

Secondary Publication



Cordero, José Miguel; Henn, Theresa; Holtel, Frederik; u. a.

Developing Effective and Value-Aligned AI Tools for Journalists : 12 Critical Questions to Reflect upon

Date of secondary publication: 18.09.2026

Version of Record (Published Version), Article

Persistent identifier: urn:nbn:de:bvb:473-irb-117254x

Primary publication

Cordero, José Miguel; Henn, Theresa; Holtel, Frederik; u. a. (2026): Developing Effective and Value-Aligned AI Tools for Journalists : 12 Critical Questions to Reflect upon, in: Journalism Practice, London: Routledge, Taylor & Francis, Vol. 20, No. 9, pp. 2776–2795, doi: 10.1080/17512786.2025.2465894.

Legal Notice

This work is protected by copyright and/or the indication of a licence. You are free to use this work in any way permitted by the copyright and/or the licence that applies to your usage. For other uses, you must obtain permission from the rights-holders.

This document is made available under a Creative Commons license.



The license information is available online:

<https://creativecommons.org/licenses/by/4.0/legalcode>

Developing Effective and Value-Aligned AI Tools for Journalists: 12 Critical Questions to Reflect upon

José Miguel Cordero^{a†}, Theresa Henn^{b†}, Frederik Holtel^{c†}, José Á. Sánchez Gómez^{d†}, Diego Arenas^e, Andrea Šipka^{e,f} and Sebastian Vollmer^{e,f}

^aDepartment of Computer Science, University of Chile, Santiago, Chile; ^bDepartment of Information Systems and Social Networks, University of Bamberg, Bamberg, Germany; ^cDepartment of Economics, Rhineland-Palatinate Technical University at Kaiserslautern-Landau, Kaiserslautern, Germany; ^dDepartment of Statistics, University of California at Riverside, Riverside, CA, USA; ^eDepartment for Data Science and its Applications, German Research Center for Artificial Intelligence (DFKI), Kaiserslautern, Germany; ^fDepartment of Computer Science, Rhineland-Palatinate Technical University at Kaiserslautern-Landau, Kaiserslautern, Germany

ABSTRACT

Artificial intelligence (AI) is playing an increasingly prominent role in journalism, with emerging tools like large language models (LLMs) and other AI applications gaining traction. This paper explores the current landscape in adopting AI tools for the information gathering and story development phase in journalistic practice. We identify two contemporary challenges for implementing AI tools in journalism: effectiveness and value alignment. While the first challenge includes the issues of generalizability, costs, user-friendliness, and accuracy, the latter refers to adherence to journalistic values, independence from external actors, transparency, and data privacy issues. We reflect upon these challenges based on the development process of prototyping an open-access AI tool for classifying documents and prioritizing information-rich pages in collaboration with the Organized Crime and Corruption Reporting Project (OCCRP), a globally recognized investigative journalism network. Based on the literature and our experience, we derive twelve critical questions that we deem important to guide the development of future AI tools for journalism.

ARTICLE HISTORY

Received 13 May 2024
Accepted 2 February 2025

KEYWORDS

Journalistic AI; generative AI; investigative journalism; computational journalism; document classification; machine learning

Introduction

Artificial intelligence (AI) tools have undeniably transformed the work of journalists (e.g., de Haan et al. 2022; Helberger et al. 2022; Stray 2019), leading to a “structural transformation of making news and engaging with the audience” (Helberger et al. 2022, 1606). In recent years, the field has been profoundly reshaped in particular due to the widespread public use of large language models (LLM) such as OpenAI’s ChatGPT and Meta’s LLaMA (Pavlik 2023; Shi and Sun 2024). These generative AI models enable journalists to

CONTACT Theresa Henn  theresa.henn@uni-bamberg.de

[†]These authors contributed equally to this work.

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

automate tasks related to content creation in unprecedented ways (Nishal and Diakopoulos 2024). Already before LLMs, many non-generative models had been developed, aiding journalists in addressing challenges such as linking records, automating classification tasks, and extracting meaningful information from heterogeneous documents (Quinonez and Meij 2024; Stray 2019). In general, AI tools are quickly increasing their presence in the journalistic practice, realizing their potential to “disrupt the media industry” (de-Lima-Santos, Yeung and Dodds 2024, 1).

Despite the significant attention that AI models have received in the past years, many journalists remain skeptical, and several journalistic institutions are hesitant to adopt these models (Cools and Diakopoulos 2024; Nishal and Diakopoulos 2024; van Dalen 2024), leading to limited implementation, especially in countries outside of Europe and the US (Gondwe 2023; Pinto and Barbosa 2024). While these tools have been available for several years, research has identified multiple reasons that hinder or complicate their successful implementation in journalism (e.g., de Haan et al. 2022; Stray 2019).¹

We argue in this article that the issues associated with the adoption of generative and non-generative AI tools can be summarized in two conflicting challenges. The first challenge relates to the *effectiveness* of AI tools for journalists. Research indicates that the widespread adoption of AI tools is often constrained by their limited generalizability, which hampers their ability to support the creation of diverse and in-depth stories (Stray 2019). Additional obstacles include high costs, particularly for small news organizations (Borchardt 2022; Stray 2019), and the lack of user-friendly interfaces, which may require specialized computational expertise and thus limit accessibility for many journalists (Borchardt 2022; Stray 2019). Although recent publicly available generative AI models have overcome many of these challenges, they still experience accuracy problems. In particular, these models are prone to hallucinations (Opdahl et al. 2023; Quinonez and Meij 2024; Shi and Sun 2024), which in turn can contribute to the spread of misinformation.

The second challenge in adopting AI tools in journalism pertains to *value alignment*, encompassing concerns related to the values embedded in the tools, as well as issues of independence, transparency, and privacy (de Haan et al. 2022; Gutierrez Lopez et al. 2023; Lin and Lewis 2022; Opdahl et al. 2023; Stark and Diakopoulos 2016). In this context, tool ownership is a critical issue, as many AI tools today are owned by “outside elite actors, such as tech companies” (de Haan et al. 2022, 1776). Often, these organizations do not disclose the tools’ internal details and might design them primarily to maximize engagement and revenue rather than aligning them with journalistic values (Borchardt 2022; de Haan et al. 2022; Diakopoulos 2019; Pinto and Barbosa 2024; Simon 2022). Furthermore, privacy and confidentiality have become especially prevalent concerns in the case of generative AI tools as these models can expose classified and sensitive data (Carlini et al. 2021; de-Lima-Santos, Yeung, and Dodds 2024).

Consequently, contemporary journalism faces the challenge of developing AI tools that are not only effective by addressing the issues related to the first challenge but also value-aligned with regard to the second challenge. Following Helberger et al. (2020, 1607), who proclaimed that “[t]he technological change AI brings may be inevitable; how AI reshapes journalism is not,” we aim to contribute to this line of research by asking: *What are critical questions to consider when developing AI tools that address the needs and values of the journalistic community?*

To address this question, we provide a detailed exploration of the two challenges touched upon above. This exploration is not only informed by reviewing the literature, but also by our experience of developing an AI tool for journalists. More concretely, we prototype an exemplary, open-access, and easy-to-use AI tool that aids journalists in automatically classifying and prioritizing pages in large document databases. To develop our tool, we collaborated with the Organized Crime and Corruption Reporting Project (OCCRP), a worldwide investigative reporting network of independent media centers and journalists in over 40 countries and one of the most important players in international investigative journalism. We use the insights obtained during the development process to derive twelve critical questions that can guide the development of effective and value-aligned AI tools for journalists in the future.

