RESPONSIBLE RESEARCH AND INNOVATION OF RECOMMENDATION SYSTEMS

A Research Paper
In STS 4600
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science

By
Param Damle

April 17, 2024

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

ADVISOR

Gerard Fitzgerald, Department of Engineering and Society

Introduction

Behind each action we take is a network of hidden algorithms designed to influence our desires. Companies have increasingly earned greater profits by predicting user behavior with models trained on large quantities of data, a process ingrained into the technology that we interact with every day. The scope of these systems is immense—anything from the rental you book on Airbnb to what you watch on Netflix to everyday Google searches is factored into constructing a profile of you across various databases (Jeckmans et al., 2013). As platforms such as these grow more ubiquitous in countries like the United States, user data monetization is dominating other revenue streams within the cloud services space (Dessemond, 2020).

The collection, retention, and distribution of this data is just the tip of the iceberg in terms of how it can be used for profit, but there's deeper concern in the less regulated domain of using this data for recommending products, people, and programs of action. For example, targeted recommendations generated Google $147 billion in revenue in 2020 and account for more than 35% of Amazon's sales, but data regulations involving informed consent rarely extend to the systems that use the data. When outcry over recommendation privacy first started in the late 2010s and Apple introduced an opt-out feature for having one's data factored into personalization, 96% chose to withdraw their personal information (Küçükgül et al., 2022). Not only is it unclear to most users that data is being collected in this manner when they visit an inconspicuous website, it is less clear that the information is being leveraged to instantaneously influence their behavior. After less than 20% of users reported privacy violations, of which only half knew how to adjust their settings to counter the intrusive software, Hauptman et al. suggested that rebuilding trust between the public and data conglomerates will be crucial to the future of the software systems we've come to rely on (2011).

To address this issue more proactively, the Responsible Research and Innovation methodology highlights how key stakeholders, including engineers, legislators, and citizen advocates, can mitigate the problematic aspects of how these systems are designed and expanded moving forward. Of these groups, I aim to show that the engineers directly engaged in the research and development (R&D) of these recommendation engines are best poised to anticipate the repercussions of engineering decisions and adjust the system design to minimize harm. Upon establishing the significance of the R&D team's part in this process, it remains to be seen whether this role is being fulfilled, or at the least acknowledged within the technical community. Thus, I seek to examine how the proliferation of recommendation systems has affected technical advocacy for and protection of user data privacy in these systems. Although the specifications of these systems require ever-increasing sources of data to boost performance, I expect the consideration of social implications to have increased attention towards ethical concerns and implementations of privacy-conscious mechanisms.

Societal Role of Recommendation Systems

The invention of social media in the 1990s produced a chronological "timeline" biased heavily in favor of spam, requiring developers to implement their own guardrails for recommending content (Meserole, 2022). As these hard-coded rules didn't scale, developers incorporated machine learning and behavioral modeling. Although this evolution seems straightforward, the implications on the communities involved are not. Understanding why these systems are as pervasive and intrusive as they are requires analysis of all actors, from those designing the system (engineers) and prescribing its greater goal (executives) to those affected by it (users) and those with a capacity to change it (lobbyists and legislators).

Framed using Hughes' lens (1987), these recommendation engines are technological systems that underwent a series of innovations to produce more accurate predictions at the cost of gathering a greater breadth and depth of personal information. We are now at the stage where various entities (in the e-commerce and social media spaces, primarily) are competing to give you the best recommendations by leveraging even more of your data. This competition incentivizes each company to transfer (copy and adapt) the most effective methods from its rivals, such as how both Facebook and Twitter invested in replicating TikTok's successful strategy of "inventorying" data from users across their platform to influence the timeline of any one user (Meserole, 2022). Given the size of the user bases, these applications and their underlying models are embedded in the way people shop for goods, consume content, and connect with each other. Companies like Google, Amazon, and Facebook are among the largest drivers to the US economy, and other firms riding their algorithmic coattails hire a significant portion of the software labor force. Under Hughes's framework, these recommendation algorithms have acquired momentum—their ubiquity and perceived necessity allows them to grow without legitimate challenge. It's difficult to argue against data farms if every major store and media company relies on them for traction, and it'll get increasingly difficult as data collectors consolidate this power and make using their data addictive (DeLeon, 2019). This is compounded by the influence platforms like Facebook have on people who don't use their service but live in a region where Facebook is popular (Kotliar, 2021).

