

Who are the publics engaging in AI?

Public Understanding of Science

2024, Vol. 33(5) 634–653

© The Author(s) 2024



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/09636625231219853

journals.sagepub.com/home/pus



Renée Sieber  and **Ana Brandusescu** 
McGill University, Canada

Abigail Adu-Daako 
University of California, Berkeley, USA

Suthee Sangiambut 
McGill University, Canada

Abstract

Given the importance of public engagement in governments' adoption of artificial intelligence systems, artificial intelligence researchers and practitioners spend little time reflecting on who those publics are. Classifying publics affects assumptions and affordances attributed to the publics' ability to contribute to policy or knowledge production. Further complicating definitions are the publics' role in artificial intelligence production and optimization. Our structured analysis of the corpus used a mixed method, where algorithmic generation of search terms allowed us to examine approximately 2500 articles and provided the foundation to conduct an extensive systematic literature review of approximately 100 documents. Results show the multiplicity of ways publics are framed, by examining and revealing the different semantic nuances, affordances, political and expertise lenses, and, finally, a lack of definitions. We conclude that categorizing publics represents an act of power, politics, and truth-seeking in artificial intelligence.

Keywords

citizen, engagement, human-in-the-loop, participant, participation, public, stakeholder

1. Introduction

A curious problem emerged as we investigated public participation in artificial intelligence (AI). As expected, authors reported that AI was opaque, complex, and required a high degree of numeracy. Comprehension was out of reach; not only to the general public but also to the majority of researchers and policymakers (Burrell, 2016). Achieving greater public understanding of AI was important because of massive investments in AI by both the public and private sectors and because of AI's impacts upon our daily lives. Consequently, there was considerable interest in assessing public

Corresponding author:

Renée Sieber, McGill University, 805 Sherbrooke Street West, Montreal, QC H3A 0B9, Canada.

Email: renee.sieber@mcgill.ca

perceptions of AI, for instance, by socio-demographics or literacy levels (Zhang and Dafoe, 2019). However, something was still missing. Were we talking about the public? The citizen? The stakeholder? All terms appeared throughout literature that spoke to participation vis-à-vis AI but terms connoted slightly different affordances, inclusions, and exclusions. The public functioned as individual citizens questioning automated decision-making or as consumers of AI services (e.g. Simonofski et al., 2017). Deployment of AI engaged members from the public in roles ranging from contributing code to analyzing model outcomes and functioning as data sources (Sloane et al., 2020). AI deployment continues to differentially affect members of the public (Benjamin, 2019). We argue that, alongside interrogations of who is left out and who is affected, we should examine the semantic, ontological, and epistemological levels of who this public, citizen, or human is in AI.

Explication of impacts, harms, and biases represents a multidisciplinary effort (Dwivedi et al., 2021; Raley and Rhee, 2023). Discourses tend to originate from computer science, philosophy, and law, but rarely social science. Diversity complicates concrete conceptions of those who are neither policymakers nor researchers who study participation and participants. Without examining the ontological or epistemological, it may be left to AI developers to interpret what is meant by citizen and those affected. Problems in identifying members of the public predates AI. Fung (2006) is well-known for his model of citizen participation vis-à-vis the state and he specifically brings the public into governance (Fung, 2015). However, a definition of the public is absent.

Concrete definitions of the participant/citizen/public in AI are important because of which characteristics of individuals are deemed important. Choosing “participant” infers some act of participation. The term participation itself is problematic; participation is “an infinitely malleable concept, [which] can be used to evoke—and to signify—almost anything that involves people” (Cornwall, 2008: 269). Calling someone a stakeholder assumes a significant interest. Use of citizen as a placeholder term infers legal rights. The AI literature, compared to prior literature on Information and Communications Technology (ICT)-related engagement, introduces a role for humans in the specific case of human-in-the-loop (HITL). A lack of grounding definitions also obscures political power. Whereas exploitation and accountability are of enhanced concern to the AI industry, researchers, and regulators; common use of abstract governance framings and vague notions of publics can entrench and obfuscate existing power structures (Whittaker, 2021). Reducing such abstractions becomes crucial to countering potential co-opting of the AI discourse.

No term fully encompasses participant/citizen/public. We choose the term publics to capture various ways they are considered in the AI literature from these diverse disciplines. Our object of study therefore is the characterization of actor and not the act within this broad literature scan. We first examine past literature reviews of public participation¹ in AI, where researchers and practitioners likely describe publics. Next, we outline a mixed methodology for a structured literature analysis on publics in AI. We present our findings on 12 types of publics. We conclude by discussing the epistemological and ontological implications of these varied terms.

2. Past literature reviews on publics in public engagement in AI

No literature directly examines the subject of publics regarding AI. We assumed that, of all the literature concerned about AI impacts, literature considering AI-related participation would contain the most refined conceptualizations of various publics. Table 1 summarizes the key focus of past literature reviews and any mentions of publics.

These literature reviews identified participation as integral to a fair, transparent, and inclusive AI; however, authors rarely defined publics beyond generic terms such as public and citizen. Participation could be services based with individuals as beneficiaries or victims (de Sousa et al., 2019); transaction-based, where individuals engaged in dialogue with government (Zuiderwijk et al., 2021); or

Table 1. Instances of definitions and characterizations of “publics” in key literature reviews on the use of AI for public participation.

Authors	Focus	Main findings related to participant/public/citizen/ human in AI
Jobin et al. (2019)	Assessed 84 AI ethics guidelines and principles created by private companies, research institutions, and public sector organizations.	Mentioned participants or the public. Main focus was stakeholders, which included civil society, government, researchers, and developers, for example, with their role being to “act with integrity.” Emphasis also on educated/uneducated publics.
Kwecko and Botelho (2018)	Described collective intelligence in smart cities using three approaches: technological, human resources, and citizen-related governance.	Primarily referred to citizen and population; also included human, individual, community, social group, and people (e.g. a “dialogue between the citizen and the city of information and knowledge”). No definitions provided.
Reis et al. (2019)	Compiled impact of AI on public administration, shifting from resource efficiencies to citizen involvement.	Used term citizens who become “co-producers and sharers of knowledge” to improve decision-making. Presumed AI will demand citizens to become “future-ready innovators” whose special and flexible skills can adapt to work with smart machines.
Savaget et al. (2019)	Focused on civil society empowerment and political participation through AI, although they find articles mentioning policymaking or politics avoid political participation in or influence of AI over democratic regimes.	Mentioned attributes of citizens and civil society groups and how they can use AI to counter government. Included recommendations for sustaining independent AI. Discussed crowdsourced contributors (including coders) who are exclusively online and unknown.
Sharma et al. (2020)	Extolled AI for effective governance, which simplifies bureaucratic processes and improves service delivery and citizen-state interactions.	Preferred term citizen, who is the recipient of state benefits (threats to individual privacy remain concerning). Government can use technology to “keep citizens safe.”
Simonofski et al. (2017)	Considered citizen participation in smart cities, where smartness sometimes is equated with AI.	Referred to citizens by their affordances, democratic participants, co-creators, and users of information and communications technologies.
de Sousa et al. (2019)	Discussed AI in the public sector, by country and government function (e.g. healthcare, public administration, decision-making).	Presented mostly positive and a few negative impacts on citizens and communities: “AI aggregates value for citizens in several areas, from infrastructure to public security.” Promoted AI as technology to sense community well-being.
Zuiderwijk et al. (2021)	Described AI in public sector governance, including participative governance through public consultation to support AI standards.	Listed main actors, who are “individuals, citizens, organizations, and systems of organizations in public, private, and nonprofit sectors,” although these are not defined. Mentioned stakeholders.

