

Engaging the Margins: Children and Aging Populations

Michael Pham

13 April 2023

Understanding the Representation and Representativeness of Age in AI Data Sets

Authors: Joon Sung Park, Michael S. Bernstein, Robin N. Brewer, Ece Kamar, Meredith Ringel Morris
Publication Date: 30 July 2021

Abstract

Ensuring diverse representation in AI training data sets is crucial for developing inclusive AI systems that work effectively for a wide range of users. While efforts have been made to create AI data sets well-balanced across race, gender, socioeconomic status, and disability status, age representation has been relatively neglected. This paper examines the representation of older adults (65+ years) and oldest-old adults (85+ years) in publicly available facial analysis data sets. They found that older adults are significantly under-represented, with only a few data sets explicitly including these two age brackets. As the global population is aging rapidly, these findings call for greater attention to age representation in data sets used by AI and the development of procedures for documenting age metadata to ensure inclusivity for older users.



Literature Background

- Studies have shown that race, gender, socioeconomic status, and disability status can negatively impact the performance of AI systems such as facial recognition or natural language processing systems.
- For instance, the Gender Shades study demonstrated how commercial AI systems for binary gender classification based on appearance often fail for women of darker skin color. Subsequent efforts to update the models with a more balanced training data set significantly reduced the error rate.
- Many older adults will experience physical and cognitive impairments, and AI-infused assistive technologies and tools present opportunities to support the needs of older adults, helping them age on their own terms. Examples include smart navigation, robotics, smart homes, and medical recommender systems. AI can also augment existing technologies and online communities that older adults use, such as facial recognition systems for unlocking phones and natural language interfaces, making technologies easier and more natural to use for those with low computer literacy.

Population Targeted

- The population targeted in this study consists of older adults in AI datasets, specifically those aged 65 and older (older adults) and those aged 85 and older (oldest-old adults). The study examines 92 face datasets to understand how age is recorded and the representation of older generations. Unique traits or characteristics of this population include:
 1. Growing in size: The older adult population is increasing, with more than 15% of the U.S. population being 65 years old or older. By 2050, this proportion will be matched globally. The oldest-old population is expected to see the greatest rate of increase.
 2. Underrepresented in AI datasets: The study found that older adults are underrepresented in AI datasets. Only five datasets in the study explicitly documented the closed age intervals of their subjects, including older adults, while only one included oldest-old adults.
 3. Inconsistent age documentation: The study found that only 24 of the datasets include any age-related information in their documentation or metadata. There is no consistent method followed across these datasets to collect and record the subjects' ages.

Research Questions

1. Is the age-related information of the subjects included in the metadata of the data set or its documentation? If so, how was the age binned?
2. What was the process for annotating the subjects' age?
3. What was the goal of creating the data set, and how did this interact with whether age was included in the metadata?
4. Is the older adult population (aged 65+) represented proportionally to the population at large? Does this representation, or lack thereof, extend to the oldest-old adults population (aged 85+)?

Major Hypothesis and Underlying Arguments

- The major hypotheses are:
 1. Older adults are under-represented in AI data sets.
 2. The methods used to collect and record subjects' ages in AI data sets are inconsistent.
 3. A lack of representation of older adults in AI data sets might lead to AI-infused technologies not working well for this fast-growing population.
- The underlying arguments in this study include:
 1. A diverse representation of different demographic groups in AI training data sets is essential for ensuring that the models will work for a large range of users.
 2. AI fairness and inclusion efforts advocate for creating AI data sets that are well-balanced across race, gender, socioeconomic status, and disability status.
 3. Age should be considered an important dimension in inclusive AI.
 4. Inclusive AI systems that work for a diverse group of users require the representation of diverse populations in the data used to train and test ML models.

Target Technology Being Evaluated

- The target technology being evaluated in this study is the representation of age in AI training datasets, specifically focusing on facial analysis systems.

Key Aspects of the Research Design

1. **Research Design:** The study is designed as a quantitative analysis to investigate the representation of older adults in AI datasets, specifically in facial analysis systems.
2. **Sample Size:** The study examines 92 face image datasets based on 277 academic publications. Out of these, 31 datasets were publicly downloadable and had clear terms of service, and they were downloaded and inspected for age information in their metadata.
3. **Data Collection:** The datasets were collected from a list of face image datasets previously compiled by Scheuerman et al. The list was generated by gathering research papers from ACM and IEEE, filtering them using specific keywords and publication periods. The researchers then analyzed the documentation and metadata of these datasets to extract information on age representation, annotation schemes, and the goals behind creating the datasets.

Key Aspects of the Research Methods

- The researchers analyzed publicly-available information about 92 face data sets to understand how they codify age and whether older generations are represented. The study was guided by the four main research questions.
- The researchers used a list of 92 face image data sets based on 277 academic publications compiled by Scheuerman et al. to study the representation of age in these data sets. They analyzed the documentation of all 92 image data sets and also inspected the metadata of 31 data sets that were still publicly available and complied with their institution's IRB data set onboarding process.
- The analysis focused on whether age-related information was included in the data sets, the representation of older adults by bracket, the goal of the data set and how it interacted with the inclusion of age-related information, and the age annotation scheme. The researchers iteratively developed a codebook to codify their data based on these aspects.