Although this classifies as an indirect impact on the user, advocates have found greater success in targeting direct impacts such as a breach of privacy (breadth, depth, or lifetime) and unethical recommendations (revealing connections, funneling users into extremist echo chambers, and so on). The extensive privacy concerns have led to an increase in digital rights advocacy over recent years, as well as technical changes to the systems, such as randomization, data anonymization, and end-to-end encryption when transmitting personal information. The problem of echo chambers has real world consequences such as the spread of COVID misinformation, far-right extremism, and perhaps an insurrection against one's government (Meserole, 2022). As such events pose a direct threat to society, the range of concerned parties extends from institutions like governments and corporations to individuals who make personal decisions regarding their own data and engagement with recommendation services.

Governments around the world have begun noticing and responding to these challenges. The General Data Protection Regulation (GDPR), deployed in 2018, placed heavy scrutiny and restrictions upon systems that processed and distribution of data collected from citizens of the European Union. Not only did this require stringent initial assessments of these systems' technology and its sociotechnical impact, but also fostered considerable investment in manpower and infrastructure to ensure continual compliance with the ethical guidelines imposed by the Regulation (Li, 2019). The United States government recently addressed its share of user information concerns by restricting the sale of personal data to the "labyrinthine ecosystem of vendors and sellers" in foreign adversarial nations (Volz & Wei, 2024).

Advocates often extend the responsibility of mitigating such threats to the source of this technology and expect the corporations instituting these systems to proactively address their ethical consequences. Despite this expectation, there's incentives for companies to pursue projects that collect data more aggressively and, potentially, nefariously. Fundamentally, this derives from the largest predictor of machine learning performance—not a model's hyperparameters, but the quality of the training set and the features derived from it (Wang & Shah, 2021). This has created a "data imperative" for maximizing information collection, whereby algorithms can craft the most specific profile for each user (Seaver, 2021). Specificity provides more personalized recommendations, leading to a uniquely memorable customer experience, which in turn produces more loyal sources of revenue for a company (Vas, 2021). Thus, without external regulation or internal policy, tech companies are incentivized to build systems that are increasingly intrusive and polarizing, as this will yield insightful data and more engaged users, convertible into increased revenue from recommendations.

Who Should Address These Problems?

Rather than diluting accountability by placing obligation upon the corporation itself, it can be more productive to look at individual actors involved with recommendation systems. Although a variety of stakeholders hold legitimate concern over the implications of recommendation technology, the uneven influence some groups have over future directions of these systems corresponds with uneven accountability for ensuring that their harmful effects are minimized. As technology flows through the R&D process, it passes from upstream groups (e.g. product managers establishing the vision for a system) to downstream groups (e.g. users interacting with the system or legislators reacting to its consequences). Anticipating problems in an upstream technology assessment is limited by the difficulty of forecasting undeveloped technology, while regulating them downstream after widespread deployment might be too late (Fisher et al., 2006). Lange et al. dubbed the latter a "pacing problem": the rate of innovation generally outpaces regulations, spawning gray areas in unimplemented or ambiguous guidelines (2023). As government regulations lag behind the threats posed by rapid innovation, those within the technology organizations bear increasing responsibility to institute guardrails against unethical system design (Zhang et al., 2014). That is, these researchers, engineers, and decision-makers need to fundamentally shift the current techno-economic prioritization into a more techno-socio-economic view (Rogerson & Bynum, 1996). The aforementioned issues of upstream unpredictability and downstream lag can thus be avoided if course corrections are performed *during* the iterative development of a technology.

Such “midstream modulation” is an example of Responsible Research and Innovation (RRI), a paradigm whereby stakeholders provide insight and alter innovation directions throughout the technological process to ensure the resulting product properly embeds itself into and advances the long-term goals of society (Fisher et al., 2006). RRI has gained increasing attention and policy relevance since 2011, particularly with the European Commission Science in Society Programme’s emphasis on orienting innovation towards sustainable purposes, institutionalizing responsiveness to ethical concerns, and making responsibility a collective action (Owen et al., 2012). This framing, through its intentional generality, could imply that all groups, from executives to engineers to lawmakers to end users, share a similar burden (and perhaps a similar ability) to address the greater problems posed by recommendation systems.