Our work contributes to the literature on the role of AI tools in journalism. First, we consolidate the contemporary discussion of challenges in the implementation of AI in journalism which previous literature has identified (see also Simon 2023; Stray 2019). Second, we advance the discussion by exploring the two challenges drawing from the practical experience of developing an AI tool prototype. We thereby follow a practice-based research methodology, which “emphasizes the study of real-world problems and practices, rather than solely theoretical or abstract concepts” (Fridman, Krøvel, and Palumbo 2023, 4). Our findings offer practical assistance for academics and practitioners alike who seek to create effective AI tools and align the developed product with journalistic values. We view our approach as addressing concerns that the “current discussion on [AI] guidelines ... lacks a crucial element: how these principles translate into action for news companies” (de-Lima-Santos, Yeung, and Dodds 2024, 16). Finally, we also reflect on the role that institutional factors can play in the development of such tools for the journalistic community.

This paper is structured as follows: In our theoretical background, we provide an overview of the status of AI tools within journalism by elaborating on both challenges mentioned above. Our methodology section provides a description of the development of our prototype’s application to OCCRP’s data platform Aleph and its performance results. In our discussion section, we explore how our prototype relates to the debate concerning effective and value-aligned AI tools and derive twelve critical questions for future AI tool development for the journalistic community. Finally, in our conclusion, we summarize our results, highlight the limitations of our approach, and outline potential future research on effective and value-aligned AI tool integration for the journalistic community.

Theoretical Background

The literature (Lin and Lewis 2022) distinguishes three stages where AI enters journalistic practice: (1) information gathering, (2) selection and production, and (3) distribution and consumption. In our study, we focus on AI tools in the first and second stage. However, note that AI tools have become an integral part of each stage of the journalistic process (see, e.g., de Haan et al. 2022; Opdahl et al. 2023; Shi and Sun 2024). In the following sections, we will delve deeper into the issues related to effectiveness and value alignment, drawing on insights from the literature.

Effective AI in Journalism

We identify four issues related to effectiveness that impede and complicate the successful adoption of AI tools in journalism.

The first issue concerns the *generalizability* of AI tools. Research indicates that journalism tasks are often tailored to the specific story being reported on (Stray 2019), resulting in the development of story-specific AI tools (e.g., Poston, Rubin, and Pesce 2015). While these tools may be effective for particular cases, they inherently lack broader applicability and necessitate the collection and curation of new datasets for each story. Conversely, tools that are general enough to be applicable beyond a single story but specific to the problems of journalism (such as fact-checking or document analysis tools) are much less adopted (de Haan et al. 2022, 1783).

Second, developing and implementing AI tools often involves high costs. This issue includes various sub-aspects. First, AI development expertise is costly to obtain in the labor market (Broussard et al. 2019; Hamilton 2016), as newsrooms compete with tech companies and other well-funded major players for hiring and retaining skilled talent (Simon 2022, 1838). Illustratively, by May of 2021, the average yearly U.S. salary for news analysts, reporters, and journalists was \$63,230, while the average software developer salary was nearly twice as much, at \$120,990 (Bureau of Labor Statistics 2021). Second, other inputs for training sophisticated AI tools, such as computational resources and data, are expensive (Simon 2022, 1838). For example, it took investigative reporter Julia Angwin and her team at ProPublica over a year to collect the data required to publish the story “Machine Bias” (Angwin et al. 2016). Third, given that “[t]he financial outlook for most of the news industry internationally remains difficult” (Simon 2022, 1838), especially for local newsrooms (Borchardt 2022), integrating AI tools specifically developed for the journalistic practice will likely continue to be challenging.

Third, AI tools specifically designed for journalists often lack *user-friendly* design because they are frequently developed in-house by newsrooms. Unlike major tech companies, which can leverage vast user data to optimize their platforms through methods like A/B testing (see, e.g., Lin and Lewis 2022, 1636), tools for newsrooms and journalists typically rely on much smaller user bases, which limits their ability to achieve similar levels of optimization. This smaller user base and the overall lower availability of resources for user interface optimization introduce friction for the user, as they impair the user experience, making adoption less likely. Several studies recognize that journalists are more inclined to use user-friendly tools, such as social media and online search engines, than tools designed explicitly for the journalistic research process (de Haan et al. 2022).

Recently, generative AI tools such as ChatGPT have become publicly available online. These new models are highly successful in mitigating issues related to effectiveness, as they are cost-effective (Castillo-Campos, Varona-Aramburu, and Becerra-Alonso 2024), designed with user-friendly interfaces (Pinto and Barbosa 2024), and versatile across a range of text generation tasks (Shi and Sun 2024). However, these LLMs have problems regarding a fourth issue which we call *accuracy*. As it is well known, generative AI models like ChatGPT are prone to hallucinations, i.e., often produce answers that contain outright false or misleading information (Breazu and Katsos 2024; Opdahl et al. 2023; Shi and Sun 2024). While accuracy is a concern for all AI models, hallucinations

are particularly problematic in journalism, as false information can have serious consequences if not identified before publication (Nishal and Diakopoulos 2024).

Hallucination issues are among the top reasons for newsrooms' reluctance to adopt generative AI models. A 2023 study in Brazil, for example, has found that no newsroom in the country had formally adopted ChatGPT (Pinto and Barbosa 2024). Due to the challenges, Bloomberg has reverted to developing and training their own AI models in-house (Quinonez and Meij 2024) to minimize hallucination errors by utilizing more tailored and specific training data. While large news organizations like Bloomberg can afford to develop in-house AI models, smaller newsrooms often lack the resources and capacity to undertake such measures (Borchardt 2022; Shi and Sun 2024, 584; Simon 2022, 1838). This demonstrates that despite the benefits of LLMs, these models can fall short, and vast amounts of resources are still required to adapt these technologies to satisfy the specific needs of journalistic organizations.

Value-aligned AI in Journalism

We identify four issues pertaining to journalistic values, collectively termed value alignment, which impede and complicate the successful adoption of AI tools in journalism.

First, a variety of journalism scholars agree that *journalistic values* such as reporting accuracy, relevance, and diversity (Lin and Lewis 2022) "serve as the ethical backbone of newsrooms" (de-Lima-Santos, Yeung, and Dodds 2024, 4). As an endeavor grounded in these values, journalism is frequently described as the "fourth estate" within democracies (Helberger et al. 2020). In this context, it is crucial to recognize that AI systems are not value-neutral, as they inherently embody their designers' and developers' value judgments, biases, and perspectives. If news organizations and journalists do not incorporate their ideas and values into AI tools, the choices in AI design will be made by other stakeholders like big tech companies (Broussard et al. 2019; Simon 2023). However, the values and interests of these actors might not align with those of journalists (de Haan et al. 2022). For instance, big tech companies tend to prioritize increasing user engagement over other goals by emphasizing personalization and virality.