“maker”-based, with citizens as contributors who could optimize AI processes (Reis et al., 2019; Savaget et al., 2019). Articles such as Vanolo (2016) emphasized a rights-based approach, where publics should have the right of access to information or a right to privacy (European Union, 2022). Simonofski et al. (2017) provided among the few reviews that identified a broad range of publics, for example democratic participants and knowledge co-creators.

Reviews focused on AI applied in the public sector (e.g. Simonofski et al., 2017), citizen participation in smart cities since smartness is associated with technologies like AI (e.g. Kwecko and Botelho, 2018), distillations of AI ethics (Jobin et al., 2019), implications of AI in governance (Zuiderwijk et al., 2021), and political empowerment through AI (Savaget et al., 2019). Characterization of publics appeared to vary by object of study. For example those conducting smart city research conceived of publics as individual citizens interacting with a technologically enabled locality. Jobin et al.’s (2019) analysis of 84 guidelines of ethical AI drew on national or global entities. These large-scale entities conceived of publics as stakeholders who could include almost everyone: private sector actors, government employees, and AI designers. Those who study AI governance (e.g. Zuiderwijk et al., 2021), Gurumurthy and Chami (2019) argued, tended to depoliticize the role of the publics. This could incentivize policy makers to seek out publics without political agendas and or knowledge of AI. One missing voice is the computer/data scientist, the builder of AI (e.g. Zuiderwijk et al., 2021 explicitly excluded them). AI developers may invoke various publics but receive little guidance from these literature reviews.

Gurumurthy and Chami’s (2019) critique also implies that political orientation affects choice of publics, for instance, framing publics as part of neoliberalism or a participatory democracy. In the first, a co-producer or “future-ready innovator” invokes an AI that aggregates value for and from citizens (Reis et al. 2019). In a participatory democracy, citizens actively participate in decision-making processes and have the potential to influence policy outcomes. We unpack the implications in the choice of representations.

3. Methods

A literature analysis encompassing AI and all references to various publics generated millions of potential articles, especially when the goal was a cross-disciplinary overview. To accommodate this breadth, we conducted a structured literature analysis using a combination of automated keyword querying and manual review, following Sacha et al. (2016). We outline this process in Figure 1. The process comprised the following seven steps: (1) determine the essential reading list, (2) create initial filter keywords, (3) develop search strings, (4) collect a corpus of articles (e.g. input search strings to collect articles), (5) classify the corpus, (6) refine the corpus (<https://osf.io/6tjb8/>) and identify articles to include and exclude, and (7) interpret a refined corpus.

Search string generation

We implemented a series of filters to arrive at our final corpus. We first developed a set of keywords that would form the search strings. This was the *search string generation* portion of the structured literature analysis (Figure 1). To achieve these keywords, we analyzed an essential reading list of 77 articles, which represented core debates on participation and AI (Step 1). Using these articles to seed the keywords, and ultimately search strings, gave us confidence that our searches would return results representative of similar discourses across the broader literature. The 77 articles were analyzed algorithmically using a bigram model, which approximates the probability of a word given all the previous words by using only the conditional probability of one preceding word. After filtering out frequent low-value words (e.g. “and”), we selected the 100 most frequent words.

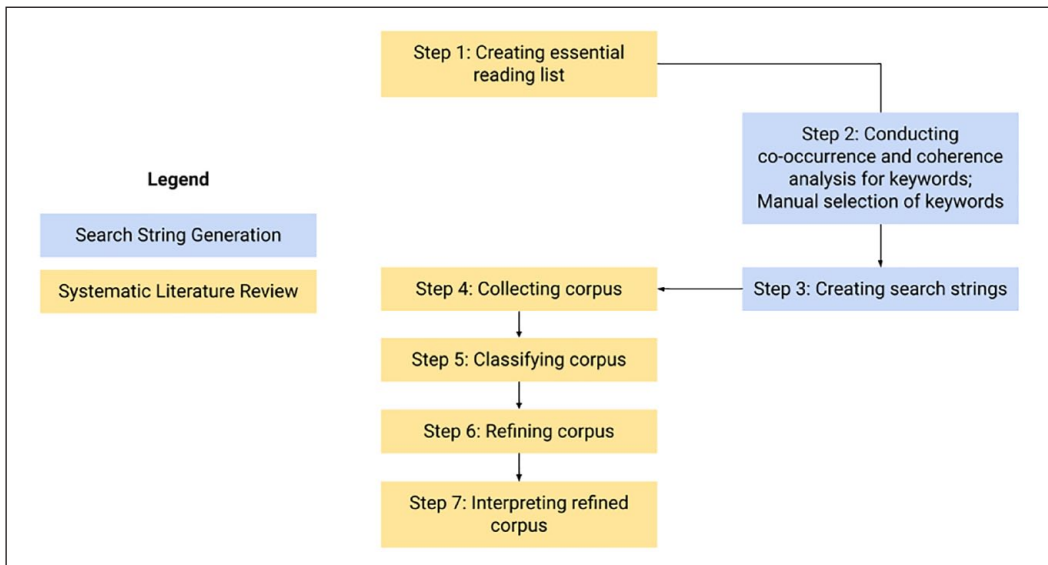


Figure 1. Overview of the structured literature analysis for corpus processing and interpretation.

These 100 keywords were manually classified into the following two categories: technology keywords and participation keywords. We employed VOSviewer 1.6.16² to visually inspect word co-occurrence and find word clusters, especially frequent phrases. Manual classification resulted in 16 technology keywords (e.g. “machine learning”), 27 participation keywords (e.g. “democratic participation”), and eight mixed keywords (e.g. “data governance”; Step 2). From these keywords, we developed 36 search strings (Step 3).

Search strings combined technology phrases including “Artificial Intelligence,” “Machine Learning,” “Algorithmic/Algorithm(s),” “Automated Decision(s).” Participation phrases included “Public Participation,” “Citizen(s)/Civic Participation,” “Political Participation,” “Public Engagement,” “Citizen(s)/Civic Engagement,” “Political Engagement,” “Democracy/Democratic,” “Governance/government,” and “Data Governance.” Searches returned both peer-reviewed articles and gray literature. We retained gray literature (e.g. reports and non-peer-reviewed articles) due to their popularity as a medium among both academic and non-academic authors. Search strings were used to query Google Scholar, which returned 3533 articles (step 4).

Systematic literature review

Step 5 (Figure 1) consisted of classifying our corpus. We chose to classify articles (Statistics Canada, 2005) to provide an overview of the multidisciplinary of our corpus (Figure 2). The two dominant categories were computer science (452 articles) and public administration and public policy (483 articles). Law contributed 245 articles. Articles offering technical methodologies and applications to multiple sectors and fields also are included (primarily health/medicine, communication/media studies, political science, sociology). Overall our corpus contained articles from 1969 to 2022, noting that the majority of the articles were published from 2010 to present (2467 of 2656 articles).