However, even involvement of a diverse array of stakeholders will suffer from a time lag between design decisions and the interface between a prototype and non-technical actors. The inherently downstream positioning of non-technical actors suffers from the Collingridge Dilemma: ethical issues are best addressed upstream when the effects of a technology are most difficult to predict (von Schomberg, 2011). The Precautionary Principle addresses this ambiguity by urging organizations to enact safeguards against potentially threatening technology, even if this eventuality isn’t certain at the moment. Relating this to information systems, the UK Information Commissioner’s Office advised that Privacy Impact Assessments (PIAs) should occur as early as possible, which would necessarily be during the R&D of the system itself (Wright et al., 2011). Intuitively, this indicates that the researchers that interface directly with the systems and shape R&D directions are most capable of factoring in ethical design decisions, making them optimal candidates to implement any necessary midstream modulation.

Complicating this notion, a study from Google, which is a major player in recommendation systems, indicated that big tech companies rely on autonomous engineering teams with few non-technical considerations to drive innovation (Lange et al., 2023). This study mandated a "role obligation shift", whereby sociotechnical considerations are formalized in researchers' job description, so that they actively gather information regarding potential ethical implications, deliberate all relevant tradeoffs for each design decision, and translate these into concrete alterations to the product. Sociotechnologist Walter Maner emphasized this, stating that integrating information ethics into the requirements of professionals will avert a catastrophe (1996). In the same guidelines, Maner suggests that innovations accompanying the 21st century will require equally innovative ethical studies, implying a research component to the ethical implications of recommendation systems.

Analyzing Attention to Ethical Concerns

Indeed, policy advisors have frequently advocated for the formalization of RRI within the research process (Stahl, 2013). At its core, this reflects a desire to embed ethical due diligence into research culture, as practical studies have shown the necessity of shaping cultural norms to encourage technical actors to perform more holistic sociotechnical reflection (Lange et al., 2023). For the subset of computer science revolving around big data and algorithms, as is for academic disciplines in general, such culture is usually defined by (or, at least, reflected in) conferences (Koch, 2021). While critics may argue that research and industry conferences only explicitly involve the technical and economic aspects of the applicable subject matter, prior work has demonstrated that introducing ethical study doesn't detract from the research process, rather it adds value to the resultant work (Fisher & Mahajan, 2006). So, to what extent is the research community surrounding recommendation systems adequately responding to concerns posed by new advances in the field?

To assess this, I thematically analyzed the corpus of proceedings from The ACM Conference Series on Recommender Systems (RecSys) between 2007-2023 for mentions of ethical concerns, data privacy, and user protection. Prior work by Bernd Carsten Stahl has demonstrated the value of distributed analysis of ethical discourse within publications on emerging information and communication technologies (2011).

I manually downloaded every submitted paper and workshop document as a PDF from the ACM Digital Library (Association for Computing Machinery, 2024) and employed Stavrakis's extraction method (2023) to pull raw text from each PDF. These full-document texts were embedded into numerically valuable data via the following process:

1. For consistency, the document was converted fully to lowercase characters and all non-alphabetic (a-z) characters were removed.
2. Tokens were generated by splitting each document's full text by whitespace.
3. All stop words (common English words not associated with a particular thematic matter, e.g. "in" and "this") in this list of tokens were removed.
4. To allow matching of words that carry similar meaning but are conjugated differently (e.g. "create" and "creation"), a Snowball stemming algorithm was employed to "stem" off the ends of words (Porter, n.d.).
5. Not only were single tokens considered, but *bigrams* of 2 successive tokens were collected to capture information about relationships between neighboring words (Jurafsky & Martin, 2024).

6. Each document's representation as a list of stemmed tokens (and bigrams) was converted into a numerical vector representation via a TF-IDF transformation that processed the frequencies with which tokens appeared in each document (Simha, 2021). Tokens with a frequency of less than 0.5% or greater than 50% were discarded to ensure that the features of the transformed dataset were relevant to comparison.