Moreover, research indicates that these journalistic values can often be compromised in the context of AI models due to biased training data. This is a particular issue in the context of LLMs like ChatGPT, as these models are trained on vast datasets, particularly web-based data, which is inherently biased. Consequently, they tend to reinforce existing stereotypes, perpetuate prejudices, and further marginalize underrepresented groups and regions of the world (Breazu and Katsos 2024; Gondwe 2023; Opdahl et al. 2023; Shi and Sun 2024). Since the output of generative AI models resembles human-written text, biases in the underlying data can have serious implications, as detecting these biases can be challenging (Castillo-Campos, Varona-Aramburu, and Becerra-Alonso 2024). In a journalistic context, the spread of such biases and misinformation is particularly concerning (Gondwe 2023). Thus, to overcome this issue and incorporate journalistic values into AI tool development, Diakopoulos (2019, 181) suggests that news organizations should consider "designing and building the technology themselves so that it better aligns their values with the implementation of the technology."

Second and relatedly, tool ownership and *independence* are crucial issues related to the adoption of AI tools in journalism. The literature has pointed out the importance of

developing AI tools in journalism that are independent of major tech corporations such as Google, Meta, and OpenAI (e.g., de Haan et al. 2022; de-Lima-Santos, Yeung, and Dodds 2024; Simon 2023). On the one hand, journalists might have unique operational needs, necessitating specialized software tailored to their specific domain. Generalist “one-size-fits-all” solutions, like widely applicable search engines (de Haan et al. 2022; Nishal and Diakopoulos 2024), could be inadequate for journalists’ more specific needs. On the other hand, when tools are not owned by journalistic organizations, changes in the models and tools do not have to be reviewed and approved by journalists, effectively leaving them out of control (Simon 2023). For instance, tech corporations often modify their algorithms, as exemplified by the constant changes in Google’s PageRank algorithm or modifications of social media feed algorithms. When such alterations occur, they inevitably impact the work of journalists, yet their ability to influence or mitigate these effects is limited (Simon 2022, 1843–1844).

This issue has implications for journalism’s role in democracy. If major journalistic non-profit organizations and newsrooms are not able to avoid “infrastructure capture” (Simon 2022, 1835) by actors such as big tech corporations, the integrity of quality journalism could be at stake (Helberger et al. 2022). Thus, by increasingly relying on AI tools that are owned by big tech companies, such as search engines and generative AI models like ChatGPT,

news organizations risk becoming even more tethered to platform companies in the long-run, potentially limiting their autonomy and, by extension, contributing to a restructuring of the public arena as news organizations are re-shaped according to the logics of platform businesses. (Simon 2022, 1833)

One response to this issue is to foster “effective collaboration between journalists and AI specialists” (de-Lima-Santos, Yeung, and Dodds 2024, 15). In particular, collaboration of newsrooms and academic researchers can provide major benefits, specially to local organizations (see Pinto and Barbosa 2024; Rinehart and Kung 2023).

Third, the literature raises the concern that AI tools are often “black boxes” that lack *transparency* (de-Lima-Santos, Yeung, and Dodds 2024), impeding a comprehensive understanding of their internal processes (de Haan et al. 2022; Diakopoulos, Zhang, and Salway 2013). This issue is inherent to various machine learning techniques, such as deep learning, and is particularly pronounced for LLMs like ChatGPT, which consist of billions of parameters. This enormous size makes it extremely challenging to gain a thorough understanding of how such models arrive at specific outputs or decisions (Simon 2022). In addition to this lack of technical transparency, much of the code and algorithms utilized by prominent tech companies are proprietary and closely guarded secrets inaccessible to external scrutiny. This lack of transparency impedes external parties from evaluating thoroughly whether biases and errors occur, potentially biasing the journalistic research process for writing stories (de Haan et al. 2022). Only when source code is accessible and organizations can independently inspect an AI tool’s internal architecture is it possible to effectively assess whether it aligns with journalistic values.

The fourth issue centers around *data privacy* and confidentiality. For example, the data or queries shared with online AI tools are often preserved and exploited by the companies that own them (de-Lima-Santos, Yeung, and Dodds 2024), in many cases with the purpose

Table 1. Overview of the challenges we identified.

Challenge	Issues
Effectiveness	Generalizability Costs User-friendliness Accuracy
Value alignment	Journalistic values Dependency Transparency Data privacy

of model retraining. While the preservation of this data may be problematic in general, it is especially concerning when the shared information is sensitive, which is frequently the case in the journalistic context. While all online services pose risks in this regard, generative AI tools are especially vulnerable to this type of privacy concern. Recent work has identified the risk that generative AI can memorize training data and share it with users with the right choice of prompts (Carlini et al. 2021). Thus, journalistic AI tools must have a clear guideline on which data may be preserved for retraining, as well as mechanisms to ensure that the privacy of journalists and the confidentiality of their data are safeguarded in cases where sensitive information is involved.

In summary, journalists face two main challenges when adopting AI tools in their practice (see Table 1). Regarding effectiveness, the development of AI tools can be costly, and the resulting product can often lack features such as user-friendliness and general applicability, which impedes their successful adoption. While some recent generative AI models overcome these problems, they introduce new issues, such as accuracy problems due to model hallucinations. Thus, no AI tool fully embodies the ideal features of effectiveness. Second, AI tools face challenges regarding their alignment with journalistic values. These issues are especially pronounced when these AI tools are owned by large tech companies, which complicates the integration of journalistic values into these models. Generative AI models face particular issues related to value alignment, as they are especially affected by issues related to data privacy and confidentiality.

Methodology

Problem Statement and Research Context

To address the two challenges identified within the literature, we discuss the prototyping of an AI tool for the automatic classification and prioritization of document pages from large databases. We developed our prototype in cooperation with OCCRP, one of the key organizations in the global field of investigative journalism. Since 2011, through OCCRP’s investigative work of unearthing evidence of crime and corruption, 7.5 billion dollars of assets have been seized, and more than 380 official investigations were initiated, leading to 672 government actions and more than 580 arrests (OCCRP 2023). We created our prototype to be integrated into their investigative data platform Aleph, which depicts an archetype of investigative journalism data platforms.

Developing the AI tool prototype presented in this paper is motivated by the following data problem experienced by OCCRP. Their platform, Aleph, provides investigative journalists across the globe access to over 350 million documents of leaked, scraped, and

otherwise collected data (equaling roughly 37 terabytes). Documents come in various file extensions (like pdf, jpeg, or tiff), are of varying page lengths, originate from more than 140 countries, and are thus multilingual (Aleph 2023). Since the archive of documents hosted by Aleph is updated by adding large databases of documents at once (e.g., data leaks), it is currently impossible to systematically capture the document type of all documents in Aleph. Furthermore, due to the heterogeneity and size of the archive, Aleph hosts an unknown variety of document types. Consequently, journalists must manually explore large document databases to extract the valuable document types they need for developing stories, which can be highly time-consuming. This paper's prototype tackles this issue by classifying documents according to their information value and document type, laying the foundation for working with these documents more efficiently. Thus, due to its document diversity, the unknown number of document types, and the large amount of data, Aleph represents a suitable data platform for training and applying our AI tool prototype.

Also, note that given OCCRP's use of highly sensitive data, integrating AI tools from external providers like Google or OpenAI into their pipeline raises confidentiality and privacy concerns. Thus, collaborating with researchers to develop AI tools depicts an approach that overcomes this issue, underscoring the value of partnerships between journalistic organizations and researchers, as highlighted in previous studies (de-Lima-Santos, Yeung, and Dodds 2024; Pinto and Barbosa 2024).

