SSOAR

Open Access Repository
www.ssoar.info

Constants and Variables: How Does the Visual Representation of the Holocaust by AI Change Over Time

Urman, Aleksandra; Makhortykh, Mykola; Ulloa, Roberto; Sydorova, Maryna; Kulshrestha, Juhi

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:
Urman, A., Makhortykh, M., Ulloa, R., Sydorova, M., & Kulshrestha, J. (2023). Constants and Variables: How Does the Visual Representation of the Holocaust by AI Change Over Time. *Eastern European Holocaust Studies*, *1*(2), 365–371. https://doi.org/10.1515/eehs-2023-0055

Nutzungsbedingungen:
Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
https://creativecommons.org/licenses/by/4.0/deed.de

Terms of use:
This document is made available under a CC BY Licence (Attribution). For more Information see:
https://creativecommons.org/licenses/by/4.0

gesis Leibniz-Institut für Sozialwissenschaften

Mitglied der Leibniz-Gemeinschaft

Diese Version ist zitierbar unter / This version is citable under:
https://nbn-resolving.org/urn:nbn:de:0168-ssoar-98467-2

Open Forum

Aleksandra Urman*, Mykola Makhortykh, Roberto Ulloa, Maryna Sydorova and Juhi Kulshrestha

Constants and Variables: How Does the Visual Representation of the Holocaust by AI Change Over Time

https://doi.org/10.1515/eehs-2023-0055
Received November 13, 2023; accepted November 13, 2023; published online November 27, 2023

Today, we can access an unprecedented volume of information about the Holocaust, which is freely available online. It includes historical collections digitised by heritage institutions such as the Shoah Foundation or the Arolsen Archives (Shandler 2022; Stone 2017), reference pages regarding Holocaust events or personalities in online encyclopaedias e.g. Wikipedia (Makhortykh 2017; Wolniewicz-Slomka 2016), and reflections of individuals visiting Holocaust memorials and sharing their experiences via social media, for example Instagram (Hinckley and Zühlke 2022; Zalewska 2017). Unfortunately, this "post-scarcity" (Hoskins 2011) ecosystem of Holocaust memory also includes a multitude of online content offering inaccurate information about specific details of the Holocaust, or promoting a rather one-sided view of this complex historical phenomenon (Grabowski and Klein 2023). In some cases, such content not only gets details incorrectly, but actually propagates antisemitism or denies the Holocaust (Allington 2017; Guhl and Davey 2020).

To cope with the abundance of Holocaust-related information and also to be able to filter out content spreading hate and distorting historical facts, we need new technological solutions. Even if we would like to follow the analogue practices of careful selection and moderation of content used to represent the Holocaust (e.g. Holtschneider 2014; Hansen-Glucklich 2014), these practices are not feasible for digital platforms that must process billions of existing web pages and integrate

***Corresponding author: Aleksandra Urman**, University of Zurich, AND 2.60, Andreasstrasse 15, Zürich 8050, Switzerland, E-mail: urman@ifi.uzh.ch
Mykola Makhortykh and Maryna Sydorova, University of Bern, Bern, Switzerland. https://orcid.org/0000-0001-7143-5317 (M. Makhortykh)
Roberto Ulloa, University of Konstanz/GESIS – Leibniz-Institut für Sozialwissenschaften, Konstanz, Germany. https://orcid.org/0000-0002-9870-5505
Juhi Kulshrestha, Aalto University, Aalto, Finland

Open Access. © 2023 the author(s), published by De Gruyter on behalf of the Babyn Yar Holocaust Memorial Center. This work is licensed under the Creative Commons Attribution 4.0 International License.

information about new ones. Under these circumstances, we cannot rely on human curators anymore, and instead, have to adopt automated solutions that are capable of processing, filtering, and ranking information about the Holocaust in a few milliseconds. These solutions are increasingly powered by artificial intelligence (AI) systems,[1] which take into consideration many factors, such as content features or user engagement with specific content sources in order to prioritise information regarding specific subjects, including the Holocaust.

Search engines, for instance Google or Yandex, are one of such AI-driven solutions which have a major impact on how individuals and societies learn about the Holocaust. By crawling the Internet and processing crawled data, search engines decide what information sources go first when individuals enter search queries dealing with different aspects of the Holocaust (Makhortykh, Urman, and Ulloa 2021; Pfanzelter 2015). While doing so, search engines have to make choices: for instance, if a user is searching for general information about the Holocaust (e.g. "What is the Holocaust" or just "Holocaust"), then what should the first source be? Should it be a Wikipedia article or a museum exhibition webpage? A related question is whether news websites offering information about the recent developments regarding Holocaust memory shall be prioritised over websites of historical institutions? And if the user is searching from a specific location (e.g. Germany), should websites of local or foreign institutions appear in the top search results?