This vector embedding stored information about semantic relationships between words, as the meaning of each word affects which documents it appears in and how often it appears in them. Figure 1 shows how different semantic relationships could be encoded in a 3-dimensional space, which is an oversimplification since the actual dimensionality of the embedding is in the dozens of thousands.

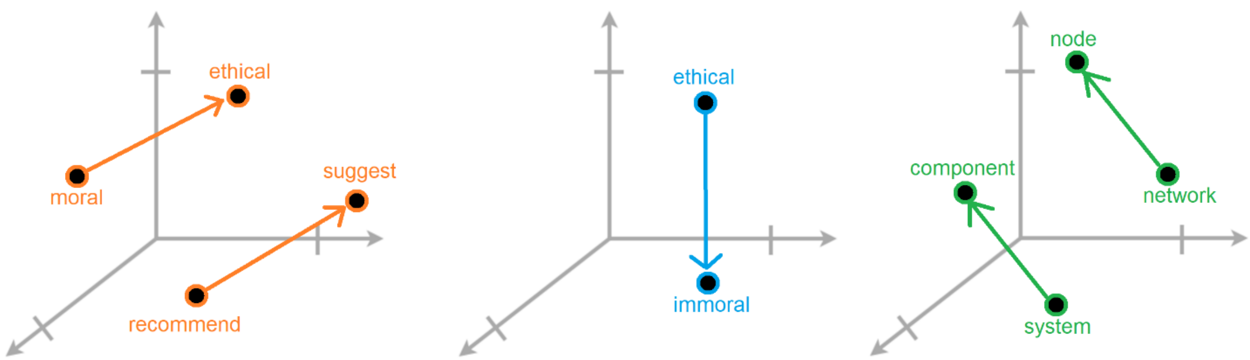

Figure 1. Potential semantic relationships (synonyms on the left, antonyms in the middle, and component relationships on the right) visualized in a 3-dimensional embedding space.

The vector representation for any set of documents (e.g. by year, by conference session, etc) was derived by averaging the vectors representing the documents in the set. By obtaining and aggregating these numerical embeddings, I could compare documents (or sets thereof) using metrics as simple as the Euclidean distance between vectors or the inner product of vectors (Pawar & Mago, 2018). This comparison was performed both using targeted and unsupervised visualization. Through this comparison, I expected to witness an increase in references to privacy topics within the average paper with respect to conference year. If this hypothesis were accepted, RecSys could serve as a model technical field where attention to social issues has, at minimum, kept pace with increased attention and innovation of the technology. This would mean that coverage within academic circles is not a reverse salient with respect to the growing technological momentum of recommendation systems. If my claim were rejected, however, my research would highlight a need for the RecSys community to invest more resources into addressing the societal consequences of such systems.

Results

When collecting the required texts, I noticed that the earliest available conference proceedings (RecSys 2007) already included a workshop dedicated to Privacy and Trust. While this suggests at least some scholarly attention to sociotechnical issues from the inception of the field, the numerical results provide a clearer picture.

Unsupervised learning clusters data without associated labels, allowing observers to glean insight on the numerical relationship between data points. For the experiment below, I took the large-dimensional sparse dataset (where a document vector had a dimension for each token in the entire RecSys corpus) and reduced it to a visualizable two dimensions using Latent Semantic Analysis (LSA), which decomposes high-dimensional data to a subset of principal components that explain the most variance across the dataset (Gopalakrishnan, 2020). Each document from RecSys (2007-2023) is plotted below with respect to these two principal components, and is colored by the year of the proceedings, to visualize whether the change in the societal role of recommendation systems over time has affected the academic discourse surrounding the topic enough to show separation along the two dimensions with maximal variance in the data.