Automatized Document Classification with AI Tool Prototype

This subsection provides an overview of how our AI tool prototype (1) automatically distinguishes between document types and (2) prioritizes between information-rich and information-poor pages. A visual summary of the different process steps within our AI tool is depicted in Figure 1. Note that the AI tool prototype's code is freely accessible at the following GitHub repository: <https://github.com/DSSGxDFKI/occrp-document-classifier?tab=readme-ov-file>.

For our AI tool prototype, we chose an image-based classification approach. Our reasons for this decision are twofold. First, the documents in OCCRP's database are multilingual, yet

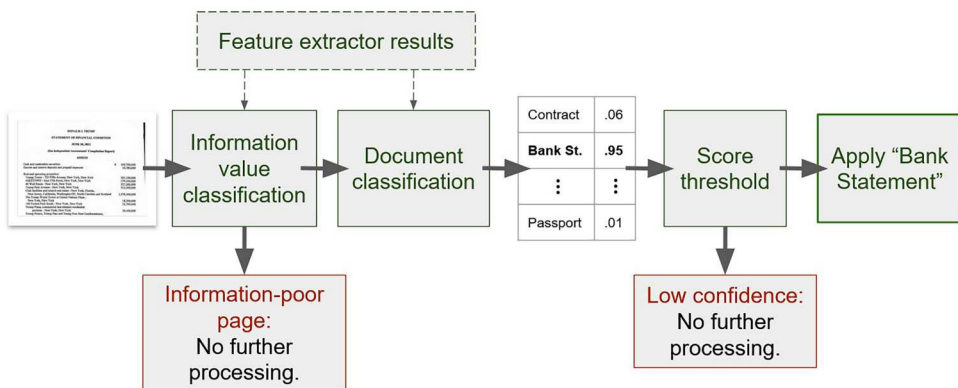


Figure 1. Visualization of the processing steps of our prototype.



Figure 2. Example of documents hosted on Aleph. (a) An exhibit from a US federal court archive (b) A picture of evidence of an International Criminal Tribunal for the former Yugoslavia hearing (c) A photo of a newspaper article.

training data for this variety of languages is scarce. Hence, a language-based classification approach may only provide reliable classifications for documents in the languages most present in the training data and perform poorly in languages not present. Conversely, the visual structure of documents belonging to a given document type often follows a “pre-defined form or standard set of components” (Shin, Doermann, and Rosenfeld 2001, 232) and is unaffected by language heterogeneity. Thus, image-based classification approaches tend to be more generalizable. Second, the format of the documents in OCCRP’s database is highly heterogeneous, as illustrated in Figure 2. This heterogeneity may pose a challenge when extracting the written content of documents.

The first processing step of our AI tool prototype is to preprocess the document data. More specifically, we automatically split the document into individual pages, transform each page into JPG image format, and resize the page image to a consistent pixel resolution of 227×227 .

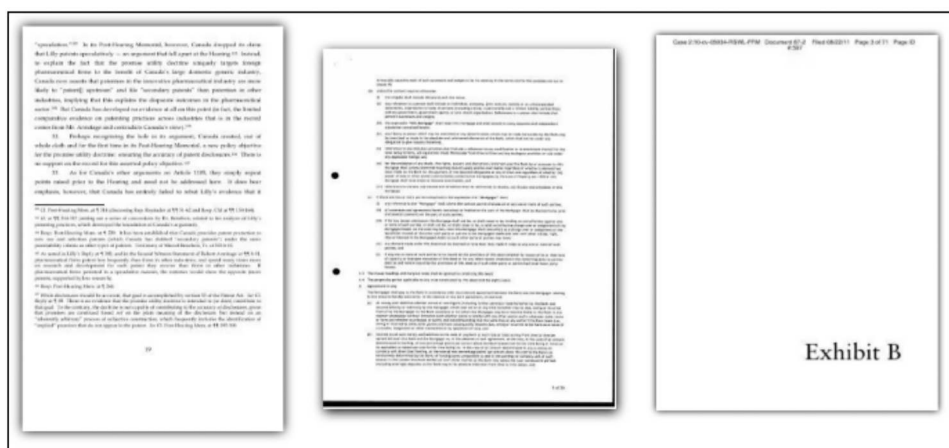


Figure 3. Examples of information-poor pages.

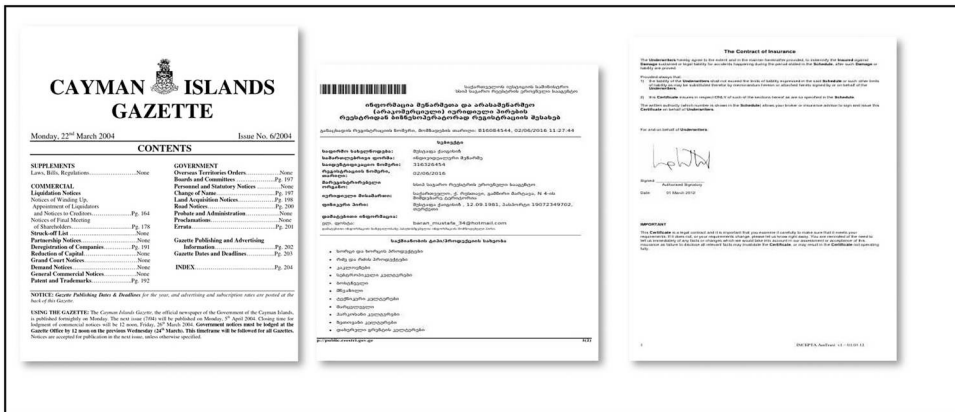


Figure 4. Examples of information-rich pages.

After preprocessing, the second step of our prototype is to identify information-rich pages with an information value classification model. For this classification, we define an information-poor page as a page containing long blocks of text or near-empty pages (see examples of information-poor pages in Figure 3). All other pages, like a person's bank details, are considered information-rich (see examples of information-rich pages in Figure 4). If a document page is identified as information-rich in this step, it will continue to the following classification step. If a page is deemed information-poor, it is omitted in the subsequent steps.

Once information-rich pages are identified, the third step in our AI tool prototype is to classify each page according to its document type. To this end, we train a classification model that classifies nine different document types useful for the journalistic community: bank statements, company registries, contracts, court documents, gazettes, invoices, passports, receipts, and shipping receipts. In this step, each page is assigned nine numeric scores, where each score corresponds to a different document type. The scores represent how well the given page matches each document type, where a higher value for the score of a specific type corresponds to a higher likelihood that the page belongs to this document type.

As previously discussed, the number of document types in the Aleph database is unknown due to its high document heterogeneity. Therefore, it is essential that our classification model recognizes if a document does not belong to one of the predefined classes. This scenario of "dynamic, open environments where some new/test documents may not belong to any of the training classes" (Shu, Xu, and Liu 2017, 2911) is widespread in complex datasets and referred to in the literature as the "open-world assumption." In our last step, we incorporate this consideration in our prototype by contrasting each page's highest score to a threshold value that controls for the quality of our classified document type and can dynamically adjust when adding more document classes. This safety thresholding step ensures that document pages are only classified if there is strong evidence that they are part of one of the document classes the model was trained on. This improves model performance and prevents previously unseen document classes from being wrongfully classified. For more information regarding the control threshold, see Shu, Xu, and Liu (2017).