Answering these questions is not a trivial task either for the AI behind the system, which is expected to give the answer to its users, or the system designers developing the AI. The situation is further complicated by changes in what information sources and interpretations the search engines prioritise according to the searched topics. These changes reflect the evolving nature of what the AI powering the search engines sees as particularly relevant in relation to a specific subject. Without accounting for the changes in relevance,[2] the selection of information provided by the search engines can easily become outdated, in particular regarding rapidly evolving topics (Ulloa et al. 2023). However, the changing perceptions of what is or is not relevant make AI systems even less transparent, and also stress the importance of understanding how the relevance of specific pieces of information about the Holocaust for AI changes over time. In other words, what are the constants and variables in the AI vision of the Holocaust?

1 These systems also include generative forms of AI which emerge as an alternative to traditional non-generative AI systems (e.g. search engines; Kelly et al. 2023). However, even generative forms of AI still depend on the prioritisation of the most relevant sources to generate answers; an example of it is perplexity.ai which supplements generated responses with image retrieval.

2 For the more detailed discussion of the concept of relevance in the context of web search, see Mao et al. (2016), Sundin, Lewandowski, and Haider (2022).

To answer this question, we conducted a series of AI audits in 2020 and 2021, aiming to investigate how the visual representation of the Holocaust by search engines evolves over time. The recent addition to the field of algorithm auditing (Mökander 2023), AI audits investigate the performance of AI-driven decision-making systems, such as search engines. In the course of the audits, we were particularly interested in how the perception of relevance changes between search queries in Latin script (i.e. "Holocaust"; same spelling in English and German languages) and in Cyrillic script (i.e. "Холокост"; same spelling in Russian and Bulgarian languages). Our interest in comparing the two was due to the profound differences in Holocaust memory practices in Western Europe (including Germany and the UK) and Russia;[3] hence, we wanted to know whether these differences translate into different perceptions of the Holocaust by AI.

For the practical implementation of the audits, we used a virtual agent-based auditing approach[4] (for a detailed discussion of the method, see Ulloa, Makhortykh, and Urman 2022). For consistency, we deployed our agents using a set of IP addresses located in Germany. We collected data from three search engines – Bing, Google and Yandex – which are among the most frequently used search engines in Western and Eastern Europe. For each search engine, we programmed our agents to enter in Latin and Cyrillic scripts a selection of search queries, which dealt with the different aspects of the Holocaust. For each of the queries, the agents retrieved the top 50 image search results (for consistency, we used .com versions of each search agent). These images were then attributed with the help of authoritative sources (e.g. the United States Holocaust Memorial Museum collections) to determine whether they are related to the Holocaust, whether these are historical evidence or recent images (e.g. photos of contemporary memorials), what aspect of the Holocaust are shown (e.g. liberation of the camps), and from which Holocaust site these images are from.

To illustrate how the perception of relevance in relation to the Holocaust by AI systems changes over time, we share some preliminary observations regarding the outputs for the "Holocaust" query in the Latin script. We find this case particularly interesting because, on occasion of such a general query, AI systems powering search engines have a particularly broad choice. Specifically, we were interested in what facets of the representation of the Holocaust remain stable constants across time, and which variables are subject to change from the viewpoint of the AI.

3 For studies discussing these differences, see Dreyer (2020), Hennebert and Sawkins (2022), Konkka (2023).

4 The virtual agent-based audits rely on simulating human behaviour to generate inputs for the AI system and then record the outputs which then can be analysed. For instance, in the case of search engine audits, this auditing approach usually relies on simulating the process of entering search queries, scrolling the result page to load the results, and then saving the HTML of the result page.

In concrete terms, we looked at what aspects of the Holocaust were shown by top image search results, and which Holocaust sites these images came from. In the case of Holocaust aspects, we observed substantial changes in relevance for Google and Bing. In 2020, the two search engines focused on images of liberated camps, thus reproducing the common pattern in the representation of the Holocaust in the post-WWII period. However, in 2021, these images became substantially less common in top search outputs. Instead, for both search engines, we observed more modern photos showing contemporary Holocaust sites; in the case of Bing, there was also an increase in images showing the evidence of mass murder (contrasted to the focus on the images of deportations in 2020). In contrast to the changes in relevance of Holocaust-related content on Google and Bing, in the case of Yandex, we observed a rather stable prioritisation of images of contemporary Holocaust sites between 2020 and 2021, with rather few historical photos.

In contrast to aspects of the Holocaust, which were treated by two out of three search engines as variables which are subject to change in relevance, we found that the selection of Holocaust sites from which the images were coming was rather constant. With the exception of Bing in 2020, where around 45 % of images came from Auschwitz-Birkenau, between 60 % and 80 % of outputs, independently of the engine and the year, were related to Auschwitz. While there was some variation in the visibility of other sites among the individual search engines – for instance, Bing giving more visibility to images from Bergen-Belsen, and Google with Yandex prioritising content from Buchenwald – the prevalence of Auschwitz-Birkenau was a constant.

