AN OVERVIEW OF THE RESEARCH & RESEARCHERS:
Topic Modeling GenAI Ethics

Sirichada Wattanasiritanawong (Tingting)
swattan@cmkl.ac.th

Abstract

With the emergence of new technology, ethical principles must be established to guide research, interest, and investment toward beneficial yet moral purposes. In the space of generative artificial intelligence, a literature review must be done to assess research areas that align with the public's ethical considerations, such as originality and the value of self-expression. This paper will model the highly discussed topics using BERTopic, Top2Vec, and LDA from 427 most relevant research articles about generative AI ethics from Crossref, and visualize the findings of the topic model as a topic co-occurrence graph alongside mapping the references from the 427 articles into a reference network graph to address people's concerns over its creativity, legality, and economic impacts. Ultimately, the results indicate the need of more researchers to investigate these ethical dilemmas and point to possible future research that can be done in domain-specific GenAI applications.

Keywords: Generative AI, AI ethics, BERTopic, Top2Vec, LDA, network graph, co-occurrence graph

---

1. Introduction

It is no secret that generative artificial intelligence (GenAI) such as ChatGPT and Midjourney have surprised the world. Seemingly, it can write Shakespearean-level writing with illustrations to accompany it to full-length essays for a history assignment, all in the blink of an eye. So caught up, we rarely wonder if it is necessary or worth the opportunity cost.

Although, as the world adapts to this technology, several concerns and questions have been raised about this technology. These concerns range from the impact of GenAI on academic work in schools to the deeper implications of what entity owns the generated output, between the GenAI company, the machine learning (ML) model, the user using the model, or the author of what is inside the model.

Artificial intelligence (AI) has many beneficial use cases and GenAI is no different. It can write, it can draw– it can even win art competitions (Roose, 2022). However, it would be immoral to turn a blind eye to the potential harm it can bring as well. For instance, creating fake documents and photographs to support a political agenda or to spread misinformation (Ferrara, 2024). Additionally with the introduction of GenAI, job postings for writing and coding have decreased by 21% while image creation postings have decreased by 17% (Demirci et al., 2024, pg. 1).

Therefore with the emergence of new technology, ethical principles must be established to guide research, interest, and investment toward beneficial yet moral purposes. In the space of generative artificial intelligence, a literature review must be done to assess research areas that align with the public's ethical considerations, such as originality and the value of self-expression. The purpose of this paper is to map out where the most value and research have been done on the ethics of AI that address people's concerns– such as creatives, consumers, and businesses– and to lay a foundation for future research directions in these fields.

1.1 Literature Review

AI is a subdomain of the computer science field focused on developing a technological system capable of demonstrating "intelligence" through making decisions or predictions (Stryker & Kavlakoglu, 2024). GenAI relies on a type of AI called deep learning, or AI that mimics the neuron layers in humans' brains to complete more complex tasks, such as identifying features in an image (Stryker & Scapicchio, 2024). Because of this increased complexity, GenAI has more functions and applications in comparison to traditional machine learning approaches.

The way GenAI can do this is through a process called training. While all AI models go through one or more forms of training, GenAI can adapt the things that it learned from the data set to new situations. For this reason, GenAI is sometimes given the name General Purpose Technology (GPT) to distinguish it from traditional model generations that are known as Weak or Narrow AI (Stryker & Kavlakoglu, 2024). This gives GenAI more leeway to perform a variety of tasks, which has been a popular topic for many. To give a couple of examples, character.ai is a website that allows users to chat with a variety of chatbots based on beloved characters or celebrities, and Canva– an online platform for graphic design– has utilized AI to make video editing more efficient for its users (Hall, 2024).

Depending on the use case, GenAI has several types of training methods. For instance, image GenAIs can be trained from generative adversarial network (GAN) or diffusion models. GAN consists of two models: a generator and a discriminator. The generator works to create images out of noise, and the discriminator determines if the generator's output is real or AI-generated (Weng, 2019). As for the diffusion method, it consists of a single model that goes through forward and backward diffusion. Forward diffusion is when the model adds noise to a real image from its dataset– otherwise known as a sample– while backward diffusion is adding noise to a random noise sample until it resembles a sample (Bergmann & Stryker, 2024).

In recent years, diffusion models have gained more traction over GAN as the more efficient image generator. A reason for this is due to the extensive process and resources needed to train two models at once with GAN, resulting in diffusion models having higher training stability and control over the output quality. Despite this, GAN is still widely used along with diffusion models for its ability to create better images in a realistic style (Peng, 2024).

With the growth of GenAI in various mediums, there has been a trend of AI usage that requires discussions on ethics and related considerations. Related considerations are such as the case of *Théâtre D'opéra Spatial*, a photography submission by Jason M. Allen for the Colorado State Fair's annual fine art competition in 2022 (Roose, 2022). After winning first place in the photomanipulation category, Allen went on to register a copyright request for his submission at the United States Copyright Office on September 21, 2022. Allen explained that his work process consisted of “numerous revisions and text prompts at least 624 times to arrive at the initial version of the image”, and then “after Midjourney produced the initial version of the Work, he used Adobe Photoshop to remove flaws and create new visual content and used Gigapixel AI to ‘upscale’ the image, increasing its resolution and size” (Wilson et al., 2023, p. 2). The court dismissed Allen’s application, deeming that the majority of the work was non-human and authored by Midjourney and Gigapixel AI, therefore the *Théâtre D'opéra Spatial* could not be copyrighted. In short, it is evident that copyright and the question of authorship are pressing topics of discussion, especially within image GenAIs.

Midjourney Image **The Work**

Figure 1. Comparison between the final generated output by Midjourney and the final result after edited by Jason M. Allen (Wilson et al., 2023, p. 6)

Another widely discussed concern is GenAI’s impact on the economy. GenAI’s functionalities extend beyond automation with Weak AI, possibly threatening to displace high-skilled workers too (Cazzaniga et al., 2024). Because of this, adaptability is vital as college-educated and younger workers may find it easier to adjust to this new technology. (Cazzaniga et at., 2024). Additionally in the U.S., strikes have been initiated in the entertainment industry– specifically, the movie and video game industry– against the usage of AI. For instance, the 2023 and 2024 SAG-AFTRA strikes.