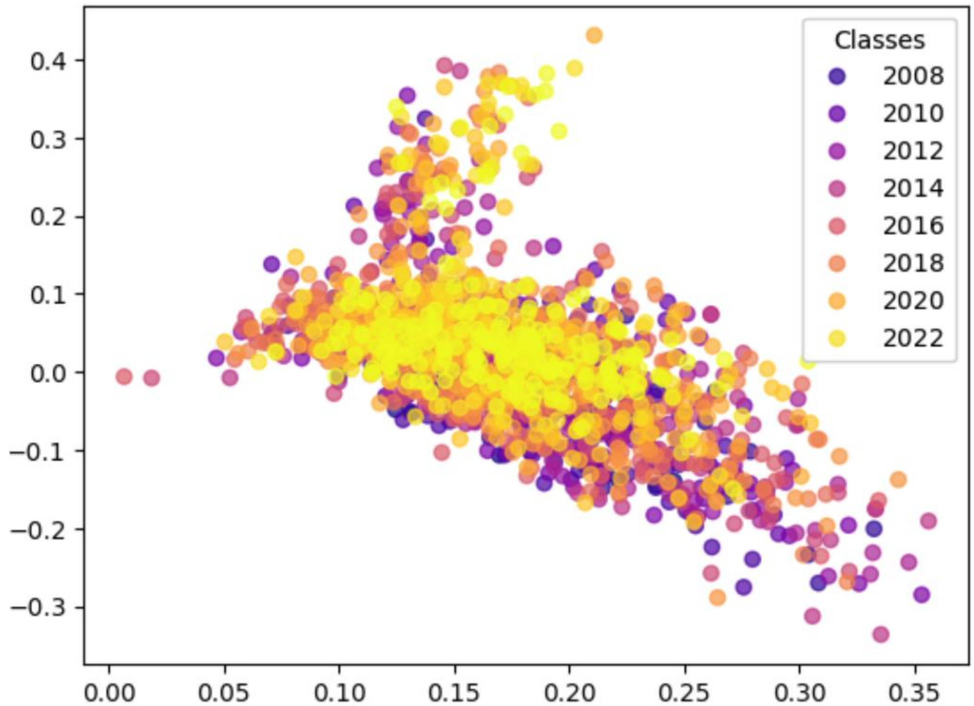

Figure 2. LSA results for entire RecSys corpus colored by proceedings year.

As shown in Figure 2, the lighter colored dots tend to focus closer to the center than the darker colored dots, which indicates that earlier proceedings (2008-2016) showed greater variance in their thematic matter (to the extent it was captured by my particular embedding scheme) than the recent conferences (2017-2023). This could be cause for concern as greater homogeneity in thinking will stifle innovation and limit the likelihood of underrepresented voices breaking through the mainstream.

Pursuing an approach better targeted at my hypothesis, I also generated a graph of the similarity scores between keyword strings related to sociotechnical issues and the entire corpus of an entire year's RecSys proceedings. These keyword strings were generated by providing prompts of the following template to Google's generative chatbot Gemini:

Generate a list of space-separated keywords that relate to _________ recommendation systems. Don't adhere to grammar or sentence structure, just generate a list of words.

The blank was replaced by the prompt in Table 1 for each chosen theme.

Theme	Prompt
Privacy	*privacy and ethical concerns in*
Society	*ethical concerns around the societal and political impact of*
Economy	*economic ethics and fairness as big tech scales up its*
Equity	*equity and access issues related to*
Sustainability	*sustainability and long-term impacts of*

Table 1. Prompts used to generate keyword strings from Google Gemini.

The result of each query was approximately 200-300 keywords that could easily be inputted into my vectorizer as if they were themselves the document text of a journal article. As a control, I also downloaded a list of the 3000 most popular English language words and randomly sampled a sequence of 250 of these words (about the length of the generated keyword lists) to get a similarity score between each year's corpus and general English vernacular (Education First, n.d.). The plotted similarity scores below are divided by this control to yield a ratio that represents how similar the thematic keywords are to a particular year's corpus, relative to how similar that corpus is to any random collection of usual English words.

