

Building and Employing New Digital Resources for the Study of US Scientific Advisors

Devon Von Arx, Jordan Traylor, and Kenneth M. Evans

Rice University's Baker Institute for Public Policy
Houston, Texas USA

Context for Project

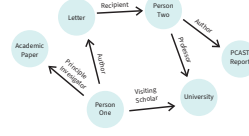
The White House Scientist and Science Policy Dynamic Digital Archive (DDA) is an online heritage collection hosted by Rice University's Woodson Research Center Special Collections and Archives. The DDA collects, preserves, and publishes archival materials related to the history of the White House Office of Science and Technology Policy (OSTP) and the lives of the scientists involved in it. OSTP advises the president on policy issues related to science and innovation, and coordinates the US effort in scientific research and development.

Since 1990, OSTP has managed two policy councils: the President's Council of Advisors on Science and Technology (PCAST) and the National Science and Technology Council (NSTC). The DDA serves as a clearinghouse for all records published by these councils, as well as the personal collections of senior OSTP staff and PCAST members, including personal communications, draft policy documents, and internal White House memoranda.



Policy makers in the United States executive branch increasingly rely on the OSTP to decide on a wide range of public policy challenges that depend on scientific data and analysis, from improving public health and strengthening the national economy to nuclear nonproliferation and advancing global diplomacy. The project's goal is to create a relational database that allows for the analysis of the role and impact of individual scientific advisors and advisory bodies involved in the United States national science, technology, and innovation policymaking by tracing decision making in policy development and performing large-scale text analysis of themes, content, and rhetoric contained in the inclusive records.

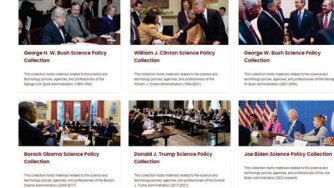
Goal of Database



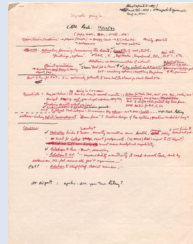
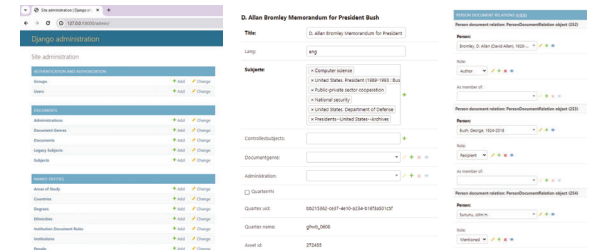
Non-Relational Database



Presidential Science Policy Collections



Relational Database

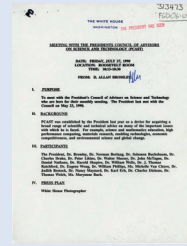


The DDA contains a wealth of previously unpublished and undigitized records from various libraries and archives, including the Library of Congress, the George H. W. Bush Presidential Library, and the Woodson Research Center. This document contains handwritten notes of John H. Garbutt, President Bill Clinton's first science advisor and co-chair of PCAST. It catalogs their first meeting on December 11, 1992, at President-elect Clinton's residence in Little Rock, Arkansas.

Methods



The database was built in Django, a high-level Python web framework, which allows for the seamless creation of a relational digital archive. It is easy to use, and easily expandable. Through Django, the DDA offers a user-friendly interface for policy makers and researchers to navigate and derive insights using its extensive archival records, photographs, and oral history interviews.



The DDA catalogs all PCAST materials online, such as meeting agendas, planning documents, and policy reports, dating from the council's creation in 1990 through the present day. This document contains the agenda for a PCAST meeting with President George H.W. Bush, issued by his science advisor and PCAST chairman, physicist D. Allan Bromley, held on July 27, 1990, in the White House.

Code

The database is organized around three core models: documents, people, and institutions. Each model has distinct fields; for example, the document model includes fields like title, issue ID, and language. Certain fields, such as language within the document model, can themselves be models, relating different tables to each other. However, these primary models feature relational models that establish connections with other primary models. For instance, the PersonDocument model uses foreign keys to define how the relationship between person and document works. The model keeps track of one such as authority, recipient, or mention within a document. This design enhances researchers' ability to analyze the DDA more effectively.

Document Model



People Model



Relation Model



Future Directions

Integrate OCR, LLM, and vision models to automatically extract the name, relations, and other relevant information from raw data (PDF, images, etc.) in order to automatically prefill those fields for the user to improve usability. If this system proves sufficiently accurate, we may allow it to automatically populate the whole database with minimal supervision, which would make the system highly efficient and scalable.

Develop better ways to search and navigate the database (fuzzy searching, page seek, etc.)

Develop a generic tool for non-technical archivists to quickly and automatically archive all their data.

Acknowledgements

This work was funded by Rice University's Fondren Fellows Program. We would also like to acknowledge the support of the Science and Technology Policy Program at Rice University's Baker Institute for Public Policy, the Woodson Research Center Special Collections and Archives at Fondren Library, Rice University's Center for Research Computing, and the Science History Institute. We also want to thank John Mulligan, Amanda Focke, Kirin Matthews, and Neal Lane.



The DDA also contains a rich collection of photographs that offer a window to understanding PCAST's role inside the White House. In this photograph, Vice President Dan Quayle issues in President George H.W. Bush's PCAST members at a press event in the Eisenhower Executive Office Building. Notable members include Walter Masaryk (farther right) and Bernadine Healy (farther third), who both left PCAST in 1993 to serve as Directors of the National Science Foundation and the National Institutes of Health, respectively.



The DDA serves as a permanent repository for the personal papers of PCAST members. In this photo, for example, digitized materials from the collection of Neal Lane, science advisor to President Clinton from 1995-2001 and Director of the National Science Foundation from 1993-1998. Neal Lane (farther right) is pictured speaking to President Clinton in the Oval Office. Also pictured are L. to left Smith, policy director for Lane, Thomas Kati, deputy director of the National Economic Council, and Jack Stewart, White House Press Secretary.