We chose to group articles based on title, not author(s) disciplines. For instance, articles from the conferences on Fairness, Accountability and Transparency in Socio-Technical Systems were categorized as computer science because proceedings were published by the Association for

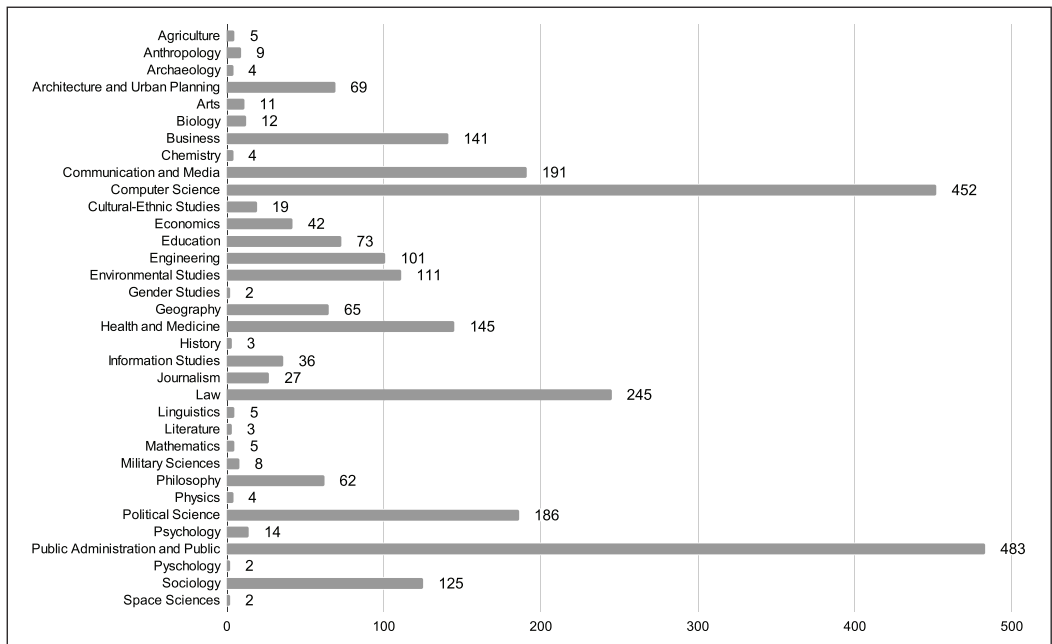


Figure 2. Diversity of corpus. Classification count of articles per Classification of Instructional Programs (CIP) Canada 2000 categorization. The X-axis corresponds to the count of articles and the Y-axis to the classification categories.

Table 2. Results of step 6 in the structured literature analysis.

Step name	Criteria/methods applied	Number of articles reviewed
Search results	Removal of duplicates and irrelevant articles: 877	3533
Articles screened by title and metadata	Availability in English; format is a journal article, conference paper, report or gray literature piece; focus is on AI/ML	2656
Articles screened on abstract	Focus on AI/ML and public participation	160
Full-text articles assessed for eligibility	Focus on AI/ML and public participation	138
Articles included for analysis	Review of full articles	121

AI/ML: artificial intelligence/machine learning.

Computing Machinery. We argue that those articles will invariably be guided by the ontologies and epistemologies of the association. We acknowledge the limitations of assigning publications to specific fields.

In Step 6 (Figure 1), we refined our corpus. After removing duplicates and irrelevant articles, we conducted manual expert filtering (Sacha et al., 2016) of 2656 article titles and abstracts. We made judgment calls to filter articles based on the level of relevance to some form of AI and participation (with cross-validation of each choice done by two of the authors). Table 2 shows the substeps of

Table 3. Examples of included and excluded articles based on their titles.

Title	Decision	Reasoning
Participation is not a design fix for machine learning	Include	Explicitly mentions ML and participation
Public participation—Its challenges in Hong Kong and some suggestions	Exclude	Not discussing technology elements of public participation
Artificial intelligence and the future of humans	Exclude	Too broad
Public participation in urban design based on information technology	Exclude	Does not specifically mention AI/ML, nor gives impression this will be focus

AI/ML: artificial intelligence/machine learning

manual filtering conducted in Step 6, which resulted in a final corpus of 121 articles. Table 3 shows some examples of titles we included or excluded.

To analyze the 121 articles (Step 7), we coded articles according to words related to publics (e.g. stakeholder, citizen, consumer). We collectively inferred definitions and affordances of what constituted publics. We compared instances across the corpus and compared corpus definitions of publics to definitions found in our essential reading list. Each substep was cross-validated by at least two authors.

4. Results

We draw on our methods to uncover the numerous ways that publics were discussed in our corpus. Our object of study is the actor, not the act. We follow DiSalvo (2009), who argued that we cannot identify or understand the publics' actions until they are first constituted as a public. In our results, we seek to explicate the blurriness of publics and implied affordances of their "publicness."

Table 4 shows the wide variety in how publics were described in our corpus. In the aggregate, they could be the public or community. As individuals, they included citizens (unarticulated, although presumably of a state/jurisdiction), consumers of public services, or residents of a smart city.

Reliance on the most common terms

Similar to the section "Past literature reviews on publics in public engagement in AI," authors in our corpus utilized the terms citizen or public generically to describe individuals or groups outside the public and private sectors. The challenge of finding exact definitions in the corpus suggests that authors either assumed that terms required no definition because they were universally understood or definitions were left to AI practitioners (e.g. AI developers or policymakers) to operationalize.

The term public vis-à-vis AI often appeared in polls, which gauge attitudes, expectations, opinions, and levels of support of the technology (Ipsos, 2022; Zhang and Dafoe, 2019). Polls have offered insight into socio-demographic differences in reactions to AI (Zhang and Dafoe, 2019). Contrary to being purely descriptive, the public has frequently functioned as a consenting or dissenting body, serving as a foundation for affirming the trustworthiness of AI (Ipsos, 2022). The public also could prioritize aspects of AI ethics. Kieslich et al. (2022) conducted a survey to rank public perception of AI ethical design principles, with accountability as the most important of the AI design principles, closely followed by fairness, security, privacy, accuracy, and explainability. The public as supplier of feedback offers a way for governments to manage trade-offs in AI ethics principles, yet public polling has limits. Cows (2020) argued that pollsters' attempts to "strive for

Table 4. Differences in how publics are perceived as engaged in AI.

Term	Working definitions	Example authors	Authors' framings
Public	Non-formalized collection of people often residing in a defined geographic bounding.	Balaram et al. (2018) and Wilson (2022)	Assumes a universal, instrumentalized, and homogenized understanding without acknowledging the existing diversity. Embeds geographic bounds (nation, city) (especially in smart cities literature). Overlaps with the non-expert/lay public. Can reduce public input to feedback, which is amenable to natural language processing.
Citizen	Legally a person who belongs to a country and has associated rights.	Mehr (2017), Vanolo et al. (2016), and Warne et al. (2021)	Implies a relationship with the state. Can exclude non-citizens like refugees, undocumented, and stateless people. Includes a sample of representatives, bounded by geography and demography, in a participatory process like a citizen jury.
Participant	Individual involved in various degrees in the decision-making process (e.g. part of the design process, critiquing implementation).	Simonofski et al. (2017)	Includes individuals acting as representatives of civil society. May also include representatives of civil society organizations. Conflates actor with act and often assumes active deliberation. Can also be passive (e.g. when individuals' data are harvested).
Consumer	End-user or recipient of a product of an AI system or platform. Also a person purchasing and/or accessing services from an online marketplace of service providers.	Darmon (2019)	Frames the individual in a market relationship with the state. Relegates the state to that of supplier of services and individuals as clients or users of services. Excludes individuals who cannot consume (e.g. homeless, poor). Assumes that people can access knowledge about AI systems (i.e. the "educated" consumer), which can increase inequity.
Co-creator, co-producer, worker	Individual who provides feedback, shares information, and contributes knowledge to the development of AI systems.	Reis et al. (2019), Shadowen et al. (2020), and Sloane et al. (2020)	Assumes individuals are active partners in AI system development. Implies an equality in knowledge and status and a flattening of hierarchical decision-making. In some cases, treats individual as crowdworker of data or services (e.g. annotating training data or moderating toxic content).
Human-in-the-loop (HITL)	Computational models in which a human agent exercises some degree of meaningful influence on the AI system.	Brennan-Marquez et al. (2019) and Wiethof and Bittner (2021)	Refers primarily to workers, who may or may not be compensated. Includes government officials and decision-makers who provide oversight. Could be a supplement to the machine, an extension of the machine, or a compensation for limitations of automation.