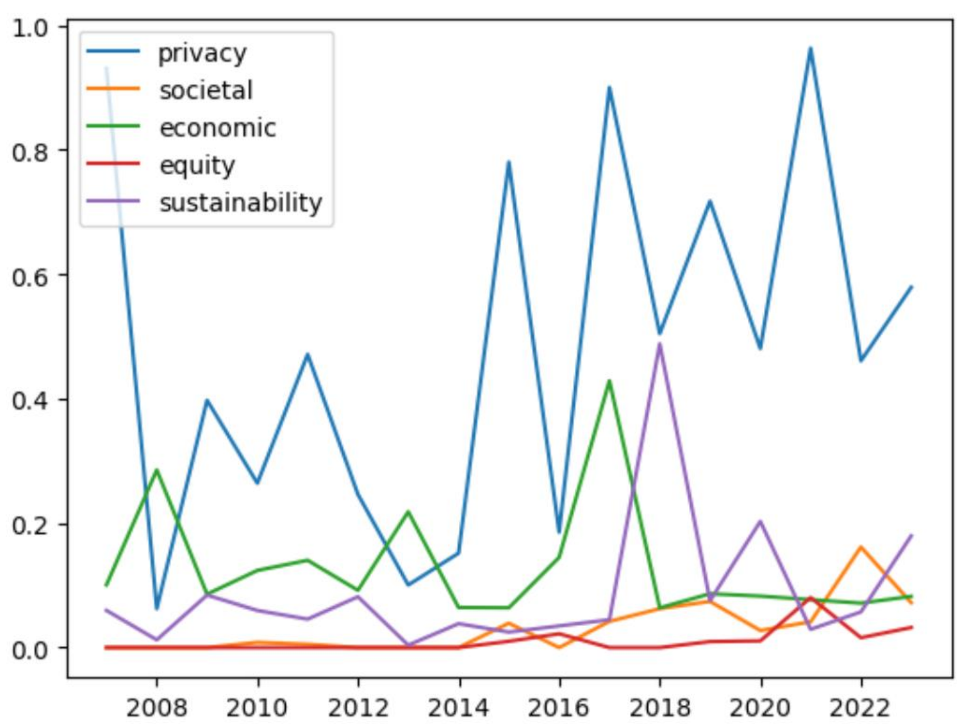

Figure 3. Trend over time of similarity of RecSys conference proceedings to thematic keyword strings, normalized by similarity to a control group of random English words.

Here I found mentions of various sociotechnical concerns *have* increased in recent years, from an average of 0.3 in the early 2010s to an average of 0.6 in the early 2020s. In addition, mentions of sociopolitical and equity-related issues also increased in the most recent proceedings after maintaining a near-0 score prior to 2014. The fact that the score has maintained below 1 is not necessarily relevant to my hypothesis, as words specifically referencing technical details about privacy and protection will certainly be less frequent than common English words in any corpus, even one focused exclusively on the relevant issues. These findings validate my hypothesis as the increased impact of recommendation systems has apparently brought with it a proportional attention to its sociotechnical implications, at least within the research community.

Conclusion

As recommendation systems play an increasing role in online interaction, commerce, and democracy, they pose concerns such as user privacy, declining public trust, echo chamber incubation, and algorithmic discrimination (Jannach & Zanker, 2022). While these concerns involve many societal groups, engineers are best poised to proactively address potential issues and employ design decisions that explicitly protect users. Current implementations of midstream reviews of sociotechnical implications suffer from a lack of integration with the daily design decisions of the researchers developing the system (Lange et al. 2023). This report sought to assess whether the recommendation systems technical community is adequately addressing and improving upon these shortcomings.

Upon examining the technical literature from the ACM Conference on Recommendation Systems, I found a notable increase in attention to the societal impacts of these systems that tied into a steady increase in their adoption. This research serves as a litmus test of the industry's acknowledgement of threats to user privacy and other rights, indicating that at the very least, professionals *are* adequately covering societal considerations in technical literature.

While this study does not assert that it will be sufficient to implement protections and overhaul contemporary project designs to a value-first approach, proportional coverage is a promising start. Addressing these concerns beyond awareness, the European Commission plans to institute mandatory PIAs, which it found to effectively manage user privacy risk at scale (Wright 2011). Eventually, cohesion between technical and ethical decision making will poise innovation as a means to "democratizing and empowering" end, rather than an "enslaving or debilitating" one (Rogerson 1996). Even the United Nations describes upcoming international advancement as innovation fueled by a "cloud economy", where public welfare is tied to the systems researchers and engineers deploy on the internet (UNCTAD, 2013). Ensuring recommendation systems remain conscious of the full range of social implications will provide an ethically sustainable path to global development.