Data Overview

Data is necessary to train our AI tool prototype and evaluate its performance appropriately. We use a combination of publicly available datasets and newly curated datasets sourced from the Aleph document archive. Through cooperation with OCCRP, we received data from their data platform Aleph corresponding to the nine document types mentioned above.

To improve the quality of our classification results, we performed several cleaning and preprocessing steps on the data provided by OCCRP. First, through several rounds of open coding, we developed a document taxonomy describing the characteristic visual features of each document type. Based on this shared understanding of each document's visual structure, we were able to define and label each document type consistently. We conducted the same process for creating an additional taxonomy to define information-rich and information-poor pages.

Based on our document type taxonomy, we manually selected 100 documents per type from the original data provided by OCCRP. As a result, we curated 900 documents from our nine distinct document types. We refer to this dataset as the "document type dataset." Additionally, we extracted pages from the selected 900 multi-page documents fitting our definition of information-rich and information-poor pages. As a result of this cleaning process, we received 900 information-rich pages (100 pages per document type) and 1233 information-poor pages. We call this dataset the "information value dataset."

Since image-based approaches need large training datasets to generate accurate classification results, we decided to boost our data with image augmentation techniques and pre-training. Image augmentation is a commonly used method that "enhance[s] the size and quality of training datasets" (Shorten and Khoshgoftaar 2019, 1) by, e.g., random zooming, rotating, changing width and height, and modifying the brightness and angle of the images (as an example, see Figure 5).

Another technique to enhance the performance of image-based classification approaches is the pre-training on large publicly available image datasets (Afzal et al. 2017). Pre-training helps an image-based model learn standard image features across contexts and thus aids the model in enhancing its understanding of the image characteristics on which it is trained. This is especially useful when working with limited training data for image classification. We used the widely known Ryerson Vision Lab Complex Document Information Processing (RVL-CDIP) dataset² for pre-training, previously considered in multiple works (Afzal et al. 2017; Kölsch et al. 2017). The RVL-CDIP dataset contains scans of 400,000 document pages, divided into 16 document types. This dataset of 37 GB is highly curated and available in a consistent image format. Therefore, no further processing steps were necessary to use this data to train our AI tool. An exemplary overview of the different classes present in the RVL-CDIP dataset can be seen in Figure 6.

Model Architecture and Training the AI Tool Prototype

For our AI tool prototype's information value assessment and document classification steps, we use convolutional neural network (CNN) models, widely applied in the literature for the visual classification of images and documents (Afzal et al. 2017). We use the



Figure 5. Illustration of three different augmentations.

EfficientNetB4 model architecture, as it performed best among the five architectures we considered.

The first step in our training process is to pre-train the CNN model architecture with the RVL-CDIP dataset, enhancing the prototype's overall performance. The pre-trained models provide the basis for the further training of our two classifiers: the information value classifier and the document type classifier. We retrain the information value classifier on 60% of the observations in the information value dataset and the document type classifier on 60% of the document type dataset. The remaining 40% of each dataset is used for testing and validation. To enhance the variety of images used for training, all document page images used in this training process are augmented as described in the section “Data Overview.”

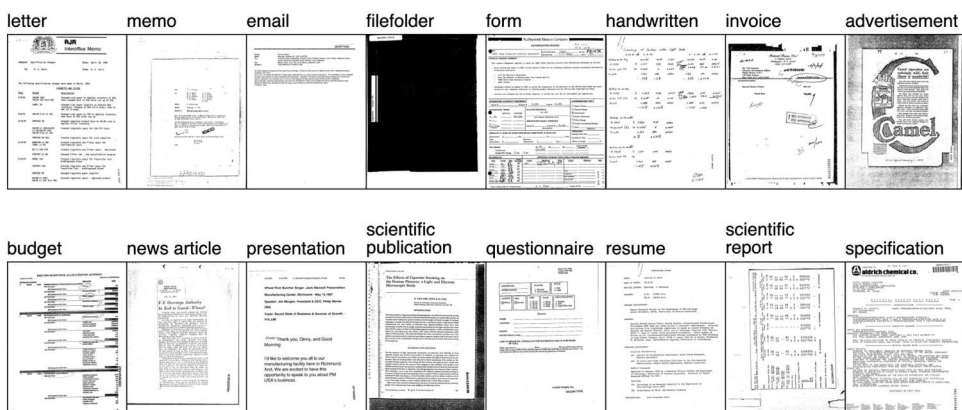


Figure 6. Exemplary overview of RVL-CDIP dataset (Harley, Ufkes, and Derpanis 2023).

Table 2. Accuracy levels of the nine document classes.

Document class	Level of accuracy (in %)
Bank statements	83.33
Contracts	41.37
Company registries	83.33
Court documents	83.33
Gazettes	86.67
Invoices	70.00
Other	100.00
Passports	86.67
Receipts	96.67
Shipping receipts	90.00

Document Classification Results

With the EfficientNetB4 architecture we obtain an overall accuracy of 91.71%. Thus, out of 100 documents, nearly 92 are correctly classified as information-rich or information-poor pages and are subsequently labeled with their correct document class. However, we observe disparate accuracy results depending on the document class (see Table 2). Receipts, gazettes, and passports are classified correctly in nearly all cases, whereas contracts are classified accurately in only about 41 out of 100 cases. One explanation for this observation could be that contracts are quite heterogeneous in their composition since they are utilized in diverse application fields (e.g., financial or human resources contexts). However, since all other classes receive accuracy levels above or equal to 70.00%, and the overall performance is comparable with previous results in this context (e.g., Afzal et al. 2017), our AI tool generates sound results useful for investigative journalists.

Discussion

As demonstrated in the previous section, our prototype can effectively assist journalists in extracting valuable information from large datasets of unlabeled documents. While the previous section focused on the technical development and results of our prototype, the following discussion explores the insights gained from our development experience, particularly concerning the two identified challenges above, thereby addressing our research question.

Creating Effective AI Tools in Journalism

First, when considering the issue of *generalizability* (Stray 2019), our prototype is a valuable tool for writing multiple stories since it currently classifies nine different document types that are important to journalists. As our prototype can be easily extended when trained with new data, it can be updated for new and more general tasks. If retrained on new document data, our AI tool's integrated threshold (Shu, Xu, and Liu 2017) for document label decisions constitutes a significant advantage as it dynamically adjusts to new data, rendering our tool applicable to new contexts.

Second, through our prototype development process, we encountered several strategies that we deem particularly useful for newsrooms with limited financial resources. More specifically, we found that pre-training our prototype with general datasets like

the RVL-CDIP dataset can significantly increase its classification results. Thus, we suggest that organizations like OCCRP could use these publicly available and free-of-charge datasets when creating new AI applications based on computer vision. Moreover, we observed that image augmentation is an effective method to easily enlarge small datasets that can then be used to train AI tools, yielding good results. Lastly, we believe it is essential to explore innovative options for data gathering that go beyond augmentation and pre-training. One promising avenue could be a crowdsourced “human-in-the-loop” approach. Organizations like OCCRP could integrate an interactive labeling system into their data platforms that allows users to give feedback on the model’s predictions (for similar applications, see, e.g., Nadj et al. 2020). This would allow journalists to take part in the creation of meaningful data and improve the cost-effectiveness of the development process of AI tools.

