

eGovernment Research and Trends in the US

Yigal Arens

Director, Digital Government Research Center

Director, Intelligent Systems Division

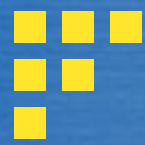
Information Sciences Institute

University of Southern California



Outline

- Overview of the DGRC
- US Federal eGovernment approach
- Example of research problem
- Digital government research in the US
- Unique example of university-state collaboration
- Initiative for international eGov research collaboration



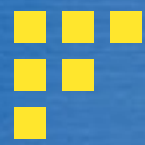
Overview of the DGRC

- Founded in 1999
- Collaboration of:
 - USC's Information Sciences Institute, and
 - Columbia University's Computer Science Dept
- Center funding awarded since it was founded:
 - Research: \$8 million
 - Community building and outreach: \$2.4 million
 - Annual conferences, electronic newsletter, websites
- Recent focus: Building an international e-gov/DG community and research collaborations



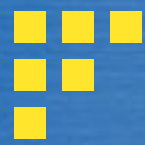
DGRC Support for the Digital Government Community

- Annual academic conference
 - 6th annual conference held in Atlanta
<http://dgrc.org/dgo2005>
 - Next year, San Diego, Calif.: May 21-24, 2006
- A community newsletter, *dgOnline*
 - To subscribe, go to <http://dgrc.org>
 - Or give me your card!
- A repository for e-gov/DG educational programs, curricula, syllabi
- A *Digital Government Society*
 - Search for “jochen scholl washington”
 - The link is at the bottom of his home page



DGRC Senior Researchers

- At USC
 - Eduard Hovy, Director of Research
 - Genevieve Giuliano
 - Jose-Luis Ambite
- At Columbia University
 - David Waltz



Principal Research Directions

- Data integration
 - EDC: 50,000 data tables aligned under large ontology/ metadata — EIA, Census (with Columbia University)
 - ARGOS: Dynamic composition of web services for transportation planning — CALTRANS (with USC's SPPD)
 - SiFT: Statistical data alignment method — California EPA
- Ontology creation through text harvesting & analysis
 - Build terminology 'lexicons' to model domains — EPA
 - Build city regulation support tool — EU's QUALEG
- eRulemaking
 - Sophisticated text analysis of public comments (clustering, opinions, stakeholders, etc.) — DOT (with CMU, Pitt)



The Federal Government and e-Government

The U.S. Government invests nearly \$60 billion annually in information technology



E-Government Act of 2002

President Bush's Memo

- Vision guided by three principles...
 - Government should be citizen-centered,
 - Results-oriented, and
 - Market-based.
- ... woven into the five Government-wide reform goals of the Administration's Management Agenda:
 - Strategic management of human capital,
 - Budget and performance integration,
 - Competitive sourcing,
 - Expanded use of the internet and computer resources to provide government services, and
 - Improved financial management.



E-Government Act of 2002

Highlights (1)

- Codifies Office of Management and Budget's role: E-Administrator and Office of E-Government
 - Requires annual report to to OMB
- Codifies existence of CIO Council
- Requires agencies to support cross-agency initiatives, e.g.,
 - E-Rulemaking, Geospatial One-Stop, E-Records Management, E-Authentication (esp. E-signatures) and Disaster Management
 - Enterprise architecture



E-Government Act of 2002

Highlights (2)

- OMB dialogue with state, local, and tribal governments, general public, private, and the non-profit sectors to find innovative ways to improve the performance of governments in collaborating on the use of information technology to improve the delivery of government information and services
- Standards for:
 - Categorizing and indexing government information
 - Agency Web sites
- Statistical Confidentiality and Data Sharing



Federal Government Focus

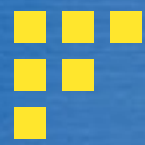
- Improve productivity
- Control IT costs
- Implement the responsibilities of E-Gov Act of 2002
- Improve cyber security
- Build an effective IT workforce



Second Annual Report 12/04

Highlights

- In 2003, 3.4 M taxpayers filed with IRS Free File program
- 350K business tax forms filed electronically
- Disaster management services used in 34 disasters, 226 exercises (Disasterhelp.gov)
- 65 M visits to USAJobs.com, 600K on-line resumes created
- Eliminated redundancies in grants management, human resources, financial management, and more



Second Annual Report 12/04

The Federal Government is delivering results through expansion and adoption of electronic government principles and best practices in managing information technology, providing timely and accurate information to the citizens and government decision makers while ensuring security and privacy.

