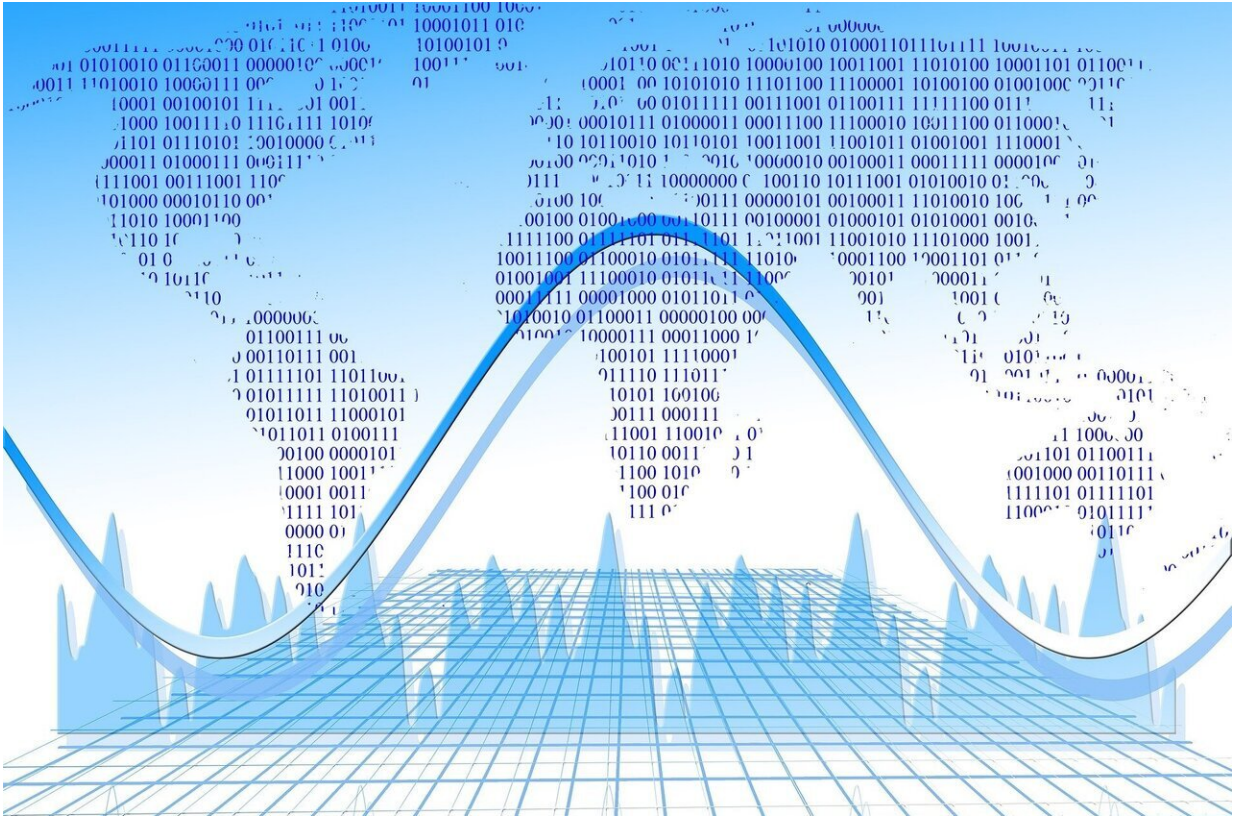


How big data is reshaping aging research and education

March 17 2020, by Katharine Gammon



Credit: CC0 Public Domain

The age of big data is here: The world has created more data in the past two years than in the entire previous history of the human race. USC Leonard Davis School of Gerontology researchers are dissecting treasure

troves of information—from sources as diverse as brain scans and the human genome—to fuel groundbreaking research on improving how we age, and to reshape gerontology education to enable future scientists to make an impact in a changing field.