Third, the level of *user-friendliness* of an AI tool is essential for its adoption by a broad range of journalists and practitioners with varying technical skills (Simon 2023). To address this issue, we advocate that journalistic organizations could use their existing infrastructure and incorporate tools as functionalities of these interfaces. For instance, OCCRP’s Aleph is widely used by journalists worldwide as a user-friendly interface to search large databases of documents. Thus, integrating an AI tool like the one presented in this paper into Aleph’s functionalities enables seamless usage while leveraging the already present infrastructure.

Fourth, model *accuracy* is a crucial aspect of any AI model. Our prototype is not designed to generate text and, therefore, does not experience hallucinations. Nevertheless, our development process has provided us with valuable insights into enhancing accuracy in other regards. First and foremost, using a well-curated dataset and ongoing model fine-tuning, our prototype has demonstrated impressive performance, achieving an overall accuracy of over 91%. Second, our integrated classification threshold serves as a “technical safeguard” for our prototype, filtering out data that it cannot classify based on its training set. This mechanism helps prevent erroneous results that may occur due to insufficient or inappropriate training data.

Creating Value-aligned AI Tools

First, our prototype addresses the integration of *journalistic values* such as reporting accuracy, relevance, and diversity in several ways. In terms of accuracy, our prototype demonstrates great performance in correctly classifying documents, as outlined above. With respect to relevance, our first-page classification model effectively differentiates between important and non-important information. Regarding diversity, we have created a prototype that can be easily adapted to new document classes and based on its open-source code, can be inspected, modified, and extended to adhere to organization-specific requirements and preferences.

To incorporate these journalistic values into AI tools, we consider it essential to directly engage with journalists (see, e.g., de Haan et al. 2022 for the merit of conducting interviews with journalists)—the end users of the products. We initiated this process in our project by consulting OCCRP-associated journalists who already use Aleph. Our inquiries aimed to gain insights into their typical workflows, understand how the tool integrates into their tasks, identify appreciated features, and discern any functionalities they

found lacking. Moreover, regular communication with OCCRP staff, who possess a nuanced understanding of the needs of journalistic end users, provided valuable guidance in shaping the tool's design decisions. Soliciting feedback from journalists in areas like document classification enables the incorporation of their interpretations into subsequent tool output. By keeping journalists in the loop, the responsibility for subjective decisions is shared rather than resting solely with developers, enabling both parties to leverage their expertise effectively. This collaboration could foster a "human-in-the-loop" approach, as previously mentioned, where journalists participate in the data labeling process and are thereby able to address issues related to data bias. Hence, we agree with de-Lima-Santos, Yeung, and Dodds (2024, 15) who state that the "successful adoption of AI technologies in newsrooms hinges on effective collaboration between journalists and AI specialists."

Second, our project uncovered principles related to *independence* (see, e.g., de Haan et al. 2022; Simon 2022, 2023). As a major non-profit journalistic organization, OCCRP is arguably more committed to traditional journalistic values than big tech companies. By developing a prototype for OCCRP, we help to reduce journalists' dependence on these corporations.

Generative AI tools such as ChatGPT are, in many cases, owned by a major tech company, such as OpenAI, and integrated into the products of other big tech firms like Microsoft. Since these tools are not open-source, journalists who rely on them are constrained to accept any alterations in design or availability. As Pinto and Barbosa (2024, 336–337) argue, for journalistic organizations in Global South countries, such developments increase the dependency on technology provided by actors in the Global North. In contrast, our openly available prototype provides flexibility, allowing adoption and modification by any journalistic organization into its data processing pipeline.

Third and relatedly, our prototype exhibits high *transparency* compared to many tools commonly relied upon by journalists and other users. In contrast to proprietary tools, including ChatGPT, our prototype is open-source (see GitHub repository). The provided transparency allows anyone to scrutinize the tool's training process and behavior, enabling the identification of potential biases and errors. Organizations can then customize the tool to align with the specific needs of their user base. Thus, this transparency facilitates error detection and supports the development of future functionalities, rendering the software more generalist (Stray 2019).

Fourth, we addressed *data privacy* concerns during the development of our prototype in two ways. First, as our prototype is open-source, it can be set up on a local machine. Journalistic organizations can, therefore, process sensitive documents without the concern that those would be shared with any external actor. Second, as a discriminative rather than a generative tool, our prototype does not generate new data, ensuring that no information is inadvertently shared during its use. This privacy concern associated with LLMs does, therefore, not apply to our prototype.

Implications for the Role of Organizations

The effectiveness and value alignment of AI tools depend heavily on the institutions that are involved in their development and ownership. While large journalistic organizations like Bloomberg News (Quinonez and Meij 2024) have the resources to develop value-

aligned and highly effective AI tools, many small news organizations lack this infrastructure. Therefore, small newsrooms are limited to adopting AI tools owned by large technology companies like OpenAI, Facebook or Google, whose models suffer major shortcomings in their alignment with journalistic values.

With this context in mind, we believe that large journalistic non-profit organizations such as OCCRP, ICIJ, and AP have a unique opportunity to develop and share effective, value-aligned AI tools with the journalistic community. By lending their expertise and collaborating with these non-profit organizations, we consider that academic researchers can play a pivotal role in developing these tools.

Non-profit organizations have great potential for creating effective AI tools. First, these organizations often host vast databases, like OCCRP's Aleph, which can be exploited to develop sufficiently generalist tools that are useful for multiple reporting stories. Second, data from these databases can be exploited cost-effectively for training. As exemplified through our work, even small sets of labeled data can provide accurate modeling results, which can be gradually improved as more data becomes available. Third, some of these organizations presently host platforms with friendly user interfaces and have significant user traction (e.g., Aleph 2023). Major journalistic non-profit platforms can more easily integrate AI tools into their interfaces, allowing their current users to enjoy the benefits of user-friendliness.

Major journalistic non-profit organizations are also exceptionally well suited to create tools aligned with journalists' values. First, these organizations' direct contact with journalists facilitates them obtaining first-hand suggestions from relevant professionals to consider in their AI tool development process. For example, the Associated Press Local News AI initiative (Rinehart and Kung 2023) gathered survey results from more than 200 news professionals in the US and used the provided feedback to design five AI tools for small newsrooms. Second, the independence of non-profit organizations

Table 3. Critical questions for effective and value-aligned AI tool development in journalism.

Challenge	Sub issue	Critical questions
Effective AI	Generalizability	1. Is the AI tool applicable to multiple contexts?
	Costs	2. Are there open-source datasets, AI models, or software provided by, for instance, journalistic organizations that can be used to build or enhance the AI tool?
	User-friendliness	3. Can a crowd-sourced approach be applied to enhance the quality and quantity of the data used for training the AI tool?
Value-aligned AI	Accuracy	4. Is the AI tool designed in a user-friendly way, making it useful for users with different levels of technical skills?
		5. Does the AI tool consistently produce reliable and accurate results in line with journalistic standards?
	Journalistic values	6. Are there technical or human safeguards in place for the AI tool that minimize erroneous results, and, to the extent possible, mitigate their impact?
		7. Does the AI tool reflect journalistic values such as reporting accuracy, relevance, and diversity?
	Independence	8. Were journalists involved and consulted in the development process?
		9. To what extent do journalists and newsrooms own the AI tool they use (data, code, infrastructure)?
	Transparency	10. How are journalists affected by changes to components that they do not own?
	Data Privacy	11. Is the AI tool's source code publicly available and well-documented, making it easy to understand and open to scrutiny?
		12. Have precautions been taken to prevent or mitigate the risk of the model sharing sensitive information from users with actors external to the journalistic organization?

allows them to align their tools with the needs of journalists, avoiding conflicts of incentives commonly faced by other actors such as tech companies.