*“Expanding E-Government”, Executive Office of the President,
December 2004*

- Nothing about eParticipation, eDemocracy, ...
- Nothing about research – only “improvement”, “innovation”



Example of DGRC Research

The eRulemaking Problem

- Government rule writers seek public input
- During rule-writing process, they contend with:
 - Tens of thousands of comments (long: like a 2,000-page book from General Motors; and short: like 75,000 emails due to an environmental group's campaign)
 - By law, they must read and respond to each comment when writing the regulations
- What can we do?
 - Use NLP technology for sophisticated text analysis:
 - Cluster comments
 - Find (near-) duplicates
 - Identify interest groups
 - Identify opinions and cluster
 - Relate comments to reg draft
 - Summarize comments



Research Project Development

- 2003: Formed initial group (2 CS, 2 Poli Sci)
- Initial study: small pilot SGER grant
 - Meetings in Washington with 40 rule writers and 40 representatives of public interest groups (e.g., unions)
 - Meetings with Dept of Transportation officials
- Early 2004: Assembled coalition of partners for a new research effort (DoT, universities, etc)
- Oct 2004: Large project funding commenced
 - 3-year plan: every year, deploy CS techniques in new workbench
 - Evaluation: Poli Sci effectiveness measures at agencies

USEPA-OAR-2002-0056

- **Description:** Proposed national emission standards for hazardous air pollutants; and, in the alternative, proposed standards of performance for new and existing stationary sources: electric utility steam generating units.
- **Source:** U.S. Environmental Protection Agency
- **Proposed rule:** [Summary](#), [Complete text](#), [EPA Regulatory Actions page](#)
- **The EPA's Mercury Rule Web page**
- **Data - The Official Docket:**
 - Format: Individual PDF files, mostly of scanned comments
 - [Table of contents](#) (Microsoft Excel)
 - [Individual files](#) [restricted access]
 - [One zip file](#) (956 MB) [restricted access]
- **Data - All Email Received:**
 - Format: [Individual files](#) of ASCII text, organized by date delivered, then date posted [restricted access]
 - Format: [One big file](#) of ASCII comments, separated by CTRL/L characters (143 MB, 1.8 GB uncompressed) [restricted access]
- **Data: Mass mail totals (supplied by the EPA)**
 - [Word, pdf](#)
- **Tools:**
 - [Search engine](#)
 - [Near-duplicate detection](#)
- **Modification history:**
 - Initial post, August 13, 2004.
 - Full docket, September 21, 2004.

(Jamie Callan, CMU)

Near-Dup Detection Example

Your recent proposal of weak air pollution standards to address mercury emissions from coal-fired power plants fails to protect our families and leaves communities at risk from mercury pollution that comes from coal-fired power plants. I urge you to strongly enforce the Clean Air Act, require power plants to use existing and proven effective mercury-reducing technology to keep more mercury out of our air and water, and prohibit pollution trading among plants that would leave some communities at risk for more pollution than others.

Toxic mercury contamination is a serious threat to public health and the environment. This is an important issue with long-term consequences. However, the Bush administration will only hold public hearings in three U.S. cities. I urge you to expand the number of public hearings you are holding so that the public has a sufficient opportunity to voice its concerns.

The power plants in New Jersey HAVE pollution controls in place that are stricter than the mercury emission pollution standards proposed by the Bush Administration. The mercury standards can and should be raised NOW, not discussed and studied in only a few cities! If you go forward and hold public hearings, the public hearing should be more numerous and should be held in cities on the East Coast, where we have serious problems with the mercury pollution from out-of-state power plants.

I understand that my comments may be placed on your public website.

Sincerely,

Your recent proposal of weak air pollution standards to address mercury emissions from coal-fired power plants fails to protect our families and leaves communities at risk from mercury pollution that comes from coal-fired power plants. I urge you to strongly enforce the Clean Air Act, require power plants to use existing and proven effective mercury-reducing technology to keep more mercury out of our air and water, and prohibit pollution trading among plants that would leave some communities at risk for more pollution than others.