The 2023 strike resulted in an updated agreement between The Screen Actors Guild - American Federation of Television and Radio Artists (SAG-AFTRA) and Alliance of Motion Picture and Television Producers (AMPTP) with a section dedicated to the use of AI, such as protecting actors from their likeness being “digitally replicated” without their explicit permission (“SAG-AFTRA TV/ Theatrical Contracts 2023”, 2023), while the 2024 SAG-AFTRA video game strike is still ongoing. Digital replications refer to “a replica of [an actor’s] voice and/or likeness that is created using digital technology, such as [AI]” (“Digital Replicas 101”, n.d.). Similarly, the 2024 strike argues for greater transparency on digital replicas and the protection of creators’ works in games (*We’re Fighting for the Survival of Video Game Performers*, 2024). As a result, these strikes represent not only the effect of GenAI on employment but also shape the future of the entertainment industry due to GenAI’s applications in the workforce.

Initially, GenAIs like large language models (LLMs) and GANs were trained on works scraped from the Internet for the past years but they may be running out of fresh training data by 2026 at the soonest and 2032 at the latest (Villalobos et al., 2022). Instead, some have opted for training on AI-generated training data but results suggest that this can lead to model collapse, which is when models degrade in quality due to their previous model's training (Shumailov et al., 2024). Therefore, it is probable that the growth of GenAI models has peaked with the stagnant quality of new training data. With regard to ethical considerations, this means that billions of works in datasets are left uncredited and are repeatedly transformed without the creator's knowledge. This sparks a necessity for companies to acknowledge and give credit to the authors of the work that is used for the training data, as their expressions of emotions into art are being leveraged for commercial (or related) applications without consent.

On account of these concerns, this literature review will provide a meta-analysis of all of these ethical considerations as they pertain to GenAI in the established scientific literature, seeking to identify the most discussed topics, alongside potential gaps in the ethical discussion in the usage of GenAI as it potentially gains popularity and momentum.

2. Methodology

2.1 Research Design

In order to find the most discussed topics in GenAI ethics, topic modeling algorithms can be used to identify and quantify these topics. Topic model is part of the natural language processing (NLP) field, which is where large language models (LLM) stem from. Other applications of NLPs are text summarization and sentiment analysis, all of which are techniques used within popular LLMs like ChatGPT.

Topic modeling is a text mining technique that summarizes large sets of data into sets of keywords. These sets of keywords then represent the common topics that a cluster or group of data is about (Murel & Kavlakoglu, 2024). A few example algorithms of topic modeling are Word2Vec, BERTopic, and latent semantic analysis (LSA). Depending on the algorithm, the model can apply to large sets of singular words in documents.

2.2 Materials

The topic modeling algorithms used in this paper will be BERTopic, Top2Vec, and Latent Dirichlet Allocation (LDA). Beginning with LDA, it is a classic topic modeling algorithm that was first published in a research paper in 2003. It is a generative probabilistic model and was introduced as an upgrade to latent semantic indexing (LSI) due to its ability to embed more complex models and train a wider variety of data, such as continuous data (Blei et al., 2003). On the other hand, there are the newer models BERTopic and Top2Vec, which were released in 2022 and 2020 respectively.

Top2Vec directly seeks to solve LDA's problem of needing to create custom stop-word lists, stemming or lemmatization to generate its topics by using joint document and word semantic embedding (Angelov, 2020). This has the bonus of the model considering the order and meaning of the words when generating the topics. Similarly, BERTopic can consider the order and meaning of words in its algorithm, since BERTopic is a model built upon the Bidirectional Encoder Representations from Transformers (BERT), a language representation model released by Google (Devlin et al., 2018). The cluster centroid is the center of a topic cluster and it is where Top2Vec generates the keywords for its topics. However, BERTopic argues that this isn't always the case. Therefore, BERTopic can generate its topics based on semantic embedding like Top2Vec but without considering the cluster centroid (Grootendorst, 2022).

In summary, these algorithms are chosen due to their ability to train on datasets with long documents and to compare the results of the topics generated as they excel in different situations.

2.3 Data Collection Method

The dataset that this paper uses is queried from Crossref API. Crossref is a nonprofit organization established in 1999 for researchers to record their publications through a digital object identifier (DOI) in order to maintain consistent reference linking between research papers (Pentz, 2020). One of the services that Crossref offers is a database of published research DOIs and related metadata records that can be accessed through its REST API.

Using this API, the following parameters are used to screen eligible research paper metadata for the journal dataset. This journal dataset will be trained by topic modeling algorithms and result in a co-occurrence graph that represents the conceptual relations, meaning the relation between the topics in this field.

```
params = {
   "filter": "type:journal-article",
   "query.bibliographic": (
       "Generative AI OR artificial intelligence OR GenAI "
       "AND "
       "Ethics OR concerns OR impact OR effect "
       "AND "
       "Art OR Artist OR creator"
   ),
   "sort": "relevance",
   "rows": 1000
}
```

Figure 2. Code block of the parameters used in the Crossref API call

The "filter" is used to screen out other types of publications that are not journal articles and the "query.bibliographic" is for returning bibliographic information, such as the title or ISSN of the article (*Crossref Unified Resource API*, n.d.). The statement after the query refers to the "query.bibliographic" to only return results that have these words in their bibliographic information. These words were chosen because they are different variations of keywords for research discussing the ethics of GenAI. Ethics is generally synonymous with concern, impact, and effect, while generative AI, artificial intelligence, and

GenAI narrow the field of research. Moreover, "art artist creator" are filter words that is added to keep the majority scope of the research in generative visual art. Therefore, the final query is that the API should only return papers with these words: (Generative AI OR artificial intelligence OR GenAI) AND (ethics OR concerns OR impact OR effect) AND (art OR artist OR creator). Lastly, the query will be sorted by most to least relevance based on Crossref's relevance score, which should sort the first 1,000 returned papers to be the most relevant.

```
'DOI': item['DOI'] if 'DOI' in item else 'N/A',
'Title': item['title'][0] if 'title' in item and item['title'] else 'N/A',
'Abstract': item['abstract'] if 'abstract' in item else 'N/A'
```

Figure 3. Code block of metadata extracted from the return to API call

Finally, the metadata that is returned from the API call will be extracted in another code segment for the DOI, title, and abstract, as seen in Figure 3. From here on, the data preprocessing phase will begin. This includes changing variations of AI and GenAI into the same variation and removing papers that are unrelated to the purpose of this paper.

