,* vol. 27, 2014.

[33] F. J. García-Peñalvo, F. Llorens-Largo, and J. Vidal, "The new reality of education in the face of advances in generative artificial intelligence," *RIED: Revista Iberoamericana de Educación a Distancia,* vol. 27, 2023.

[34] W. M. Lim, A. Gunasekara, J. L. Pallant, J. I. Pallant, and E. Pechenkina, "Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators," *The International Journal of Management Education,* vol. 21, no. 2, p. 100790, 2023.

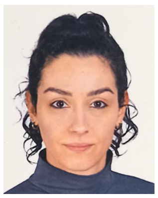
[35] L. Tredinnick and C. Laybats, "The dangers of generative artificial intelligence," ed: SAGE Publications Sage UK: London, England, 2023, p. 02663821231183756.

[36] B. Meyer. (2022). *What Do ChatGPT and AI-based Automatic Program Generation Mean for the Future of Software.* Available: https://bit.ly/3LyAJLj

[37] O. Hazzan. (2023). *ChatGPT in Computer Science Education.* Available: http://bit.ly/3WYTxpv

Francisco José García-Peñalvo

He received the degrees in computing from the University of Salamanca and the University of Valladolid, and a Ph.D. from the University of Salamanca (USAL). He is Full Professor of the Computer Science Department at the University of Salamanca. In addition, he is a Distinguished Professor of the School of Humanities and Education of the Tecnológico de Monterrey, Mexico. Since 2006 he is the head of the GRIAL Research Group GRIAL. He is head of the Consolidated Research Unit of the Junta de Castilla y León (UIC 81). He was Vice-dean of Innovation and New Technologies of the Faculty of Sciences of the USAL between 2004 and 2007 and Vice-Chancellor of Technological Innovation of this University between 2007 and 2009. He is currently the Coordinator of the PhD Programme in Education in the Knowledge Society at USAL. He is a member of IEEE (Education Society and Computer Society) and ACM.



Andrea Vázquez-Ingelmo

Andrea Vázquez-Ingelmo received her bachelor's degree in Computer Science from the University of Salamanca (USAL), in 2016, her master's degree in Computer Science (USAL) in 2018, and her Ph.D. degree in Computer Science (USAL) in 2022. She is a member of the Research Group of Interaction and eLearning (GRIAL) since 2016. Her area of research is related to human-computer interaction, software engineering, data visualization, and machine learning applications.