Gerontology is ready to take on the newest tools—in part because the field has always involved [big data](#) sets, says Mireille Jacobson, a microeconomist and associate professor of gerontology at the USC Leonard Davis School. For example, her work has relied on large population data sets—and in a way, that hasn't changed, she says. "It's mostly that more and more data is available."

Jacobson works with data from Medicare and other publicly available databases to understand how health insurance affects the well-being of older people. For example, an analysis of Medicare data found that receiving Medicare benefits can help reduce financial stress in people over 65. She also researches [health care providers](#) and how they make care decisions in response to various outside factors, including new screening recommendations and drug shortages.

"The effort to digitize and make everything available electronically is a new thing," she adds.

Jacobson is part of a group of gerontology researchers at the USC Leonard Davis School who are diving into vast sets of data in order to better understand aging and the lifespan. Their work has important implications for training students and for creating better datasets, which can help researchers better understand individual risk factors, identify the role of genes in disease and develop more precise interventions.

Moving Across Disciplines

Em Arpawong, research assistant professor of gerontology and director

of the Gerontology Bioinformatics Core, looks to bring together diverse information to better understand how genetic and environmental components interact to result in different health outcomes in older adults. Her current work integrates the use of both genomewide and twin and family modeling approaches from large datasets representing hundreds of thousands of individuals over many decades, such as the [U.S. Health and Retirement Study](#) and the [Project Talent Aging Study](#), both of which span decades of follow-up with tens of thousands of participants.

Arpawong says studying aging is unique in the field of health research because there is so much that happens early in life that impacts a person's trajectory later.

"I take a lifespan developmental approach to study effects of earlier life conditions on later-life health, including genetics, behaviors and contextual factors such as socioeconomic status and family adversity," she says, "and this requires putting together a lot of data pieces."

Working on projects like creating an index of frailty, developing a genomewide scan for depressive symptomology in older adults, calculating how genetic and environmental factors contribute to aging-related cognitive changes and assessing the stability of MRI markers for dementia takes a lot of skills in different areas—and collaboration. There's an extra layer of complexity when researchers have to translate findings from animal studies of genetic markers to humans.

"The focus of my work with the Bioinformatics Core is the translation, or collaborating with people on the translation, of their findings from model systems such as mice into human population data," says Arpawong.

It's a bit of a circular process: Often, the researchers use data from

humans to look at the impact of the findings from the animal model systems. Once they find some things in human data, they circle back and run those experiments in the animal models to see whether there are some causal mechanisms. This data-driven exploration opens up many new ways to understand aging, because it's not possible to do these types of translational and integrative gerontology studies solely in humans, given our long lifespans and vastly different living environments.

"[This translation] has become a bigger part of the work here in Gerontology that's been fascinating and is helping to accelerate the pace of research findings across disciplines that have traditionally functioned more independently," Arpawong says.

The work has become naturally collaborative, involving many different investigators with diverse backgrounds.

"There's a lot of crossover in different departments and multiple benefits of working with folks from the Dornsife College, Keck School of Medicine and the Information Sciences Institute, including from psychology to computational biology," explains Arpawong. "It really is a tangled web all over USC. It just points to the whole transdisciplinary nature of this work. You need to be talking with and working with a lot of people to make sure that you're moving in the right direction."

Arpawong recently used diverse datasets to find the connection between genetics and verbal memory. She found that a genetic marker of Alzheimer's disease wasn't alone—there is a second gene that plays a role specifically in effects on aging-related memory ability.

Big data has also changed the way that people collaborate, says Eileen Crimmins, USC University Professor and AARP Professor of Gerontology. No single researcher can know all the parts of a project, she explains.

"There are many more large multidisciplinary groups where everybody has one specialization and nobody knows it all," she says. "So there's a lot more of trusting people that they do know it and they can do it right."

All that data requires newer solutions to housing and transferring it, especially when working with different researchers around the globe.

"The scale is much bigger in what we have to deal with and [in] the frequency and the need for transferring these things, maintaining data security, and then having the tools available to do this," says Arpawong. "A lot of the data analysis that we need to do requires coding in different programming languages that some of the more common statistical software doesn't have the capacity for, and housing the data in ways that go beyond one's own hard drive."