Strengths and Potential Weaknesses of the Methods Used

- Strengths

1. Large sample size: The study examines 92 face datasets, providing a solid foundation for assessing age representation across a wide range of datasets.
2. In-depth analysis: The researchers focus on multiple aspects, such as metadata, age brackets, dataset goals, and age annotation schemes, which offers a comprehensive understanding of the representation of age in these datasets.
3. Methodical approach: The researchers use a systematic and organized method to compile and analyze the datasets, which helps maintain the study's rigor and consistency.
4. Utilizing existing resources: The study leverages a previously compiled list of face image datasets, making efficient use of existing research to further explore age representation.

Strengths and Potential Weaknesses of the Methods Used

- Potential Weaknesses:
 1. Limited to facial analysis systems: The study focuses solely on facial analysis systems, which may limit the generalizability of the findings to other AI domains.
 2. Incomplete dataset access: Out of the 92 datasets, only 31 were available for download and inspection, which could potentially affect the accuracy and completeness of the findings.
 3. Age annotation inconsistencies: The study reveals inconsistencies in age annotation schemes, making it challenging to compare and contrast datasets directly.
 4. Potential selection bias: The study relies on the dataset list compiled by Scheuerman et al., which may introduce selection bias if that list does not represent the full range of facial analysis datasets.

Major Findings

- A majority of the face image data sets (74%) did not include any age-related information.
- Of the data sets that included age information, 20% of the documentation and 33% of the data sets included or mentioned raw age or date/year of birth. The rest aggregated the age information without consistent standards.
- The methods for annotating subjects' age can be categorized into three: recording age provided by subjects themselves (58%), inferring age using other data sources (21%), and estimating age by observing appearance (13%).
- Of the 24 data sets with age-related information, 58% did not specify the reason for collecting the subjects' ages. However, for those that did, the most common reasons were connected to supporting age-related classification or analysis tasks (60%).
- The representation of older adults (65+ years) in face data sets is not proportional to the population at large, with under-representation being evident in the age ranges provided. Only one data set explicitly included any representation of the oldest-old adults (85+ years).
- Some data sets were particularly skewed towards the younger population in their 20s, which could be an artifact of convenience sampling, as 25% of the data set documentation that included age-related information mentioned drawing subjects from a university undergraduate population.

Discussion Points or Implications

1. Under-representation of older adults: The study found that less than half of the datasets with documented maximum age included at least one person older than 65, and only one dataset explicitly had at least one person older than 85. This under-representation raises concerns about the generalizability of AI tools for the aging population.
2. Challenges of age representation: The study notes that age distribution is fast-changing due to increasing life expectancies, and the way people age is also changing. Additionally, intersectionality should be considered when creating age-representative datasets, as older adults may be under-represented in certain intersectional demographics.
3. Need for a standardized approach: The study found inconsistencies in age categorization and documentation in the datasets. It suggests collecting raw ages if possible, or adhering to age categories used in large-scale censuses for standardization. The annotation procedure should be clearly documented, especially when age is inferred from appearance.

Personal Thoughts

- The results were believable
- The methods were valid
- Possible extension: How well does AI perform for older adults given that they are under-represented?

“If Alexa Knew the State I Was in, It Would Cry”: Older Adults’ Perspectives of Voice Assistants for Health

Author: Robin N. Brewer

Publication Date: 28 April 2022

Abstract

This paper investigates older adults' perceptions of AI-powered voice assistants in personal health management within long-term care settings. The researchers interviewed 10 older adults living in an assisted living community in the U.S. to explore their values and preferences concerning AI for health. The findings revealed that older adults value technologies that generate and share positive and relational health information. The study contrasts this preference with existing deficit-based health tracking technologies and discusses how researchers, developers, and designers can adopt better approaches to AI-driven health for older adults and historically marginalized populations. The paper aims to understand older adults' perceptions of AI-powered personal health management tools and propose design recommendations for technologies that support their visions of AI-driven health.



Literature Background

- In the past couple decades, personal health technologies have advanced significantly, with research showing that personal health data is often shared collaboratively between individuals and their medical providers. Voice technologies are seen as more accessible for older adults compared to screen or keyboard-based interactions. Studies have shown that older adults appreciate the accessibility and natural language interaction provided by smart speaker voice interfaces. However, concerns remain over conversation structure, privacy, and reliance on technology.
- Older adults often use technology to supplement their health needs and are increasingly exploring how voice assistants can be a part of their health ecosystem. Voice assistants are seen as ideal for engaging with and sharing health information, providing personalized health data, tracking daily activities, and offering a holistic health summary. Conversational agents can reduce health literacy barriers and stigma of disclosing personal information. Trust, health, and voice technologies intersect, and older adults may perceive voice interfaces as more trustworthy than younger adults. However, challenges include understanding, verifying, and trusting health information from a voice assistant, comprehension conflicts, and designing for increased autonomy.