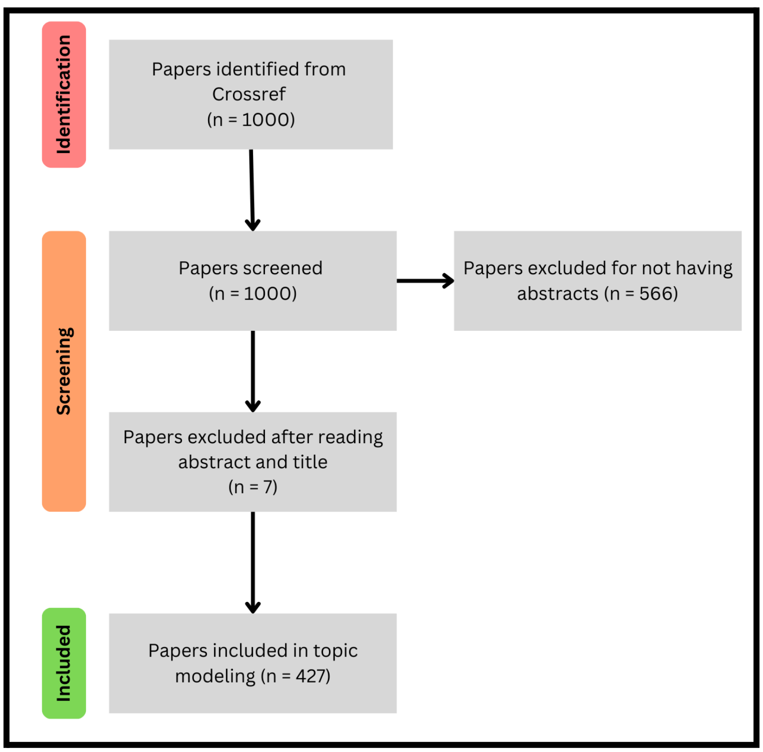

Figure 4. Diagram of data collection and preprocessing pipeline for journal dataset

The second dataset is the reference dataset. Unlike the previous dataset, this dataset will be used to present a network graph showing the links in academia. Using Crossref again, the API was called to find the references of the 427 papers, which resulted in 8,528 references. After preprocessing the data, the dataset consists of 7,645 papers and 8,283 references.

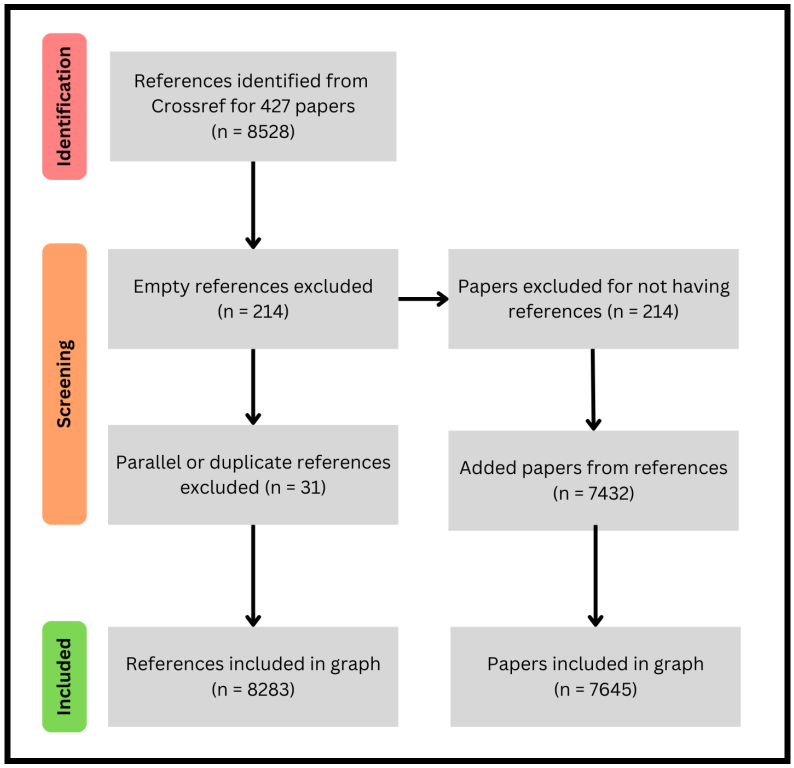

Figure 5. Diagram of data collection and preprocessing pipeline for reference dataset

2.4 Data Analysis Techniques

Once the topics are finalized, the next step is to visualize the data. In this paper, two visualization methods will be used: network graph and co-occurrence network. Both stem from the concept of graph theory and consist of the same features: nodes and edges. Nodes represent entities, such as individuals or concepts, while edges represent the relationship between those nodes or entities. Generally, graphs are undirected or directed. Undirected graphs refer to edges where the structure of node relationships doesn't matter, unlike directed graphs where the edges are either incoming or outgoing relations (Majeed & Rauf, 2020).

An example of this is with social media and how each user and who they follow can be represented in a social network graph. Each user can be represented by a node. If Node A had an outgoing relation to Node B, that would be that User A is following User B. On the flip side, if User A has an incoming relation from Node B, it would represent User A being followed by User B.

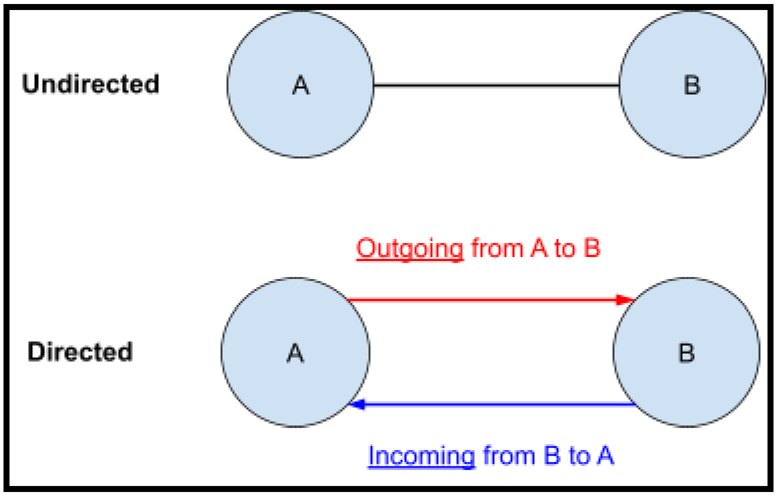

Figure 6. Comparison between undirected and directed nodes

The applications of graph theory will be used to create a network graph to show the links of research in terms of citations and references and a co-occurrence graph will map out the discussions in the GenAI ethics area, called research graph and topic graph respectively. Network graphs, also known as social network graphs, are generally used to depict the connections between people in a network, like social media (Majeed & Rauf, 2020). However, in this case, the network will be academia. Meanwhile, the co-occurrence graph will represent the similarity score of each topic relation as given from the BERTopic similarity matrix (Grootendorst, n.d.).