While non-profit organizations possess data, infrastructure, and incentives for value alignment, the computational training and maintenance of AI tools can constitute a challenge. In this regard, academic researchers can intervene by sharing their technical knowledge for these tasks. While our prototype serves as an example, similar partnerships can be found elsewhere. The Associated Press Local News AI initiative (Rinehart and Kung 2023) created collaborations between university researchers and small newsrooms to aid the development of AI tools for journalists across the US. In Brazil, a review of the literature found that academia is the third-largest contributor to AI adoption in journalism (Pinto and Barbosa 2024). We hope academic researchers and non-profit journalistic organizations can further recognize this opportunity and seek collaborations in developing AI tools for a more robust and independent journalistic practice.

Critical Questions for the Development of AI Tools in Journalism

Based on these observations, we derive the following twelve questions that we deem important when developing further effective and value-aligned AI tools for journalists (see Table 3). Likewise, these questions can be used to evaluate the usefulness of existing AI tools and their alignment with journalistic values.

Conclusion

In this article, we develop a set of critical questions that may guide future development to create effective and value-aligned AI tools in journalism. Our analysis is grounded in the development process of an AI tool prototype designed to categorize documents by type and to distinguish between information-rich and information-poor pages within a journalistic context. Regarding the effectiveness aspect of AI tools for journalism, we suggest that tools should be generalist, cost-effective, user-friendly, and accurate. Concerning the value alignment aspect, we propose that the tools should reflect journalistic values, be independent of companies that do not prioritize journalistic values, be transparent, and address data privacy issues. Regarding the type of organization, we argue that major journalistic non-profit organizations like OCCRP are in the best position to lead the development of such tools.

We see several avenues for future research. First, regarding our prototype's technical advancements, it seems promising to create a document classifier based on large language models able to handle multiple types of inputs (i.e., PDFs and images). For example, given that GPT4o has demonstrated exceptional performance in other domains, it might be incorporated into a document classification AI tool. While pursuing such an approach might result in high effectiveness (being accurate while not costly to set up), note that this would raise several of the concerns related to value alignment outlined in this paper (being non-transparent and dependent on a large tech firm). Our objective in this study was not to create a competitive product but to draw general insights from the experience of developing a concrete, effective, and value-aligned AI tool prototype. We believe that the insights about which aspects to consider and which critical questions to ask when developing an AI tool for the journalistic community are generally applicable.

Our main contribution, which is on the conceptual level, remains valid despite the most recent technological developments.

Second, we see great potential to more rigorously analyze institutions' role in developing successful AI tools in journalistic practice (for a similar discussion, see Simon 2023). The results of our discussion show that major journalistic non-profit organizations like OCCRP could be the leading actors in implementing AI tools in journalism. As a final remark, we want to re-emphasize the role that researchers can play in developing AI tools for journalism. There are great opportunities for non-profit journalistic organizations to collaborate with interested academics. If researchers and journalistic organizations work together, it might be possible to overcome the "gravitational force" that pulls news organizations into the "orbit" of major tech companies (Simon 2022, 1850).

Notes

1. For recent work on design strategies that should be considered when implementing AI tools in the journalistic workflow, please see Gutierrez Lopez et al. (2023).
2. For accessing the dataset, see <https://adamharley.com/rvl-cdip/>

Disclosure Statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported as part of the Data Science for Social Good fellowship hosted by the Rhineland-Palatinate Technical University Kaiserslautern-Landau in cooperation with the German Research Centre for Artificial Intelligence (DFKI GmbH).

Data Availability Statement

The data supporting this study's findings are available from the Organized Crime and Corruption Reporting Project's data platform Aleph. Restrictions apply to the availability of these data, which were used under license for this study. Data is available from the authors upon request with the permission of the Organized Crime and Corruption Reporting Project.

References

- Afzal, Muhammad Zeshan, Andreas Kölsch, Sheraz Ahmed, and Marcus Liwicki. 2017. "Cutting the Error by Half: Investigation of Very Deep CNN and Advanced Training Strategies for Document Image Classification." In *Proceedings of the 14th International Conference on Document Analysis and Recognition*, 883–888. Kyoto, Japan: IEEE.
- Aleph. 2023. "OCCRP Aleph: The Global Archive of Research Material for Investigative Reporting." *Aleph*. <https://aleph.occrp.org>.
- Angwin, Julia, Jeff Larson, Surya Mattu, and Lauren Kirchner. 2016. "Machine Bias." In *ProPublica*. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
- Borchardt, Alexandra. 2022. "Go, Robots, Go! The Value and Challenges of Artificial Intelligence for Local Journalism." *Digital Journalism* 10 (10): 1919–1924. <https://doi.org/10.1080/21670811.2022.2149584>.