(Continued)

Table 4. (Continued)

Term	Working definitions	Example authors	Authors' framings
Lay person, non-technical expert, unprofessionalized citizen	Regular or ordinary person, who offers unbiased, authentic opinions, speaks from heart, and is not agenda driven. Requires guidance from technical experts.	Cardullo and Kitchin (2019) and Park et al. (2015)	Excludes professionalized citizens (e.g. activists with technical and policy knowledge) or public interest lobbyists. Also excludes professional associations and industry lobbyists. Can privilege individuals based on their lack of (computational or policy) knowledge. Often characterized by "voting and venting" behaviors. Amenable to literacy initiatives.
Impacted individual or community	Marginalized individual or group impacted by AI systems.	Bondi et al. (2021), Henman (2020) and Vogl et al. (2020)	Includes minorities, marginalized, poor, oppressed, vulnerable, disadvantaged, and social groups. Can rob agency (e.g. ability to participate) from individuals because they are treated as subjects of AI harms. Can invoke civic paternalism because it leaves open who decides (e.g. AI developers or employers).
Community	Group of people who can be geographically bound or collectively of share values, interests, and/or consumption patterns.	Kone (2021)	Can erase geographic bounding of community. Can leave undefined concepts of collective, shared values, or interests.
Stakeholder	Individual, representing a group, who has an interest as defined by an external party. Consults at various stages of the process to identify context-specific needs.	Sloane et al. (2020), Jobin et al. (2019), Wong (2018), Abdala et al. (2020) and Spielkamp (2019)	Includes communities of practice like data scientists, experts, and industry lobbyists. Includes affected individuals as well as decision-makers, AI developers, and standard-setting experts. Can obscure role of civil society when using the term "multi-stakeholder" in reference to stakeholder engagement. Assumes a neutral party determines who possesses a legitimate interest. Leaves open how inclusive and meaningful the engagement is.
Resident	Person who can afford the rent or purchase price to live in a particular area or district.	Cardullo and Kitchin (2019)	Includes renters or homeowners, with a preference toward the latter. Usually excludes the homeless, poor, and uneducated who do not have access to residency. Is geographically bounded.
Data subject	Usually derived from European Union's GDPR, refers to any living individual/legal person whose personal data is collected, held, or processed by an entity.	Jabbari Jahromi et al. (2019) and Kalluri (2020)	Defines individuals as the data collected about them. Can frame individuals along specific roles (e.g. conflating individuals with data producers or as citizens with the right to be forgotten).

GDPR: General data protection regulation.

overall representativeness” (p. 108) in assessing perceptions of AI, may homogenize the public into a single entity and further what he called a democratic deficit.

Far more effort in the literature is devoted to characterizations of individuals (i.e. citizen, public, resident, democratic participant, and also civil society). Where a definition of citizen occurred, it was largely drawn from what publics afford AI deployment (Mehr, 2017). Characterizations could evoke aspects of citizenship, those civil, political, and social rights and responsibilities that accompany membership in a national community (Vanolo, 2016). Instead of a static definition derived from legal status, Vanolo (2016) preferred a cosmopolitan definition where citizenship would recognize “membership in all cross-cutting political communities, from the local to the global” (p. 8). Conversely, Cardullo (2020) expressed citizenship in socio-technical systems as digital rights “expressed, inhibited, or enhanced through digital platforms or electronic voting; civic rights shaped and regulated deeply through algorithmic processing” (p. 62). The concern with a rights-based approach was the inability to keep pace with technical disruptions and private sector power (Fukuda-Parr and Gibbons, 2021).

Predominance of the neoliberal

Figure 2 shows the importance of the disciplines of public administration and public policy in our corpus. Unsurprisingly, many of those authors situated citizens in relation to the speed and efficiency that AI promises for public sector service delivery (Battina, 2017; Mehr, 2017). We found that, reinforced by Cardullo and Kitchin (2019), AI could easily remake citizens who have rights and are entitled to receive services, into consumers who select government services through a market. In a market-based framing, public engagement becomes transactional and expressed through individual actions like purchases, feedback, or boycotts (Wilson, 2022). As governments adopt AI systems for service delivery, the citizen transitions into an impersonalized user, with government-to-citizen interactions focused on “resolving inquiries,” “trouble-shooting network inactivity,” and “enabling a better customer experience” (Darmon, 2019: n.p.). A market-based framing extends to citizens as co-creator and co-producer in AI deployment and policymaking (Shadowen et al., 2020). Neologisms like co-producers often conceal demotions of publics from democratic participants engaged, for instance, in social justice, to participants-as-worker (Sloane et al., 2020).

Publics also may be valued for their ability to contribute to collective intelligence (Verhulst et al., 2019), a networked zeitgeist of productive knowledge. These buzzwords envisage individuals as “future-ready innovators who uniquely combine specialization and flexibility” (Reis et al., 2019: 242), which suggests ideal participants will embody entrepreneurial values contained in an innovation discourse. Collective intelligence is emblematic of the inevitable overlaps in any discrete definitions of publics. For Simonofski et al. (2017), a participant could manifest neoliberalism (consumer, co-producer) as well as a participatory democracy (citizen, resident, community). There are inevitable limits to making definitions of terms discrete, that can obscure power differentials, including resources and skills like literacy. HITL is one such example.

STEM confidence in human-in-the-loop

In contrast to prior reviews (e.g. Zuiderwijk et al., 2021), we included articles from core disciplines such as computer science and engineering, which play a central role in dictating the AI discourse. When opening up the corpus to AI developers, one primary concept emerges, HITL.

HITL arose from human factors and human-computer interfaces research and was originally described as “a decision-making system in which the initial triage or categorization of cases is performed by a machine, but a human agent exercises some degree of meaningful influence”

(Brennan-Marquez et al., 2019: 749). HITL positions publics as essential components of AI systems, a form of human–computer interaction in which the system is not fully automated.

Generic terms like human in HITL can be disciplinarily dependent. To policymakers, the human is a decision-maker who ensures a degree of oversight over the AI's results (Kuziemski and Misuraca, 2020). For Binns (2022), “each [human] is situated in widely varying social and structural conditions with different powers, pressures, motivations, and agendas affecting how they exercise judgment in ways which are unlikely to conform to expectations of the algorithm's designers” (p. 9).

Despite this reference to publics, HITL models can misrepresent human oversight, where a person appears to be in the loop but actually is not. Such was the case in Poland, where an AI profiling system was found to have less than 1% of AI-made decisions questioned by human clerks, despite instituting an HITL advisory body. In addition, the system was found to be unconstitutional for its potential to discriminate, despite being subject to the law (Kuziemski and Misuraca, 2020). A misalignment between appearance and actuality of human involvement could create a false sense of public trust. The human's role is potentially marginal in most decisions but we are reassured by the appearance of human oversight.

We found science, technology, engineering, and mathematics (STEM) professionals to conceive of the human as a worker, given microtasks in platforms like Mechanical Turk, or an uncompensated contributor, for instance, in answering reCaptcha. The work tends toward the mechanical, distanced in a crowdsourcing platform with anonymous contributors who collectively train, classify, and refine data in AI systems. The humans may not be acknowledged or adequately compensated. They are not overseers (Binns, 2022) but can be reduced to dehumanized ghost workers, for instance, labeling social media content to train classification systems (Gray and Suri, 2019; Roberts, 2019). The complication is that approaches like HITL can become conflated with participatory and inclusionary discourses and enable a kind of “participation-washing” (Sloane et al., 2020), for instance, with content moderators.

Notwithstanding these concerns, researchers have found utility in expanding HITL to community-in-the-loop (Häußermann and Lütge, 2022) or society-in-the-loop (Rahwan, 2018). These authors suggested we move beyond the origins of HITL, which employ individuals' input to optimize AI systems and toward embedding societal values and stakeholder deliberations into AI systems. Similar to crowdsourcing and its inception as anonymous outsourcing (Howe, 2008), unacknowledged implications and assumptions can linger. These definitions continue to frame humans *within a system*, where humans are prone to exploitation as labor. What of humans outside the loop?

Those left out

Even though expansions to HITL promise greater inclusion of actors, there are still publics who are excluded. This finding cuts across our categories.

An overreliance on the uncomplicated member of the public reflects what Cardullo (2020) termed the “general citizen.” For Cardullo (2020), the general citizen “is largely framed as white, male, heterosexual, ablebodied and middle class” (p. 65). This general usage of citizen invokes and is in direct contrast with an “‘absent citizen’, referring to all those diverse communities that hold identities, values, concerns, and experiences.” To counter the generic use of citizen, Vanolo (2016) recommended the inclusion of the marginalized, minoritarian or even subaltern, who are deprived of credibility and agency. Benjamin (2019), referred to Our Data Bodies project (<https://www.odbproject.org>) and termed the left out “[t]hose who are typically watched but not seen” (p. 189). Molnar and Gill (2018: 7) classified some who are subject to AI surveillance as “people on the move,” where “even well-intentioned policymakers are sometimes too eager to see

new technologies as a quick solution to what are otherwise tremendously complex and intractable policy issues.” These policy decisions are seen in how refugees, for example, are being used as testing subjects for biometric systems.

Sieber and Brandusescu (2021) conducted a workshop during which AI ethics researchers and data scientists had difficulty establishing a common definition for the citizen, while noting many terms to be exclusionary, especially when people affected by the system fell outside the description bracket or jurisdiction. Workshop attendees suggested expansions beyond a traditional demographic, for instance, to include youth and the elderly, and to incorporate those who function to counter the state or big tech (e.g. journalists, whistleblowers). To the term citizen, attendees suggested the addition of the formerly incarcerated, undocumented, stateless, disabled, and wards of the state. Arguably, these publics should have a stake, which leads us to the term stakeholder.

Sheer breadth of actors

We found a significant variety of actors in our corpus, among them stakeholders. Stakeholders were ubiquitous in the literature on AI ethics but similarly lacked definition. The category has long been argued to have power due to its conceptual breadth (Phillips et al., 2003). Phillips et al. (2003) continued that “The term means different things to different people and hence evokes praise or scorn from a wide variety of scholars and practitioners” (p. 479).

For many AI researchers, the term stakeholder was a catch-all for publics, including nearly anyone who expressed an interest in the AI decision-making process (Jobin et al., 2019). Stakeholders included those with political power: decision-makers, government officials and industry lobbyists. Stakeholders also included the disempowered: impacted groups, vulnerable, victims, targets, the profiled, or those who were discriminated against (e.g. Fukuda-Parr and Gibbons, 2021). Individuals afforded stakeholder status functioned as stand-ins for their identity group, which could reinforce a performative check-box action on the part of a facilitator. Designating who did and did not have an official stake could diminish the role of community organizations, who themselves could provide representatives (Bondi et al., 2021). Definitional breadth diluted the influence of affected individuals and hid significant imbalances in resources (Sambuli, 2021).

Part of the breadth of actors stems from jurisdictionally dependent terms like resident, although people outside particular jurisdictions might still be affected or implicated. Publics could be labeled as subjects of AI systems (Jabbari Jahromi et al., 2019). In data protection and privacy laws, the data subject is any individual, for example, under the European Union’s General Data Protection Regulation (EU GDPR, 2016). Anyone is potentially subject to AI harm, not just the marginalized. As with stakeholders, this generalization could render the concept relatively meaningless and exploitable by those with power and knowledge (Cofone, 2020). The concept of data subjects results in a datafication of publics (Micheli et al., 2020; Posada et al., 2021). Ultimately, individuals could become a series of pixels that are everywhere (i.e. data is repurposable and shareable) and nowhere (i.e. publics are dematerialized). The data subject suggests a position of vulnerability (“we are subject to”); however, authors also noted possibilities for empowerment (Kalluri, 2020), including the ability to contest decisions (Almada, 2019) and gain expertise.

Affordances of expertise

The most detailed explication of publics in the corpus referenced levels of expertise, which acknowledged the important affordance of knowledge and skill development. Unlike other definitions of publics (e.g. public, consumer), researchers like Havrda (2020: 440) provided detailed characterizations of relevant expert stakeholders: governments, well-being researchers, AI developers, communication

experts, video game developers, practitioners and local decision-makers, and activists. These technical experts, “create efficiencies and use accreted knowledge to tackle issues in which citizens may have little experience or knowledge” (Cardullo and Kitchin, 2019: 4). Expertise creates a binary of experts and non-experts, which is prevalent in knowledge-heavy AI. In this formulation, technical experts and facilitators control who is the expert and which expertise is valued. Those without technical expertise may find the value of their lived experiences diminished and their role lessened.

Contrast the above with labels of the public such as non-professionalized, non-technical, or lay persons. Labels align with sentiment in the AI ethics and the human–computer interaction research community, where a fundamental predicate for public engagement in AI is an increase in technical literacy. This includes skills and competencies required for the “learner” (who ranges from a school student to the broader public; Long and Magerko, 2020; Ng et al., 2021; Ylipulli and Luusua, 2019) and the role that public spaces and libraries play in developing public AI literacy (Watkins, 2020). Finland’s highly ambitious public education program on AI targets a public whose affordance can shift from uneducated to varying levels of expertise. The temptation of these literacy efforts can be the formation of an AI-ready workforce (Wilson, 2022), where humans educated under national training programs may be viewed as economic agents more than democratic actors. Occasionally, an educated public is not always desired by the state. Policymakers may want a lay public: individuals who provide authentic feedback (Verhulst et al., 2019) without preconceived positions or political agendas. The lay public can simultaneously imply a characterization of which types of expertise are and are not valued as well as strategies to ameliorate deficits.

5. Discussion

Our corpus uncovered numerous terms to describe publics. Rather than focus on only one term, we considered the diversity and changing dynamics of these terms, especially since a diverse corpus introduced new terms. What is revealed is that those who write in the AI literature present fundamental ontological shifts and epistemological shapings in what we consider to be publics in AI.

Ontological shifts

How do we define the public vis-à-vis AI ontologically? Concepts are intertwined with affordances (e.g. a participant is one who participates). Avoiding explication of terms by shifting to different terms only introduces new challenges. Relying on the term citizen embeds complex and potentially unarticulated associations to citizenship, which can create whole categories of absent citizens (Cardullo, 2020). We are left with assumptions of homogeneity, representation, identity, or (economic) value generation. Not only do we see a flattening within categories but across categories, for instance, creating equivalencies between human and public. Flattening blurs power dimensions and elides movement from one category to another (e.g. from consumer to citizen or from local citizen to global citizen).

What definitively is a citizen? A citizen may or may not be linked with human and digital rights. Reliance on human rights frameworks like GDPR remains problematic because frameworks are frequently vague and aspirational, presume universality, and display the unevenness of legally binding compliance with those rights (focusing on process as opposed to outcome, cf., Miller and Redhead, 2019). Smuha (2021) found rights-based frameworks to be too Western and individualistic and to impose a narrow unit of analysis. For example, an American or European approach emphasizes individual rights to privacy, resulting in individuals as the primary data subjects who receive and benefit from governance and regulatory regimes, as opposed to communities and collective rights to privacy (e.g. First Nations Information Governance Centre, 2020). Relying on

Western frameworks can be too narrow because they position rights in relation to the state instead of the private sector, the latter being the primary developer of AI systems (Smuha, 2021). Viewing the connection of public with properties and affordances as inexorable can render members of the public always in the service of external forces, whether these forces manifest in the AI developer, government convener, or the algorithm itself.