These graphs will be visualized through Gephi, a network analysis and visualization program. For the research graph, each node will represent a research paper and the edge will represent the papers referencing one another. As for the topic graph, the node will be a generated topic and an edge is equivalent to how frequently the topics are written together.

3. Results
3.1 Generated Topics

Each journal dataset is trained on the three topic model algorithms, resulting in the tables below.

Topic	Count	Name
-1	115	ai, genai, research, ethical
0	108	ai, ethical, ethics, moral
1	60	students, genai, learning, education
2	40	art, ai, content, music
3	28	business, marketing, genai, customer
4	25	healthcare, medical, ai, care

5	18	gan, data, network, adversarial
6	15	genai, text, extraction, language
7	14	ethical, healthcare, ai, principles
8	11	security, cybersecurity, cyber, threats

Figure 7. Generated topic model results from BERTopic

Topic	Count	Name
1	210	ai, ethics, ethical, algorithms, technologies, human, technology, generative, implications, moral, innovation, article, academic, decision, approaches, considerations, concerns, explores, based, data, risks, enhance, digital, future, challenges, models, between, paper, approach, critical, model, machine, within, genai, analysis, assessment, about, training, studies, regarding, study, understanding, impact, insights, science, research, development, efficiency, society, security
2	6	ai, algorithms, generative, technologies, enhance, ethical, efficiency, technology, innovation, ethics, analysis, data, studies, implications, considerations, study, approaches, approach, decision, generated, digital, which, paper, impact, between, integration, benefits, genai, human, assessment, risks, article, within, explores, also, insights, however, design, across, machine, potential, tasks, text, findings, development, generation, understanding, framework, review, higher
3	62	ai, educational, generative, academic, enhance, algorithms, learning, assessment, technologies, education, studies, ethics, study, ethical, students, genai, technology, integration, training, which, including, explores, paper, approaches, level, benefits, understanding, impact, implications, data, development, these, review, text, digital, between, regarding, as, future, generated, human, article, has, concerns, with, by, approach, decision, methods, literature
4	61	algorithms, generative, ai, data, enhance, technologies, model, analysis, approaches, models, approach, study, paper, technology, digital, framework, implications, assessment, studies, machine, accuracy, generated, which, training, human, ethical, insights, impact, findings, generation, ethics, learning, decision, text, understanding, efficiency, risks, based, design, academic, systems, system, challenges, concerns, with, application, review, context, explores, considerations
5	35	ai, generative, art, algorithms, technologies, creation, design, innovation, digital, technology, generated, ethical, article,

		explores, paper, enhance, approaches, ethics, approach, human, implications, models, data, potential, process, model, machine, development, studies, efficiency, future, considerations, decision, based, risks, study, generation, literature, methods, challenges, learning, understanding, impact, using, analysis, existing, within, educational, text, with

Figure 8. Generated topic model results from Top2Vec

Topic	Count	Name
0	37	tools, learning, use, applications, future, art, issues, students, concerns, impact
1	39	learning, challenges, development, approach, technologies, based, new, education, also, within
2	43	technologies, moral, use, model, also, healthcare, development , impact, issues, content
3	25	development, learning, use, technology, issues, using, concerns, challenges, healthcare, technologies
4	105	use, learning, technology, students, new, education, tools, art, development, future
5	49	technology, students, model, development impact, work, models, technologies however, use
6	34	use, based, using, impact models, applications, de, challenges development, technology
7	34	model, development, social, based results, learning, models, used field, technology
8	37	learning, issues, development, work use, literature, healthcare, moral impact, challenges
9	31	learning, healthcare, technology digital, impact, education, challenges music, implications, design

Figure 9. Generated topic model results from LDA

The next step is to determine the most optimal topic results for the topic graph. This is done by rating each topic across three categories: Is the topic related to GenAI? Does the topic describe an ethical impact of some kind? And lastly, does that impact concern creators and the arts? The ratings are given to each category on a scale of 1-3. 1 is for not mentioned, 2 for somewhat or not explicitly mentioned, and 3 is for explicitly mentioned and relevant. Below are the results of the rating process of the topics of each algorithm.

Topic	Is it related to GenAI?	Does it describe an impact?	Does it have an impact on creators and arts?
1	2	1	0
2	2	2	0
3	2	2	2
4	2	2	2
5	2	0	0
6	2	0	1
7	2	0	1
8	2	2	0
Average Score	2	1.125	0.75
Total Score	2 + 1.125 + 0.75 = 3.875		

Figure 10. Determining the quality of topic model results from BERTopic

Topic	Is it related to GenAI?	Does it describe an impact?	Does it have an impact on creators and arts?
1	2	1	0
2	2	0	0
3	2	1	0
4	2	0	0
5	2	1	2
Average Score	2	0.6	0.4
Total Score	2 + 0.6 + 0.4 = 3		

Figure 11. Determining the quality of topic model results from Top2Vec

Topic	Is it related to GenAI?	Does it describe an impact?	Does it have an impact on creators and arts?
1	1	2	2
2	2	2	0
3	2	2	0
4	1	2	2
5	2	0	0
6	2	1	0
7	2	2	0
8	0	2	0
9	2	2	0
Average Score	1.4	1.5	0.4
Total Score	1.4 + 1.5 + 0.4 = 3.4		

Figure 12. Determining the quality of topic model results from LDA

A thing to note here was that "GenAI" was added as a stopword to LDA because, unlike the other two algorithms, LDA calculates its topics from how frequently the words appear together as probabilities. For this reason, GenAI was added to the stopword list to improve the quality of the LDA results, and the GenAI rating category was rated on the relation of the topic to technology and AI instead.

After all topics are rated by each algorithm, the category is given an average score by adding up all of the ratings per category and dividing by the number of topics. Finally, the averages are summed to calculate the total quality score of each algorithm to determine the "best" topic model result, which is the BERTopic topics results at a score of 3.4.

