



UNIVERSITY LIBRARIES
VIRGINIA TECH.

Ensuring Scholarly Access to Government Archives and Records

A Collaboration of Virginia Tech and the
National Archives and Records Administration

William A. Ingram
Assistant Dean and Dir. of IT
University Libraries
Virginia Tech

Sylvester A. Johnson
Executive Director, Tech for
Humanity Initiative
Virginia Tech

NARA Town Hall—May 19, 2021



Ensuring Scholarly Access to Government Archives and Records

- Sponsored by the Andrew W. Mellon Foundation
- University Libraries in collaboration with Virginia Tech Center for Humanities and the National Archives and Records Administration (NARA)
- Collaborative planning grant for using AI and machine learning to increase the scale and effectiveness of archival programs
- 12-month project; extended due to COVID-19
- Online workshops in April/May 2021



Project Team



Bill Ingram
University Libraries



Sylvester Johnson
Center for Humanities



Pamela Wright
NARA

Students: Bipasha Banerjee, Shu Pan, Stephanie Ramirez,
Angelica Rivera, and Jenna Slezak

Special thanks: Kira Dietz, Joe Forte, Dominique Francesca,
Alex Kinnaman, Alan Munshower, and James Tuttle

Workshop Methods

Devising transdisciplinary collaboration



Workshop Process, Methods, and Approach

- Plenaries & breakout rooms
- Keynote talks
- **Shared ideation**
- Conversational feedback
- Intersections of expertise
- Shared notes Google Docs



Artificial Intelligence in the Archives

- Automatic categorization
- Summarization
- Extraction of bibliographic metadata
- Named entity identification and disambiguation
- Facial recognition



A small green plant with several leaves is growing out of a crack in a concrete surface. The background is a blurred, light-colored wall with some faint, brownish lines. The text is overlaid on the image in a white, sans-serif font.

Workshop Goals

Avoiding Algorithmic Bias

Who Is Making Sure the A.I. Machines Aren't Racist?

When Google forced out two well-known artificial intelligence experts, a long-simmering research controversy burst into the open.



“In the nearly 10 years I’ve written about artificial intelligence, two things have remained a constant: The technology relentlessly improves in fits and sudden, great leaps forward. And bias is a thread that subtly weaves through that work in a way that tech companies are reluctant to acknowledge.”

Metz, C. (2021, March 15). Who Is Making Sure the A.I. Machines Aren't Racist? *The New York Times*. <https://www.nytimes.com/2021/03/15/technology/artificial-intelligence-google-bias.html>

Avoiding algorithmic bias



How will an AI classify these images?
Which one depicts a peaceful protest?