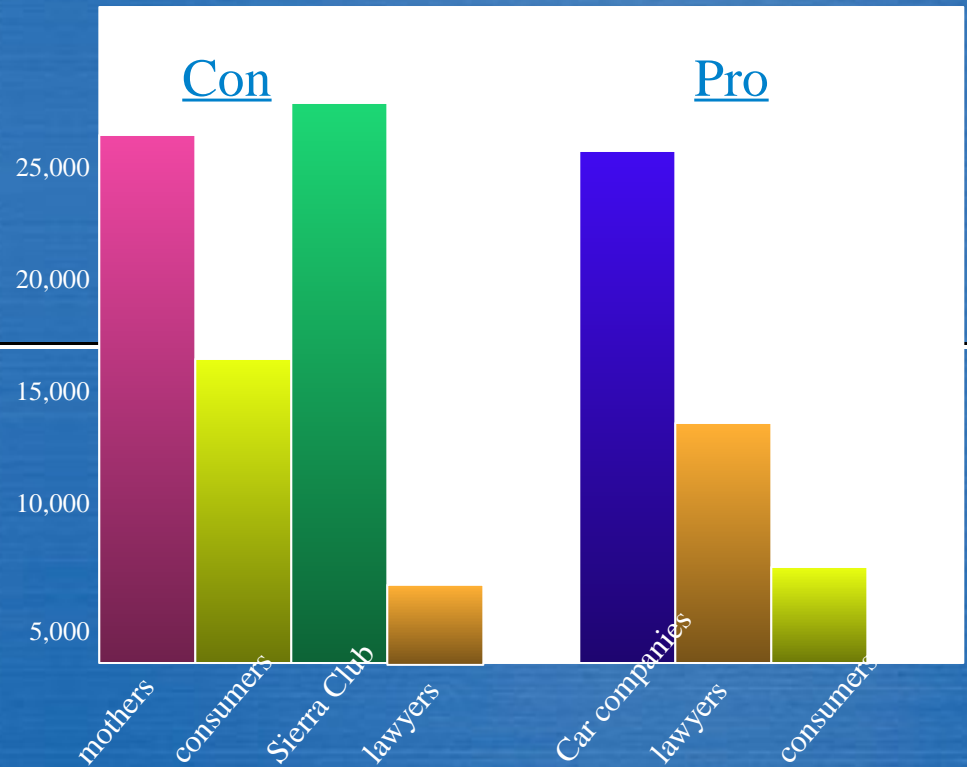
Please eliminate all mercury emissions now, immediately, and make every polluter pay for the cleanup with every penny they have. I have lost two wonderful people to brain tumors in the last two months and I've spent the past year chelating the mercury, lead, cadmium and arsenic from my own body because my body turned against me.

I developed internal cysts, which ruptured filling me with blood and adhesions. First week into a new job left me completely uninsured. I lost the job and had to declare bankruptcy. After two years of hospitals, doctors, pills, we found that my heavy metal levels were off the charts. I began chelating treatments. Surprisingly, the adhesions, and free bleeding stopped and I'm almost clean. However, I am unable to meet my goal.

It seems I have to stop ingesting the mercury to get and stay clean and I can't do it without you. My environment, your environment is a cesspool. Perhaps if you force companies who poison us to stop completely, we would have a chance. Do it. Make them stop at all costs, by any means necessary and make them pay for it because I'm broke. We must stop mercury emissions now.

The Regwriter Workbench

- Comments grouped by opinion and writer type
- Comments grouped by topic and cross-related



Par 2.2(a1)

- Con:
 - 150, 818: "impossible to maintain"
 - 272: "too expensive for elderly"
- Pro:
 - 169, 213, 391, 392, 394: "already being done in Alaska"
 - 18: "extend to children"

A diagram showing several snippets of comment text. Red boxes highlight specific phrases, and red arrows point from these boxes to the corresponding items in the list on the left. The snippets are as follows:

- Snippet 1: Xxx xx xxx xx x xxx, x xxx, Xx xxx x xxx x, xxxxxx x
- Snippet 2: Xxxx x xx xxxx x, X xx xxxx x xxxx x, xxx, Xxxx xxxx x xxxx, xxx xxx, Xxx xxx xxxxxxx x, xxx xx x, Xx xx xxxx x
- Snippet 3: Xxx xx xxx xx x xxx, xx, xxxx x xxx x, xxxxx x, xxx x xx xxx x, Xx xxxx x xxxx x, xx xxxx x xxxx, xx xxx, x xxx xxxxxxx x, x xx x, xx xxxx x
- Snippet 4: Xxx xx x xxx, x xxx x, x xxx x, xx xxxx x
- Snippet 5: Xxxx xxxx x xxxx, xxx xxx
- Snippet 6: Xxx xxx xxxxxxx x, xxx xx x, Xx xx xxxx x



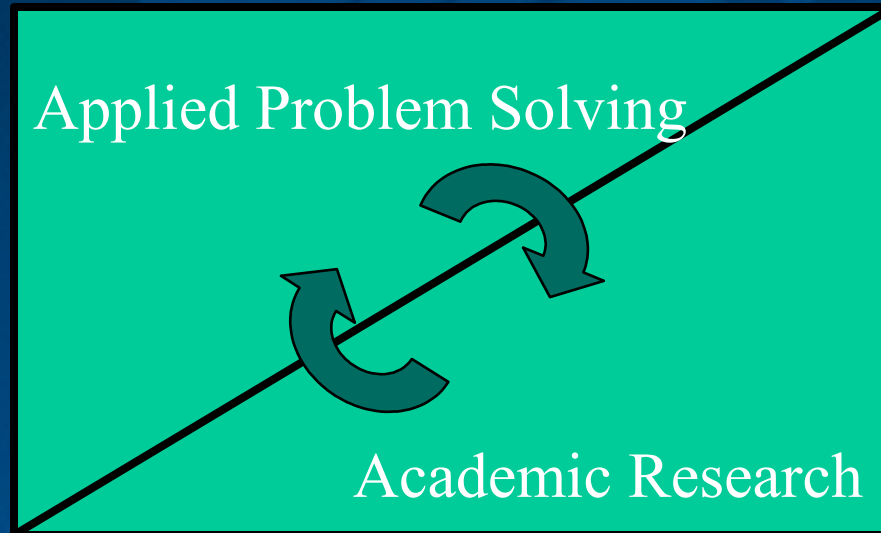
History of DG Research in US (NSF-Funded)

- 1997: Initial NSF workshop
 - “R&D Opportunities in Federal Information Services”
 - Organized by USC and Columbia University
- 1999: NSF “Digital Government Research” program initiated
 - Will run at least until 2007
 - Total of \$30 million awarded so far
- Affiliated programs and awards, e.g.,
 - Biodiversity (2002)
 - Digital Archiving (2004)
 - Information Technology Research (ITR) awards

A Model for University-State Collaboration

Center for Technology in Government
The University at Albany (SUNY)
Sharon Dawes, Director