3.2 Topics Graph

Before visualizing the dataset into a graph, the edge table was constructed from BERTopic's similarity matrix, which gave a similarity score to each relationship

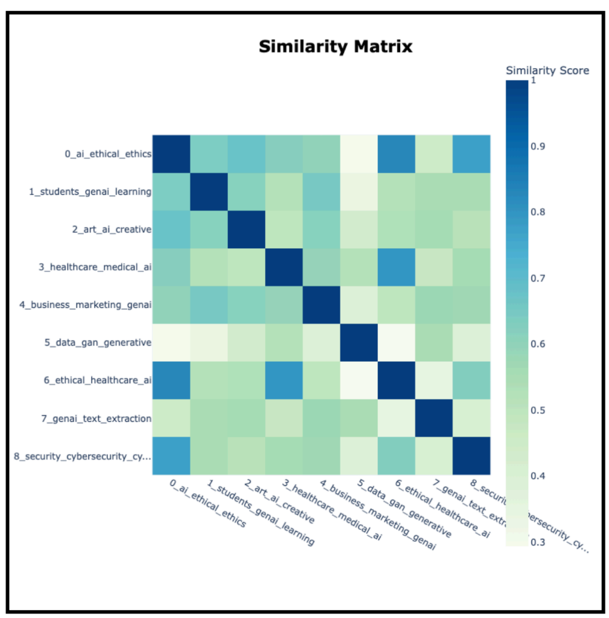

Figure 13. Heatmap of similarity score between topics

between 8 topics on a scale of 0.3 to 1. This resulted in the heatmap seen in Figure 13.

After that, the topic graph was constructed through Gephi. Each color node represents a generated BERTopic topic from 0-8. Topic -1 was not considered because it is assigned to papers that are uncategorized or could not be assigned a topic, unlike the other nine topics.

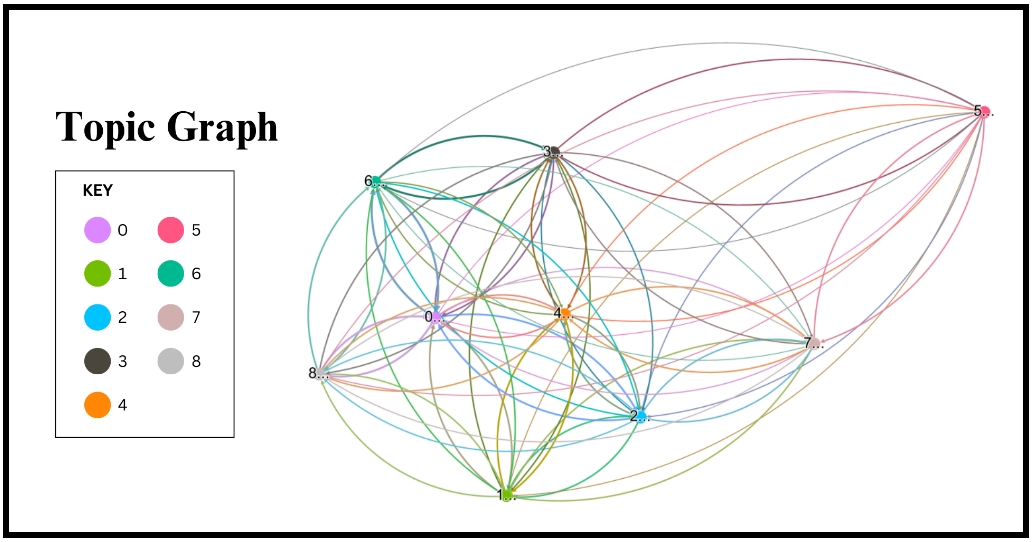

Figure 14. Relationship between BERTopic topics and topics similarity score

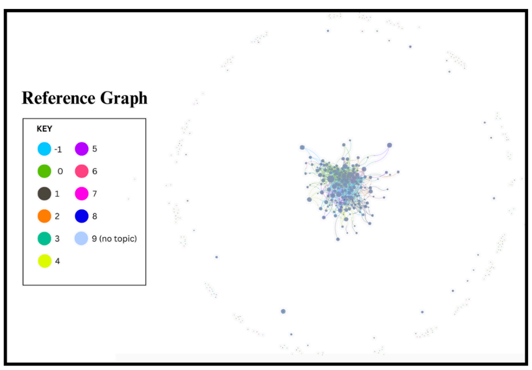

Figure 15. Relationship between topic-assigned nodes and its references

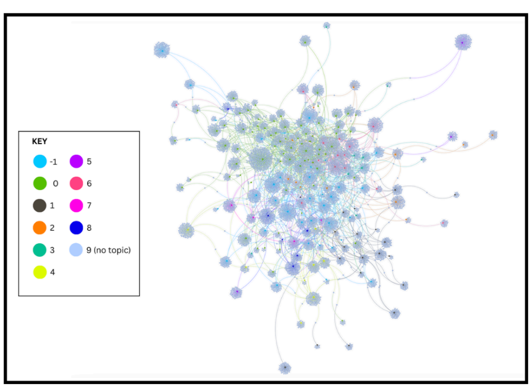

Figure 16. A closer look at the center cluster of the references graph

The thickness of the edges differs based on the similarity between the nodes. For instance, there is a thicker outgoing arrow from topic 6 to 3 than the outgoing arrow from topic 6 to 5.

3.3 Research Graph

Figure 15 is the visualization of references as a network graph. Nodes or papers that were assigned a topic -1 to 8 from BERTopic are color-coded depending on the topic. Other nodes or references that were queried later during the data preprocessing for the reference dataset are assigned "topic 9" to represent that they are part of the reference list to the papers in the topic dataset, and not included in the topic graph representation.

Additionally, a more focused look of the center cluster is shown in Figure 16. Unlike the previous figure, this figure ignores the outliers and surrounding networks that are not part of the main center.

4. Discussion

4.1 Interpretation of Results

After developing and finalizing the graphs, both graphs can finally be analyzed and evaluated. Beginning with the topic graph (Figure 14), all but one node seems to be clustered together. The outlier node is topic 5– gan, data, network, adversarial (Figure 7)– and it is further supported by what has the lowest sum rating score of 2 across three categories (Figure 10), which suggests that it is the least relevant topic to the discussion on ethical GenAI. That being said, most of the outlier nodes in the reference graph are from topic -1 and in fact, the majority of topic 5 references are found in the centroid as seen by the assigned purple node color.