Desired Outcomes

Devising a feasible and sustainable solution



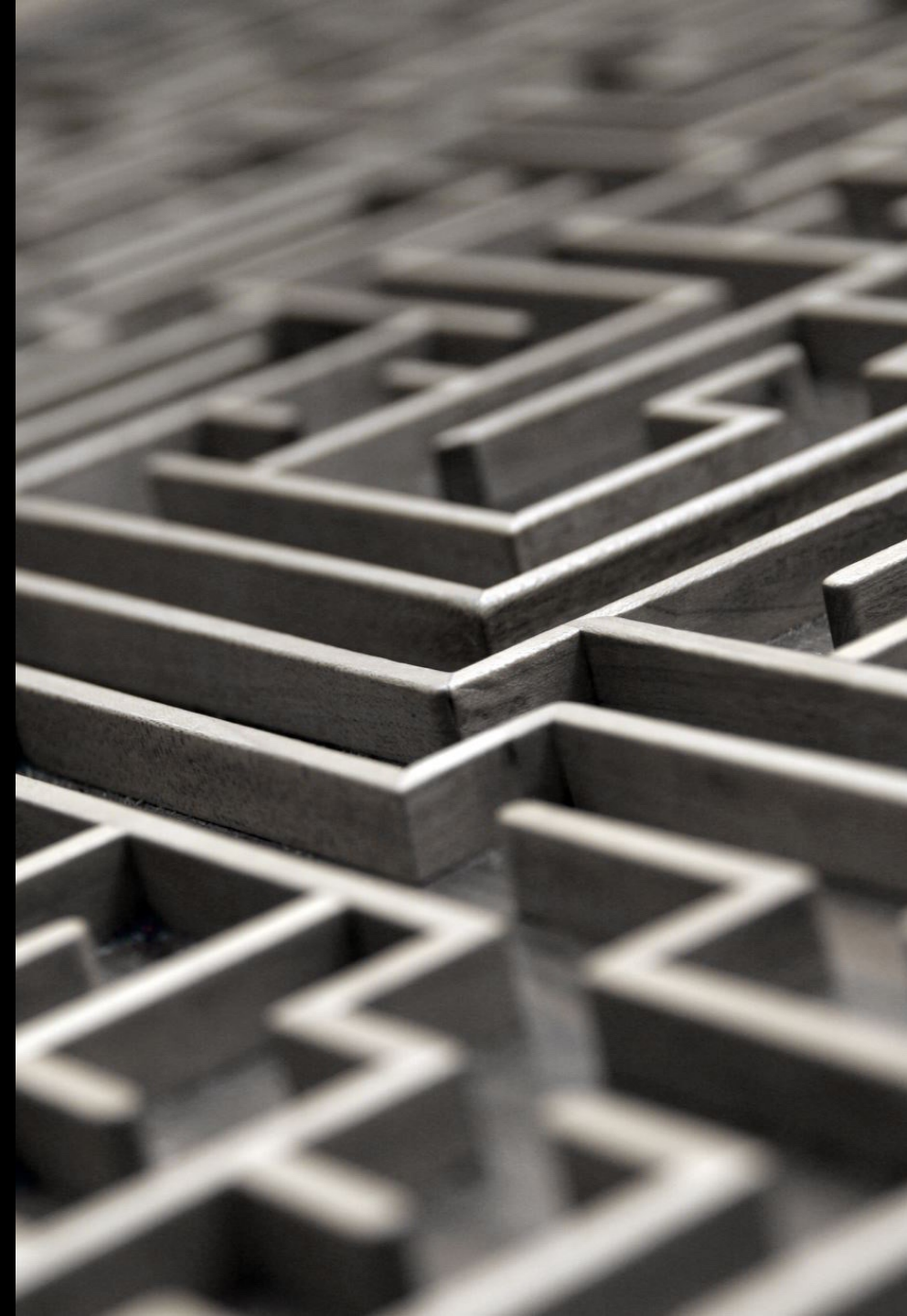


Feasibility and Sustainability

- Algorithmic design
- Human-Computer Interaction
- Financial/value-creation factors
- Capital infrastructure
- Data, analytics, & equity

Devising Feasible and Sustainable Solutions

- The overall point of this workshop series is to devise a workplan for using computational tools and technologies to increase the effectiveness of the NARA catalog.
- We want to stress the importance of devising a feasible and sustainable operating model that avoids replicating previous work and optimizes the likelihood of long-term viability.



Thank you!!

[William A. Ingram](#)

Assistant Dean and Dir. of IT
University Libraries
Virginia Tech

[Sylvester A. Johnson](#)

Executive Director, Tech for
Humanity Initiative
Virginia Tech

<https://lib.vt.edu/computational-archives-workshop>



Ensuring Scholarly Access to Government Archives and Records by [William A. Ingram](#) and [Sylvester A. Johnson](#) is licensed under [CC BY-SA 4.0](#)



Image Credits

- Meilan Solly. Facial Recognition Software Is Helping Identify Unknown Figures in Civil War Photographs. *Smithsonian Magazine*. Retrieved January 12, 2021 from <https://www.smithsonianmag.com/smart-news/facial-recognition-software-helping-identify-unknown-figures-civil-war-photographs-180970863/>
- Tyler Merbler from USA. 2021. DSC09254-2. Retrieved January 12, 2021 from [https://commons.wikimedia.org/wiki/File:2021_storming_of_the_United_States_Capitol_2021_storming_of_the_United_States_Capitol_DSC09254-2_\(50820534063\).jpg](https://commons.wikimedia.org/wiki/File:2021_storming_of_the_United_States_Capitol_2021_storming_of_the_United_States_Capitol_DSC09254-2_(50820534063).jpg)
- Rozina Sabur, Patrick Sawyer, and David Millward. 2020. Why are there protests over the death of George Floyd? *The Telegraph*. Retrieved January 12, 2021 from <https://www.telegraph.co.uk/news/0/us-america-riots-george-floyd-death-protests/>
- Ansgar Koreng. 2018. *Deutsch: Eine alte Pentacon Six TL Mittelformatkamera mit dem Carl Zeiss Jena Biometar 2.8/80 Standardobjektiv.English: Vintage Pentacon Six TL medium format camera with the standard prime lens Carl Zeiss Jena Biometar f/2.8 80mm mounted on.* Retrieved January 12, 2021 from https://commons.wikimedia.org/wiki/File:Pentacon_Six_TL_1803101902_ako.jpg