Center for Technology in Government's Dual Focus on E-gov research



Applied Research Projects in New York State

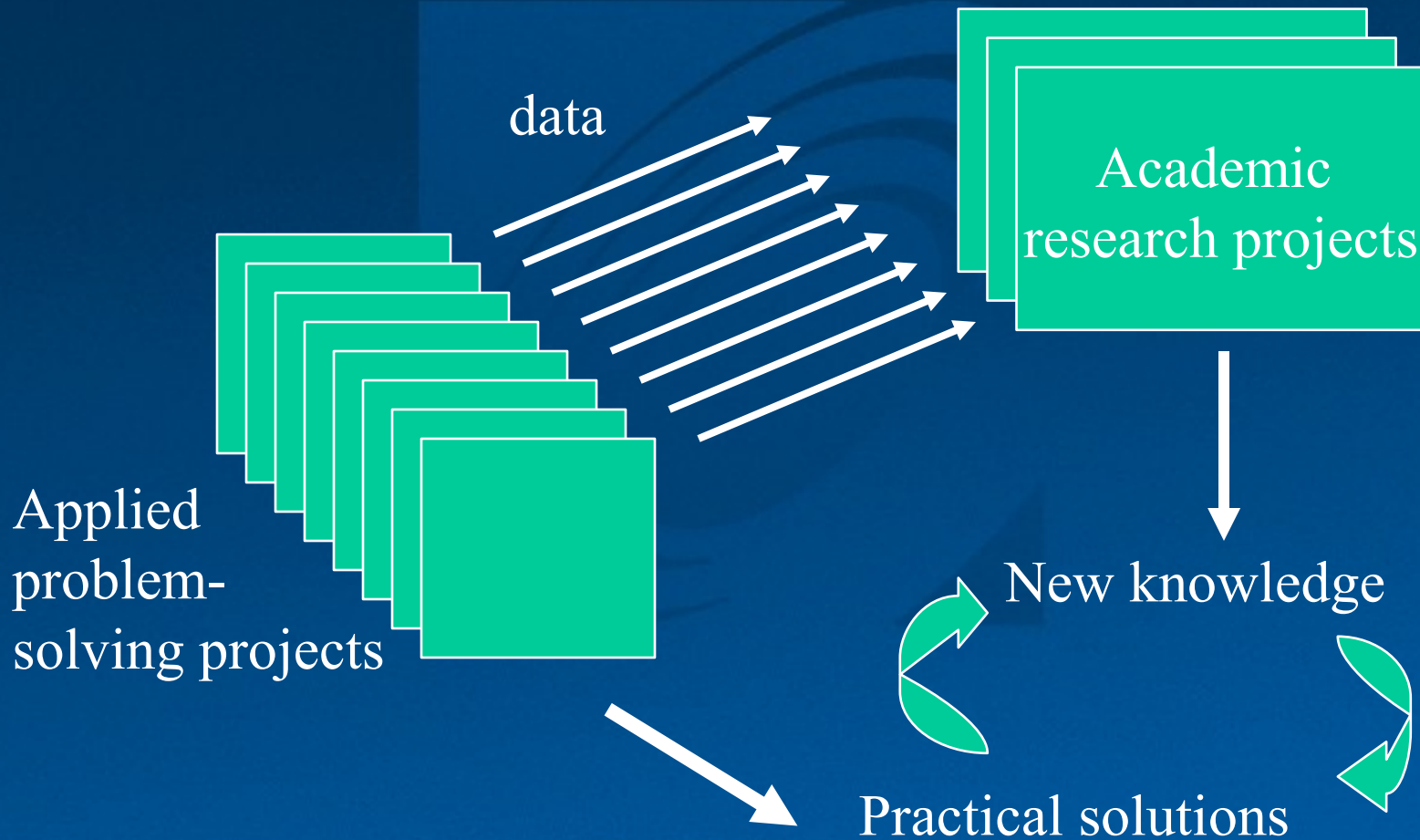
- **XML Testbed** (5 agencies experimenting with and evaluating XML as a management tool for web content and workflow)
- **IT Workforce Skills Assessment** (survey of all IT professionals in state government re: existing skills and training needs compared to a CIO survey of future workforce needs)
- **Use of parcel data** (study of current and potential uses of parcel data, value to various users, and benefits & barriers to data sharing across organizations, sectors, and levels of government).
- **State-local Gateway Prototype** (internet gateway and associated policies, practices, and technologies designed to explore a unified "back office" linking state and local governments)

Applied Research Tools

- Problem definition
- Stakeholder analysis
- Current practice reviews
- Surveys & interviews
- Process analysis
- Modeling
- Prototyping
- Evaluation



Action Research





A Short History of NSF Support for US-EU DG/e-Gov

Generally ad hoc over the Past 3 Years

- Two research projects
 1. Belize, Dominican Republic;
 2. Canada, Belgium
- DG research workshop: U.S.–Hungary '04
- DG researcher travel to international conferences with a DG/e-Gov focus
 - DEXA EGOV 2004/5, eChallenges 2003/4/5, IST 2004
- International DG/e-Gov panels and workshops at dg.o2003/4/5 (and next year at dg.o2006!)
- Two Information Technology Research supplements for working with EU E-Gov IST projects
 - With IntelCities, Qualeg



New Initiative for International eGov Research Collaboration

“Building and Sustaining an International Digital Government Research Community of Practice”

- Four years of support, \$1 million total
 - Small advisory board, including U.S. and non-U.S. leaders in position to support research in this field
 - Benchmark study to understand which countries and international organizations are investing in DG/E-Gov research, topics, and research results
 - 3 – 4 focused working groups meeting twice a year (U.S. researchers supported) on DG/e-gov topics proposed by community; prepare joint research papers
 - Summer Institute for doctoral candidates studying the DG/e-gov field

Thank You!