

# Understanding the Representation and Representativeness of Age in AI Data Sets



Joon Sung Park

Michael S. Bernstein

Robin N. Brewer

Ece Kamar

Meredith Ringel Morris

AIES '21

65+ years old



**15%**

of the US population today,  
of the world population by 2050.

**RQ:** How, and how well is older adult population (65+) represented in popular AI data sets?



**92**

We studied:  
face data sets from from  
Scheuerman et al. 2020.

Of the 92 data sets...

**26%**

mentioned collecting  
some form of age  
metadata

**53%** collected  
age-related data  
in-person

**20%** inferred age from  
other metadata

**13%** estimated from  
subjects' appearance

**58%**

did not specify the  
reason for including  
age metadata

**Only five**

contain at least one  
person older than 65

Ending notes:

How can we better represent  
older adults in AI?

What data sets should we  
consider beyond face data sets?

What are the normative  
questions around people's age  
and AI?