While we are still finishing the analysis of the outputs to the "Holocaust" query in Cyrillic for 2021, we expect that it will likely follow its own set of constants and variables. Based on 2020 data, we observed substantive variation compared with the outputs for the query in the Latin script, in particular regarding the visibility of specific Holocaust aspects. Notably, we observed a higher number of images in top search results showing contemporary Holocaust memory sites for Google and Yandex (a trend which was less pronounced on Google in response to the Latin query) and substantially fewer images showing deportations of Jews on Bing. Similarly, we observed some variation in the visibility of individual Holocaust sites, with some sites absent in the search results for the Latin query appearing in response to the Cyrillic query; examples included Ravensbrück for Google and Mauthausen for Bing. At the same time, for all three engines, the prevalence of images from Auschwitz-Birkenau also remained constant also for the Cyrillic query.

These empirical insights contribute to our understanding of how the representation of the Holocaust by AI can change over time, but also raise a number of questions regarding our expectations about constants and variables in what the AI sees relevant for the Holocaust. The core question is whether we expect AI to

reiterate the constants in how the Holocaust is remembered and represented, including the large gap in awareness about Holocaust sites which feature more or less prominently in the popular culture (e.g. Claims Conference 2020), or can we strive for it to advance the current state of popular knowledge about the Holocaust? The answer to this question will influence how we approach other related questions: for instance, shall the representation of the Holocaust by the AI change over time and if yes, then what should the rate of such change be? Shall the AI systems be able to decide on such changes themselves, or should human stakeholders shape their decisions? Furthermore, what can be the impact of the changes in the AI's perception of relevance regarding Holocaust-related changes in Holocaust memory and education practices?

Similarly challenging are the questions regarding how diverse or fair the perception of information relevance by AI should be, and to what degree these concepts are applicable to the case of the Holocaust. There is extensive debate about the importance of embedding the values of fairness and diversity into the AI system design (e.g. Helberger, Karppinen, and D'acunto 2018; Robert et al. 2020; Chi, Lurie, and Mulligan 2021; Madaio et al. 2022). Still, so far, it rarely relates to the AI systems used in the context of genocide memory. To enable such a relation, it would first be needed to define what is meant by diverse or non-diverse (or fair and unfair) Holocaust memory: is, for instance, the higher visibility of images coming from Auschwitz-Birkenau in the outputs of search engines an indicator of the lack of diversity in what AI sees as information relevant for the Holocaust in general? Can the unequal perception of relevance regarding specific Holocaust sites or aspects by AI systems be viewed as a form of unfairness? How can we define the meaning of fairness and diversity in the context of Holocaust memory, and what groups of stakeholders shall be responsible for these definitions? Neither these, nor the earlier questions have easy and available answers. However, we argue that trying to answer them is essential for the future of Holocaust remembrance, which is increasingly shaped by AI, already in the present.

References

Allington, William. 2017. "Holocaust Denial Online: The Rise of Pseudo-Academic Antisemitism on the Early Internet." *Journal of Contemporary Antisemitism* 1 (1): 33–54.

Chi, Nicole, Emma Lurie, and Deirdre K. Mulligan. 2021. "Reconfiguring Diversity and Inclusion for AI Ethics." In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, 447–57. New York: ACM.

Claims Conference. 2020. "First-Ever 50-State Survey on Holocaust Knowledge of American Millennials and Gen Z Reveals Shocking Results." *Claims Conference*. https://www.claimscon.org/millennial-study/ (accessed November 10, 2023).

Dreyer, Nicolas. 2020. "Genocide, Holodomor and Holocaust Discourse as Echo of Historical Injury and as Rhetorical Radicalization in the Russian-Ukrainian Conflict of 2013–18." In *The Holocaust/Genocide Template in Eastern Europe*, edited by Ljiljana Radonić, 63–82. London: Routledge.

Grabowski, Jan, and Shira Klein. 2023. "Wikipedia's Intentional Distortion of the History of the Holocaust." *The Journal of Holocaust Research* 37 (2): 133–90.

Guhl, Jakob, and Jacob Davey. 2020. *Hosting the "Holohoax": A Snapshot of Holocaust Denial Across Social Media*. London: The Institute for Strategic Dialogue.

Hansen-Glucklich, Jennifer. 2014. *Holocaust Memory Reframed: Museums and the Challenges of Representation*. New Brunswick: Rutgers University Press.

Hennebert, Solveig, and Isabel Sawkins. 2022. "Beyond the Normative Understanding of Holocaust Memory: Between Cosmopolitan Memory and Local Reality." In *Youth and Memory in Europe: Defining the Past, Shaping the Future*, edited by Félix Krawatzek, and Nina Friess, 339–48. Berlin: De Gryuter.