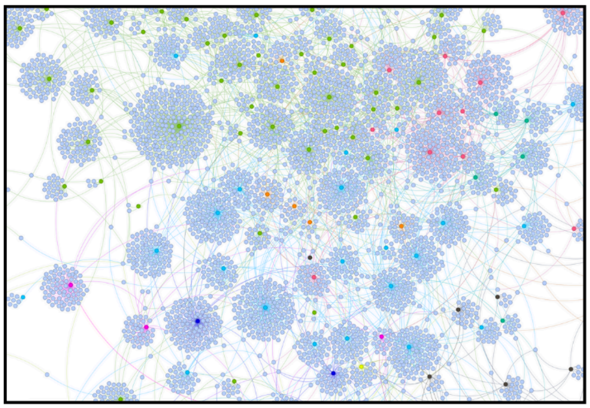

Figure 17. A closer look at the cluster without labels

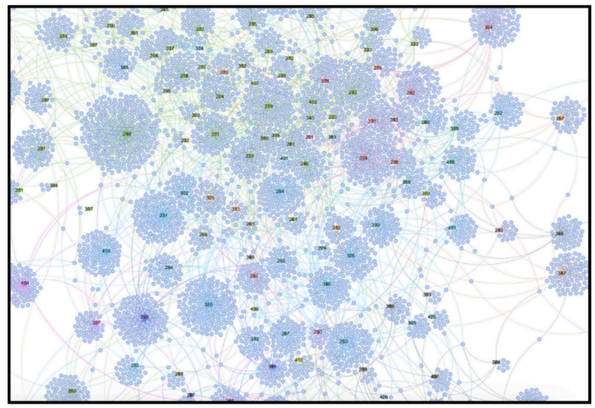

Figure 18. A closer look at the cluster with labels

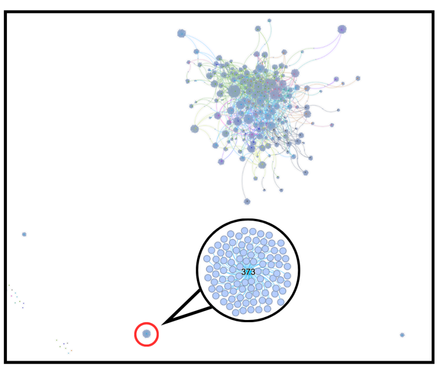

Figure 19. Largest outlier network cluster (circled in red) with reference to the centroid, and zoomed in view of the cluster

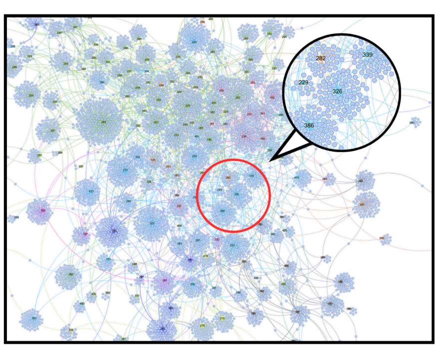

Figure 20. A network of blue nodes inside the centroid cluster with zoomed in view of cluster

Surrounding the centroid are several smaller networks. The largest minor centroid is node ID 373 (Figure 19) or DOI 10.1093/bjrai/ubae012, a paper on applications of GenAI in radiology (Zhong & Xie, 2024). Comparing node ID 373 to the blue nodes inside the centroid, these blue clusters– for instance, 229, 326, 386, and 339 (Figure 20)– are more focused specifically on AI ethics. To give an example, node 229 or DOI 10.1007/s43681-022-00221-0 is a paper proposing sustainable AI (Bogani et al., 2022), and node 386 or DOI 10.1007/s11023-024-09694-w, which is another paper reviewing the studies on GenAI ethics (Hagendorff, 2024). All these suggest that, at least for assigned topic -1 nodes, nodes closer or within the centroid cluster are more relevant to ethics on AI.

Moving on, the most frequently cited node or the node with the highest in-degree is node 529 or 10.1038/s42256-019-0088-2. This is an article cited, including citations outside of the dataset, 1993 times in total (Jobin et al., 2019) and 26 times in the dataset. Other high in-degree nodes are 672 and 660. The DOIs for the two nodes are

10.1007/s11023-020-09517-8 and 10.1007/s11023-018-9482-5, and they are cited in the dataset 17 and 14 times respectively. Furthermore, a majority of the papers citing these nodes are from topic 0. Therefore, it is apparent that many researchers use these three papers as their baseline for discussing ethics in broad AI.

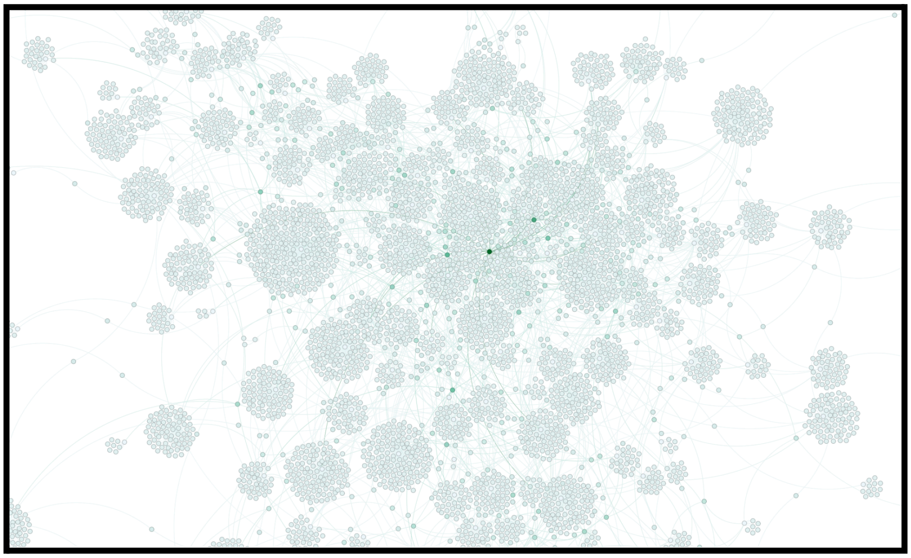

Figure 21. Reference graph with node color ranked by highest to lowest in-degree

Back to the topic graph, topics 0 and 6 are close and have the thickest edge or highest similarity score. In terms of the topic description, the graph suggests that the ethics of AI and AI applications in text are similar to one another, and perhaps should be a topic approached together. In fact, in the references graph, they are largely clustered next to each other, suggesting that there is an overlap in the research being done for the two topics.

