Exploring the Integration of Data-intensive Analytical Skills in Public Affairs Education

WORKSHOP REPORT
Workshop Report:
Exploring the integration of data-intensive analytical skills in public affairs education

Friday, May 9th, 2014
Center for Technology in Government
Albany, New York

Prepared by:
Sharon Dawes
Natalie Helbig
Sreekumar Nampoothiri

©2014 The Research Foundation of State University of New York
The Center grants permission to reprint this document provided this cover page is included.
Workshop Participants

David Andersen
University at Albany

Steven Arquitt
Millennium Institute

George Berg
University at Albany

Frank Boscoe
New York State Department of Health & University at Albany

Laurence Brooks
Brunel University

Sharon Dawes, Workshop Co-Chair
Center for Technology in Government

Anand Desai
The Ohio State University

Radboud Duintjer Tebbens
Kids Risk, Inc.

J. Ramon Gil-Garcia
Center for Technology in Government

Wil Gorr
Carnegie Mellon University

Teresa Harrison
University at Albany

Yvonne Harrison
University at Albany

Natalie Helbig, Workshop Co-Chair
Center for Technology in Government

Marijn Janssen
TU Delft

Erik Johnston
Arizona State University

Chris Koliba
University of Vermont

Rey Koslowski
University at Albany

Evert Lindquist
University of Victoria

Spiro Maroulis
Arizona State University

Karen Mossberger
Arizona State University

Sreekumar Nampoothiri, Workshop GA
Center for Technology in Government

Andrew Nicklin
Open NY, Executive Chamber New York State

Nicholas O’Brien
NYC Mayor’s Office of Data Analytics

Theresa Pardo
Center for Technology in Government

Karl Rethemeyer
University at Albany

Elliot Rich
University at Albany

Nur Sat
Hitit Uni-Turkey

Genie Stowers
San Francisco State University

Giri Kumar Tayi
University at Albany

Maria Wimmer
University of Koblenz-Landau
Introduction
Every year about 3,000 of New York City’s one million buildings erupt in a major fire.¹ The city Fire Department (FDNY) has jurisdiction over more than 300,000 retail spaces, restaurants, and high-occupancy towers, and proactively inspects a sample of about 25,000 each year as a prevention measure.² But what if FDNY could target the most vulnerable buildings for an inspection instead of just a random sample? In 2013 that question was put to the newly created Mayor’s Office of Data Analytics.³ The answer was found in using a state-of-the-art DataBridge warehouse that connects data systems in more than 40 different City departments that have information relevant to building safety. FDNY’s Risk Based Inspection System (RBIS) uses this data to consider about 60 different factors that make a building more likely to catch fire – such as age, electrical issues, location of sprinklers, and vacancy. The system uses an algorithm that assigns each of 330,000 buildings with a risk score that is then used to direct FDNY fire prevention resources toward the highest risk buildings in order to reduce the number and severity of fires.⁴

As this one example shows, public employees, at every level, are being touched by changes and advances in data and technology. Whether setting direction for analysts doing the number crunching, creating and testing algorithms, redesigning work processes based on analysis or explaining to the public what exactly the government is doing with data, government professionals need new knowledge and skills in data management, analysis, communication, and use.

---

³ Story was presented by Nicholas O’Brien at the May 9, 2014 workshop and further elaborated using published documentation.
⁴ Dwoskin (2014)
While much is being said and written about big data and data science, much less attention has been given to the skills required of the current and next generation of public managers, policy analysts, and informed citizens who are expected to use new data resources and tools effectively. To begin to address this gap, on May 9, 2014, the Center for Technology in Government at the University at Albany hosted a one-day workshop to explore the integration of data-intensive analytical skills in public affairs education. The event represented the convergence of two streams of activity in the United States and Europe on the topics of policy informatics and policy modeling developed over the past several years.

The workshop was motivated by several trends and interests: (1) recognition that complex societal challenges and related public policy problems impose ever-increasing demands on public management capabilities; (2) an expectation that these capabilities can be improved by careful use of expanding data and information, new technology developments, and advances in analytical approaches; and (3) a desire as educators to understand what the next generation of government professionals and researchers need to know about the uses and value of the emerging field of policy informatics and how university public affairs programs can prepare them.

We adopted policy informatics as the best way to describe the focus of our discussion, following the definition created by Arizona State University’s Center for Policy Informatics: “the trans-disciplinary study of how computation and communication technology leverages information to better understand and address complex public policy and administration problems and realize innovations in governance processes and institutions.”

Public administration and public policy curricula need to confront these trends and develop ways to train professional analysts and managers to understand and address them. This report summarizes the workshop activities and results, and offers a set of next steps for continuing to advance efforts to bring these important new skills into public affairs education.

---

**Workshop Funding & Support**

This workshop was an activity of eGovPoliNet related to building an international community of researchers and practitioners (Policy Community) in ICT solutions for governance and policy modeling. eGovPoliNet is a multi-institutional consortium sponsored by the European Commission FR7 Framework Program. The workshop was co-sponsored by the Center for Technology in Government (University at Albany) under a grant from the National Science Foundation (Grant # 054069).

---

For more information on the projects visit: http://www.ctg.albany.edu/projects/egovpolinet and http://www.policy-community.eu/

For more information visit: https://cpi.asu.edu/
About the Workshop

Thirty-two participants interested in the intersections of technology and information in public policy and management represented a wide range of academic and professional backgrounds including e-government, business and public administration, comparative politics, and operations research among many others. They worked in more than a dozen policy domains ranging from agriculture to criminal justice to public health to transportation and urban planning. Individuals at the workshop possessed many kinds of specific expertise they wanted to share including analytic methods, building data repositories, understanding cultural differences and complexity in research and practice, interacting with practitioners, and innovative teaching strategies. They also expressed interest in a variety of topics they hoped to learn more about related to policy informatics including data visualization, innovative teaching methods, international comparisons, balancing theory and practice, and leveraging technology.

Participants represented the following schools: Arizona State University, Brunel University, Carnegie Mellon University, Delft University of Technology, Ohio State University, San Francisco State University, University at Albany, University of Koblenz, Germany University of Vermont, and University of Victoria. Experts from government and non-profit organizations represented New York State, New York City, Kid Risk, Inc. and the Millennium Institute.

The workshop had four objectives:

- Understand the analytical needs of policy makers and program managers
- Share approaches to educating students in the types, uses, and limitations of policy informatics
- Explore new methods for policy informatics education
- Consider curriculum recommendations for public affairs schools

Two morning panels and a keynote talk introduced experiences and ideas that informed afternoon small group discussions. The first panel focused on Policy Challenges and set the stage for the workshop by illustrating the kinds of issues our students will face as
they enter careers in government or government-related research. The second panel focused on Preparing to Meet the Challenges, highlighting examples of innovative teaching in public affairs education to stimulate discussion about how faculty can employ creative methods and tools in the classroom such as modeling, data analytics, and other non-traditional ways of assessing public problems and considering solutions. The keynote talk, Visualization, Informatics, and Teaching Policy Analysis and Management, by Evert Lindquist provided an opportunity to delve into the history of visualization in research and practice across different disciplines, and to reflect on the needs of public sector leaders, the new ways they are consuming information, and how visualization can and cannot be a useful tool.

Data, Technologies & Context:
How the practice of government is changing

Government continues to be the residual claimant of intractable public problems like air quality, public health or sustainable development. However, the context of public policy and management is changing with the advent of more and different kinds of data and the technologies and tools that can be brought to bear on them. For example, in Chicago Is Your Big (Friendly) Brother, Susan Crawford, former Special Assistant for Science, Technology, and Innovation Policy in the Obama Administration describes the City of Chicago’s Array of Things project. Building on advances in nanotechnologies and sensors to gather data about the city, the Array of Things project will place up to 400 sensors within the city in an experiment to detect light, sound, air quality, and other ‘things’ that affect quality of life. This data will then be made available, in open formats, for a variety of users free of charge. The hope is that having data about the things we cannot always see, or that government does not currently collect, will spur entrepreneurs to develop innovative mobile applications that can help citizens walking down a street learn there is ice ahead or that one neighborhood has higher levels of pollen than another so asthma sufferers can avoid situations that could have adverse and costly health effects. Crawford concludes that, “Investing time and money in data makes sense, and it is changing how local government works.”

In Philadelphia, Mayor Michael Nutter just announced the creation of an Open Innovation Lab for City Employees. Director of the lab, and Chief Innovation Officer, Adel Ebeid, described how the lab will run in 90-day increments, prompting employees to consider a range of topics, from geo-spatial analysis, to public health, poverty and economic development. The goal is to get public employees working together and learning new skills to solve complex problems.

Non-profit agencies working on behalf of communities around the world are also using new tools and methods and engaging with government practitioners. Moving from a local to national and international scale, the Millennium Institute (MI), a global non-profit founded in 1983 focuses attention on global sustainability issues. MI makes accessible system dynamics modeling tools and other analytic techniques to help national leaders, especially in developing countries, use systems thinking and tools to analyze and understand the interconnectedness among economic, social, and environmental factors,

---

and issues of environmental sustainability, peace and security. The Millennium Institute has created Threshold 21 (T21), a dynamic simulation tool designed to support comprehensive, integrated, long-term national development planning. Using various models and computational tools, T21 supports comparative analysis of different policy options, and helps users to identify the sets of policies that tend to lead toward desired development goals. MI has created more than 15 unique, customized T21 models in countries such as Malawi, Mozambique, Bangladesh, United States, and Italy. In Jamaica, for example, MI worked with the government to create a custom T21 model focused on economic growth and industrialization. The tool helped leaders envision more holistic approaches to sustainable growth.

At the global level, polio eradication has been a major public health goal for decades. In 1988, when 350,000 children per year were being paralyzed by the polio virus, the World Health Assembly committed to eradicating polio by the year 2000. Today, the US Center for Disease Control (CDC) alone spends more than $100 million annually along with significant public employee resources to achieve polio eradication. CDC maintains high standards for developing evidence-based policies and cost-effective use of its resources. Consequently, in 2001, CDC collaborated with the non-profit Kid Risk Project to use a range of computational and modeling techniques to develop integrated analytical models to evaluate the global risks, benefits, and costs of policy choices for polio eradication. For more than a decade, leaders at the World Health Organization (WHO) and CDC have benefited from the intelligence and evidence generated by the Kid Risk Project. To date, the annual burden of the disease has been reduced by more than 99%, to less than 2000 cases of paralysis annually. But critical questions remain: “Is total polio eradication worth the continuing investment? What would happen if we stopped investing in eradication? How can we best prevent/control outbreaks in polio-free countries?” “What, if any, vaccine should we use after global polio eradication?” Dr. Bruce Aylward assistant director-general of Polio, Emergencies and Country Collaboration at WHO said, “This work has been fundamental to so much of what’s happened in the polio eradication program over the last few years, and it has helped to support many of our decisions over the last decade to bring the world much, much closer to one where future generations will never know the terror of this disease.”

These types of initiatives, from New York City’s municipal fire prevention to national strategies for sustainable growth and development, to global health challenges can be supported, evaluated, and improved by government data collection, and by public, private, and civic analysis and data use everywhere in the world. Unfortunately, popular excitement about the prospects for newly available data and tools often overshadows an appreciation for their limitations, gaps, and risks.

On the technological front, computational and simulation tools are becoming simultaneously more sophisticated, easier to use, and increasingly available. However, big questions remain about suitability, cost, and usability of the tools and the capability and skills to choose and use them effectively. Data visualization techniques, for example, can expand our ability to display and disseminate complex

---

10 Millennium Institute Website: http://www.millennium-institute.org/index.html
11 For more information on the model see: http://www.millennium-institute.org/integrated_planning/tools/T21/
14 Story was presented by Dr. Duintjer Tebbens at the May 9, 2014 workshop and further elaborated using published documentation.
temporal and spatial information and communicate evidence, but they can also contain biases that distort important aspects of data trends, impacts, evidence and meaning. Models rest on key assumptions that are not always made explicit or tested against current or historical evidence.

On the data front, innovations in information collection, analysis, synthesis, and dissemination are changing the type, amount, and quality of policy-relevant information. But, are these data fit for the many types of uses for which we need them? Data quality is a many-faceted concern that involves understanding and evaluating factors such as accuracy, granularity, timeliness, and comprehensiveness against a given use. For example, data on whole populations of interest (e.g., all Medicaid claims), or data collected from social media sites (e.g., mentions of illness or disease outbreaks on Twitter) or from monitoring devices such as scanners and sensors, are becoming more readily available but they confuse traditional distinctions between data samples and populations. Administrative data reflects the policy and organizational context for its creation, and although it can be quite valuable for different kinds of analysis, it may not be described or managed in ways that make it valid for additional uses. Perhaps most important, new information policies are needed for governing in a data-rich, interconnected world.

**Preparing Government Professionals for a New Context**

The examples above illustrate the new contexts government professionals at all levels will face in the coming years. They will need to understand and manage different types of data, create new organizations or programs focused on analytics, select tools to support computation or public engagement, and engage in data-intensive policy analysis and evaluation. Policy informatics, as defined above, can act as a useful bridge between research and practice in the areas of technology, data, analytics, and policy analysis and management. Academics need to study current and emerging practices in their research in order to shape curricula in public policy and management to meet the needs of the practice community. In turn, the practice community can take advantage of insights from research to help address pressing public problems and should be able to rely on our graduate programs to produce professionals who are well-versed in advanced design and analytical approaches to public sector governance and management.

**Challenges for Practice**

The presenters and workshop participants discussed the ways that current data, analytical, and technical capabilities are changing the nature of government as a ‘user of data.’ Several themes emerged from the presentations: (1) public problems vary widely in content and complexity and thus have different data and analytic needs, (2) stakeholders play, and will continue to play, various roles in the administration of program and policy making process, and (3) public affairs graduates, regardless of specialty or career goal, will need to work with a variety of technical and policy specialists.

**Finding and using relevant data**

Finding "good" data that will answer important practice and policy questions and does not cost a lot has been a perennial problem for government. Government is one of the largest collectors of data on a vast range of topics, but data are usually collected for specific purposes related to programmatic needs or

---

compliance with rules or statute. The data are dispersed across different departments and protected by various laws regarding collection, access, management, and use. Today new sources of data from outside of government (e.g., social media networks, sensor data, or text) can often be combined with government data. Practitioners need to know how to find data sources across departmental and jurisdictional boundaries, understand the limitations of the data they find in respect to its intended use, and assess whether the use of new data sources is feasible or to what extent combining different classes of data produces the kinds of analysis they need. And, they will need to answer these kinds of questions while operating within and balancing the legal and ethical parameters of appropriate use of government information.

**Applying tools and analytic techniques that fit the situation**

Complex policy problems can be big or small, broad or narrow, acute or chronic, unique or ubiquitous. Different analytic tools, computational techniques, and technologies will fit different situations. The data and tools available to address an immediate crisis tend to be those readily at hand, even though they are probably incomplete or otherwise flawed. By contrast, developing a major piece of legislation is likely to occur over a longer time period with the opportunity to search out or even collect appropriate data and apply a variety of analytical techniques to test different policy choices. Some techniques, such as visualization, pattern matching, or geospatial analysis may be applicable in many kinds of policy domains and settings. Others, such as the integrated models Millennium Institute uses to forecast the challenges of sustainable development are tuned to a certain kind of problem that demands complex understanding of different scenarios now and in the future. Practitioners will need a diverse ‘tool kit’ of tools and techniques and have a good understanding of the benefits and limitations of each.

**Communicating and engaging with a range of stakeholders**

Workshop speakers emphasized the importance of being ‘good communicators,’ particularly about translating messy, complex problems into more meaningful and manageable areas for discussion with stakeholders and leaders. Translating the data, analysis, and models in ways that ‘keep the right balance of detail’ for the decisions at-hand, but conveys the limitations, assumptions, and holes in the data is an increasingly important skill. The recent open government movement, combined with new governance practices, promotes engagement of a range of stakeholders as essential to the policy making process. Policy modeling tools are very useful for engaging experts and lay stakeholders in the design and implications of various policy options. This kind of engagement helps create understanding, creativity, and buy-in. However, explaining to average citizens the efficacy of different policies or the limitations of data and technology in making choices is different than explaining it to experts in the field or to legislators. Workshop participants emphasized that the ability to communicate clearly and meaningfully in these different situations is an essential skill for responsible policy informatics work. Public affairs graduates need to be able to identify and address questions about the ensemble of technologies, data, and policies so that they are better able to manage new programs, innovations, and experts who use these technologies.

**Working across specialties**

Using technologies, analytics, or modeling to address problems requires the ability to assemble and work in multi-disciplinary teams. Public managers, data analysts, subject matter experts, and policy makers need to work together in situations where their different kinds of knowledge and expertise can
be focused jointly on problems. Some actors will have more technical expertise in coding, mathematics, visualizations, modeling, and technology, others in policy relevant information, or organizational and implementation considerations. Policy informatics takes all these views into consideration and thus helps not only to identify the different kinds of expertise that are salient to a problem but to see how they can complement or conflict with one another in various policy scenarios.

**Challenges for Teaching**
Policy informatics provides an opportunity to teach students the importance of understanding data in the broader social, economic, and political context. Workshop discussions focused on four areas for consideration: (1) the importance of teaching for ‘the real world’, (2) providing students with a broad appreciation for data and information, (3) providing access to robust tools and technologies, and (4) finding ways to connect and balance policy informatics competencies with core curriculum requirements.

**Finding real cases, using real data, and dealing with realistic levels of complexity**
Effective use of data and computational tools for problem solving demands attention to situations, assumptions and dynamics that reveal the complexity of the problem and the suitability of different interventions. Speakers urged that analytical projects and assignments should use real world situations, not made up problems, and apply existing available data, not artificially constructed data sets. The open government data movement provides an opportunity to do this as large numbers of datasets are being cataloged and released for public use. Most of these data resources are rich in content, but they have also limitations and flaws that students must learn to address in their analyses. Cases can come from any level of government and any policy domain, as long as they are reasonably good representations of the interests and conflicts that are at play. These kinds of activities and resources prepare students for the likely issues they will face when they take positions in government and will give them a realistic dress rehearsal for the work they are training to do and the challenges it will inevitably hold. For faculty, the challenge is to teach the principles and tools of analysis without relying on simplified cases or sanitized data that give students a false comfort of a straightforward application that leads to a “best” answer.

**Imparting a broad appreciation for the role of data in public policy and management**
While some students will want to develop strong technical expertise, all should be able to discern what types of analysis and sources are appropriate in various contexts. Every public affairs graduate should be a discriminating consumer of data, a critical audience for data analysis, and a trusted steward of data resources. Graduates who go into government positions will inevitably be in some way responsible for the quality and management of data in their own agencies. Workshop participants noted that policy informatics has its own policy components, including ethics and legal frameworks, or as one workshop participant put it, “We need policies for data, not just data for policies.” Participants began a discussion of the importance of identifying the threshold knowledge and skills that all students should acquire in their degree programs, emphasizing the need to connect and integrate these newer demands with existing core competencies and traditional classes.
Acquiring robust tools and technologies
Students who want a specialization in policy informatics face another challenge in the cost and accessibility of robust tools. The challenge is three fold: first, tools (or licenses) that can handle large, realistic data sets with a good selection of features can be costly. Second, the tools and techniques for using them are constantly changing, thus requiring the ability to upgrade and branch out to different packages or features. Third, very few public affairs faculty have the knowledge or skills to teach about or with these tools, or room in the teaching schedule to devote whole courses to the topic. Open source programs for classroom use such as R or Quantum GIS (QGIS) can provide good teaching tools, but even these may be beyond the reach of individual students or departmental resources. Team teaching, cross-listing with statistics, math, or business courses, and cross-disciplinary courses are all possible ways to deal with this challenge. It is also helpful to remember that many government agencies also lack the resources for high-end tools, so affordable options in the classroom may well be the right choice for long term usefulness in practice.

Advocating for changes in curriculum
Demand for students with data-intensive skills is on the rise in many fields. