






REVIEW ARTICLE



<https://doi.org/10.1057/s41599-023-01999-y>

OPEN

Age-related bias and artificial intelligence: a scoping review

Charlene H. Chu^{1,2,3,4}, Simon Donato-Woodger¹, Shehroz S. Khan^{3,5},
Rune Nystrup⁶, Kathleen Leslie⁷, Alexandra Lyn⁸, Tianyu Shi^{1,9},
Andria Bianchi^{3,10,11}, Samira Abbasgholizadeh Rahimi^{12,13,14} &
Amanda Grenier^{2,15,16}

There are widespread concerns about bias and discriminatory output related to artificial intelligence (AI), which may propagate social biases and disparities. Digital ageism refers to ageism reflected design, development, and implementation of AI systems and technologies and its resultant data. Currently, the prevalence of digital ageism and the sources of AI bias are unknown. A scoping review informed by the Arksey and O'Malley methodology was undertaken to explore age-related bias in AI systems, identify how AI systems encode, produce, or reinforce age-related bias, what is known about digital ageism, and the social, ethical and legal implications of age-related bias. A comprehensive search strategy that included five electronic bases and grey literature sources including legal sources was conducted. A framework of machine learning biases spanning from data to user by Mehrabi et al. is used to present the findings (Mehrabi et al. 2021). The academic search resulted in 7595 articles that were screened according to the inclusion criteria, of which 307 were included for full-text screening, and 49 were included in this review. The grey literature search resulted in 2639 documents screened, of which 235 were included for full text screening, and 25 were found to be relevant to the research questions pertaining to age and AI. As a result, a total of 74 documents were included in this review. The results show that the most common AI applications that intersected with age were age recognition and facial recognition systems. The most frequent machine learning algorithms used were convolutional neural networks and support vector machines. Bias was most frequently introduced in the early 'data to algorithm' phase in machine learning and the 'algorithm to user' phase specifically with representation bias ($n = 33$) and evaluation bias ($n = 29$), respectively (Mehrabi et al. 2021). The review concludes with a discussion of the ethical implications for the field of AI and recommendations for future research.

A full list of author affiliations appears at the end of the paper.

Introduction

Artificial intelligence (AI) has been defined as ‘intelligent agents that receive percepts from the environment and take actions that affect that environment’ (Russell and Norvig, 2020). Powered in part by vast increases in data collection, including the 2.5 quintillion bytes (2,500,000 Terra Byte) of data that the world’s online activities generate each day (Devakunchari, 2014), AI technologies are often deployed on the promise of increased efficiency and increased objectivity (Drage and Mackereth, 2022). However, the ability of AI to provide impartial judgement has been called into question (Howard and Borenstein, 2018). As most commercial AI systems depend on data collected from a multitude of public and private sources (such as Twitter, open-source datasets), societal inequities arising from prejudiced beliefs, actions, and laws may be reflected in these systems. For example, a widely used algorithm for population health management in the United States underestimated the health risks of Black patients because they have reduced access to health care due to issues related to systemic racism (Obermeyer et al., 2019). AI technologies may be as biased as the data on which they are trained unless this bias is conscientiously addressed (Mehrabi et al., 2021). With the ubiquity of AI systems and applications in our everyday lives, accounting for fairness has gained significant importance in designing and deployment of such systems (Zou and Schiebinger, 2018). It is crucial to ensure that these AI does not reinforce inequities or discriminatory behaviour toward certain groups or populations. ‘Algorithmic fairness’ has emerged to counter and explore algorithmic bias and harm with the literature predominantly focused on race and gender (Dawson et al., 2019; Center for Democracy and Technology, 2018; The Royal Society, 2017; Future of Privacy Forum, 2017; Mehrabi et al., 2021), with little attention to age. The purpose of this review is to broaden the discourse by focusing on age-related bias.

The acceleration of technological advancement and the growing scope of AI has created a sense of urgency to examine the implications these technologies may have on a globally aging population (Fenech et al., 2018). Ageism is an implicit age-related bias conceptualised as (1) prejudicial attitudes toward older people and the process of ageing; (2) discriminatory practices against older people; and/or (3) institutionalised policies and social practices that foster the attitudes and actions in relation to (1) and (2) (Datta et al., 2015). The notion of *digital ageism* is used to refer to the extension of ageism into the realm of the design, development, deployment, and evaluation of technology, and how AI and related digital and socio-technical structures may produce, sustain or amplify systemic processes of ageism (Billette et al., 2012; Chu et al., 2022a, 2022b, 2022c, 2022d; Nyrup et al., 2023). Factors that can be considered to contribute to digital ageism involve excluding older adults from the development or design processes (Ashley, 2017; Neary and Chen, 2017; Sourdin and Cornes, 2018), replicating uneven power dynamics between older and younger people, and result in algorithms and/or products that are not optimised for them. Such processes may further negatively impact older adults’ desire to use digital technologies or services (Giudici, 2018; Guégan and Hassani, 2018; Rosales and Fernández-Ardèvol, 2019), thereby generating less training data to better understand the needs of this demographic (WHO, 2022a) and further perpetuate digital ageism.

A previous literature review (Rosales and Fernández-Ardèvol, 2019) investigated ageism in big data reporting on biased samples and tools. Building on this work, a recent conceptual piece on digital ageism (Chu et al., 2022b) describes how older adults are often grouped in large image-based datasets, such as ‘50+’ or ‘60+’ in comparison to other age ranges with smaller categories, such as by decade (Anderson and Perrin, 2017; Tsai et al., 2015).

This process of labelling data may reflect the ageist beliefs that older adults are a homogenous group rather than a diverse demographic (Chu et al., 2022b; Rosales and Fernández-Ardèvol, 2016, 2019). It may also reflect the ageist assumption linking aging with mental and physical decline, as is reflected in marketing of health-monitoring technologies as imperatives of achieving health and success rather than as instruments for adaptation (Culter, 2005; McDonough, 2016; Neven, 2010).

This scoping review is, to our knowledge, the first comprehensive review on age-related bias that explicitly focuses on AI and includes grey literature, such as legal and guidance documents. We have based our examination on the framework developed by Mehrabi et al. which identifies different types of bias that can emerge during the machine learning life cycle. Mehrabi et al.’s domain-specific framework distinguishes between biases that can occur in three machine learning phases (2021): data to algorithm (data origin), algorithm to user (modelling), and user to data (deployment), providing a comprehensive approach to identifying biases and analysing sources of bias. Biases that originate from the ‘data to algorithm’ phase refer to biases that are present in the training and testing data used to develop algorithms, which may be perpetuated in the outputs of algorithms (Mehrabi et al., 2021). The ‘algorithm to user’ phase refers to biased algorithmic outputs, which may in turn modulate and introduce bias into user behaviours (Mehrabi et al., 2021). Lastly, the biases that originate from the ‘user to data’ phase refer to the inherent biases present in users that will be reflected in the data users produce. If this data ends up in a training dataset for future algorithms, it will result in data to algorithm biases, completing a self-sustaining ‘data, algorithm, and user interaction’ feedback loop reinforced by AI (Mehrabi et al., 2021). We apply Mehrabi’s framework to obtain a greater understanding of *how* and *where* along the machine learning pipeline ageism is encoded or reinforced, in order to motivate and inform future solutions. Further, we extracted the societal, legal, and ethical implications to understand the implications of age-related bias in AI systems from the grey literature.

This scoping review aims to answer the following research questions:

1. What is known about age-related bias in AI technology?
2. What literature exists on the extent of age-related bias in AI systems?
3. How do AI systems encode, produce, or reinforce age-related bias?
4. What are the societal, legal, and ethical implications of age-related bias in AI systems?
5. What is the state of knowledge on older people’s experiences of age-related bias in AI systems?

The first three questions take a broader approach to explore the nature of digital ageism followed by a more focused examination of the literature available in order to provide a comprehensive overview.

Methods

Protocol and registration. A scoping review is well-suited to explore the intersection of ageism and AI, or digital ageism, as it allows synthesis and analysis of the scientific landscape and grey literature (Arksey and O’Malley, 2005). In this section, we will describe the two separate searches conducted in academic and grey literatures (Chu et al., 2022a). This review is foundational because older adults comprise one of the fastest-growing demographics of the global population (Datta et al., 2015; WHO, 2021), but there has yet to be a comprehensive examination of the presence of digital ageism (Chu et al. 2022c). This review follows

the six-step methodological framework developed by Arksey and O'Malley (Arksey and O'Malley, 2005), supplemented by Levac et al. (2010), using the format from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) (Tricco et al. 2018). The protocol for this scoping review has been published in JMIR research protocols and registered in the Open Science Framework database, under the doi: 10.17605/OSF.IO/AMG5P (Chu et al., 2022a).

Eligibility criteria. To ensure depth and breadth, articles were included if they were published in English and reported on 'AI' (i.e., algorithms that predict or classify data), 'bias,' and terms related to 'age' (aging, older, demographic). The term 'AI' has existed for over 50 years (Smith et al. 2006), so the publication date was not limited by the year of publication. Exclusion criteria included papers with non-human subjects (e.g., animals, agriculture), theses, dissertations, conference abstracts, non-peer reviewed conference proceedings, perspectives/editorials, books, book chapters, and letters to editors.

Information sources/search. An information specialist helped develop the search strategy in Scopus and then translated it to five other databases (Web of Science, CINAHL, EMBASE, IEEE Xplore, and ACM digital library). The robust search strategy included the terms: 'machine learning', 'artificial intelligence', 'algorithms', 'neural networks', 'deep learning', 'algorithmic bias'; 'biased', 'discrimination', 'ageism'; 'age', and 'older people'. Additionally, following guidance from an information specialist and from experts on the research team/their networks, a comprehensive search of grey literature was undertaken. The same search terms were used to search the grey literature, which included: OpenGrey, Google Scholar, Google search engine, as well as relevant websites and organisations. In this review we included 235 websites (e.g., WHO, AlgorithmWatch). The first 200 results of Google Scholar and Google searches were reviewed for additional academic and grey literature sources. PDFs were downloaded and reviewed for content related to our topic. The academic literature searches were completed in January 2022 and grey literature searches were completed in February 2022.

To help address our fourth research question about the societal, legal, and ethical implications of age-related bias in AI systems, we also conducted a search for relevant legal information. The legal scholars on the research team applied a simplified search strategy to the legal databases WestlawNext Canada and CanLII to capture relevant legal literature in various domains including employment law, human rights law, and health law. An iterative search strategy for legislation, constitutional documents, and jurisprudence (court cases) was also conducted. The legal search focused on sources relevant to the Canadian legal context to maintain feasibility given the variation in legal systems across international jurisdictions and the complex web of national and subnational legal sources impacting AI. Societal and ethical implications relevant to a broader international audience are drawn from this legal review and included in the discussion and priorities for future research.

Selection of sources of evidence. Following the academic search, these citations were uploaded to Covidence systematic review software (Veritas Health Innovation, Melbourne, Australia) and duplicates were removed. The titles and abstracts of the academic articles were screened by two independent reviewers according to the eligibility criteria. Once the abstract screening was complete, the full text of each article was reviewed by two independent reviewers to judge the article's relevance to the research questions. Biweekly meetings were held between February 2022 to

December 2022 to discuss the progress of the charting, extraction, and analysis process. Disagreements were resolved via discussion or by having the first author (CC) act as a third reviewer.

Data extraction. Data were extracted from academic papers and grey literature included in the scoping review by two independent reviewers (SDW, TS) using an excel spreadsheet developed by the reviewers (Chu et al., 2022a). The data extracted included details such as the aim of the study, location, publication year, study design, type of AI, purpose of the AI, data source, AI approach, performance measure, key findings relevant to the research questions, and information related to the legal, social, and ethical implications. Any disagreements between reviewers were resolved through discussion, referring to the original text, and with a third reviewer at regular team meetings. Authors of papers were contacted to request missing or additional data, where required.

Data analysis and presentation. Results of the academic literature search are reported graphically and with supplementary tables. The narrative accompanying the tables further describes the body of academic literature. Findings are reported according to the framework of bias in machine learning by Mehrabi et al. (Mehrabi et al., 2021) that identifies the machine learning life cycle encompassing 'data to algorithm' (data origin), 'algorithm to user' (modelling), and 'user to data' (deployment). Findings are presented in this order to identify the most common phases where bias is encoded, produced, or reinforced to address the research questions. A narrative summary of the legal, social, and ethical implications is presented separately.

Results

The PRISMA diagram is presented in Fig. 1. The academic search returned a total of 14611 academic publications. After removing duplicates, the abstracts of the remaining 7903 academic publications were screened resulting in the removal of 7595 publications during abstract screening. From these, 306 full texts were screened, and 49 academic publications were included in the review. From the grey literature search that was done separately, 235 records were found and 25 were included.

Characteristics of the included studies from the Academic Literature Search.

Figure 2 indicates the countries of the first author of the included articles, most of which were from the U.S.A ($n = 17$) (Amini et al., 2019; Biswas and Rajan, 2020; Culotta et al., 2016; Dev and Phillips, 2019; Diaz et al., 2019; Dinges et al., 2005; Helleringer et al., 2019; Kim et al., 2021; Klare et al., 2012; Lanka et al., 2020; Liang et al., 2019; Smith and Ricanek, 2020; Srinivasan et al., 2020; Strath et al., 2015; Tokola et al., 2014; Wang and Kambhamettu, 2015; Zhao et al., 2020), followed by China/Hong Kong ($n = 7$) (Kuang et al., 2015; Li et al., 2017; Liu et al., 2015; Pei et al., 2017; Sun et al., 2020; Wang and Kambhamettu, 2015; Zou et al., 2016), and Spain ($n = 4$) (Clapés et al., 2018; Principi et al., 2019; Rodriguez et al., 2017; Rosales and Fernández-Ardévol, 2019) and the United Kingdom ($n = 4$) (Abdurrahim et al., 2018; Georgopoulos et al., 2020a, 2020b; Todd et al., 2019). The publication dates range from 2005 to 2021, and there has been an increase in the number of publications each year (see Fig. 3). The most common study aims and applications of AI included in the review were age recognition ($n = 20$) (Alashkar et al., 2020; Alexander and Logashanmugam, 2016; Bekios-Calfa et al., 2011; Clapés et al., 2018; Georgopoulos et al., 2020b; Helleringer et al., 2019; Jung et al., 2018; Kuang et al., 2015; Li et al., 2017; Liu et al., 2015; Luu et al., 2009; Pei et al., 2017; Rodriguez et al., 2017; Smith and Ricanek, 2020; Smith-Miles and Geng, 2020; Tian et al., 2020;

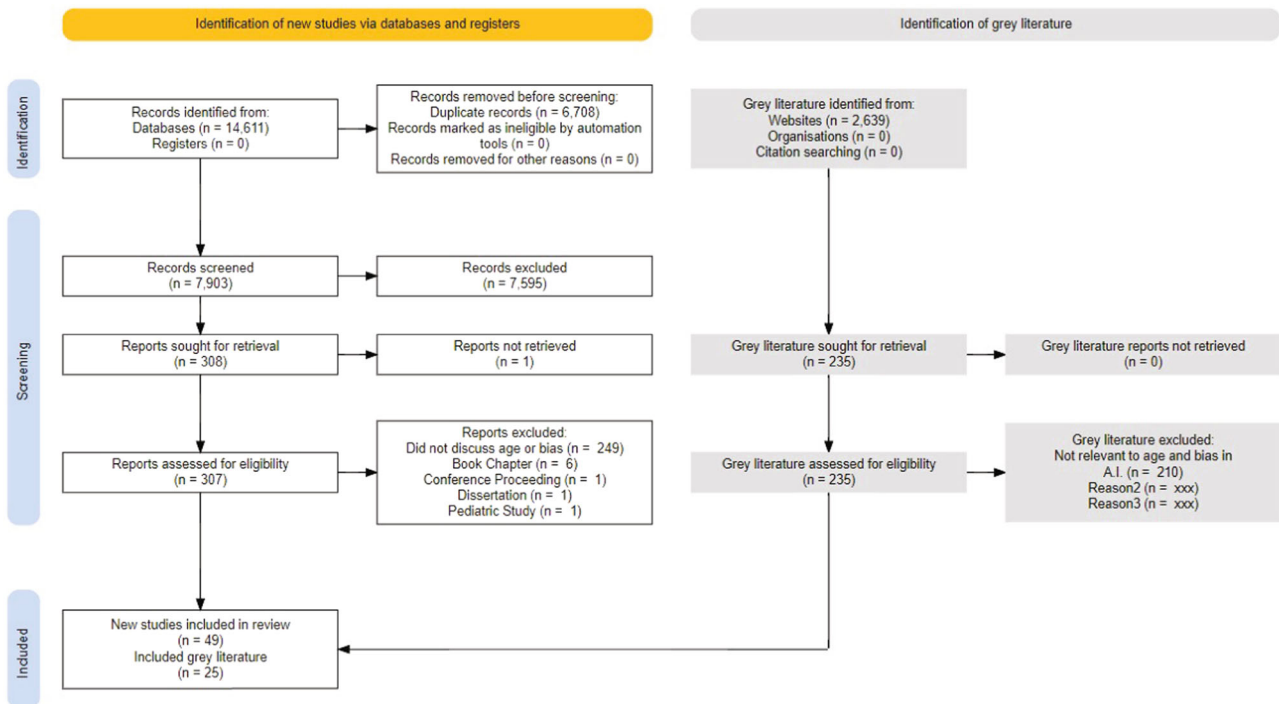


Fig. 1 PRISMA Flowsheet for Academic and Grey Literature Searches.

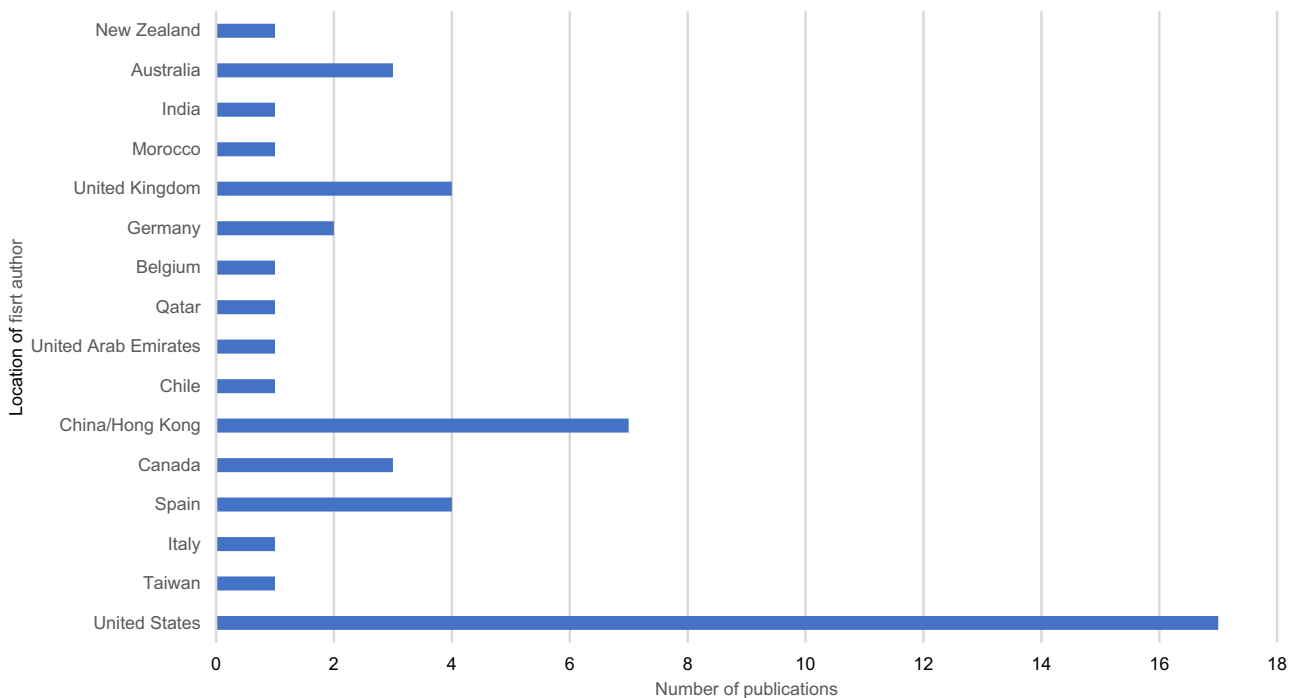


Fig. 2 Country of first author.

Tokola et al., 2014; Wang and Kambhamettu, 2015; Ye et al., 2018; Zou et al., 2016), facial recognition ($n = 6$) (Abdurrahim et al., 2018; Amini et al., 2019; Georgopoulos et al., 2020b; Jung et al., 2018; Klare et al., 2012; Terhörst et al., 2020), emotion recognition ($n = 5$) (Alashkar et al., 2020; Dinges et al., 2005; Kim et al., 2021; Principi et al., 2019; Taati et al., 2019), and gender recognition ($n = 5$) (Alashkar et al., 2020; Georgopoulos et al., 2020b; Jung et al., 2018; Rodriguez et al., 2017; Smith and Ricanek, 2020) (see Fig. 4). Figure 5 graphically plots the frequency of technical approaches and software indicating that the

most used types of AI algorithms were convolutional neural networks ($n = 17$; (Abderrahmane et al., 2020; Alashkar et al., 2020; Amini et al., 2019; Clapés et al., 2018; Drozdowski et al., 2020; Georgopoulos et al., 2020b; Helleringer et al., 2019; Kuang et al., 2015; Pei et al., 2017; Principi et al., 2019; Rodriguez et al., 2017; Smith and Ricanek, 2020; Taati et al., 2019; Tian et al., 2020; Wang and Kambhamettu, 2015; Ye et al., 2018; Zhao et al., 2020), and support vector machines ($n = 8$) (Helleringer et al., 2019; Lanka et al., 2020; Luu et al., 2009; Smith-Miles and Geng, 2020; Strath et al., 2015; Tokola et al., 2014; Wang and Kambhamettu,

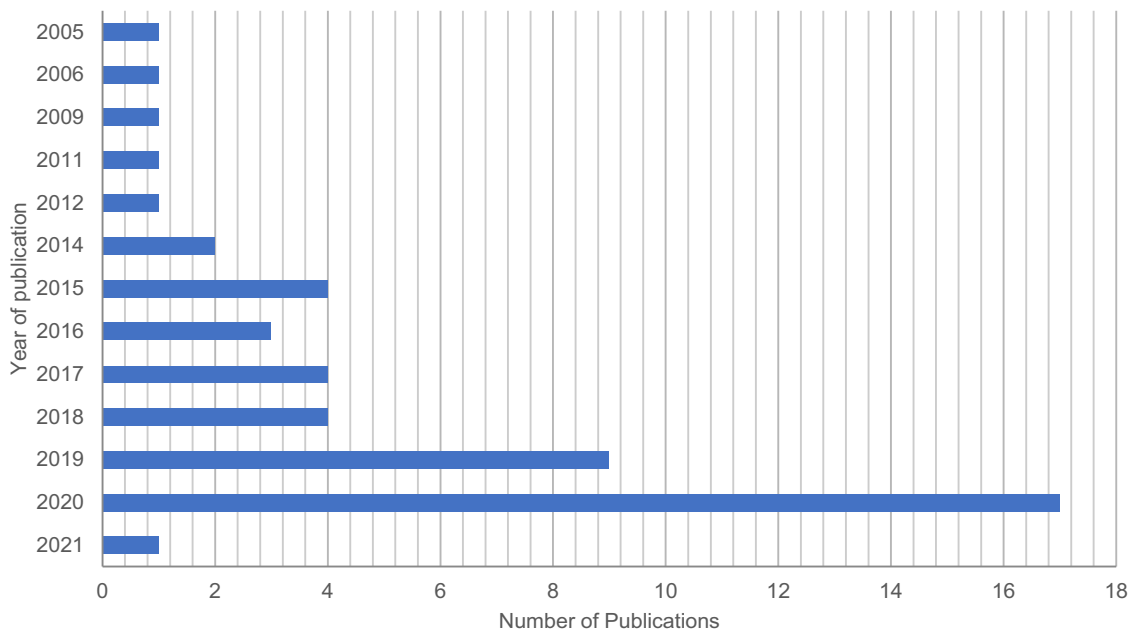


Fig. 3 Number of publications per year.

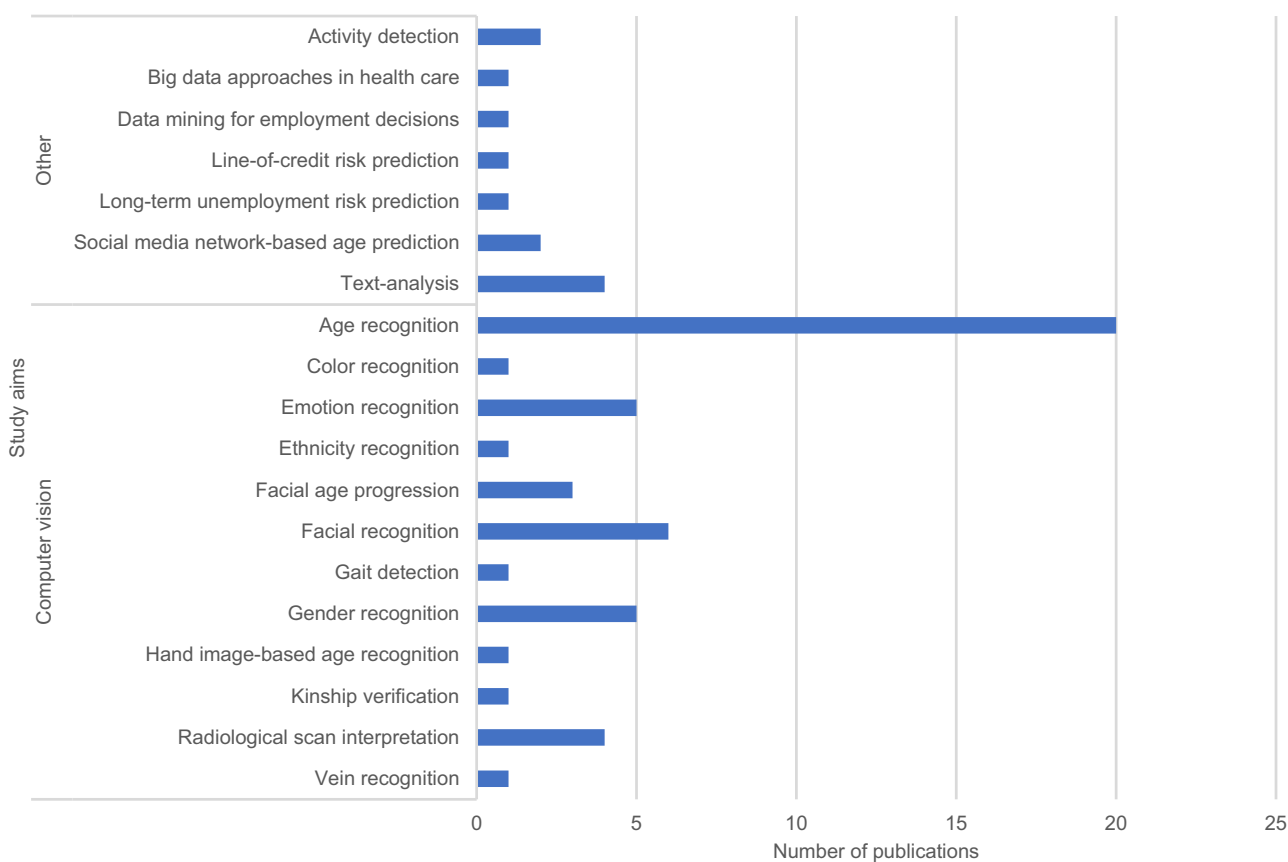


Fig. 4 Purposes of machine learning systems.

2015; Zou et al., 2016). Figure 6 shows frequency of databases used in the academic studies, and that the most used databases were MORPH ($n = 16$) (Abdurrahim et al., 2018; Alexander and Logashanmugam, 2016; Georgopoulos et al., 2020a; Georgopoulos et al., 2020b; Kuang et al., 2015; Li et al., 2017 Liu et al., 2015; Rodriguez et al., 2017; Smith and Ricanek, 2020; Smith-Miles and Geng, 2020; Sun et al., 2020; Terhörst et al., 2020; Tian et al.,

2020; Wang and Kambhamettu, 2015; Xie and Hsu, 2020; Zou et al., 2016), FG-Net ($n = 15$) (Abdurrahim et al., 2018; Alexander and Logashanmugam, 2016; Bekios-Calfa et al., 2011; Georgopoulos et al., 2020a; Georgopoulos et al., 2020b; Kuang et al., 2015; Li et al., 2017 Liu et al., 2015; Luu et al., 2009; Smith-Miles and Geng, 2020; Tian et al., 2020; Tokola et al., 2014; Wang and Kambhamettu, 2015; Xie and Hsu, 2020; Zou et al., 2016),

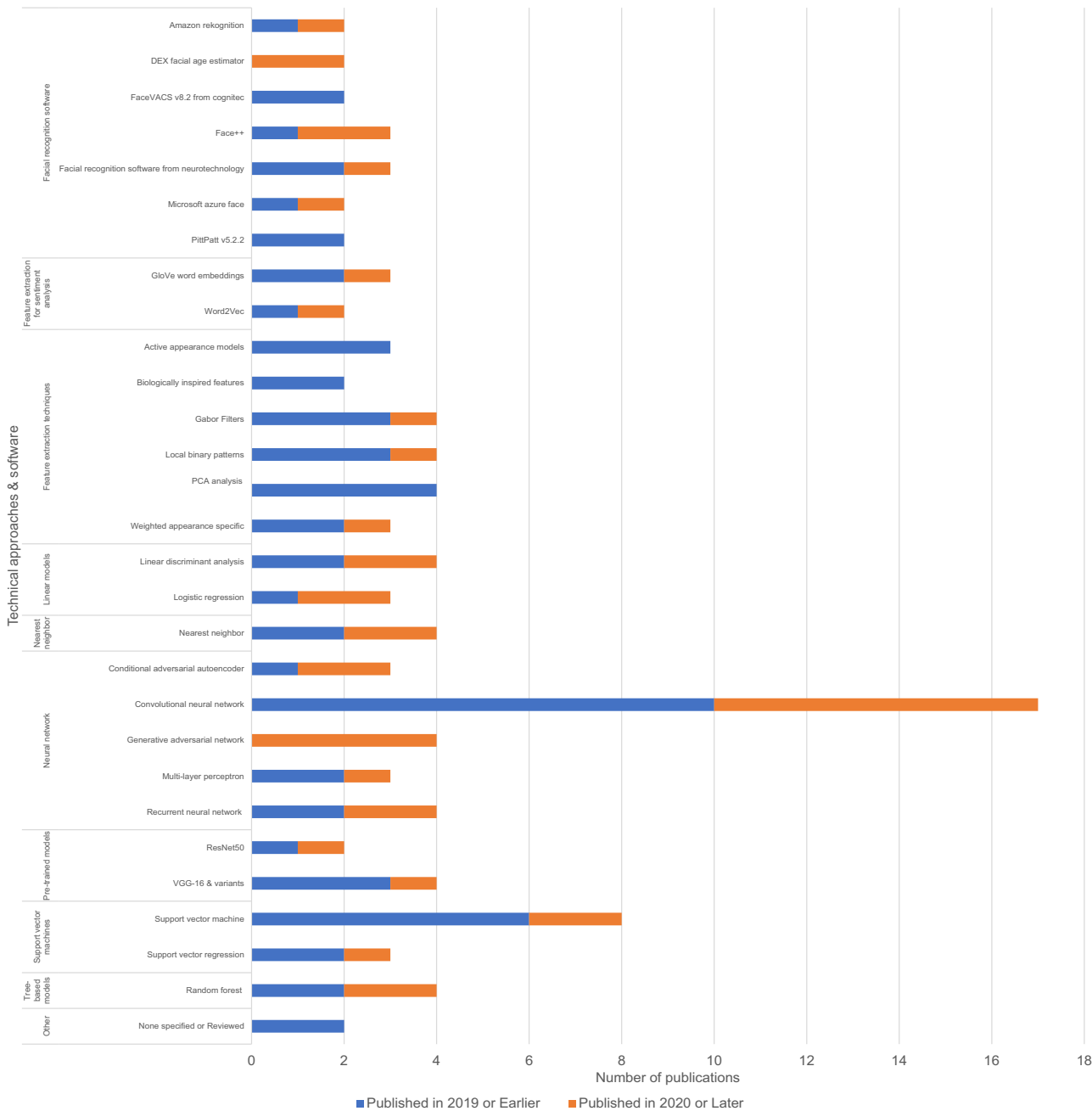


Fig. 5 Frequency of technical approaches and software (includes all software and technical approaches mentioned in >1 publication).

and Adience ($n = 5$) (Alashkar et al., 2020; Kuang et al., 2015; Rodriguez et al., 2017; Terhörst et al., 2020; Ye et al., 2018). The characteristics of the included academic literature are summarised in Supplementary Table S1 (online).

Review findings from the academic literature. The publications discussed in this section are the 49 academic publications from our search of academic databases. In total, nine types of biases in the machine learning life cycle were identified (Mehrabi et al., 2021). Five biases in the ‘data to algorithm’ phase (representation bias, aggregation bias, measurement bias, omitted variable bias, linking bias), one bias in the ‘algorithm to user’ (evaluation bias), and three interrelated biases in the ‘user to data’ phase (historical bias, content production bias and social bias). Figure 7 outlines the frequencies of all the biases identified in this review. 35 academic studies contained more than one type of bias. The types of

bias from Mehrabi’s framework (Mehrabi et al., 2021) that were either present or discussed in each paper are summarised in Supplementary Table S2 (online).

Data to algorithm. This section presents biases related to the ‘data to algorithm’ (data origin) phase of the machine learning life cycle from Mehrabi’s framework found in the 49 academic publications. These biases stem from the issues within the data that influence the machine learning algorithms (Mehrabi et al., 2021). The data to algorithm biases found in the academic literature include representation bias, measurement bias, omitted variable bias, aggregation bias, and linking bias.

Representation bias. Representation bias occurs when the dataset underrepresents or misrepresents subsets of the population resulting in a non-representative dataset (Mehrabi et al., 2021;

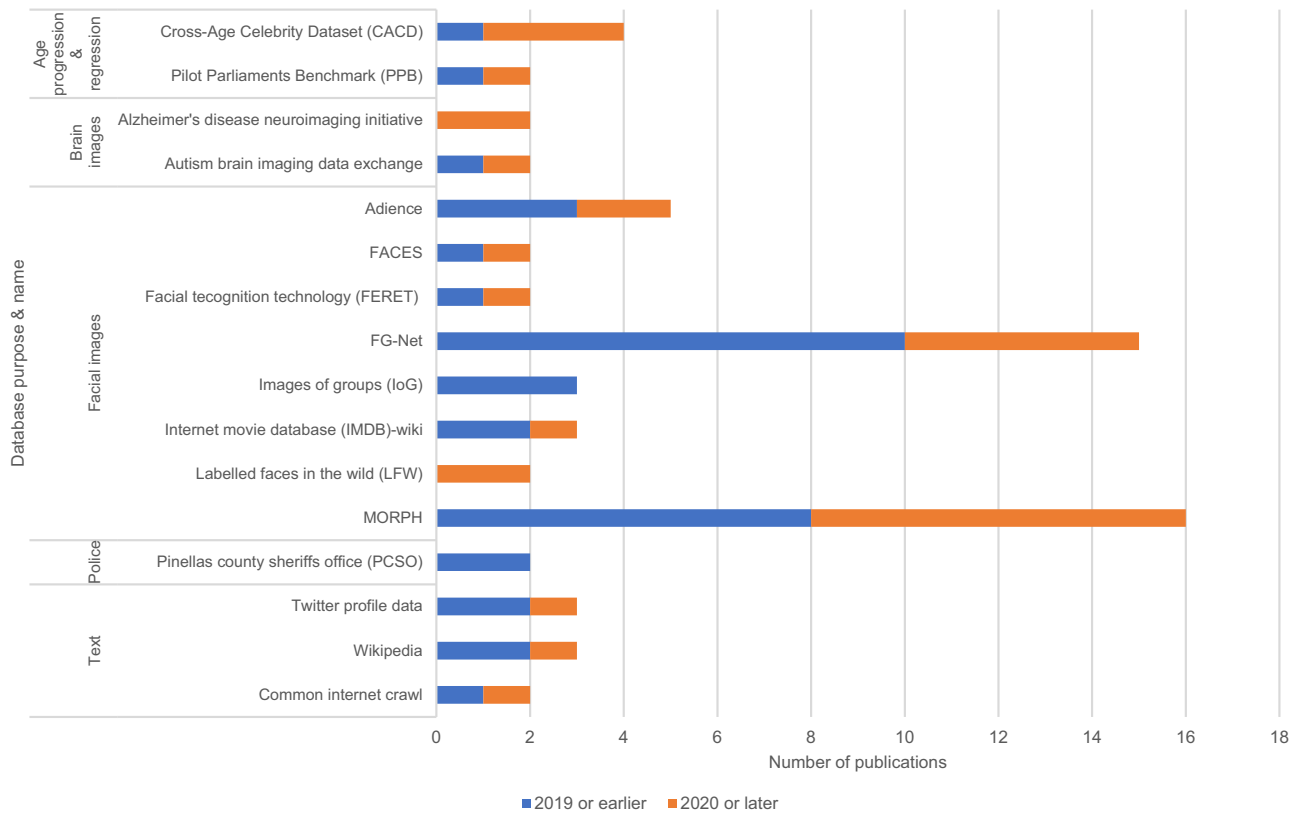


Fig. 6 Frequency of database utilisation by year (includes all databases used in >1 paper).

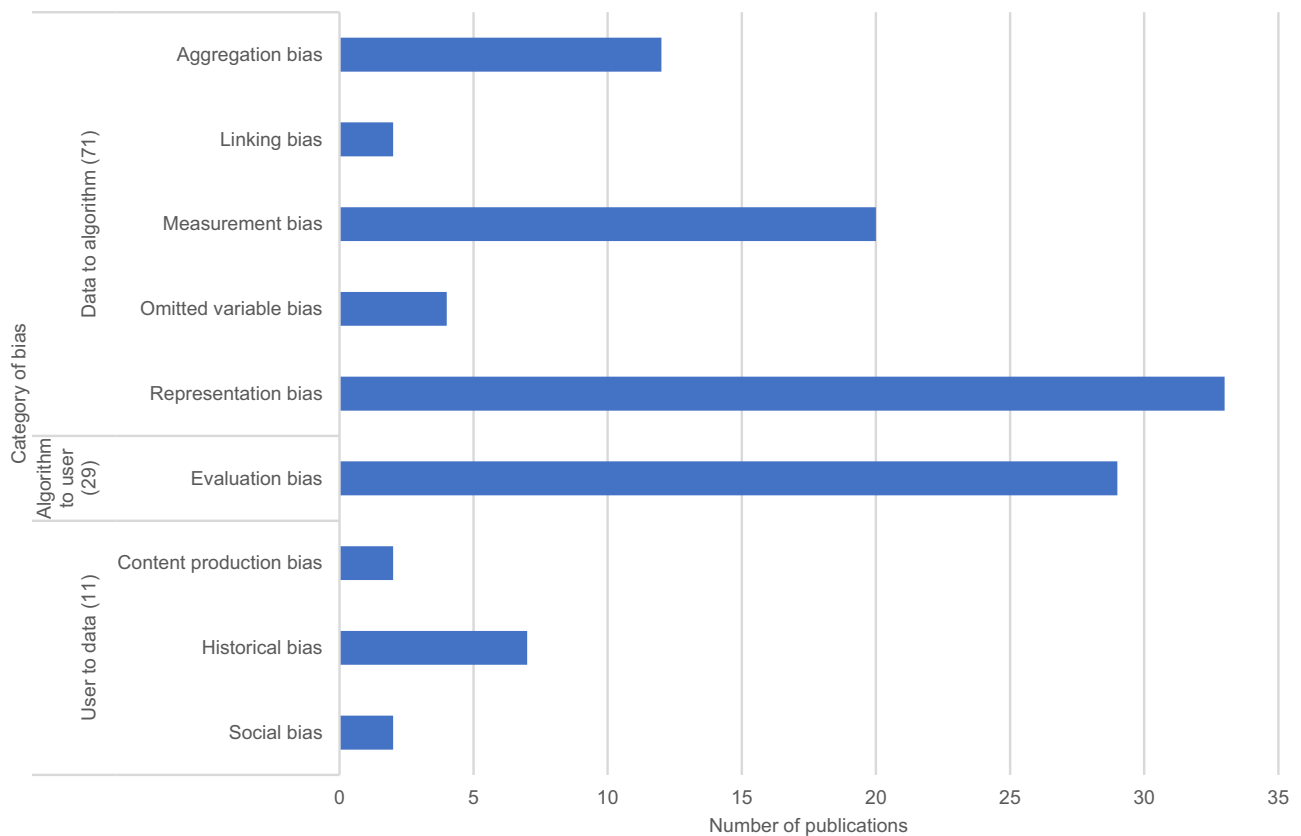


Fig. 7 Frequency of bias (per Mehrabi's framework).

Suresh and Gutttag, 2021). This was one of the most common forms of bias, with 33 of the 49 academic papers either demonstrating or discussing representation bias. For example, authors of one paper described the number of subjects 60 years old and older or younger than 16 as ‘too few to be considered for [AI] training as it is commonly acknowledged that imbalanced data will degrade the learning and result in biased estimation’ (p.4) (Xie and Hsu 2020). Of these papers, authors of 23 studies either discussed or addressed the presence of representation bias (Abderrahmane et al., 2020; Abdurrahim et al., 2018; Bekios-Calfa et al., 2011; Clapés et al., 2018; Georgopoulos et al., 2020a; Georgopoulos et al., 2020b; Jung et al., 2018; Klare et al., 2012; Kuang et al., 2015; Liang et al., 2019; Liu et al., 2015; Pei et al., 2017; Principi et al., 2019; Rosales and Fernández-Ardèvol, 2019; Smith and Ricanek, 2020; Smith-Miles and Geng, 2020; Strath et al., 2015; Sun et al., 2020; Terhörst et al., 2020; Tian et al., 2020; Tokola et al., 2014; Xie and Hsu 2020; Zou et al., 2016). Separately, authors of 10 other papers demonstrated representation bias but did not acknowledge that it may have affected their results (Alashkar et al., 2020; Alexander and Logashanmugam, 2016; Dinges et al., 2005; Diraco et al., 2017; Drozdowski et al., 2020; Li et al., 2017; Luu et al., 2009; Rodriguez et al., 2017; Wang and Kambhamettu, 2015; Ye et al., 2018).

The risk for representation bias was due to imbalanced datasets that significantly underrepresented older adults. From the papers, the most commonly used or discussed databases for age recognition (for which demographic data was available) were MORPH ($n = 16$) (Abdurrahim et al., 2018; Alexander and Logashanmugam, 2016; Georgopoulos et al., 2020a; Georgopoulos et al., 2020b; Kuang et al., 2015; Li et al., 2017; Liu et al., 2015; Rodriguez et al., 2017; Smith and Ricanek, 2020; Smith-Miles and Geng, 2020; Sun et al., 2020; Terhörst et al., 2020; Tian et al., 2020; Wang and Kambhamettu, 2015; Xie and Hsu, 2020; Zou et al., 2016), FG-Net ($n = 15$) (Abdurrahim et al., 2018; Alexander and Logashanmugam, 2016; Bekios-Calfa et al., 2011; Georgopoulos et al., 2020a; Georgopoulos et al., 2020b; Kuang et al., 2015; Li et al., 2017; Liu et al., 2015; Luu et al., 2009; Smith-Miles and Geng, 2020; Tian et al., 2020; Tokola et al., 2014; Wang and Kambhamettu, 2015; Xie and Hsu, 2020; Zou et al., 2016), Adience ($n = 5$) (Alashkar et al., 2020; Kuang et al., 2015; Rodriguez et al., 2017; Terhörst et al., 2020; Ye et al., 2018), the IMDB-Wiki ($n = 3$) (Jung et al., 2018; Principi et al., 2019; Smith and Ricanek, 2020), and Images of Groups ($n = 3$) (Bekios-Calfa et al., 2011; Rodriguez et al., 2017; Tokola et al., 2014). Two databases discussing or demonstrating representation bias appeared in one paper each: the 11 K hands (Abderrahmane et al., 2020), and the UvA NEMO Smile and Disgust Databases (Pei et al., 2017).

Efforts to mitigate representation bias. While most researchers did not address the bias in datasets, eight studies identified age-imbalances present in their selected datasets and corrected the imbalance by modifying the training datasets to balance the representation of specific age categories ($n = 8$) (Abderrahmane et al., 2020; Clapés et al., 2018; Georgopoulos et al., 2020a; Jung et al., 2018; Liang et al., 2019; Smith and Ricanek, 2020; Taati et al., 2019; Zou et al., 2016). Zou et al. used images from the MORPH database to balance the FGNet database, although their balanced dataset still underrepresented the 60–69 and 70–79 demographics (Zou et al., 2016). Abderrahmane et al., applied data augmentation techniques to images from the 11 K Hands dataset, for the purpose of giving each age demographic equal representation, although the nature of these augmentation techniques was not specified (Abderrahmane et al., 2020). Georgopoulos et al., applied digital aging techniques to existing images in

the MORPH and CACD databases to generate a dataset three to four times the size of their existing dataset, without the need to search for more images: all additional images were variations of images from the original datasets. They proposed a novel, generative adversarial network architecture for data augmentation that can capture and reproduce fine-grained aging patterns in images (Georgopoulos et al., 2020a). As a result, their model can diversify the existing distribution of a biased dataset to mitigate the evaluation bias. Smith and Ricanek supplemented their data with extra training on images of older adults, children, women, and dark-skinned individuals from MORPH and IMDB-Wiki (Smith and Ricanek, 2020). Clapes et al., improved the accuracy of their activity prediction model by training it on each individual demographic, finding that performance improved for the under-represented demographics within their dataset (Clapés et al., 2018). Jung et al., responded to the under-representation of older adults in their dataset by creating a new dataset called the 100-celebrities dataset, which was balanced for demographic variables such as ethnicity, gender, and age (Jung et al., 2018). Taati et al., introduced new images of older adults, with and without dementia in their dataset, and found that while facial landmark detection improved on images of healthy adults, it did not improve significantly for older adults affected by dementia (Taati et al., 2019). Lastly, Liang et al., found age-related bias in a neuroimaging model that underestimated the ages of older adults, and attempted to address representation bias by balancing their dataset (Liang et al., 2019). When that mitigation strategy was ineffective, Liang et al., made statistical adjustments to the model’s algorithm itself, which indicated that algorithmic bias, and not representation bias was the source of the issue. The important point made here by Liang et al., is that even when addressing representation bias in a dataset, algorithmic bias may still persist and requires additional strategies (Liang et al., 2019).

Overall, various approaches exist to create more equitable training and benchmark datasets, to reduce the risk of both representation and evaluation bias (discussed later); however, the efficacy of these strategies appears to vary based on each model’s purpose and the data being used.

Measurement bias. Measurement bias is when the data collected inaccurately reflects the variable of interest, and can relate to how the data is selected, used, and measured (Mehrabi et al., 2021). In our results, 20 out of 49 academic papers demonstrated or discussed measurement bias. Of these, 15 tested computer vision methods and demonstrated measurement bias by returning higher mean average errors (MAE) for age or facial recognition for older adults’ demographics in comparison to younger age groups (Alashkar et al., 2020; Bekios-Calfa et al., 2011; Clapés et al., 2018; Georgopoulos et al., 2020a; Georgopoulos et al., 2020b; Jung et al., 2018; Kuang et al., 2015; Lanka et al., 2020; Liang et al., 2019; Pei et al., 2017; Srinivasan et al., 2020; Taati et al., 2019; Xie and Hsu, 2020; Zhao et al., 2020; Zou et al., 2016). In their study about detecting age-related gait changes, Begg and Kamruzzaman significantly improved their model’s performance by modifying the number of selected features used in their model which demonstrates the impact that data representation can have on performance (Begg and Kamruzzaman, 2006). Zhao examined a model designed to predict the risk for long-term unemployment, and found a higher risk of false omissions for older adults (Zhao, 2020). Two additional papers discussed mitigation strategies against measurement bias: one using a variational auto-encoder to re-weight latent variables (e.g., age) and then re-balance the model’s training data during the training phase (Amini et al., 2019) the other using protected attributes in the models to mitigate bias against groups of people, such as bias based on race or marital status (Biswas and Rajan, 2020).

Aggregation bias. Aggregation bias occurs when conclusions are drawn about individuals based on observations about a larger group (Mehrabi et al., 2021). Aggregation bias was demonstrated or discussed in 12 papers (Alashkar et al., 2020; Bekios-Calfa et al., 2011; Georgopoulos et al., 2020a; Georgopoulos et al., 2020b; Helleringer et al., 2019; Jung et al., 2018; Kuang et al., 2015; Rodriguez et al., 2017; Rosales and Fernández-Ardèvol, 2019; Terhörst et al., 2020; Tokola et al., 2014; Ye et al., 2018). In 11 of these cases, aggregation bias was due to datasets which aggregated older adults under arbitrarily large labels, such as ‘60+’ while comparatively assigning other younger age groups more narrow labels. The most common dataset in our review that demonstrated aggregation bias toward older adults was the Adience dataset, which groups older adults into a single category called ‘60+’, appeared in 5 papers (Alashkar et al., 2020; Kuang et al., 2015; Rodriguez et al., 2017; Terhörst et al., 2020; Ye et al., 2018). The Images of Groups (IoG) dataset, which aggregated all adults aged 37–65 into a single category appeared in 3 papers (Bekios-Calfa et al., 2011; Rodriguez et al., 2017; Tokola et al., 2014). Mehrabi notes that aggregation bias can occur in the absence of representation bias, and vice versa (Mehrabi et al., 2021). Older adults can be accurately represented, but still aggregated into a group that erases demographic diversity. For example, the 100 Celebrities dataset, which was used in one paper, aggregated all older adults aged 55+ into a single category, despite otherwise affording older adults more equal representation with other age groups (Jung et al., 2018). One paper reviewed big data development methods and discussed how the over-aggregation of older adults in big data methods leads to reduced accuracy that does not occur with other groups (Rosales and Fernández-Ardèvol, 2019).

Georgopolous et al., used meta-data to mitigate aggregation bias when examining facial images in FG-NET, where all older adult images were pre-classified into a ‘50 years old+’ category. The aggregation of wide age ranges fails to fully capture the heterogeneity of older adults’ facial features and results in less accurate predictions. The same authors also provide a collection of training data known as meta-data that they collected themselves to augment their FG-NET dataset (Georgopoulos et al., 2020a). The meta-data provides richer information about under-represented older adults, which in turn mitigates aggregation bias and improves prediction accuracy.

Omitted variable bias. Omitted variable bias occurs when a variable is omitted from an AI model resulting in the model developing an incomplete understanding of the data (Mehrabi et al., 2021). Four academic studies demonstrated or discussed omitted variable bias that indirectly affected age: all four were focused on face, age, or emotion recognition (Abdurrahim et al., 2018; Kim et al., 2021; Rodriguez et al., 2017; Xie and Hsu, 2020). For example, Xie and Hsu removed the data of underrepresented demographics, namely adults >60 years old and children <16 years old, from their model to enhance the accuracy for younger age groups (Xie and Hsu, 2020). In their study about age-identification, Kim et al., discussed that the absence of other data in their dataset contributed to gender becoming a confounding factor in age recognition (Kim et al., 2021). Additionally, two other studies described that age and gender recognition models estimate age and gender based on a combination of various aspects of subject appearance, such as age, ethnicity, makeup, gender (Abdurrahim et al., 2018; Rodriguez et al., 2017). Discussion of omitted variable bias raised the challenges in addressing bias in AI given the ‘black-box’ nature of many algorithms, where there is a question of ‘unknown unknowns’. The authors discussed how the interaction of these covariates in their model makes it difficult to determine which, if any, variables have been

omitted, and how it may not be possible to pre-emptively account for every variable in their model. This lack of clarity presents a challenge for researchers and developers attempting to mitigate omitted variable bias (Mehrabi et al., 2021).

Linking bias. Linking bias refers to correlations that AI draws about particular users based on the characteristics of other users linked to them via social media networks (Mehrabi et al., 2021). Only two papers discussed linking bias related to age: Rosales and Fernández-Ardèvol (Rosales and Fernández-Ardèvol, 2019) identified that linking bias could result from differences in social media use between older adults and the general population. Insofar as older adults are more likely to use social media to interact with younger users (e.g., their children, grandchildren), the age-prediction strategies that are accurate for younger demographics are less accurate for older users. Culotta et al., found that age prediction based on language output was more accurate than age based on their social network (Culotta et al., 2016).

Algorithm to user. ‘Algorithm to user’ biases occurs when algorithmic outcomes that are biased affect user behaviour. Biases in this phase are related to the algorithms themselves, through the results they generate, to influence the biases of users of those algorithms (specifically age-related bias for the purpose of this review) (Mehrabi et al., 2021). In the review of the academic publications, evaluation bias was the only bias found in this phase.

Evaluation bias. Evaluation bias occurs when inappropriate evaluation benchmark data are selected to assess machine learning models (Mehrabi et al., 2021). Models are optimised on their training data and their quality is often measured against benchmarks; however, the use of inappropriate and misrepresentative benchmark data will result in models that only work well in specific groups of the population (Mehrabi et al., 2021; Suresh and Gutttag, 2021). This review identified 29 articles that demonstrated evaluation bias (Abderrahmane et al., 2020; Abdurrahim et al., 2018; Alashkar et al., 2020; Alexander and Logashanmugam, 2016; Bekios-Calfa et al., 2011; Clapés et al., 2018; Drozdowski et al., 2020; Georgopoulos et al., 2020a; Georgopoulos et al., 2020b; Helleringer et al., 2019; Jung et al., 2018; Klare et al., 2012; Kuang et al., 2015; Li et al., 2017; Liang et al., 2019; Liu et al., 2015; Luu et al., 2009; Pei et al., 2017; Principi et al., 2019; Rodriguez et al., 2017; Smith and Ricanek, 2020; Smith-Miles and Geng, 2020; Sun et al., 2020; Terhörst et al., 2020; Tian et al., 2020; Tokola et al., 2014; Wang and Kambhamettu, 2015; Xie and Hsu, 2020; Zou et al., 2016). In these papers, evaluation bias was the result of researchers using the same unrepresentative benchmark data for testing and training. A complete list of the datasets and their respective frequencies can be found in Fig. 5. Machine learning models can develop biases that are present in unrepresentative benchmarks because the biases pass through the testing phase undetected (Mehrabi et al., 2021).

User to data. Seven out of 49 academic publications indicate biases arising from the ‘user to data’ (deployment phase) of the machine learning life cycle. Mehrabi described these biases as the ones attributed to user behaviours that are then reflected in the data they generate. The data can in turn influence algorithms through the data-to-algorithm phase, completing the feedback loop (Mehrabi et al., 2021). Three biases from the user to data phase of Mehrabi’s framework were present in the academic literature: content production bias, historical bias, and social bias (Mehrabi et al., 2021).

Content production bias. Content production bias emerges from structural, lexical, semantic, and syntactic differences in the contents generated by users (Mehrabi et al., 2021). Two articles reflected upon sentiment analysis in language-processing and found that phrases associated with advanced age and older adults were associated with more negative sentiments than phrases associated with younger adults (Dev and Phillips, 2019; Diaz et al., 2019). Dev and Phillips conducted a sentiment analysis on first names and found ‘older names’ carried more frequent associations with negative words compared to ‘younger names’ (Dev and Phillips, 2019). Diaz et al., conducted a multiphase study where they examined the explicit encoding of age in sentiment analysis across 15 different sentiment analysis tools. Results indicated that sentences containing ‘young’ adjectives were 66% more likely to be perceived favourably, compared to otherwise identical sentences swapped out with ‘old’ adjectives (Diaz et al., 2019). Corpus-based tools that are supervised and trained on labelled text were more likely than unsupervised lexicon-based tools to denote this bias. Additional results of regression analysis showed implicit coding of age where ‘old’ adjectives were 9% less likely to be scored positively and 3% times more likely to be scored negatively compared to control adjectives, while ‘young’ adjectives were 6% likely to be scored negatively and 9% more likely to be scored as positively as the control adjectives (Diaz et al., 2019). The application of AI techniques to publicly-generated text highlights how language reflects ageism within our society (Dev and Phillips, 2019; Diaz et al., 2019).

Historical bias. Historical bias occurs when existing biases and socio-technical issues in the world are captured during the data generation process and reflected in the data (Mehrabi et al., 2021). In our review, seven articles discussed historical bias (Berendt and Preibusch, 2014; Dev and Phillips, 2019; Diaz et al., 2019; Rosales and Fernández-Ardèvol, 2019; Rozado, 2020; Todd et al., 2019; Zhao, 2020). Three of these articles studied sentiment analysis in word-embedding models, demonstrating how ageism and ageist language is captured in the corpus (Dev and Phillips, 2019; Diaz et al., 2019; Rozado, 2020). For example, Dev and Phillips used Wikipedia as their dataset, studying the GLoVe embedding tool. The paper focused on bias against gender and ethnicity, but the researchers also found age-related bias against first names (such as ‘Ruth’ or ‘Horace’) that were encoded as ‘old names’ defined as names used more frequently at the turn of the 20th century while detecting positive associations with names encoded as ‘young names’ (such as ‘Aaron’ or ‘Miranda’) more frequently found in the late 20th century (Dev and Phillips, 2019). Similarly, Diaz et al., also found older names were correlated with negative labels, and younger names correlated with positive labels, labels being words used in either a positive or negative context (Diaz et al., 2019). Another study examined popular word-embedding models, and found that terms related to advanced age, such as ‘elderly’, ‘old’, ‘aging’, ‘senior citizen’, and ‘old age’ correlated more strongly with negative terms compared to youth-centric terms, such as ‘young’ and ‘youthfulness’ (Rozado, 2020).

The remaining four studies (Berendt and Preibusch, 2014; Rosales and Fernández-Ardèvol, 2019; Todd et al., 2019; Zhao, 2020) that reflected historical bias varied in nature and aim. Zhao’s research on algorithmic prediction of unemployment risk discussed advanced age as being a contributing factor for long-term unemployment (Zhao, 2020). In their literature review of social media and mobile applications related to ageism, Rosales and Fernández-Ardèvol explored the existing stereotypes about older adults that lead their digital exclusion, for example, older adults are perceived as being less interested and benefitting less from digital technology, and how this results in age-related bias being perpetuated in big data (Rosales and Fernández-Ardèvol,

2019). In their study about the use of discrimination-aware data mining, Berendt and Preibusch found that rejecting an older applicant applying for a job on account of their age was perceived as more acceptable compared to bias based on nationality and sex in hiring decisions (Berendt and Preibusch, 2014). Finally, in their discursive paper about data gathering practices from older adults in healthcare, Todd et al., discussed the historical challenges conducting studies on older adults, including the fact that older adults had historically been excluded from medical research. They discussed these historical disadvantages in the context of machine learning providing the possibility to gather data from routine interactions with the health care system (Todd et al., 2019).

Social bias. Social bias relates to our judgement and the perceptions of others or ourselves (Mehrabi et al., 2021). Two articles displayed or discussed social bias, highlighting that AI has the potential to aid people in identifying and correcting social biases. Berendt and Preibusch suggest that using ‘Discriminatory Aware Data Mining’ software could positively influence hiring managers to increase their own self-awareness of potential biases (Berendt and Preibusch, 2014). Next, Zhao found that older adults are at higher risk than the general population for long-term unemployment (defined by the OECD as unemployment lasting greater than 12 months) despite the affected older individuals looking for work (Zhao, 2020). This may indicate that older adults might face challenges to compete in the labour market compared to their younger contemporaries.

Results of the grey literature: societal, legal, and ethical implications. The second of our searches was concerned with understanding what is known about the societal, legal, and ethical implications of digital ageism. Thus, in our review of the grey literature including legal sources, we extracted and narratively synthesised information on the societal, legal, and ethical implications of age-related bias in AI systems. Our search generated 2639 results, of which 235 received a full-text screening, and 25 of these were included in this review (Fig. 1). In this section, we describe the results of the grey literature ($n = 25$) supplemented by relevant findings from the academic literature search ($n = 49$) to provide additional context. Understanding these broader implications helps demonstrate the real-world impacts of this bias and potential legal and policy approaches to regulate this fast-changing area of technology.

Societal implications. The grey literature ($n = 25$) broadly discussed the potential of AI to create significant economic and social benefits as well as how it also poses ethical and societal risks related to algorithmic biases and harms. Age was typically listed as one of the demographic variables that is a source of bias, but few sources discussed the societal implications in-depth ($n = 17$) BasuMallick, 2019a, 2019b; Chin, 2019; Freedom House, 2021; Constine, 2017; Engler, 2020; Fischer, 2021; Ghosh, 2020; Ho et al., 2021; AITrends, 2019; Kariuki, 2021; Lee, 2016; Leufer, 2021; Jansen et al., 2020; MacCarthy, 2021; Margetts and Dorobantu, 2019; Wang, 2017). Three sources discussed age-related bias related to employment; two were based on results from an Indeed.com survey conducted in 2017 that 43% of workers in the technology sector worry about losing their jobs due to age (Blank, 2021; Kinnard, 2018), and the third was a statement from Craig Mokhiber from the United Nations calling for increased digital education for older adults (Windegger, 2018). The three sources identified the impacts of ageism more broadly in society and how technologies and AI may lead to discrimination, which Mokhiber terms ‘artificial intolerance’ against older adults (Windegger, 2018). Similarly, as discussed above in

our review of the academic literature, three sentiment analysis papers revealed a consistent finding of negative associations with older adults and aging in the English language, which affects how older adults internalise ageist self-perceptions, as well as how they are treated by society (Dev and Phillips, 2019; Diaz et al., 2019; Rozado, 2020). Notably, these papers used similar methods and data mined from ubiquitous social media and information platforms, such as Twitter and Wikipedia which reflect the general attitudes of society at large regarding ageism and the potential impact of these ageist attitudes in society.

Ethical and legal implications. The impact of digital ageism on the everyday lives of older adults is not well understood based on the grey literature (Chu et al., 2022a; Mehrabi et al., 2021). From an ethical perspective, a critical examination is needed to understand what factors, if any, make age-related bias algorithmically distinct from other forms of discrimination, and what methods have proven effective for counteracting age-related bias (Chu et al., 2022a; Nyrup et al., 2023; Rosales and Fernández-Ardèvol, 2016). Our grey literature results indicated several sources that briefly mentioned ethical concerns related to disinformation, harm, and transparency related to AI and older adults (Constine, 2017; Druga et al., 2021; Ham, 2021; Kantayya, 2021; Windegger, 2018). Similar ethical concerns related to trust, justice, and transparency were also found in the academic literature (Berendt and Preibusch, 2014; Rosales and Fernández-Ardèvol, 2019; Todd et al., 2019; Zhao, 2020). In their study of discrimination-aware data mining, Berendt and Preibusch reported that age-based discrimination was not considered as problematic by individuals responsible for hiring candidates (Berendt and Preibusch, 2014). The academic literature suggests that age-based discrimination and its algorithmic harms may be overlooked and draws attention to implicit discriminatory motivations. Digital exclusion can significantly impact older adults due to ageist stereotypes, for example, the development of select social media applications based on the perceived needs of older adults (Rosales and Fernández-Ardèvol, 2019), as well as exclusion from clinical research which results in a lack of data about the effects of medications on older adults (Todd et al., 2019). Finally, the ethical implications of Zhao suggest that older adults would be less likely to receive unemployment assistance compared to the general population based on a predictive model of long-term unemployment, which demonstrates the potential socioeconomic impacts of algorithmic bias and digital ageism (Zhao, 2020).

The legal implications associated with age-related bias in AI are multifaceted and wide ranging. As a result, the legal analysis cannot be constrained to a single legal discipline, area of practice, or area of law. For example, matters relating to age-related bias in AI have the potential to engage legal issues in tort law, including medical malpractice (Froomkin et al., 2019), human rights and constitutional law (Henderson et al., 2022), and employment law (Ajunwa, 2018). Our review of the legal landscape in Canada (where the majority of the researchers on this project work and live) found that the issue of age-related bias in AI has not been directly explored by primary Canadian authorities (statutes or case law); however, our search did reveal primary sources that may be extrapolated and applied to the analysis of age-related bias in AI. These primary sources included:

1. The Directive on Automated Decision-Making from the Government of Canada (Secretariat, 2019)
2. Laws of more general application including the Canadian Charter of Rights and Freedoms (The Constitution Act, 1982) and the Personal Information Protection and Electronic Documents Acts (Personal Information Protection and Electronic Documents Act, 2000),

3. Case law examining the section 15 equality rights as guaranteed by the Charter (Gosselin v. Quebec (Attorney General), [2002] 4S.C.R. 429, 2002 SCC 84, 2002);

The absence of primary authorities is unsurprising given the significant barriers to litigating bias in AI. The *black box* phenomenon, which refers to the ‘...lack of transparency and explainability in machine learning systems in regard to how and why the algorithm interprets data and reaches its conclusion’ (Gosselin v. Quebec (Attorney General), [2002] 4S.C.R. 429, 2002 SCC 84, 2002) poses a significant evidentiary hurdle for potential litigants. This is particularly true for claims where the plaintiff bears the burden of establishing a *prima facie* case of demonstrating disparate treatment by AI (Beatson et al., 2020; Ewert *v.* Canada, 2018 SCC 30, [2018] 2S.C.R. 165, 2018). Additionally, there remains, ‘only minimal, tangential, or no legislative regulation ... governing the development of certification or AI, the deployment of AI systems and tools with adequate training or policy support, or regular auditing and oversight of AI driven activities.’ (Beatson et al., 2020) At the time of our review, Canada did not have specific AI legislation, but instead relied on laws of general application, the existing market framework, and softer policies such as data governance frameworks, policies, standards, impact assessment tools, ethics boards, and other informal declarations (Martin-Bariteau and Scassa, 2021).

Given the significant challenges facing litigants seeking legal redress for biased AI, some scholars have postulated that the existing legal landscape is inadequate to address claims of AI related discrimination (Henderson et al., 2022). Although the issue of bias in AI has been broadly explored by scholars, our research suggests that the analysis has largely been in the context of gendered (Howard and Borenstein, 2018) and racialized bias (Angwin et al., 2016), not age-related bias or the consequences of digital ageism. Some legal scholars have mentioned age-related bias in a cautionary sense or have simply identified age-related bias as a possible concern, but none of the legal literature we reviewed directly explored the legal implications of age-related bias in AI within the healthcare context or otherwise.

Discussion

This scoping review included 74 articles from the academic ($n = 49$) and grey literature ($n = 25$) about age-based bias in machine learning algorithms pertaining to older adults to explore how age-related bias and digital ageism may be created, sustained, or reinforced algorithmically. This interdisciplinary scoping review contributes a systematic and comprehensive set of results identifying nine types of bias throughout the three phases along the machine learning pipeline and could be used to target future research and critical enquiry about digital ageism. We expand on our ideas below to answer each of the research questions.

To answer the first research question pertaining to what is known about digital ageism in AI technology, we identified nine distinct biases from the machine learning bias framework (Mehrabi et al., 2021) based on 19 different AI applications that cover a range of purposes. Among these AI applications, age recognition and facial recognition were the most frequently used. Our findings crucially demonstrate that age-related bias can manifest in various forms of AI technology along different points of the machine learning pipeline, ranging from underrepresentation in datasets to AI deployment. By identifying these biases, we can begin to address and mitigate them in order to create more equitable and inclusive AI systems. Publications about age-related bias have increased over the years in the included studies; however, our grey literature review found little mention of age-related bias in AI or its risks. Additionally, this highlights the need for more attention to be given to the

intersection of ageism and technology in order to better understand and address these issues through policy and legal parameters. More recently, the issue of age-based bias in AI and digital ageism has garnered more attention by academics (Chu et al., 2022b; Neves et al., 2023; Peine and Neven, 2021; Rosales and Fernández-Ardèvol, 2019; Stypinska, 2022), developers of algorithms (Wang et al., 2023), and The World Health Organization (WHO) published a 2022 policy brief on *Ageism in Artificial Intelligence for Health* (WHO, 2022a) as a follow-up to its *Global Report on Ageism* (WHO, 2021) after the completion of this review.

The review addresses the second research question by revealing the extent of age-related bias in AI systems. We found that age-related bias most commonly arises from the ‘data to algorithm’ phase ($n = 71$), followed by ‘algorithm to user’ ($n = 29$) and ‘user to data’ ($n = 11$) phases. The majority of academic publications ($n = 35$) contained discussion relevant to more than one type of bias, resulting in a high level of publication overlap between different types of bias. Our review reports empirical findings about measurement bias in the ‘data to algorithm’ phase indicate an under-performance of AI related to the data of older adults (see Table 2). These papers demonstrated measurement bias, primarily through facial and age recognition systems, revealing a higher mean error rate for older people than younger people. Of these 20 papers in our review which either demonstrated or discussed measurement bias, 12 also discussed or demonstrated representation and evaluation bias (Alashkar et al., 2020; Bekios-Calfa et al., 2011; Chin, 2019; Clapés et al., 2018; Georgopoulos et al., 2020a; Georgopoulos et al., 2020b; Klare et al., 2012; Kuang et al., 2015; Liang et al., 2019; Pei et al., 2017; Xie and Hsu, 2020; Zou et al., 2016). This key finding supports the notion that when algorithms are trained on datasets that lack adequate representation of older adults, their future performance on older adults’ data may be sub-optimal, which may in turn intensify inequality at the level of the AI system. These inequities have their roots in the historical and social bias that are then reflected in the ‘user to data’ biases. These biases are the product of our societal ageist tendencies. The results showed a binomial distribution of the number of biases in the datasets (data origin phase) and the user (deployment phase) which demonstrates the socio-technical aspects of how ageism manifests into AI systems and technologies. Therefore, it is essential not to disregard its social dimensions as emphasised by Crawford (2021). The interpretation, reaction, and application of AI, as well as the individuals involved in these processes, are influenced by broader social and political structures and procedures.

To address the third research question about how AI systems encode, produce, and reinforce age-related bias, we identified two important and distinct phenomena contributing to digital ageism. The first is the exclusion or absence of older adults from the data that is used to build models. We observed that the most utilised datasets for the purpose of facial and age recognition programs, such as FG-Net (Han and Jain, 2014), MORPH (UNCW, 2022) and Adience (Eidinger et al., 2014), significantly under-represented older adults. This lack of data about older adults could occur for several reasons, for example sampling bias, which stems from the use of sources like personal collections of photos, police booking databases, and social media, where older adults are less likely to be represented (Rosales and Fernández-Ardèvol, 2019). The second phenomenon is the misrepresentation of age as a category and older adults’ data. Our analysis revealed the aggregation of older adult data into broad age groups that inaccurately represents the diversity of older adults. This suggests that the process of labelling of data is another source of age bias, similar to Crawford’s, 2021, ch. 4, argument that labelling contributes to race and gender bias in datasets, such as ImageNet and

UTKFace (Crawford, 2021). Of note, is that we were unable to find this labelling approach applied to younger people and were unable to locate any rationale to justify the aggregation of data from older adults in the included academic papers.

The results of the grey literature search also produced findings about the societal, legal, and ethical implications of age-related bias in AI systems to address our fourth research question. First, there are societal and ethical implications where the amplification and replication of negative associations or stigmatising practices against older adults are concerned. The idea that AI can create, sustain and reinforce negative associations that equate ‘old’ with negative valuation and ‘young’ with positive valuation is unfairly discriminatory. Such processes reflect and reinforce broader social systems and structures that associate positive features with ‘youthfulness’. Further, these implicit age-related stereotypes held by younger people, as well as in society as a whole (Diaz et al., 2019; Rodriguez et al., 2017; Rozado, 2020), can maintain negative interpretations of aging and later life. For example, interviews with technology sector workers revealed that employees over the age of 30 are often considered ‘old’ within the industry, and worry about job security (Rosales and Svensson, 2020). These ageist perspectives permeate the digital space, where they can be detected and replicated by AI, and can shape the messages and ideas that circulate with regards to aging and later life. This process, left unfettered, raises several challenges given that AI has the potential to fundamentally change the experience of aging in our increasingly digitised society.

The reinforcement of negative stereotypes or old age through technology and AI is an important ethical issue as older people already experience disadvantages (Crystal et al., 2017) and social exclusion as they age in our society (Dannefer, 2003; Grenier et al., 2022; Walsh et al., 2017). While the results of this review were focused on the machine learning pipeline, the results must be discussed within a broader societal context. Intersectionality applied to aging outlines how intersecting social locations such as gender, race, class, and sexual orientation impact aging (Calasanti and King, 2015; McMullin and Ogmundson, 2005). The growing scope of AI systems that exhibit age-based bias may reinforce the societal practices and patterns based on discriminatory ageist beliefs and stereotypes in a negative feedback loop (Zou and Schiebinger, 2018) that perpetuates digital ageism. Such technology-mediated cycles of injustice can lead to algorithmic and social harm, for example the loss of social capital for older adults or opportunity costs across different areas such as healthcare, social welfare, and housing which could lead to further disadvantages (Chu et al., 2022d). Moreover, negative connotations related to ageing are internalised by older adults, changing their self-representation and further reducing their willingness or interest to engage with technology (Chu et al., 2022d; Gendron et al., 2016), which continues to drive the negative feedback loop (Köttl et al., 2021). Adequately addressing these complex problems requires a combination of solutions on multiple levels. Examples include the development of technical solutions, such as mitigation strategies, as well as exploring theoretical insights about the nature and impacts of intersecting structural inequities. However, while there is a need for research to advance our understanding on how to obtain accurate and representative data, these goals must be balanced with ethical considerations, including older adults’ preferences related to privacy, right to refuse, autonomy, fairness, and security.

Additionally, our enquiry into the legal implications of age-related bias in AI systems via the grey literature search demonstrated a lack of jurisprudential precedent and legal scholarship relating to age-related bias in machine learning. This is a fast-evolving area and an important element of the AI ecosystem. In Canada, the federal government introduced Bill C-27, the *Digital*

Charter Implementation Act in June 2022 that proposes to enact (among other things), the *Artificial Intelligence and Data Act* to regulate AI and reduce the potential of harm and bias in AI systems (Government of Canada, 2023). Further research is needed to monitor the evolving legal landscape and the potential that older people would have to challenge age-based exclusion and/or the non-realisation of their rights.

While our review of the legal literature focused on the Canadian legal landscape, research in this field could be expanded to include international case law and international treaties, such as the UNESCO global agreement on the ethics of AI (UNESCO, 2021). Although international authorities are not binding in the Canadian context, these authorities, particularly those from common law jurisdictions, can provide overarching policy setting agendas and insight as to how global and local legal landscapes may evolve to handle issues of ageism in AI in future. Comparative work with other jurisdictions could also prove informative. In the United States, for example, a complex web of local, state, and federal legislation and policy govern AI. At the federal level, the *National Artificial Intelligence Initiative Act* became law in January 2021 (National Artificial Intelligence Initiative Office, 2021), a blueprint for an AI Bill of Rights was published by the White House in October 2022 (The White House, 2022), and the *Algorithmic Accountability Act* was reintroduced in 2022 after failing to gain support in 2019 (Morgan et al., 2022). Further, many individual US states have introduced state-level AI legislation (National Conference of State Legislatures, 2022). Despite the increased legislative efforts to ensure AI is not used against the public interest, or the interests of minority groups, a larger complex issue that remains for scholars and legislators is to determine the type, amount, and extent of governance infrastructure required for the involvement in the impact assessment of AI and the oversight during development and deployment (Government of the United Kingdom, 2022). In the European Union, the proposed *Artificial Intelligence Act* (European Parliament, 2021) has been the subject of much legal discussion and may set a new global standard for regulating AI. Because these legal reforms are meant to reduce discriminatory or biased outcomes when algorithms are used in decision-making, future reviews focused on age-based bias and AI should consider including a legal analysis to monitor jurisprudential and legislative developments in the field.

In our efforts to explore older peoples' experiences of age-related bias as per our final research question, one glaring omission from our findings was the perspectives of older adults and their experiences of age-related bias. Our academic and grey literature searches did not return any research from the perspectives of older people with regards to ageism and AI. The lack of research in this area may represent the emerging nature of digital ageism, or potentially indicate another form of exclusion of older adults as a marginalised group. To gain a better understanding of the perspectives, needs, concerns, and preferences of older adults in the development of AI technology, it is crucial to pursue inclusive and participatory research that incorporates representative voices. This requires conducting research that involves older adults not just as study participants, but also as active decision-makers and contributors.

This scoping review comprised of an academic and grey literature search has several strengths. First, we used a rigorous methodology based on the Arksey and O'Malley (Arksey and O'Malley, 2005) scoping review framework, and a comprehensive search strategy that includes interdisciplinary and discipline-specific databases in relation to Mehrabi's framework outlining bias in AI. Our protocol was registered and published in JMIR Research Protocols as an open access document (Chu et al., 2022a). Second, the grey literature and web searches were

conducted to find additional sources to supplement the academic literature in order to assess societal, ethical and legal aspects. Third, the results were interpreted by an interdisciplinary team of experts from law, computer science, engineering, health, social gerontology and philosophy. One of the potential limitations of this study is the exclusion of publications in non-English languages. Also, our review would not have captured all examples of implicit age-related bias, only in papers that mentioned the keyword 'bias'; however, the studies included in this review can serve as exemplars of implicit and explicit bias. The implicit nature of ageism in the context of the technology sector with limited ethics oversights and regulations underscores the importance of further research and policy development. One major concern is that these biases and omissions may continue to produce exclusion and push older people whose experiences are not read as 'youthful' further outside the peripheries of shared social and cultural everyday spaces, including but not limited to those mediated through technological systems.

This review highlights several areas for future research. First, more research is needed to illuminate the perspectives of older adults, as well as examining their experiences with the intersectionality of algorithmic bias related to age, race, and gender (Chu et al., 2022b) from various theoretical perspectives would increase our understanding of the impact of digital ageism. Added to this is an exploration of the mitigation strategies that could be used to reduce age-related biases found in AI. Second, this academic literature review only uncovered efforts to address representation and measurement bias, thus more research in this area is warranted to understand possible technical and methodological solutions. For example, the generation of best practices and a theoretical framework to improve the inclusion of older people is needed, as well as more research about the efficacy of the methods used to include older adults in research (Fischer et al., 2020). Third, to advance the field of digital ageism, there is a pressing need to establish a consensus on research and policy priorities, bringing together diverse transdisciplinary perspectives for collective progress and to instigate meaningful change. These should include national and international agencies, research institutes, organizations, and consortiums should collaborate in an effort to embed the identified priorities given the vast reach and scale of digital ageism. Future policy directions should include determining effective technical, design and governance strategies to address issues of representation and misrepresentation of older adults (data to algorithm), poor evaluation and model building (algorithm to user), and societal and structural ageism that generates public data (user to data) - or some combination thereof. Fourth, to address algorithmic under-performance, a conscious systematic effort is needed which should involve developing and evaluating technical and methodological approaches to mitigate digital ageism. Based on the review findings, digital ageism is an extension of the negative societal devaluation and representation of older people in the realm of technology and its effects (Chu et al., 2022b). The broader critical and scholarly discussion about the societal structures that contribute to ageism should continue to include the relationship between technology and age as it intersects with social locations, processes of social exclusion and inequitable power dynamics within social structures, in institutional processes, and the everyday lives of older people (Nyrup et al., 2023). Finally, future research should include societal discourse about older adults and AI and how the stereotypes and biases are proliferated to create ageist norms as this would provide important context for research about digital ageism. Given our increasing aging population, there exists an ethical imperative to ensure that societal and technological advances occur in a way that takes into meaningful consideration the lives of older adults. Future advancements in this area need to be interdisciplinary,

involve a wide range of insights from older people, and understood within the broader and place-based social, cultural and political contexts within which they occur.

Conclusion

The findings of our scoping review and grey literature searches advance the discourse about age-related bias in AI systems, digital ageism, and the related societal, ethical and legal implications. The foundational knowledge gained through this study has identified various challenges and areas of future research, policy and action to address age-related bias along the machine learning pipeline. Future research and efforts can use these empirical findings to further advance theory and practice to address digital ageism, including but not limited to, strategies to mitigate the biases. The results underscore the need for better representation of older people and their perspectives, as well as the development of protective mechanisms concerned with ethics, privacy and legal rights. Similar to other forms of algorithmic bias, digital ageism is deeply entangled with societal biases and wider structural inequalities. A concerted multifaceted interdisciplinary effort will be needed to begin to effectively address them.

Data availability

Data sharing is not applicable to this research as no data were generated or analysed.

Received: 28 November 2022; Accepted: 27 July 2023;

Published online: 17 August 2023

References

- Abderrahmane MA, Guelzim I, Abdelouahad AA (2020) Hand image-based human age estimation using a time distributed CNN-GRU. In: Proceedings of the International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI), Sakheer, Bahrain, 26–27 Oct 2020. <https://doi.org/10.1109/ICDABI51230.2020.9325667>
- Abdurrahim SH, Samad SA, Huddin AB (2018) Review on the effects of age, gender, and race demographics on automatic face recognition. *Vis Comput* 34(11):1617–1630. <https://doi.org/10.1007/s00371-017-1428-z>
- AITrends (2019) IDC: legislation to ban use of facial recognition could restrict public sector innovation. <https://web.archive.org/web/20190711214827/https://www.aitrends.com/ai-in-government/idc-legislation-to-ban-use-of-facial-recognition-could-restrict-public-sector-innovation/>. Accessed 5 Jan 2021
- Ajunwa G (2018) How artificial intelligence can make employment discrimination worse. *The Independent*. <https://suindependent.com/artificial-intelligence-can-make-employment-discrimination-worse/>. Accessed 5 Jan 2021
- Alashkar R, El Sabbahy M, Sabha A et al (2020) AI-vision towards an improved social inclusion. In: Proceedings of the IEEE/ITU International Conference on Artificial Intelligence for Good (AI4G), Virtual Event, 21–23 Sept 2020. <https://doi.org/10.1109/AI4G50087.2020.9311049>
- Alexander J, Logashanmugam E (2016) Image based human age estimation using principle component analysis/artificial neural network. *J Eng Appl Sci* 11:6859–6862
- Amini A, Soleimany, AP, Schwarting W (2019) Uncovering and Mitigating Algorithmic Bias through Learned Latent Structure. In: Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society, Honolulu, Hawaii, 27–28 Jun 2019. <https://doi.org/10.1145/3306618.3314243>
- Anderson M, Perrin A (2017) Tech Adoption Climbs Among Older Adults. *Pew Research Center: Internet, Science and Technology*. <https://www.pewresearch.org/internet/2017/05/17/tech-adoption-climbs-among-older-adults/>. Accessed 5 Jan 2021
- Angwin J, Larson J, Mattu S et al (2016) Machine bias risk assessments in criminal sentencing. *ProPublica*. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Accessed 5 Jan 2021
- Arksey H, O'Malley L (2005) Scoping studies: towards a methodological framework. *Int J Soc Res Methodol* 8(1):19–32. <https://doi.org/10.1080/1364557032000119616>
- Ashley KD (2017) Artificial intelligence and legal analytics: new tools for law practice in the digital age. Cambridge University Press, Cambridge, 10.1017/9781316761380
- Basumallick C (2019a) 4 workplace diversity trends we can expect in 2019. <https://www.spiceworks.com/hr/engagement-retention/articles/4-workplace-diversity-trends-for-2019/>. Accessed 5 Jan 2021
- BasuMallick C (2019b) How retech can help improve gender diversity in the tech industry. *Technology signals*. <https://technology-signals.com/how-rectech-can-help-improve-gender-diversity-in-the-tech-industry/>. Accessed 5 Jan 2021
- Beatson J, Chan G, Presser JR (2020) Litigating artificial intelligence. Edmond Publishing. <https://emond.ca/ai21>. Accessed 5 Jan 2021
- Begg R, Kamruzzaman J (2006) Neural networks for detection and classification of walking pattern changes due to ageing. *Australas Phys Eng S* 29(2):188. <https://doi.org/10.1007/BF03178892>
- Bekios-Calfá J, Buenaposada JM, Baumela L (2011) Age regression from soft aligned face images using low computational resources. In Vitrià J, Sanches JM, Hernández M (Eds.) *Pattern recognit. Image Anal.* (pp. 281–288). Springer. https://doi.org/10.1007/978-3-642-21257-4_35
- Berendt B, Preibusch S (2014) Better decision support through exploratory discrimination-aware data mining: foundations and empirical evidence. *Artif Intell Law* 22:175–209. <https://doi.org/10.1007/s10506-013-9152-0>
- Billette V, Lavoie JP, Séguin AM et al. (2012) Réflexions sur l'exclusion et l'inclusion sociale en lien avec le vieillissement. L'importance des enjeux de reconnaissance et de redistribution. *Frontières* 25(1):10–30. <https://doi.org/10.7202/1018229ar>
- Biswas S, Rajan H (2020) Do the machine learning models on a crowd sourced platform exhibit bias? An empirical study on model fairness. In: Proceedings of the 28th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE '20), Virtual event, New York, 8–13 Nov 2020. <https://doi.org/10.1145/3368089.3409704>
- Blank A (2021) How AI & mindfulness can tackle age bias in the modern workplace. *Spiceworks*. <https://www.spiceworks.com/hr/hr-strategy/guest-article/how-ai-mindfulness-can-tackle-age-bias-in-the-modern-workplace/>. Accessed 5 Sept 2022
- Calasanti T, King N (2015) Intersectionality & age. In: *Routledge Handbook of Cultural Gerontology* (pp. 215–222). Routledge, Oxfordshire
- Center for Democracy and Technology (2018) Digital decisions. <https://cdt.org/wp-content/uploads/2018/09/Digital-Decisions-Library-Printer-Friendly-as-of-20180927.pdf>. Accessed 20 Jan 2021
- Chin C (2019) Assessing employer intent when AI hiring tools are biased. *Brookings Institute*. <https://www.brookings.edu/research/assessing-employer-intent-when-ai-hiring-tools-are-biased/>. Accessed 5 Jan 2021
- Chu CH, Leslie K, Shi J et al. (2022a) Ageism and artificial intelligence: protocol for a scoping review. *JMIR Res Protoc* 11(6):e33211. <https://doi.org/10.2196/33211>
- Chu CH, Nyrupe R, Leslie K et al. (2022b) Digital ageism: challenges and opportunities in artificial intelligence for older adults. *Gerontologist* 62(7):947–955. <https://doi.org/10.1093/geront/gnab167>
- Chu C, Leslie K, Nyrupe R et al (2022c) Artificial intelligence can discriminate on the basis of race and gender, and also age. *The Conversation*. <http://theconversation.com/artificial-intelligence-can-discriminate-on-the-basis-of-race-and-gender-and-also-age-173617>. Accessed 5 Sept 2022
- Chu C, Nyrupe R, Donato-Woodger S et al (2022d) Examining the technology-mediated cycles of injustice that contribute to digital ageism: advancing the conceptualization of digital ageism, evidence, and implications. In: Proceedings of the 15th International Conference on Pervasive Technologies Related to Assistive Environments (PETRA), Corfu, Greece, 29 June–3 July 2022. <https://doi.org/10.1145/3529190.3534765>
- Clapés A, Anbarjafari, G, Bilici, O et al (2018) From apparent to real age: gender, age, ethnic, makeup, and expression bias analysis in real age estimation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 18–23 June, Salt Lake City, Utah. <https://doi.org/10.1109/CVPRW.2018.00314>
- Constine J (2017) Pymetrics attacks discrimination in hiring with AI and recruiting games. *TechCrunch.Com*. <https://techcrunch.com/2017/09/20/unbiased-hiring/>. Accessed 5 Jan 2021
- Crawford K (2021) *Atlas of AI*. Yale University Press, New Haven, Connecticut
- Crystal S, Shea D, Reyes AM (2017) Cumulative advantage, cumulative disadvantage, and evolving patterns of late-life inequality. *Gerontologist* 57(5):910–920. <https://doi.org/10.1093/geront/gnw056>
- Culotta A, Ravi NK, Cutler J (2016) Predicting twitter user demographics using distant supervision from website traffic data. *J Artif Intell Res* 55:389–408. <https://doi.org/10.1613/jair.4935>
- Culter SJ (2005) Ageism & technology. *Generations* 29:67–72
- Dannefer D (2003) Cumulative advantage/disadvantage and the life course: cross-fertilizing age and social science theory. *J Gerontol B Psychol Sci Soc Sci* 58(6):S327–337. <https://doi.org/10.1093/geronb/58.6.s327>
- Datta A, Tschantz MC, Datta A (2015) Automated experiments on ad privacy settings: a tale of opacity, choice, and discrimination (arXiv:1408.6491). *arXiv*. <https://doi.org/10.48550/arXiv.1408.6491>

- Dawson D, Schleiger E, Horton, J et al (2019) Artificial intelligence: Australia's ethics framework—a discussion paper (Australia). Analysis and Policy Observatory. <https://apo.org.au/node/229596> Accessed 5 Jan 2021
- Dev S, Phillips J (2019) Attenuating bias in word vectors. arXiv. <https://doi.org/10.48550/arXiv.1901.07656>
- Devakunchari R (2014) Analysis on big data over the years. IJSRP 4(1):7
- Diaz M, Johnson I, Lazar A et al (2019) Addressing age-related bias in sentiment analysis. In: Proceedings of the CHI Conference on Human Factors in Computing Systems, Montreal, Canada, 21–26 April 2018. <https://doi.org/10.24963/ijcai.2019/852>
- Dinges DF, Rider RL, Dorrian J et al. (2005) Optical computer recognition of facial expressions associated with stress induced by performance demands. *Aviat Space Environ Med* 76(6 Suppl):B172–182
- Diraco G, Leone A, Siciliano P (2017) A radar-based smart sensor for unobtrusive elderly monitoring in ambient assisted living applications. *Biosensors* 7(4):E55. <https://doi.org/10.3390/bios7040055>
- Drage E, Mackereth K (2022) Does AI debias recruitment? Race, gender, and AI's "eradication of difference". *Philos Technol* 35(4):89. <https://doi.org/10.1007/s13347-022-00543-1>
- Drozdzowski P, Prommegger B, Wimmer G et al. (2020) Demographic bias: a challenge for fingevein recognition systems? In: Proceedings of the 28th European Signal Processing Conference (EUSIPCO), Amsterdam, Netherlands, 18–21 Jan 2021. <https://doi.org/10.23919/Eusipco47968.2020.9287722>
- Druga S, Yip J, Preston M et al. (2021) The 4As: ask, adapt, author, analyze - AI literacy framework for families. *Works in Progress*. <https://wip.mitpress.mit.edu/pub/the-4as/release/1>. Accessed 5 Sept 2022
- Eidinger E, Enbar R, Hassner T (2014) Age and gender estimation of unfiltered faces. *IEEE Trans Inf Forensics Secur* 9(12):2170–2179. <https://doi.org/10.1109/TIFS.2014.2359646>
- Engler A (2020) A guide to healthy skepticism of artificial intelligence and coronavirus. Brookings Institute. <https://www.brookings.edu/research/a-guide-to-healthy-skepticism-of-artificial-intelligence-and-coronavirus/> Accessed 5 Sept 2022
- European Parliament (2021) Proposal for a regulation of the European parliament and the council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts, 2021/0106(COD). <https://eur-lex.europa.eu/legal-content/EN/TEXT/?uri=CELEX%3A52021PC0206> Accessed 5 Sept 2022
- Ewert *v.* Canada, 2018 SCC 30, [2018] 2 S.C.R. 165, No. 37233 (Supreme Court of Canada June 13, 2018). <https://www.scc-csc.ca/case-dossier/cb/37233-eng.aspx>. Accessed 5 Jan 2021
- Fenech M, Strukelj N, Buston O (2018) Ethical, social, and political challenges of artificial intelligence in health. *Future Advocacy*. <https://rri-tools.eu/-/ethical-social-and-political-challenges-of-artificial-intelligence-in-health>
- Fischer B, Peine A, Östlund B (2020) The importance of user involvement: a systematic review of involving older users in technology design. *Gerontologist* 60(7):e513–e523. <https://doi.org/10.1093/geront/gnz163>
- Fischer M (2021) How AI can and will affect the recruiting process (and how it won't). Spiceworks. <https://www.spiceworks.com/hr/hr-strategy/guest-article/how-ai-can-and-will-affect-the-recruiting-process-and-how-it-wont/>
- Freedom House (2021) Coalition letter requests federal moratorium on the use of facial recognition technology. <https://freedomhouse.org/article/coalition-letter-requests-federal-moratorium-use-facial-recognition-technology>. Accessed 5 Jan 2021
- Froomkin AM, Kerr I, Pineau J (2019) When AIs outperform doctors: confronting the challenges of a tort-induced over-reliance on machine learning. 61 *Ariz. L. Rev.* 33 (2019) University of Miami Legal Studies Research Paper No. 18-3. <https://doi.org/10.2139/ssrn.3114347>
- Future of Privacy Forum (2017) Unfairness by algorithm: distilling the harms of automated decision-making. <https://Fpf.Org/>. <https://fpf.org/blog/unfairness-by-algorithm-distilling-the-harms-of-automated-decision-making/>. Accessed 05 Jan 2021
- Gendron TL, Welleford EA, Inker J et al. (2016) The language of ageism: why we need to use words carefully. *Gerontologist* 56(6):997–1006. <https://doi.org/10.1093/geront/gnv066>
- Georgopoulos M, Oldfield J, Nicolaou MA, Panagakis Y, Pantic M (2020a) Enhancing facial data diversity with style-based face aging. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, p 14–15. https://openaccess.thecvf.com/content_CVPRW_2020/papers/w1/Georgopoulos_Enhancing_Facial_Data_Diversity_With_Style-Based_Face_Aging_CVPRW_2020_paper.pdf
- Georgopoulos M, Panagakis Y, Pantic M (2020b) Investigating bias in deep face analysis: The KANFACE dataset and empirical study. *Image and Vision Computing* 102:103954. <https://doi.org/10.1016/j.imavis.2020.103954>
- Ghosh P (2020) What is recruitment? Definition, process, techniques, metrics and strategies for 2020. Spiceworks. <https://www.spiceworks.com/hr/recruitment-onboarding/articles/what-is-recruitment/> Accessed 5 Sept 2022
- Giudici P (2018) Fintech risk management: a research challenge for artificial intelligence in finance. *Front Artif Intell* 1:1. <https://doi.org/10.3389/frai.2018.00001>
- Gosselin v. Quebec (Attorney General) (2002) 4 S.C.R. 429, 2002 SCC 84, No. 27418 (Supreme Court of Canada December 19, 2002). <https://scc-csc.lexum.com/scc-csc/scc-csc/en/item/2027/index.do>. Accessed 5 Jan 2021
- Government of Canada (2023) The Artificial Intelligence and Data Act (AIDA)—Companion document. Innovation, Science and Economic Development Canada; Innovation, Science and Economic Development Canada. <https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act-aida-companion-document>. Accessed 8 Apr 2023
- Government of the United Kingdom (2022) Establishing a pro-innovation approach to regulating AI (CP 728). <https://www.gov.uk/government/publications/establishing-a-pro-innovation-approach-to-regulating-ai/establishing-a-pro-innovation-approach-to-regulating-ai-policy-statement>. Accessed 14 Nov 2022
- Grenier A, Phillipson C, Settersen Jr R (2022) Precarity and ageing. Policy Press & Bristol University Press, Old Park Hill, Bristol, <https://policy.bristoluniversitypress.co.uk/precarity-and-ageing>
- Guégan D, Hassani B (2018) Regulatory learning: how to supervise machine learning models? An application to credit scoring. *The Journal of Finance and Data Science* 4(3):157–171. <https://doi.org/10.1016/j.jfds.2018.04.001>
- Ham B (2021) Using machine learning to predict high-impact research. MIT Media Lab. <https://www.media.mit.edu/articles/using-machine-learning-to-predict-high-impact-research/>. Accessed 5 Sept 2022
- Han H, Jain AK (2014) Age, gender and race estimation from unconstrained face images. MSU Technical Report MSU-CE-14-5
- Helleringer S, You C, Fleury L et al. (2019) Improving age measurement in low- and middle-income countries through computer vision: a test in Senegal. *Demogr Res* 40(9):219–260. <https://doi.org/10.4054/DemRes.2019.40.9>
- Henderson B, Flood C, Scassa T (2022) Artificial intelligence in Canadian healthcare: will the law protect us from algorithmic bias resulting in discrimination? *CJLT* 19(2):24
- Ho DE, King J, Wald RC et al. (2021) Building a national AI research resource: a blueprint for the national research cloud. Stanford University. https://hai.stanford.edu/sites/default/files/2022-01/HAI_NRCR_v17.pdf
- Howard A, Borenstein J (2018) The ugly truth about ourselves and our robot creations: the problem of bias and social inequity. *Sci Eng Ethics* 24(5):1521–1536. <https://doi.org/10.1007/s11948-017-9975-2>
- Jansen P, Brey P, Fox A et al. (2020) SIENNA D4.4: ethical analysis of ai and robotics technologies (version V1.1). Zenodo. <https://doi.org/10.5281/zenodo.4068083>
- Jung SG, An J, Kwak H et al. (2018) Assessing the accuracy of four popular face recognition tools for inferring gender, age, and race. In: Proceedings of the International AAAI Conference on Web and Social Media 12/Vol. 12 No. 1 (2018): Twelfth International AAAI Conference on Web and Social Media, Palo Alto, California, 25–28 Jun 2018. <https://doi.org/10.1609/icwsm.v12i1.15058>
- Kantayya S (Director) (2021) Spotlight—Coded Bias Documentary [Documentary]. Algorithmic Justice League. <https://www.ajl.org/facial-recognition-technology>
- Kariuki D (2021) How employers are using artificial intelligence to stop bias in hiring. Spiceworks. <https://www.spiceworks.com/hr/hr-strategy/articles/how-employers-are-using-ai-to-stop-bias-in-hiring/>. Accessed 5 Sept 2022
- Kim E, Bryant D, Srikanth D, et al. (2021) Age bias in emotion detection: an analysis of facial emotion recognition performance on young, middle-aged, and older adults. In: Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society, Virtual Event, 19–21 May 2022. <https://doi.org/10.1145/3461702.3462609>
- Kinnard K (2018) Ageism is hurting your tech company's hiring more than you realize. Entrepreneur. <https://www.entrepreneur.com/leadership/ageism-is-hurting-your-tech-companys-hiring-more-than-you/318816>. Accessed 5 Jan 2021
- Klare BF, Burge MJ, Klontz JC et al. (2012) Face recognition performance: role of demographic information. *IEEE Trans Inf Forensics Secur* 7(6):1789–1801. <https://doi.org/10.1109/TIFS.2012.2214212>
- Köttl H, Gallistl V, Rohner R et al. (2021) “But at the age of 85? Forget it!”: Internalized ageism, a barrier to technology use. *J Aging Stud* 59:100971. <https://doi.org/10.1016/j.jaging.2021.100971>
- Kuang Z, Huang C, Zhang W (2015) Deeply learned rich coding for cross-dataset facial age estimation. In: Proceedings of the IEEE International Conference on Computer Vision Workshop (ICCVW), Santiago, Chile, 7–15 Dec 2015. <https://doi.org/10.1109/ICCVW.2015.52>
- Lanka P, Rangaprakash D, Dretsch MN et al. (2020) Supervised machine learning for diagnostic classification from large-scale neuroimaging datasets. *Brain Imaging Behav* 14(6):2378–2416. <https://doi.org/10.1007/s11682-019-00191-8>
- Lee NT (2016) Addressing racial bias in the online economy. Brookings Institute. <https://www.brookings.edu/blog/techtank/2016/12/01/addressing-racial-bias-in-the-online-economy/>. Accessed 5 Jan 2021
- Leufer D (2021) How AI systems undermine LGBTQ identity. Access now. <https://www.accessnow.org/how-ai-systems-undermine-lgbtq-identity/>. Accessed 5 Sept 2022
- Levac D, Colquhoun H, O'Brien KK (2010) Scoping studies: advancing the methodology. *Implementation Sci* 5:69. <https://doi.org/10.1186/1748-5908-5-69>