- Breazu, Petre, and Napoleon Katsos. 2024. "ChatGPT-4 as a Journalist: Whose Perspectives Is It Reproducing?" *Discourse & Society* 35 (6): 09579265241251479.
- Broussard, Meredith, Nicholas Diakopoulos, Andrea L. Guzman, Rediet Abebe, Michel Dupagne, and Ching-Hua Chuan. 2019. "Artificial Intelligence and Journalism." *Journalism & Mass Communication Quarterly* 96 (3): 673–695. <https://doi.org/10.1177/1077699019859901>.
- Bureau of Labor Statistics. 2021. "Occupational Employment and Wage Statistics." May 2021 Report. Bureau of Labor Statistics. https://www.bls.gov/oes/2021/may/oes_stru.html
- Carlini, Nicholas, Florian Tramèr, Eric Wallace, Matthew Jagielski, Ariel Herbert-Voss, Katherine Lee, Adam Roberts, et al.. 2021. "Extracting Training Data from Large Language Models." In *Proceedings of the 30th USENIX Security Symposium*, 2633–2650. Vancouver, B.C., Canada: USENIX Association.
- Castillo-Campos, Mar, David Varona-Aramburu, and David Becerra-Alonso. 2024. "Artificial Intelligence Tools and Bias in Journalism-Related Content Generation: Comparison between Chat GPT-3.5, GPT-4 and Bing." *Tripodos* 55:99–115.
- Cools, Hannes, and Nicholas Diakopoulos. 2024. "Uses of Generative AI in the Newsroom: Mapping Journalists' Perceptions of Perils and Possibilities." *Journalism Practice*, 1–19. <https://doi.org/10.1080/17512786.2024.2394558>.
- de-Lima-Santos, Mathias-Felipe, Wang Ngai Yeung, and Tomás Dodds. 2024. "Guiding the Way: A Comprehensive Examination of AI Guidelines in Global Media." *AI & Society*, 1–19.
- de Haan, Yael, Eric van den Berg, Nele Goutier, Sanne Kruijkemeier, and Sophie Lecheler. 2022. "Invisible Friend or Foe? How Journalists Use and Perceive Algorithmic-Driven Tools in Their Research Process." *Digital Journalism* 10 (10): 1775–1793. <https://doi.org/10.1080/21670811.2022.2027798>.
- Diakopoulos, Nicholas. 2019. "Towards a Design Orientation on Algorithms and Automation in News Production." *Digital Journalism* 7 (8): 1180–1184. <https://doi.org/10.1080/21670811.2019.1682938>.
- Diakopoulos, Nicholas, Amy X. Zhang, and Andrew Salway. 2013. "Visual Analytics of Media Frames in Online News and Blogs." In *Proceeding of the IEEE InfoVis Workshop on Text Visualization*, 15–16. Atlanta, Georgia, USA: IEEE.
- Fridman, M., R. Krøvel, and F. Palumbo. 2023. "How (Not to) Run an AI Project in Investigative Journalism." *Journalism Practice*, 1–18. <https://doi.org/10.1080/17512786.2023.2253797>.
- Gondwe, Gregory. 2023. "CHATGPT and the Global South: How Are Journalists in Sub-Saharan Africa Engaging with Generative AI?" *Online Media and Global Communication* 2 (2): 228–249. <https://doi.org/10.1515/omgc-2023-0023>.
- Gutierrez Lopez, Marisela, Colin Porlezza, Glenda Cooper, Stephann Makri, Andrew MacFarlane, and Sondess Missaoui. 2023. "A Question of Design: Strategies for Embedding AI-Driven Tools into Journalistic Work Routines." *Digital Journalism* 11 (3): 484–503. <https://doi.org/10.1080/21670811.2022.2043759>.
- Hamilton, James. 2016. *Democracy's Detectives: The Economics of Investigative Journalism*. Cambridge, Massachusetts, USA: Harvard University Press.
- Harley, Adam W, Alex Ufkes, and Konstantinos G. Derpanis. 2023. "The RVL-CDIP Dataset." <https://adamharley.com/rvl-cdip/>.
- Helberger, Natali, S. J. Eskens, M. Z. van Drunen, M. B. Bastian, and J. E. Möller. 2020. "Implications of AI-Driven Tools in the Media for Freedom of Expression." In *Artificial Intelligence–Intelligent Politics: Challenges and Opportunities for Media and Democracy*, 1–34. Nicosia, Cyprus: Council of Europe.
- Helberger, Natali, Max van Drunen, Judith Moeller, Sanne Vrijenhoek, and Sarah Eskens. 2022. "Towards a Normative Perspective on Journalistic AI: Embracing the Messy Reality of Normative Ideals." *Digital Journalism* 10 (10): 1605–1626. <https://doi.org/10.1080/21670811.2022.2152195>.
- Kölsch, Andreas, Muhammad Zeshan Afzal, Markus Ebbecke, and Marcus Liwicki. 2017. "Real-time document image classification using deep CNN and extreme learning machines." In *Proceedings of the 14th IAPR International Conference on Document Analysis and Recognition*, 1318–1323. Kyoto, Japan: IEEE.

- Lin, Bibo, and Seth C. Lewis. 2022. "The One Thing Journalistic AI Just Might Do for Democracy." *Digital Journalism* 10 (10): 1627–1649. <https://doi.org/10.1080/21670811.2022.2084131>.
- Nadj, Mario, Merlin Knaeble, Maximilian Xiling Li, and Alexander Maedche. 2020. "Power to the Oracle? Design Principles for Interactive Labeling Systems in Machine Learning." *KI-Künstliche Intelligenz* 34 (2): 131–142.
- Nishal, Sachita, and Nicholas Diakopoulos. 2024. "Envisioning the Applications and Implications of Generative AI for News Media." In *CHI'23 Workshop on Generative AI and HCI*. Hamburg, Germany: ACM.
- OCCRP. 2023. "Impact to Date." <https://www.occrp.org/en/impact-to-date>.
- Opdahl, Andreas L., Bjørnar Tessem, Duc-Tien Dang-Nguyen, Enrico Motta, Vinay Setty, Eivind Throndsen, Are Tverberg, and Christoph Trattner. 2023. "Trustworthy Journalism through AI." *Data & Knowledge Engineering* 146:102182. <https://doi.org/10.1016/j.datak.2023.102182>.
- Pavlik, John V. 2023. "Collaborating With ChatGPT: Considering the Implications of Generative Artificial Intelligence for Journalism and Media Education." *Journalism & Mass Communication Educator* 78 (1): 84–93. <https://doi.org/10.1177/10776958221149577>.
- Pinto, Moisés Costa, and Suzana Oliveira Barbosa. 2024. "Artificial Intelligence (AI) in Brazilian Digital Journalism: Historical Context and Innovative Processes." *Journalism and Media* 5 (1): 325–341.
- Poston, Ben, Joel Rubin, and Anthony Pesce. 2015. "LAPD Underreported Serious Assaults, Skewing Crime Stats for 8 Years." *Los Angeles Times*. <http://www.latimes.com/local/cityhall/la-me-crime-stats-20151015-story.html>
- Quinonez, Claudia, and Edgar Meij. 2024. "A New Era of AI-Assisted Journalism at Bloomberg." *AI Magazine* 45 (2): 187–199. <https://doi.org/10.1002/aaai.12181>.
- Rinehart, Aimee and Ernest Kung (2023, October). "Associated Press' Local News AI Initiative." Associated Press. <https://www.ap.org/solutions/artificial-intelligence/local-news-ai/>
- Shi, Yi, and Lin Sun. 2024. "How Generative AI Is Transforming Journalism: Development, Application and Ethics." *Journalism and Media* 5 (2): 582–594.
- Shin, Christian, David Doermann, and Azriel Rosenfeld. 2001. "Classification of Document Pages Using Structure-Based Features." *International Journal on Document Analysis and Recognition* 3 (4): 232–247. <https://doi.org/10.1007/PL00013566>.
- Shorten, Connor, and Taghi M. Khoshgoftaar. 2019. "A Survey on Image Data Augmentation for Deep Learning." *Journal of Big Data* 6 (1): 60. <https://doi.org/10.1186/s40537-019-0197-0>.
- Shu, Lei, Hu Xu, and Bing Liu. 2017. "DOC: Deep Open Classification of Text Documents." In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, edited by Martha Palmer, Rebecca Hwa, and Sebastian Riedel, 2911–2916. Copenhagen, Denmark: Association for Computational Linguistics.
- Simon, Felix M. 2022. "Uneasy Bedfellows: AI in the News, Platform Companies and the Issue of Journalistic Autonomy." *Digital Journalism* 10 (10): 1832–1854. <https://doi.org/10.1080/21670811.2022.2063150>.
- Simon, Felix M. 2023. "Escape Me If You Can: How AI Reshapes News Organisations' Dependency on Platform Companies." *Digital Journalism* 12 (2): 149–170. <https://doi.org/10.1080/21670811.2023.2287464>.
- Stark, Jennifer A, and Nicholas Diakopoulos. 2016. "Towards Editorial Transparency in Computational Journalism." *Computation + Journalism Symposium* 5:1–5.
- Stray, Jonathan. 2019. "Making Artificial Intelligence Work for Investigative Journalism." *Digital Journalism* 7 (8): 1076–1097. <https://doi.org/10.1080/21670811.2019.1630289>.
- van Dalen, Arjen. 2024. "Revisiting the Algorithms Behind the Headlines. How Journalists Respond to Professional Competition of Generative AI." *Journalism Practice*, 1–18. <https://doi.org/10.1080/17512786.2024.2389209>.