Education also has to match the newest developments in big data science, requiring students to be coding proficient, data driven and able to ask new questions about the science of aging. Big data has changed the game for graduate students, says Crimmins, who directs the Multidisciplinary Research Training in Gerontology Program at the USC Leonard Davis School. The program helps predoctoral and postdoctoral students become familiar with the study of aging across multiple disciplines.

"There's a lot more multidisciplinary activity because the questions really have moved," Crimmins explains.

Today's students are constantly gaining new skill sets and knowledge, from the pathology of air pollution to genetics, in addition to their foundational studies in gerontology. A lot of what people are doing is learning on the job to gain the skills that aren't taught in the classroom, Arpawong adds.

For students entering the field, understanding statistics and having a strong grasp of numeracy is crucial. In her class, Jacobson presents some "weird data"—for example, why the average of a dataset might be way higher than the median.

"If you think about average income in the U.S. as opposed to median, that might tell us something about Americans overall," she says. "In some sense, the big data that's available should force us to go back to basics and see the foundations."

Coding competency is also at the core of the tools new researchers use in the field—and sometimes that requires coding in multiple languages.

Adjusting for the Future

The availability and types of data will only increase in the future, and researchers are thinking of how to adjust their studies to make room for new information. Crimmins is a co-investigator for the Health and Retirement Study, which has been going on for about 30 years. The study is conducted every two years, half in person and half on the phone. Since 2006, the researchers have collected biomarkers from people they visit in person—metrics such as weight and height, blood pressure and a blood sample.

Those metrics will allow the researchers to make bigger associations in the existing data—and they also dwarf the other data in the analysis. "Two million markers for each person, and we have 20,000 people in the study," Crimmins says. "It really requires high-capacity computing."

Using data from the Health and Retirement Study, Crimmins has made several surprising findings. One is that people are actually having more years of good brain health after the age of 65 than was the case in the past. Another study of the data found that education gives people an

edge in their later years, helping them to keep dementia at bay and their memories intact. Collecting biomarkers from study participants also enabled USC Leonard Davis alumna Morgan Levine '08, Ph.D. '15 and Crimmins to develop a promising method to measure biological, as opposed to chronological, age. Their work resulted in findings showing that Americans may be aging more slowly than they were two decades ago.

USC has emerged as a data-producing university. A lot of work goes into encouraging the use of data by making them available and in a usable format, Crimmins says. It makes up a large part of what she does: doing service for the larger field.

"It's kind of expected in our field for those of us who are data producers, which is a big thing here at USC because we do produce a lot of the international datasets," she says.

Big data is breaking down traditional boundaries between fields, says Arpawong. The strongest outcomes will likely come from a mishmash of different data types—for example, medical billing and imaging, or genomics and environmental data.

"You need to break it down into pieces. You need people who know how to manipulate the data to get what you want—and it's very nuanced for each piece," she says. "You can get results any which way you can code something, but is it correct? And a big issue for bioinformatics is making sure people are trained in these types of data resources to make sure that they're doing what they aim to do."

USC Leonard Davis School Dean Pinchas Cohen agrees that in today's modern research environment, leveraging data from a variety of sources is as important as understanding cellular functions. In his own lab, he is leading big data–driven studies to identify previously unknown

mitochondrial genes, working to understand their functions and whether they can be targets for treatments for Alzheimer's, diabetes and other diseases.

"Instead of a one-size-fits-all mindset, the age of big data allows us to have a 21st-century approach to address disease risk and promote healthy aging with a deep understanding of an individual's risk factors," Cohen says. "Science is no longer about looking into a microscope in one's own lab; it's about looking outward to data from millions of people across the world."

Provided by University of Southern California

Citation: How big data is reshaping aging research and education (2020, March 17) retrieved 27 April 2024 from <https://techxplore.com/news/2020-03-big-reshaping-aging.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.