References

Association for Computing Machinery. (2024). *RECSYS (ACM Digital Library)* [dataset]. https://dl.acm.org/conference/recsys

DeLeon, H. (2019, April 24). The Ethical and Privacy Issues of Recommendation Engines on Media Platforms. *Towards Data Science*. https://towardsdatascience.com/the-ethical-and-privacy-issues-of-recommendation-engines-on-media-platforms-9bea7bcb0abc

Dessemond, E. G. (2020). Restoring Competition in "Winner-Took-All" Digital Platform Markets. *United Nations Conference on Trade and Development*, *40*. https://doi.org/10.18356/5ee7948e-en

Education First. (n.d.). *3000 most common words in English* [dataset]. https://www.ef.edu/english-resources/english-vocabulary/top-3000-words/

Fisher, E., & Mahajan, R. L. (2006, November). Midstream Modulation of Nanotechnology in an Academic Research Laboratory. *Proceedings of IMECE2006*. ASME 2006 International Mechanical Engineering Congress and Exposition, Chicago, Illinois, USA. https://doi.org/10.1115/IMECE2006-14790

Fisher, E., Mahajan, R. L., & Mitcham, C. (2006). Midstream Modulation of Technology: Governance From Within. *Bulletin of Science, Technology & Society*, *26*(6). https://doi.org/10.1177/0270467606295402

Gopalakrishnan, J. (2020, June 1). Latent Semantic Analysis. *Portland State University*. https://web.pdx.edu/~gjay/teaching/mth271_2020/html/17_LSA.html

Hauptman, A., Sharan, Y., & Soffer, T. (2011). Privacy Perception in the ICT Era and Beyond. In *Towards Responsible Research and Innovation in the Information and Communication Technologies and Security Technologies Fields*. Publications Office of the European Union.

Hughes, T. (1987). The Evolution of Large Technological Systems. In W. Bijker, T. Hughes, & T. Pinch, *The Social Construction of Technological Systems* (pp. 45–76). MIT Press.

Jannach, D., & Zanker, M. (2022). Value and Impact of Recommender Systems. In F. Ricci, L. Rokach, B. Shapira, & P. Kantor, *Recommender Systems Handbook* (3rd ed., pp. 519–546). Springer. https://web-ainf.aau.at/pub/jannach/files/BookChapter_RS_Handbook_Value_Impact.pdf

Jeckmans, A. J. P., Beye, M., Erkin, Z., Hartel, P., Lagendijk, R. L., & Tang, Q. (2013). Privacy in Recommender Systems. In N. Ramzan, R. van Zwol, J.-S. Lee, K. Clüver, & X.-S. Hua (Eds.), *Social Media Retrieval* (pp. 263–281). Springer. https://doi.org/10.1007/978-1-4471-4555-4_12

Jurafsky, D., & Martin, J. H. (2024). N-gram Language Models. In *Speech and Language Processing* (3rd ed.). https://web.stanford.edu/~jurafsky/slp3/

Koch, S. (2021). *Researching conferences as spaces of valuation: A pragmatic approach*. https://doi.org/10.31235/osf.io/8sxtp

Kotliar, D. M. (2021). Who Gets to Choose? On the Socio-algorithmic Construction of Choice. *Science, Technology, & Human Values*, *46*(2), 346–375. https://doi.org/10.1177/0162243920925147

Küçükgül, C., Özer, Ö., & Wang, S. (2022). *Recommender Systems under Privacy Protection*. https://doi.org/10.2139/ssrn.4138757

Lange, B., Keeling, G., McCroskery, A., Zevenbergen, B., Blascovich, S., Pedersen, K., Lentz, A., & Agüera y Arcas, B. (2023). Engaging engineering teams through moral imagination: A bottom-up approach for responsible innovation and ethical culture change in technology companies. *AI and Ethics*. https://doi.org/10.1007/s43681-023-00381-7

Li, H., Yu, L., & He, W. (2019). The Impact of GDPR on Global Technology Development. *Journal of Global Information Technology Management*, *22*(1), 1–6. https://doi.org/10.1080/1097198X.2019.1569186

Maner, W. (1996). Unique ethical problems in information technology. *Science and Engineering Ethics*, *2*(2), 137–154. https://doi.org/10.1007/BF02583549

Meserole, C. (2022, September 21). How do recommender systems work on digital platforms? *Brookings*. https://www.brookings.edu/articles/how-do-recommender-systems-work-on-digital-platforms-social-media-recommendation-algorithms/

Owen, R., Macnaghten, P., & Stilgoe, J. (2012). Responsible research and innovation: From science in society to science for society, with society. *Science and Public Policy*, *39*(6), 751–760. https://doi.org/10.1093/scipol/scs093

Pawar, A., & Mago, V. (2018). *Calculating the similarity between words and sentences using a lexical database and corpus statistics* (arXiv:1802.05667). arXiv. https://doi.org/10.48550/arXiv.1802.05667

Porter, M. (n.d.). *Snowball*. https://snowballstem.org/

RecSys – ACM Recommender Systems. (n.d.). The ACM Conference Series on Recommender Systems. Retrieved October 25, 2023, from https://recsys.acm.org/

Rogerson, S., & Bynum, T. W. (1996). Information Ethics: The Second Generation. In *The Evolving Landscape of Ethical Digital Technology* (1st ed., pp. 23–31). CRC Press.

Seaver, N. (2021). Seeing like an infrastructure: Avidity and difference in algorithmic recommendation. *Cultural Studies*, *35*(4–5), 771–791. https://doi.org/10.1080/09502386.2021.1895248

Simha, A. (2021, October 6). Understanding TF-IDF for Machine Learning. *Capital One*. https://www.capitalone.com/tech/machine-learning/understanding-tf-idf/

Stahl, B. C. (2011). IT for a better future: How to integrate ethics, politics and innovation. *Journal of Information, Communication and Ethics in Society*, *9*(3), 140–156. https://doi.org/10.1108/14779961111167630

Stahl, B. C. (2013). Responsible research and innovation: The role of privacy in an emerging framework. *Science and Public Policy*, *40*(6), 708–716. https://doi.org/10.1093/scipol/sct067

Stavrakis, G. (2023, September 21). Extracting text from PDF files with Python: A comprehensive guide. *Towards Data Science*. https://towardsdatascience.com/extracting-text-from-pdf-files-with-python-a-comprehensive-guide-9fc4003d517

United Nations Conference on Trade and Development. (2013). The cloud economy ecosystem. In *Information Economy Report* (pp. 1–14). United Nations. https://doi.org/10.18356/63479764-en

Vas, G. (2021, September 21). How Recommendation Systems Comply with Privacy Regulations. *Gravity Research & Development*. https://www.yusp.com/blog-posts/recommendation-systems-comply-with-privacy-regulations/

Volz, D., & Wei, L. (2024, February 28). U.S. to Limit Sales of Americans' Personal Data to China, Other Adversaries. *The Wall Street Journal*. https://www.wsj.com/politics/national-security/u-s-to-limit-sales-of-americans-personal-data-to-china-other-adversaries-e82a3ca5

von Schomberg, R. (2011). *Towards Responsible Research and Innovation in the Information and Communication Technologies and Security Technologies Fields*. European Commission-DG Research and Innovation. https://doi.org/10.2777/58723

Wang, A., & Shah, K. (2021, March 4). Building Riviera: A Declarative Real-Time Feature Engineering Framework. *DoorDash Engineering*.

https://doordash.engineering/2021/03/04/building-a-declarative-real-time-feature-engineering-framework/

Wright, D., Gellert, R., Gutwirth, S., & Friedewald, M. (2011). Precaution and privacy impact assessment as modes towards risk governance. In *Towards Responsible Research and Innovation in the Information and Communication Technologies and Security Technologies Fields* (pp. 83–98). Publications Office of the European Union.

Zhang, B., Wang, N., & Jin, H. (2014). *Privacy Concerns in Online Recommender Systems: Influences of Control and User Data Input*. 159–173. https://www.usenix.org/conference/soups2014/proceedings/presentation/zhang