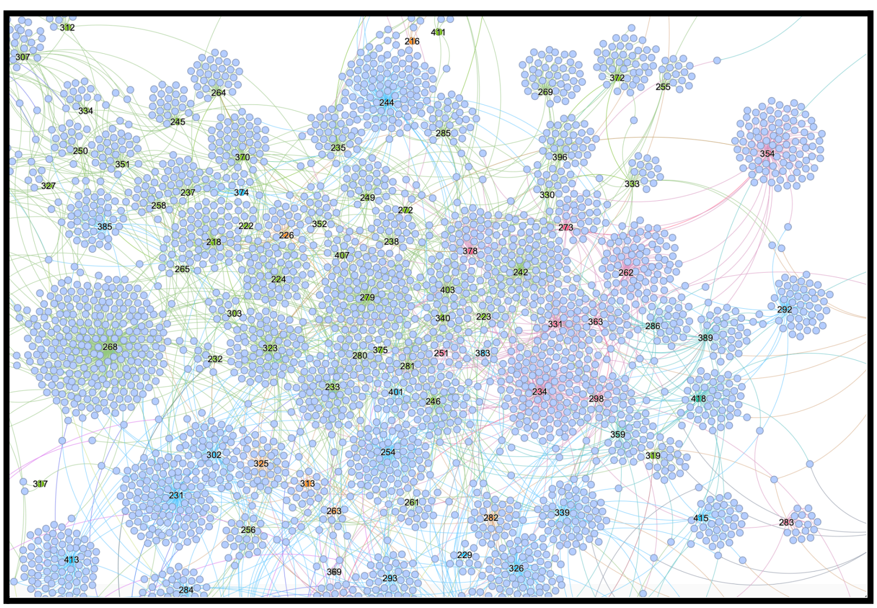

Figure 22. Close-up of where most green (topic 0) and pink (topic 6) are clustered together on the graph

On the other hand, topics 5 and 6 share the lowest similarity score. Since topic 5 or GAN, is a model generally used for visual art generation, it may link to why text and language (topic 6) are considered to be on the opposite end.

Other correlations and overlaps worth mentioning are topics 3 and 6. Topic 3 relates to marketing business, and together this links to GenAI's effects on the workplace with productivity (Cazzaniga et al., 2024). Additionally, topics 0 and 8 are in third place for the highest similarity score. Topic 0 refers to ethical AI while 8 refers to cybersecurity, and while not as densely populated and correlated as 0 and 6 are on the graph, both 3-6 and 0-8 do show some overlap in the reference graph.

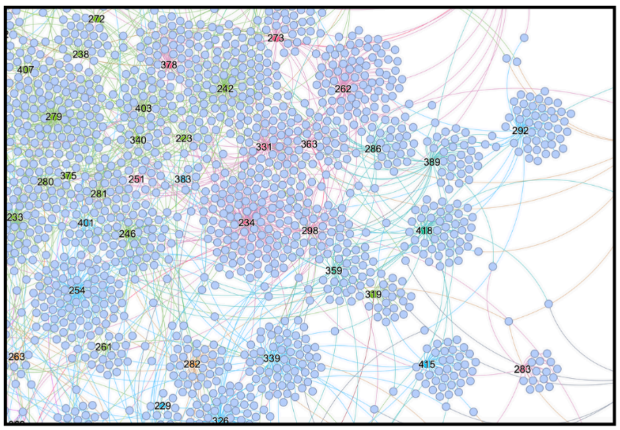

Figure 23. Cluster of pink nodes (topic 6) with turquoise green (topic 3)

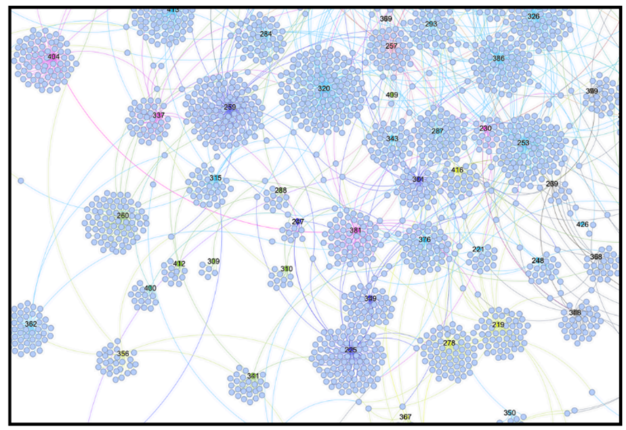

Figure 24. Cluster of dark blue nodes (topic 8) with green (topic 0)

4.2 Limitations of Method

The evaluation of this paper can be separated into three parts. First, the limitations of the datasets, models, and graphs. For the datasets, a large limitation is the number of data. Though the metadata of 1,000 papers was returned at first, after screening, there were only 427 papers left. While this is still a large number, it is still not enough to accurately represent the models of the entire research and its related fields. This possibly led to around a quarter of the topics dataset being uncategorized without a topic. Furthermore, around half of the 427 papers did not have a registered references list on Crossref, therefore decreasing the number of papers eligible for processing even further.

Topics	Total 'n' of literature review	Number of papers referencing the topic	Percentage of papers mentioning topics
-1	427	115	26.93%
0	427	108	25.29%
1	427	60	14.05%
2	427	40	9.37%
3	427	28	6.56%
4	427	25	5.85%

5	427	18	4.22%
6	427	15	3.51%
7	427	14	3.28%
8	427	11	2.58%

Figure 2. Percentage of papers associated with each topic

As for the models, most of their limitations stem from the quality of the topics. Since the size of the dataset is already limited, this impacts the quality of the generated topics because there may not be enough data to train the models. Additionally, although we did choose the "best" generated topic results to visualize in the graph, it is impossible to determine the quality of the topics. This is in part due to how difficult it is to track the accuracy of the generated topics as it would take a long time to verify each generated topic and topic assigned to the paper.

Finally, the graphs have their visualization limitations. The reference graph cannot show all node IDs or DOIs at the same time without affecting the visibility of other IDs. Furthermore, it is difficult to compare the thickness of the edge in the topic graph, therefore affecting the ability to analyze the similarity score relations.

4.3 Future Directions

A few potential areas of research can be viewed from the overlap in topics and citations in their respective graphs. Outside of ethical AI and language that seems to be rising in popularity and relevance, other links worth noting are GenAI, specifically language generation or extraction, which can impact traditional roles in the workplace, such as marketing. Additionally, the ethics of cybersecurity could be an avenue for future research.

Focusing on GenAI research, most works on the ethical discussion cite research from 10.1038/s42256-019-0088-2, 10.1007/s11023-020-09517-8, and 10.1007/s11023-018-9482-5. All three of these lay a solid foundation for a deeper understanding of this conversation. Moreover, the reference graph can be utilized to find relevant research in AI and GenAI application domains. In this case, another possible point of future direction is to replicate the study but with more metadata, such as the authors and year of publication to identify further trends in the research landscape.