Helberger, Natali, Kari Karppinen, and Lucia D'acunto. 2018. "Exposure Diversity as a Design Principle for Recommender Systems." *Information, Communication & Society* 21 (2): 191–207.

Hinckley, Samantha, and Christin Zühlke. 2022. "More Than Meets the Eye–The Intricate Relationship Between Selfies at Holocaust Memorial Sites and Their Subsequent Shaming." *Eastern European Holocaust Studies* 1 (1): 197–214.

Holtschneider, K. Hannah. 2014. *The Holocaust and Representations of Jews: History and Identity in the Museum*. London: Routledge.

Hoskins, Andrew. 2011. "7/7 and Connective Memory: Interactional Trajectories of Remembering in Post-Scarcity Culture." *Memory Studies* 4 (3): 269–80.

Kelly, Dominique, Yimin Chen, Sarah E. Cornwell, Nicole S. Delellis, Alex Mayhew, Sodiq Onaolapo, and Victoria L. Rubin. 2023. "Bing Chat: The Future of Search Engines?" *Proceedings of the Association for Information Science and Technology* 60 (1): 1007–9.

Konkka, Olga. 2023. "'Millions of Russians, Ukrainians, Belarusians, Jews, People of All Ethnicities': Teaching and Remembering the Holocaust in Russian Schools." *Holocaust Studies* 29 (1): 39–65.

Madaio, Michael, Lisa Egede, Hariharan Subramonyam, Jennifer Wortman Vaughan, and Hanna Wallach. 2022. "Assessing the Fairness of AI Systems: AI Practitioners' Processes, Challenges, and Needs for Support." *Proceedings of the ACM on Human-Computer Interaction* 6: 1–26.

Makhortykh, Mykola. 2017. "War Memories and Online Encyclopedias: Framing 30 June 1941 in Wikipedia." *Journal of Educational Media, Memory, and Society* 9 (2): 40–68.

Makhortykh, Mykola, Aleksandra Urman, and Roberto Ulloa. 2021. "Hey, Google, Is It What the Holocaust Looked Like? Auditing Algorithmic Curation of Visual Historical Content on Web Search Engines." *First Monday* 26 (10): 1–24.

Mao, Jiaxin, Yiqun Liu, Ke Zhou, Jian-Yun Nie, Jingtao Song, Min Zhang, Shaoping Ma, Jiashen Sun, and Hengliang Luo. 2016. "When Does Relevance Mean Usefulness and User Satisfaction in Web Search?" In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 463–72. New York: ACM.

Mökander, Jakob. 2023. "Auditing of AI: Legal, Ethical and Technical Approaches." *Digital Society* 2 (3): 1–32.

Pfanzelter, Eva. 2015. "At the Crossroads with Public History: Mediating the Holocaust on the Internet." *Holocaust Studies* 21 (4): 250–71.

Robert, Lionel P., Casey Pierce, Liz Marquis, Sangmi Kim, and Rasha Alahmad. 2020. "Designing Fair AI for Managing Employees in Organizations: A Review, Critique, and Design Agenda." *Human-Computer Interaction* 35 (5–6): 545–75.

Shandler, Jeffrey. 2022. "Digitizing Holocaust Memories." In *Jewish Studies in the Digital Age*, edited by Andreas Fickers, Valérie Schafer, Sean Takats, and Gerben Zaagsma, 25–42. Berlin: De Gruyter.

Stone, Dan. 2017. "The Memory of the Archive: The International Tracing Service and the Construction of the Past as History." *Dapim: Studies on the Holocaust* 31 (2): 69–88.

Sundin, Olof, Dirk Lewandowski, and Jutta Haider. 2022. "Whose Relevance? Web Search Engines as Multisided Relevance Machines." *Journal of the Association for Information Science and Technology* 73 (5): 637–42.

Ulloa, Roberto, Mykola Makhortykh, and Aleksandra Urman. 2022. "Scaling up Search Engine Audits: Practical Insights for Algorithm Auditing." *Journal of Information Science*. https://doi.org/10.1177/01655515221093029.

Ulloa, Roberto, Mykola Makhortykh, Aleksandra Urman, and Juhi Kulshrestha. 2023. "Novelty in News Search: A Longitudinal Study of the 2020 US Elections." *Social Science Computer Review*. https://doi.org/10.1177/08944393231195471.

Wolniewicz-Slomka, Daniel. 2016. "Framing the Holocaust in Popular Knowledge: 3 Articles about the Holocaust in English, Hebrew and Polish Wikipedia." *Adeptus* 8: 29–49.

Zalewska, Maria. 2017. "Selfies from Auschwitz: Rethinking the Relationship between Spaces of Memory and Places of Commemoration in the Digital Age." *Studies in Russian, Eurasian and Central European New Media* 18 (1): 95–116.