We cannot easily determine definitions because categories we assign to publics blur, which creates significant implications for affordances. Numerous articles (e.g. Darmon, 2019) conflated citizens with customers. Under a neoliberal framing, publics shift toward consumers of government services made more efficient by AI and the government as a business focused on optimization. In this shift, objections to AI are articulated, for example, through consumer boycotts instead of political elections. In an effort to label publics as co-producers or co-creators of the state, these roles can underestimate the potential for exploitation, particularly of marginalized participants (cf. Tönurist and Surva, 2017). The term data subjects obscures the fact that we are talking about data contributors, digitizing their lives. In this commodification of human as data, publics are as much suppliers as they are victims (e.g. of privacy violations). Neither can we ignore ghost workers when the emphasis shifts from citizen to consumer (Gray and Suri, 2019). These unacknowledged workers comprise the invisible members of the public constructing AI's training data and shielding our eyes from harmful content (Roberts, 2019). HITL is implicated in this shift and cannot be separated from discussions of publics as AI aids in this neoliberal shift.

Epistemological shapings

Categorizing publics represents an act of power, politics, and truth-seeking. Here, we examine which new terms are introduced in disciplines and how organizations leverage categories of publics to shift the boundaries of what is accepted as truth. In certain cases, technologies socially shape those notions. For instance, who decides that important publics are those with a stake in a technology or that all publics are humans in HITL? HITL is an interesting example of how classification systems interweave people and technology.

Classifications are not merely neutral organizational tools but have profound consequences for individuals and communities (Bowker and Star, 2000). Through acts of defining, sorting, and ranking (Suchman, 1993), classification systems can both differentiate and discard people (Liboiron and Lepawsky, 2022). Such systems embody motivations of control and disciplinary action (Suchman, 1993), while disguising their provenance by removing identifiers of things and people (Latour, 2005). AI expresses this power and control dimension with techniques like unsupervised classification, which offers a prime tool for large-scale passive harvesting of participants. The method appears objective due to automation but obscures the direct role of developers in removing words and topics and controlling selection of labels. In this, developers can use AI to influence which voices are heard and prioritized.

Researchers in our corpus appeared to normalize certain categories, such as the public, as naturally occurring entities or as the only way to define individuals or groupings of individuals. The notion of a naturally occurring public has long been revealed as socially constructed, which arises from social relations as opposed to containing inherent, immutable qualities. One hundred years ago, philosopher John Dewey and journalist Walter Lippman debated the ability of the public to engage in decision-making; however they largely agreed that the public “emerges once its members are affected by a problem . . . (the public) can never exist in the abstract, but must be thought of as entangled with an issue” (Monsees, 2019: 45). In this way, the public is not simply a set of characteristics of a group, but rather the result of relations between and within groups working on a common issue. As mentioned earlier, publics must be formalized before their actions can be

understood and analyzed (DiSalvo, 2009). In addition to the rejection of the public as a natural entity, science and technology studies researchers have rejected the public as a homogeneous entity. A lack of a single public generated calls for “hybrid forums” (Callon et al., 2011), composed of actors with diverse interests, perspectives, and forms of expertise. Here, instruments like AI-driven platforms could be used to facilitate public interactions, although those technologies also can nudge publics toward outcomes that align with powerful interests (Monsees, 2019).

Categories like citizen or the public existed before AI but they can be shaped by AI. For instance, AI decenters geography as a natural way to understand publics by disrupting the notion of a public as an aggregation of citizens delineated by jurisdiction. A jurisdiction aggregates individuals that achieve an identity as a physical community. A spatial logic remains relevant as long as publics are jurisdictionally bound (e.g. electoral districts and political representation). However, AI and its impacts often transcend such boundaries.

AI can enable the passive and unbounded enrollment of publics into systems (e.g. through crowdwork or passive harvesting of data subjects). Publics are not distinct from the technological processes in which they are involved—they are mutable cyborgs (Haraway, 1991). Automated systems increasingly intervene in social relations and individuals, such as the (conscious and unconscious) use of filtering algorithms in online searches and social networks, and prevalence of personal assistants. Gillespie (2014) argued that we risk a “calculated public,” a positive feedback loop in which algorithms present publics back to themselves and in turn shape publics’ view of themselves. For us, AI may constantly shape and reshape who is affected. Lippmann’s and Dewey’s points are therefore amplified by AI where innovations in AI create new publics.

6. Conclusion

We conducted a structured literature analysis to create a diverse corpus that revealed twelve categories of publics used by researchers and practitioners in AI. Most authors used the terms citizen and the public uncritically. A diverse corpus introduced new terms like HITL, although even authors could not agree on whether the human was a worker or decision-maker. We saw affordances of expertise and properties of politics, which cut across terms. We also examined who is included and excluded, with the implication that those without the resources or status to participate are left out. Efforts toward broader and inclusive categorizations like data subjects can elide differentials in power. Our goal was to make the implicit explicit.

Precise definitions and affordances are more than semantic exercises. Without precision, interpretations are left to readers who will apply terms to their own domain and culturally specific contexts. Definitions may be outsourced to more powerful external authorities. Impactful decisions will be made as to who, for example, is constituted as a stakeholder and whose voices are acknowledged. Uncovering the affordances remains crucial as AI delivers new promises and problems that recalculate publics.

Acknowledgements

The authors are grateful for the early work of Sam Lumley, who did some of the initial bibliographic work and Zhibin Zhang who provided technical support for the search string generation.


Funding


The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Social Sciences and Humanities Council of Canada, SSHRC 430-2020-00564.

ORCID iDs

Renée Sieber  <https://orcid.org/0000-0002-0720-3498>

Ana Brandusescu  <https://orcid.org/0000-0003-1672-5244>

Abigail Adu-Daako  <https://orcid.org/0009-0002-0397-4903>

Suthee Sangiambut  <https://orcid.org/0000-0003-4661-4422>

Notes

1. There is considerable debate in the literature on the subtle distinctions among and hierarchical implications of the terms engagement, involvement, and participation. In this article, we use the terms as they are popularly used, which is interchangeably.
2. VOSviewer: <https://www.vosviewer.com/>

References

- Abdala MB, Ortega A and Pomares J (2020) Managing the transition to a multi-stakeholder artificial intelligence governance. *Technology*. Available at: <https://www.g20-insights.org/wp-content/uploads/2020/11/managing-the-transition-to-a-multi-stakeholder-artificial-intelligence-governance-1606764749.pdf> (accessed 15 May 2022).
- Almada M (2019) Human intervention in automated decision-making: Toward the construction of contestable systems. In: *Proceedings of the Seventeenth international conference on artificial intelligence and law*, Montreal, QC, Canada, 17 June, pp. 2–11. New York, NY: Association for Computing Machinery.
- Balaram B, Greenham T and Leonard J (2018) *Artificial Intelligence: Real Public Engagement*. London: RSA. Available at: https://www.thersa.org/globalassets/pdfs/reports/rsa_artificial-intelligence—real-public-engagement.pdf (accessed 15 May 2022).
- Battina DS (2017) Research on artificial intelligence for citizen services and government. *International Journal of Creative Research Thoughts (IJCRT)* 5(2): 769–773. Available at: <https://www.ijcrt.org/papers/IJCRT1134263.pdf> (accessed 10 October 2022).
- Benjamin R (2019) *Race After Technology: Abolitionist Tools for the New Jim Code*. New York, NY: Polity.
- Binns R (2022) Human judgment in algorithmic loops: Individual justice and automated decision-making. *Regulation & Governance* 16(1): 197–211.
- Bondi E, Xu L, Acosta- Navas D and Killian JA (2021) Envisioning communities: A participatory approach towards AI for social good. In: *Proceedings of the 2021 AAAI/ACM conference on AI, ethics, and society*, Virtual Event, 30 July, pp. 425–436. New York, NY: Association for Computing Machinery.
- Bowker GC and Star SL (2000) *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT Press.
- Brennan-Marquez K, Levy K and Susser D (2019) Strange loops: Apparent versus actual human involvement in automated decision making. *Berkeley Technology Law Journal* 34: 745. Available at: https://btlj.org/data/articles2019/34_3/02_Brennan-Marquez_Web.pdf (accessed 1 May 2022).
- Burrell J (2016) How the machine “thinks”: Understanding opacity in machine learning algorithms. *Big Data & Society* 3(1): 1–12.
- Callon M, Lascoumes P and Barthe Y (2011) *Acting in an Uncertain World: An Essay on Technical Democracy*. Cambridge, MA: MIT Press.
- Cardullo P (2020) *Citizens in the “Smart City”: Participation, Co-Production, Governance*. London: Routledge.
- Cardullo P and Kitchin R (2019) Being a “citizen” in the smart city: Up and down the scaffold of smart citizen participation in Dublin, Ireland. *GeoJournal* 84(1): 1–13.
- Cofone I (2020) Online harms and the right to be forgotten. In: Cofone I (ed.) *The Right to Be Forgotten: A Canadian and Comparative Perspective*. London: Routledge, pp. 1–16.
- Cornwall A (2008) Unpacking “participation”: Models, meanings and practices. *Community Development Journal* 43(3): 269–283.