- Li Z, Gong D, Zhu K et al. (2017) Multifeature anisotropic orthogonal gaussian process for automatic age estimation. *ACM Trans Intell Syst Technol* 9(1):2. <https://doi.org/10.1145/3090311>. 1-2:15
- Liang H, Zhang F, Xin N (2019) Investigating systematic bias in brain age estimation with application to post-traumatic stress disorders. *Hum Brain Mapp* 40:3143–3152. <https://doi.org/10.1002/hbm.24588>
- Liu T, Lei Z, Wan, J et al. (2015) DFDnet: discriminant face descriptor network for facial age estimation. In Yang J, Yang J, Sun Z et al (Eds.) In: 10th Chinese Conference Biometric Recognition (CCBR 2015), Tianjin, China, 13–15 Nov 2015. https://doi.org/10.1007/978-3-319-25417-3_76
- Luu K, Ricanek K, Bui, TD et al. (2009) Age estimation using active appearance models and support vector machine regression. In: Proceedings Of the IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems, Washington DC, 28–30 Sept 2009. <https://doi.org/10.1109/BTAS.2009.5339053>
- MacCarthy M (2021) Mandating fairness and accuracy assessments for law enforcement facial recognition systems. Brookings Institute. <https://www.brookings.edu/blog/techtank/2021/05/26/mandating-fairness-and-accuracy-assessments-for-law-enforcement-facial-recognition-systems/>. Accessed 5 Sept 2022
- Margetts H, Dorobantu C (2019) Rethink government with AI. *Nature* 568(7751):163–165. <https://doi.org/10.1038/d41586-019-01099-5>
- Martin-Bariteau F, Scassa T (2021) Artificial intelligence and the law in Canada. LexisNexis. <https://store.lexisnexis.ca/en/products/artificial-intelligence-and-the-law-in-canada-skusku-cad-6810/details>. Accessed 14 Nov 2022
- McDonough C (2016) The effect of ageism on the digital divide among older adults. *Gerontol Geriatr Med* 2(1):1–7. <https://doi.org/10.24966/GGM-8662/100008>
- McMullin J, Ogmundson RL (2005) Understanding social inequality: intersections of class, age, gender, ethnicity and race in Canada. *Can J Sociol* 30(4):549–550
- Mehrabi N, Morstatter F, Saxena N et al. (2021) A survey on bias and fairness in machine learning. *ACM Comput Surv* 54(6):1–115. <https://doi.org/10.1145/3457607>. 115:35
- Morgan CS, Langlois F, Lee J (2022) U.S. HOUSe And Senate Reintroduce The Algorithmic Accountability Act intended to regulate AI. McCarthy Tetrault. <https://www.mccarthy.ca/en/insights/blogs/techlex/us-house-and-senate-reintroduce-algorithmic-accountability-act-intended-regulate-ai>. Accessed 5 Sept 2022
- National Artificial Intelligence Initiative Office (2021) Legislation & Executive Orders. <https://www.ai.gov/legislation-and-executive-orders/>. Accessed 14 Nov 2022
- National Conference of State Legislatures (2022) Legislation related to artificial intelligence. <https://www.ncsl.org/research/telecommunications-and-information-technology/2020-legislation-related-to-artificial-intelligence.aspx>. Accessed 14 Nov 2022
- Neary M, Chen S (2017) Artificial intelligence: legal research and law librarians. *AALL Spectrum* 21(5):16–20
- Neven L (2010) “But obviously not for me”: robots, laboratories and the defiant identity of elder test users. *Sociol Health Illn* 32(2):335–347. <https://doi.org/10.1111/j.1467-9566.2009.01218.x>
- Neves BB, Petersen A, Vered M et al. (2023) Artificial intelligence in long-term care: technological promise, aging anxieties, and sociotechnical agency. *J Appl Gerontol* 42(6):1274–1282. <https://doi.org/10.1177/07334648231157370>
- Nyrup R, Chu C, Falco E (2023) Digital ageism, algorithmic bias, and feminist critical theory. In: Brown J, Cave S, Mackereth K, et al., (Eds.) *Feminist AI: critical perspectives on data, algorithms and intelligent machines*. Oxford University Press, Cary, North Carolina
- Obermeyer Z, Powers B, Vogeli C et al. (2019) Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 366(6464):447–453. <https://doi.org/10.1126/science.aax2342>
- Pei W, Dibeklioglu H, Baltrušaitis T et al. (2017) Attended end-to-end architecture for age estimation from facial expression videos. *IEEE Trans Image Process* 29:1972–1984. <https://doi.org/10.1109/TIP.2019.2948288>
- Peine A, Neven L (2021) The co-constitution of ageing and technology—a model and agenda. *Ageing Soc* 41(12):2845–2866. <https://doi.org/10.1017/S0144686X20000641>
- Personal Information Protection and Electronic Documents Act, Pub. L. No. S.C. 2000 (2000) <https://laws-lois.justice.gc.ca/eng/acts/p-8.6/index.html>
- Principi R, Palmero C, Junior J et al. (2019) On the effect of observed subject biases in apparent personality analysis from audio-visual signals. *IEEE Trans Affect Comput* 12(3):607–621. <https://doi.org/10.1109/TAFFC.2019.2956030>
- Rodriguez P, Cucurull G, Gouffaux J et al. (2017) Age and gender recognition in the wild with deep attention. *Pattern Recognit* 72:563–571. <https://doi.org/10.1016/j.patcog.2017.06.028>
- Rosales A, Fernández-Ardèvol M (2016) Beyond WhatsApp: older people and smartphones. *Romanian J Commun Public Relat* 18(37):27–47. <https://doi.org/10.21018/rjcp.2016.1.200>
- Rosales A, Fernández-Ardèvol M (2019) Structural ageism in big data approaches. *Nord Rev* 40:51–64. <https://doi.org/10.2478/nor-2019-0013>
- Rosales A, Svensson J (2020) Perceptions of age in contemporary tech. *Nord Rev* 42(1):79–91. <https://doi.org/10.2478/nor-2021-0021>
- Rozado D (2020) Wide range screening of algorithmic bias in word embedding models using large sentiment lexicons reveals underreported bias types. *PLoS One* 15(4):e0231189. <https://doi.org/10.1371/journal.pone.0231189>
- Russell S, Norvig P (2020) *Artificial intelligence: a modern approach*, 4th ed. Pearson, Hoboken, New Jersey
- Secretariat, The Treasury Board of Canada (2019) Directive on Automated Decision-Making. The Treasury Board of Canada. <https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592>. Accessed 5 Sept 2022
- Smith C, McGuire B, Huang T et al. (2006) The history of artificial intelligence. University of Washington. <https://courses.cs.washington.edu/courses/csep590/06au/projects/history-ai.pdf> Accessed 5 Sept 2022
- Smith P, Ricanek K (2020) Mitigating algorithmic bias: evolving an augmentation policy that is non-biasing. In: Proceedings of the IEEE Winter Applications of Computer Vision Workshops (WACVW), Snowmass Village, Colorado, 1–5 Mar 2020. <https://doi.org/10.1109/WACVW50321.2020.9096905>
- Smith-Miles K, Geng X (2020) Revisiting facial age estimation with new insights from instance space analysis. *IEEE Trans Pattern Anal Mach Intell* 44(5):2689–2697. <https://doi.org/10.1109/TPAMI.2020.3038760>
- Sourdin T, Cornes R (2018) Do judges need to be human? The implications of technology for responsive judging. In: Sourdin T, Zariski A (eds) *The responsive judge. ius gentium: comparative perspectives on law and justice*. Springer, Singapore, 10.1007/978-981-13-1023-2_4
- Srinivasan D, Erus G, Doshi J et al. (2020) A comparison of Freesurfer and multi-atlas MUSE for brain anatomy segmentation: findings about size and age bias, and inter-scanner stability in multi-site aging studies. *NeuroImage* 223:117248. <https://doi.org/10.1016/j.neuroimage.2020.117248>
- Strath SJ, Kate RJ, Keenan KG et al. (2015) Ngram time series model to predict activity type and energy cost from wrist, hip and ankle accelerometers: implications of age. *Physiol Meas* 36(11):2335–2351. <https://doi.org/10.1088/0967-3334/36/11/2335>
- Stypinska J (2022) AI ageism: a critical roadmap for studying age discrimination and exclusion in digitalized societies. *AI Soc* 38(2):665–677. <https://doi.org/10.1007/s00146-022-01553-5>
- Sun Y, Tang J, Shu X et al. (2020) Facial age synthesis with label distribution-guided generative adversarial network. *IEEE Trans Inf Forensics Secur* 15:2679–2691. <https://doi.org/10.1109/TIFS.2020.2975921>
- Suresh H, Guttap JV (2021) A framework for understanding sources of harm throughout the machine learning life cycle. In: Proceedings of the EAAMO '21: Equity and Access in Algorithms, Mechanisms, and Optimization, New York City, 5–9 Oct 2021. <https://doi.org/10.1145/3465416.3483305>
- Taati B, Zhao S, Ashraf AB et al. (2019) Algorithmic bias in clinical populations—evaluating and improving facial analysis technology in older adults with dementia. *IEEE Access* 7:25527–25534. <https://doi.org/10.1109/ACCESS.2019.2900022>
- Terhöst P, Kolff JN, Damer N et al. (2020) Face quality estimation and its correlation to demographic and non-demographic bias in face recognition. In: Proceedings of the IEEE International Joint Conference on Biometrics (IJCB), Virtual Event, 28 Sept–1 Oct. <https://doi.org/10.1109/IJCB48548.2020.9304865>
- The Constitution Act, § 7, P.1 (1982) <https://laws-lois.justice.gc.ca/eng/Const/page-12.html>
- The Royal Society (2017) Machine learning: the power and promise of computers that learn by example. <https://royalsocietypublishing.org/~/media/policy/projects/machine-learning/publications/machine-learning-report.pdf>. Accessed 5 Jan 2021
- The White House (2022) Blueprint for an AI bill of rights. <https://www.whitehouse.gov/ostp/ai-bill-of-rights/>. Accessed 14 Nov 2022
- Tian Q, Sun H, Ma C et al. (2020) Age estimation via selecting discriminated features and preserving geometry. *KSII Trans Internet Inf Syst* 14(4):1721–1737. <https://doi.org/10.3837/tiis.2020.04.017>
- Todd O, Burton J, Dodds R et al. (2019) New horizons in the use of routine data for ageing research. *Age Ageing* 49(5):716–722. <https://doi.org/10.1093/ageing/afaa018>
- Tokola R, Bolme D, Boehnen C et al. (2014) Discriminating projections for estimating face age in wild images. In: Proceedings of the IEEE International Joint Conference on Biometrics, Clearwater, Florida, 29 Sept–1 Oct, 2014. <https://doi.org/10.1109/BTAS.2014.6996287>
- Tricco AC, Lillie E, Zarin W et al. (2018) PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 169(7):467–473. <https://doi.org/10.7326/M18-0850>
- Tsai HYS, Shillair R, Cotton SR et al. (2015) Getting grandma online: are tablets the answer for increasing digital inclusion for older adults in the U.S.? *Educ Gerontol* 41(10):695–709. <https://doi.org/10.1080/03601277.2015.1048165>
- UNCW (2022) Morph database. Office of Innovation & Commercialization. <https://uncw.edu/oic/tech/morph.html>. Accessed 5 Sept 2022
- UNESCO (2021) UNESCO member states adopt the first ever global agreement on the ethics of artificial intelligence. <https://www.unesco.org/en/articles/unesco-member-states-adopt-first-ever-global-agreement-ethics-artificial-intelligence>. Accessed 5 Sept 2022

- Walsh K, Scharf T, Keating N (2017) Social exclusion of older persons: a scoping review and conceptual framework. *Eur J Ageing* 14(1):81–98. <https://doi.org/10.1007/s10433-016-0398-8>
- Wang X, Kambhampati C (2015) Age estimation via unsupervised neural networks. In: 2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), Ljubljana, Slovenia, 4–8 May 2015. <https://doi.org/10.1109/FG.2015.7163119>
- Wang Y (2017) China is quickly embracing facial recognition tech, for better and worse. *Forbes*. <https://www.forbes.com/sites/ywang/2017/07/11/how-china-is-quickly-embracing-facial-recognition-tech-for-better-and-worse/>. Accessed 5 Sept 2022
- Wang Z, Zhou Y, Qiu M et al. (2023) Towards fair machine learning software: understanding and addressing model bias through counterfactual thinking (arXiv:2302.08018). arXiv. <https://doi.org/10.48550/arXiv.2302.08018>
- Windegger M (2018) Mr. Craig Mokhiber [Text]. In: ICHROP - International Conference on Human Rights of Older Persons, Vienna, Austria, 12–13 Nov 2018. <http://ageing.at/en/speaker/mr-craig-mokhiber>
- World Health Organization (2021) Global report on ageism. <https://www.who.int/publications-detail-redirect/9789240016866>. Accessed 5 Sept 2022
- World Health Organization (2022a) Ageism in Artificial intelligence for Health. <https://www.who.int/publications-detail-redirect/9789240040793>. Accessed 5 Sept 2022
- Xie RC, Hsu GSJ (2020) A hybrid network for facial age progression and regression learning. In: 2020 International Conference on Advanced Robotics and Intelligent Systems (ARIS), Taipei, Taiwan, 19–21 August 2020. <https://doi.org/10.1109/ARIS50834.2020.9205788>
- Ye L, Li B, Mohammed N, Wang Y et al. (2018) Privacy-preserving age estimation for content rating. In: MMSP 2018 IEEE 20th International Workshop on Multimedia Signal Processing, Vancouver, Canada, 28–31 Aug 2018. <https://doi.org/10.1109/MMSP.2018.8547144>
- Zhao Q, Adeli E, Pohl KM (2020) Training confounder-free deep learning models for medical applications. *Nat Commun* 11(1):6010. <https://doi.org/10.1038/s41467-020-19784-9>
- Zhao L (2020) Data-driven approach for predicting and explaining the risk of long-term unemployment. In Ahn Y, Wu F (Eds.) E3S Web of Conferences 214:01023, Nanjing, China, 18–20 Dec 2020. <https://doi.org/10.1051/e3sconf/202021401023>
- Zou J, Schiebinger L (2018) AI can be sexist and racist—it's time to make it fair. *Nature* 559(7714):324–326. <https://doi.org/10.1038/d41586-018-05707-8>
- Zou M, Niu J, Chen J et al. (2016) Facial age estimation with images in the wild. In: Tian Q, Sebe N, Qi GJ (eds) In: Proceedings of the 22nd International Conference, Multimedia Modelling 2016. Springer International Publishing,

Miami, Florida, USA, 4–6 Jan 2016. https://doi.org/10.1007/978-3-319-27671-7_38

Competing interests

The authors declare no competing interests.

Ethical approval

REB approval was not required because human data was not collected for this review.

Informed consent

This article does not contain any studies with human participants performed by any of the authors.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1057/s41599-023-01999-y>.

Correspondence and requests for materials should be addressed to Charlene H. Chu.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

¹Lawrence S. Bloomberg Faculty of Nursing, University of Toronto, Toronto, ON, Canada. ²Institute for Life Course and Aging, University of Toronto, Toronto, ON, Canada. ³KITE Research Institute, Toronto Rehabilitation Institute—University Health Network, Toronto, ON, Canada. ⁴Rehabilitation Sciences Institute, University of Toronto, Toronto, ON, Canada. ⁵Institute of Biomedical Engineering, University of Toronto, Toronto, ON, Canada. ⁶Leverhulme Centre for the Future of Intelligence, University of Cambridge, Cambridge, UK. ⁷Faculty of Health Disciplines, Athabasca University, Athabasca, AB, Canada. ⁸School of Business, Department of International Business, Marketing, Strategy, and Law, MacEwan University, Edmonton, AB, Canada. ⁹Intelligent Transportation Systems Centre, University of Toronto, Toronto, ON, Canada. ¹⁰Centre for Clinical Ethics, Unity Health Toronto, Toronto, ON, Canada. ¹¹Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada. ¹²Department of Family Medicine, and Faculty of Dental Medicine and Oral Health Sciences, McGill University, Montreal, QC, Canada. ¹³Mila—Quebec AI Institute, Montreal, Quebec, Canada. ¹⁴Lady Davis Institute for Medical Research, Herzl Family Practice Centre, Jewish General Hospital, Montreal, QC, Canada. ¹⁵Factor-Inwentash Faculty of Social Work, University of Toronto, Toronto, ON, Canada. ¹⁶Baycrest Hospital, Toronto, ON, Canada.

✉email: charlene.chu@utoronto.ca