Finally, to address the ethical discussion, based on the generated topics and topic graph, it can be seen that there has not been much research on the legalities and originality of GenAI yet. Even though impacts on employment have been discussed, the graphs suggest that other concerns have gone unanswered and unresearched besides the applications of GenAI. Thus, this opens up new unmarked fields of research– each field representing a specific concern– that will gain more momentum and relevance over time, as suggested by the increasing number of general ethical reviews on GenAI.

4.4 Conclusion

Ultimately, this paper has captured and modeled the current scope of research on GenAI ethics to survey what academics have found value in addressing and contributing to the discussion on GenAI ethics. While the model indicates that the research topics conducted by researchers do not answer most people's ethical concerns yet, this paper hopes to raise attention to investigating these concerns and exploring more related domains in regard to the ethics of GenAI.

---

5. Works Cited

Angelov, D. (2020). *Top2Vec: Distributed Representations of Topics*. https://doi.org/https://doi.org/10.48550/arXiv.2008.09470

Bergmann, D., & Stryker, C. (2024, August 21). *What are Diffusion Models?*. IBM. https://www.ibm.com/think/topics/diffusion-models#

Blei, D., Ng, A., & Jordan, M. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*. https://doi.org/https://dl.acm.org/doi/10.5555/944919.944937

Bogani, R., Theodorou, A., Arnaboldi, L., & Wortham, R. H. (2022). Garbage in, toxic data out: a proposal for ethical artificial intelligence sustainability impact statements. *AI and Ethics*, *4*, 1135–1142. https://doi.org/10.1007/s43681-022-00221-0

Cazzaniga, M., Jaumotte, F., Li, L., Melina, G., Panton, A. J., Pizzinelli, C., Rockall, E. J., & Tavares, M. M. (2024). *Gen-AI: Artificial Intelligence and the Future of Work*. https://doi.org/https://doi.org/10.5089/979840026254

Crossref Unified Resource API. (n.d.). Swagger. Retrieved November 9, 2024, from https://api.crossref.org/swagger-ui/index.html#/

Demirci, O., Hannane, J., & Zhu, X. (2024). Who is AI Replacing? The Impact of Generative AI on Online Freelancing Platforms. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4991774

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. https://doi.org/ https://doi.org/10.48550/arXiv.1810.04805

Digital Replicas 101. (2023). SAG-AFTRA . https://www.sagaftra.org/sites/default/files/sa_documents/DigitalReplicas.pdf

Ferrara, E. (2024). GenAI against humanity: nefarious applications of generative artificial intelligence and large language models. *Journal of Computational Social Science*, *1*, 549–569. https://doi.org/10.1007/s42001-024-00250-1

Grootendorst, M. (n.d.). *Visualization*. BERTopic; GitHub. Retrieved November 17, 2024, from https://maartengr.github.io/BERTopic/getting_started/visualization/visualization.html#visualize-topic-similarity

Grootendorst, M. (2022). *BERTopic: Neural topic modeling with a class-based TF-IDF procedure*. https://doi.org/https://arxiv.org/abs/2203.05794

Hagendorff, T. (2024). Mapping the Ethics of Generative AI: A Comprehensive Scoping Review. *Minds and Machines*, *4*. https://doi.org/10.1007/s11023-024-09694-w

Hall, B. (2024, September 24). *185 real-world gen AI use cases from the world's leading organizations*. Google Cloud Blog; Google Cloud. https://cloud.google.com/transform/101-real-world-generative-ai-use-cases-from-industry-leaders

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *9*, 389–399. https://doi.org/10.1038/s42256-019-0088-2

Majeed, A., & Rauf, I. (2020). Graph Theory: A Comprehensive Survey about Graph Theory Applications in Computer Science and Social Networks. *Inventions*, *1*, 10. https://doi.org/10.3390/inventions5010010

Murel, J., & Kavlakoglu, E. (2024, March 30). *What is topic modeling?*. IBM. https://www.ibm.com/topics/topic-modeling

Peng, Y. (2024). A Comparative Analysis Between GAN and Diffusion Models in Image Generation. *Transactions on Computer Science and Intelligent Systems Research*, 189–195. https://doi.org/10.62051/0f1va465

Pentz, E. (2020, February 2). *History*. Crossref. https://www.crossref.org/about/history/

Roose, K. (2022, September 2). *An A.I.-Generated Picture Won an Art Prize. Artists Aren't Happy.* The New York Times. https://www.nytimes.com/2022/09/02/technology/ai-artificial-intelligence-artists.html

SAG-AFTRA TV/ Theatrical Contracts 2023. (2023). https://www.sagaftra.org/sites/default/files/sa_documents/TV-Theatrical_23_Summary_Agreement_Final.pdf

Shumailov, I., Shumaylov, Z., Zhao, Y., Papernot, N., Anderson, R., & Gal, Y. (2024, June 24). *AI models collapse when trained on recursively generated data | Nature*. Nature; Nature. https://www.nature.com/articles/s41586-024-07566-y#Sec2

Stryker, C., & Kavlakoglu, E. (2024, August 16). *What Is Artificial Intelligence (AI)?* IBM. https://www.ibm.com/topics/artificial-intelligence

Stryker, C., & Scapicchio, M. (2024, March 22). *What is Generative AI?*. IBM. https://www.ibm.com/topics/generative-ai

Villalobos, P., Ho, A., Sevilla, J., Besiroglu, T., Heim, L., & Hobbhahn, M. (2022). *Will we run out of data? Limits of LLM scaling based on human-generated data*. https://doi.org/ https://doi.org/10.48550/arXiv.2211.04325

Wen, L. (2019, April 18). *From GAN to WGAN*. ArXiv. https://doi.org/10.48550/arXiv.1904.08994

Weng, L. (2021, July 11). *What are Diffusion Models?*. Lil'Log; Lil'Log. https://lilianweng.github.io/posts/2021-07-11-diffusion-models/

We're Fighting for the Survival of Video Game Performers. (2024, August 19). SAG-AFTRA. https://www.sagaftra.org/were-fighting-survival-video-game-performers

Zhong, Z., & Xie, X. (2024). Clinical applications of generative artificial intelligence in radiology: image translation, synthesis, and text generation. *BJR Artificial Intelligence*, *1*. https://doi.org/10.1093/bjrai/ubae012