- Cowls J (2020) Deciding how to decide: Six key questions for reducing AI's democratic deficit. In: Burr C and Milano S (eds) *The 2019 Yearbook of the Digital Ethics Lab*. Cham: Springer, pp. 101–116.
- Darmon E (2019) Channeling AI into government citizen engagement. *GovTech*, 26 August. Available at: <https://www.govtech.com/opinion/channeling-ai-into-government-citizen-engagement-contributed.html> (accessed 10 June 2022).
- De Sousa WG, De Melo ERP, Bermejo PHDS, Farias RAS and Gomes AO (2019) How and where is artificial intelligence in the public sector going? A literature review and research agenda. *Government Information Quarterly* 36(4): 101392.
- DiSalvo C (2009) Design and the construction of publics. *Design Issues* 25(1): 48–63.
- Dwivedi YK, Hughes L, Ismagilova E, Aarts G, Coombs C, Crick T, et al. (2021) Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management* 57: 101994.
- EU General Data Protection Regulation (GDPR): Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679>
- European Union (2022) Complete guide to GDPR compliance. Available at: <https://gdpr.eu/tag/gdpr/> (accessed 1 May 2022).
- First Nations Information Governance Centre (2020) Understanding the First Nations Principles of OCAP (ownership, control, access and possession). Available at: https://fnigc.ca/wp-content/uploads/2020/09/2be8f15f2eff14b1f122b6a26023836a_fnigc_ocap_brochure_en_final_0.pdf (accessed 1 June 2022).
- Fukuda-Parr S and Gibbons E (2021) Emerging consensus on “Ethical AI”: Human rights critique of stakeholder guidelines. *Global Policy* 12: 32–44.
- Fung A (2006) Varieties of participation in complex governance. *Public Administration Review* 66: 66–75.
- Fung A (2015) Putting the public back into governance: The challenges of citizen participation and its future. *Public Administration Review* 75(4): 513–522.
- Gillespie T (2014) The relevance of algorithms. *Media Technologies: Essays on Communication, Materiality, and Society* 167(2014): 167.
- Gray ML and Suri S (2019) *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*. Boston, MA: Houghton Mifflin Harcourt.
- Gurumurthy A and Chami N (2019) *Radicalising the AI Governance Agenda*. Global Information Society Watch. Available at: <https://giswatch.org/node/6204> (accessed 10 May 2022).
- Haraway D (1991) *Simians, Cyborgs, and Women: The Reinvention of Nature*. New York, NY: Routledge.
- Häußermann JJ and Lütge C (2022) Community-in-the-loop: Towards pluralistic value creation in AI, or— Why AI needs business ethics. *AI and Ethics* 2: 341–362.
- Havrdá M (2020) Artificial intelligence's role in community engagement within the democratic process. *International Journal of Community Well-Being* 3(4): 437–441.
- Henman P (2020) Improving public services using artificial intelligence: Possibilities, pitfalls, governance. *Asia Pacific Journal of Public Administration* 42(4): 209–221.
- Howe J (2008) *Crowdsourcing: How the Power of the Crowd Is Driving the Future of Business*. New York, NY: Random House.
- Ipsos (2022) Canadians among least likely to believe artificial intelligence will make their lives better. Available at: <https://www.ipsos.com/en-ca/news-polls/Canadians-Least-Likely-AI-Make-Lives-Better> (accessed 1 May 2022).
- Jabbari Jahromi A, Man Oram MY, Shrisankaraan VS and Trevan J (2019) Citizens as real-time emotional sensors in smart cities. In: *International conference on smart infrastructure and construction 2019 (ICSIC) driving data-informed decision-making*, Cambridge, 8–10 July, pp. 571–576. London: ICE Publishing.
- Jobin A, Ienca M and Vayena E (2019) The global landscape of AI ethics guidelines. *Nature Machine Intelligence* 1(9): 389–399.

- Kalluri P (2020) Don't ask if artificial intelligence is good or fair, ask how it shifts power. *Nature* 583(7815): 169–169.
- Kieslich K, Keller B and Starke C (2022) AI-ethics by design: Evaluating public perception on the importance of ethical design principles of AI. *Big Data & Society* 9: 1–15.
- Kone SM (2021) Smart public participation: An inquiry into the role of process and technology. In: Zaman N, Rafique K and Ponnusamy V (eds) *ICT Solutions for Improving Smart Communities in Asia*. Hershey, PA: IGI Global, pp 87–105.
- Kuziemiński M and Misuraca G (2020) AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *Telecommunications Policy* 44(6): 101976.
- Kwecko V and Botelho SSDC (2018) Collaborative intelligence in smart cities: A systematic review. *Social Networking* 7(3): 181–195.
- Latour B (2005) *Reassembling the Social: An Introduction to Actor-Network-Theory*. London, UK: Oxford University Press.
- Liboiron M and Lepawsky J (2022) *Discard Studies: Wasting, Systems, and Power*. Cambridge, MA: MIT Press.
- Long D and Magerko B (2020) What is AI literacy? Competencies and design considerations. In: *Proceedings of the 2020 CHI conference on human factors in computing systems*, Honolulu, HI, 25–30 April, pp. 1–16. New York, NY: Association for Computing Machinery.
- Mehr H (2017) *Artificial Intelligence for Citizen Services and Government*. Ash Center for Technology & Democracy, Harvard Kennedy School. Available at: https://ash.harvard.edu/files/ash/files/artificial_intelligence_for_citizen_services.pdf
- Micheli M, Ponti M, Craglia M and Berti Suman A (2020) Emerging models of data governance in the age of datafication. *Big Data & Society* 7(2): 1–15.
- Miller H and Redhead R (2019) Beyond “rights-based approaches”? Employing a process and outcomes framework. *The International Journal of Human Rights* 23(5): 699–718.
- Molnar P and Gill L (2018) *Bots At the Gate: A Human Rights Analysis of Automated Decision-Making in Canada's Immigration and Refugee System*. Citizen Lab. Available at: <https://citizenlab.ca/2018/09/bots-at-the-gate-human-rights-analysis-automated-decision-making-in-canadas-immigration-refugee-system/> (accessed 1 April 2022).
- Monsees L (2019) *Crypto-Publics: Encryption and Democratic Practices in the Digital Era*. London: Routledge.
- Ng DTK, Leung JKL, Chu SKW and Qiao MS (2021) Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence* 2: 100041.
- Park J, Blake C and Cardie C (2015) Toward machine-assisted participation in eRulemaking: An argumentation model of evaluability. In: *Proceedings of the 15th international conference on artificial intelligence and law*, New York, NY, 8–12 June, pp. 206–210. New York, NY: Association for Computing Machinery.
- Phillips R, Freeman RE and Wicks AC (2003) What stakeholder theory is not. *Business Ethics Quarterly* 13(4): 479–502.
- Posada J, Weller N and Wong WH (2021) We haven't gone paperless yet: Why the printing press can help us understand data and AI. In: *Proceedings of the 2021 AAAI/ACM conference on AI, ethics, and society*, New York, NY, 19–21 May, pp. 864–872. New York, NY: Association for Computing Machinery.
- Rahwan I (2018) Society-in-the-loop: Programming the algorithmic social contract. *Ethics and Information Technology* 20(1): 5–14.
- Raley R and Rhee J (2023) Critical AI: A field in formation. *American Literature* 95: 2.
- Reis J, Santo PE and Melão N (2019) Artificial intelligence in government services: A systematic literature review. In: *World conference on information systems and technologies*, Santiago de Compostela, 16–19 April, pp. 241–252. Cham: Springer.
- Roberts ST (2019) *Behind the Screen: Content Moderation in the Shadows of Social Media*. New Haven, CT: Yale University Press.
- Sacha D, Zhang L, Sedlmair M, Lee JA, Peltonen J, Weiskopf D, North SC and Keim DA (2016) Visual interaction with dimensionality reduction: A structured literature analysis. *IEEE Transactions on*

- Visualization and Computer Graphics* 23(1): 241–250. Available at: <https://ieeexplore.ieee.org/document/7536217> (accessed 29 April 2022).
- Sambuli N (2021) *Five Challenges With Multistakeholder Initiatives on AI—Artificial Intelligence and Equality Initiative*. Carnegie Council for Ethics and International Affairs. Available at: <https://www.carnegieica.org/blog/five-challenges-with-multistakeholder-initiatives-on-ai/> (accessed 1 April 2022).
- Savaget P, Chiarini T and Evans S (2019) Empowering political participation through artificial intelligence. *Science and Public Policy* 46(3): 369–380.
- Shadowen N, Lodato T and Loi D (2020) Participatory governance in smart cities: Future scenarios and opportunities. In: *International conference on human-computer interaction*, Copenhagen, 19–24 July, pp. 443–463. Cham: Springer.
- Sharma GD, Yadav A and Chopra R (2020) Artificial intelligence and effective governance: A review, critique and research agenda. *Sustainable Futures* 2: 100004.
- Sieber R and Brandusescu A (2021) Civic empowerment in the development and deployment of AI systems—Final report of the FAccT CRAFT Workshop. In: *ACM fairness, accountability and transparency conference*, Toronto, ON, Canada, 3–10 March. Available at: https://aifortherestofus.ca/wp-content/uploads/2021/10/Civic-Empowerment-Workshop-Report-FAccT_CRAFT.pdf (accessed 1 May 2022).
- Simonofski A, Asensio ES, De Smedt J and Snoeck M (2017) Citizen participation in smart cities: Evaluation framework proposal. In: *2017 IEEE 19th conference on business informatics*, Thessaloniki, Vol. 1, 24–27 July, pp. 227–236. New York, NY: IEEE.
- Sloane M, Moss E, Awomolo O and Forlano L (2020) Participation is not a design fix for machine learning. In: *Proceedings of the 37th international conference on machine learning*, PMLR 119, Vienna, 13–18 July.
- Smuha NA (2021) Beyond a human rights-based approach to AI governance: Promise, Pitfalls, Plea. *Philosophy & Technology* 34(1): 91–104.
- Spielkamp M (2019) Automating society: Taking stock of automated decision-making in the EU. *BertelsmannStiftung Studies*. Available at: <https://algorithmwatch.org/en/automating-society-2019/> (accessed 1 April 2022).
- Statistics Canada (2005) Classification of instructional programs, Canada 2000. Available at: <https://www150.statcan.gc.ca/n1/pub/12-590-x/12-590-x2005001-eng.pdf> (accessed 4 April 2023).
- Suchman L (1993) Do categories have politics? The language/action perspective reconsidered. In: *Proceedings of the third European conference on computer-supported cooperative work (ECSCW'93)*, Milan, 13–17 September, pp. 1–14. Dordrecht: Springer.
- Tönurist P and Surva L (2017) Is volunteering always voluntary? Between compulsion and coercion in co-production. *Voluntas* 28: 223–247.
- Vanolo A (2016) Is there anybody out there? The place and role of citizens in tomorrow's smart cities. *Futures* 82: 26–36.
- Verhulst SG, Zahuranec AJ and Young A (2019) Identifying citizens; needs by combining AI and CI. *The GovLab*, 3 October. Available at: https://thegovlab.org/static/files/publications/CI-AI_oct2019.pdf (accessed 1 May 2022).
- Vogl TM, Seidelin C, Ganesh B and Bright J (2020) Smart technology and the emergence of algorithmic bureaucracy: Artificial intelligence in UK local authorities. *Public Administration Review* 80(6): 946–961.
- Warne H, Dencik L and Hintz A (2021) *Advancing Civic Participation in Algorithmic Decision-Making: A Guidebook for the Public Sector*. Data Justice Lab. Available at: https://datajusticelab.org/wp-content/uploads/2021/06/PublicSectorToolkit_english.pdf (accessed 1 March 2022).
- Watkins T (2020) Cosmology of artificial intelligence project: Libraries, makerspaces, community and AI literacy. *AI Matters* 5(4): 14–17.
- Whittaker M (2021) The steep cost of capture. *Interactions* 28(6): 50–55.
- Wiethof C and Bittner EAC (2021) Hybrid intelligence—Combining the human in the loop with the computer in the loop: A systematic literature review. In: *International conference on information systems 2021 proceedings*, Austin, TX, 12–15 December. Available at: https://aisel.aisnet.org/icis2021/ai_business/ai_business/11 (accessed 17 May 2022).

- Wilson C (2022) Public engagement and AI: A values analysis of national strategies. *Government Information Quarterly* 39(1): 101652.
- Wong JS (2018) Design and fiction: Imagining civic AI. *Interactions* 25(6): 42–45. Available at: <https://interactions.acm.org/archive/view/november-december-2018/design-and-fiction> (accessed 1 April 2022).
- Ylipulli J and Luusua A (2019) “Without libraries what have we?” Public libraries as nodes for technological empowerment in the era of smart cities, AI and big data. In: *Proceedings of the 9th international conference on communities & technologies-transforming communities*, Vienna, 3 June, pp. 92–101. New York, NY: Association for Computing Machinery.
- Zhang B and Dafoe A (2019) *Artificial Intelligence: American Attitudes and Trends*. Oxford: Center for the Governance of AI, Future of Humanity Institute, University of Oxford.
- Zuiderwijk A, Chen YC and Salem F (2021) Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. *Government Information Quarterly* 38(3): 101577.

Author biographies

Renée Sieber has 20+ years experience working at the intersection of civic empowerment and computational technologies. She is best known for her research and practice on public participation in computerized mapping and in climate modeling, and currently researches public participation in AI.

Ana Brandusescu researches public governance of artificial intelligence pertaining to two significant factors: political power and privatization. Previously she led research and policy projects on government transparency and digital rights to advance the open data movement.

Abigail Adu-Daako is a tech policy professional with interests in content policy, online trust and safety, and responsible AI. She is passionate about shaping technology/online platform policies and standards to ensure user trust and safety.

Suthee Sangiambut has worked in open data and open government. He currently researches spatial networks of government artificial intelligence.