Population Targeted

- The population targeted in this study comprises older adults aged 65 and above living in an assisted living community in the United States. These individuals experience age-related disabilities and are more likely to develop moderate to severe health conditions. The study focuses on older adults in long-term care settings, such as assisted living, memory care, and skilled nursing homes, where they rely on a collaborative network of family members, care staff, and medical professionals to address their health needs. The research aims to understand the perceptions of these older adults regarding AI-powered personal health management tools, specifically voice assistants, and their use in long-term care communities.

Research Questions, Major Hypotheses, and Underlying Arguments

- What are the perceptions of AI-powered personal health management tools for data and sharing by older adults in an assisted living community?
- What are the design recommendations for technologies that support older adults' visions of AI-driven health?
- Older adults living in long-term care communities have unique health and data preferences that must be considered when designing AI-powered technologies for personal health management.
- Voice assistants, as an AI-powered technology, have the potential to support health-related data and personal health management for older adults living in assisted living communities.
- Older adults value positive and relational health information and may refuse negative health representations, which contrast with existing deficit-based health tracking technologies.
- Researchers, developers, and designers should engage in more positive and participatory approaches to AI-driven health for older adults and other historically marginalized populations.

Target Technology Being Evaluated

- The target technology being evaluated in this study is AI-powered voice assistants, such as Siri, Alexa, and Google Assistant, and their potential use for personal health management in long-term care settings.

Key Aspects of the Research Design

- Research design: Semi-structured interviews with a speculative inquiry exercise to explore older adults' perceptions of AI-powered voice assistants for health-related data use, reporting, and sharing.
- Sample size: 10 older adults aged between 73 and 86 years old, with an average age of 78 years old, comprising 5 males and 5 females.
- Data collection: Interviews were conducted with older adult residents living in an assisted living community that had partnered with a start-up to provide access to Amazon Alexa devices. The interviews focused on health perceptions of voice assistants and engaged older adults in scenarios to understand how these devices could be used for health and well-being.
- Location of study: The study was conducted in an assisted living community in a large city in the Northeastern region of the United States.

Key Aspects of the Research Methods

- In this study, the researchers used semi-structured interviews to understand older adults' digital health data monitoring and sharing perceptions, specifically focusing on voice assistants. The study was conducted with older adult residents living in an assisted living community in the Northeastern United States, which had partnered with a startup to provide its residents with Amazon Alexa devices. The interviews explored prior experience with voice assistants and engaged older adults in scenarios to understand how these devices could be used for health and well-being. A speculative inquiry exercise was included to understand their perceptions of health-related data use, reporting, and sharing.
- The researchers opted for this method because it is ideal for "eliciting tech futures" in contexts where desired realities have not yet been designed, where alternate realities are desired, or to critique existing realities. They recruited 10 older adults (73-86 years old) with a range of voice assistant experiences, as perspectives between novice and expert users can differ. The interviews were transcribed and analyzed using an inductive thematic analysis approach to identify patterns across the transcripts. The final set of themes included positive reporting and comprehension, among others.

Strengths and Potential Weaknesses of the Methods Used

- Strengths of the study include its focus on a neglected topic (older adults' perceptions of AI-driven health in long-term care settings), the use of semi-structured interviews to gather data, and the use of thematic analysis to identify patterns in the data. The authors also provide a clear rationale for their study and discuss the implications of their findings for researchers, developers, and designers working on AI-driven health technologies.
- A potential weakness of the study is its small sample size (10 participants), which limits the generalizability of the findings. Additionally, the sample was recruited from an assisted living community in a single city in the Northeastern region of the United States, which may limit the diversity of perspectives represented. Finally, the authors note that their sample skewed towards older, older adults, which may limit the representativeness of their findings for younger older adults.

Major Findings

- Positive Health and Aging Representations
 - The study found that older adults in assisted living communities value AI technologies that generate and share positive and relational health information. They prefer technologies that report positive aspects of their health and well-being, such as their personality and social activity. The participants expressed concerns about how the data should be presented to be useful and how to appropriately share data if someone has multiple health conditions. The study also found that older adults have comprehension concerns with voice-based health AI technologies. The study highlights the preference of older adults for positive health representations, which is in contrast to the deficit-based health tracking technologies for aging.
- Comprehending health information
 - They expressed concern over understanding health data and information from voice-based health AI technologies and were skeptical of learning new information about their health from a voice assistant. They also raised questions about the modality of AI-generated health reports and how to present them in ways that are comprehensible and trustworthy.

Discussion Points or Implications

- The study found that older adults living in assisted living communities want positive and relational health information generated and shared through AI-powered technologies like voice assistants. This preference for positive health representations is contrasted with existing deficit-based health tracking technologies for aging, and the study highlights the importance of understanding critical refusal of negative health representations. The study suggests recommendations for designers and developers of voice assistants and other AI-powered health tools, including the need for subjective and relational data about social well-being and intelligible formats of health-related data provided by voice assistants. Future research may include survey data of older adults across different living communities and co-design sessions with older adults to design health summaries generated from voice assistant data.

Personal Thoughts

- Results seem to match what I expected given my time around some old people.
- The methods were valid enough for this kind of study but should make sure the sample population is bigger and more diverse. Old people might be different in different cultures.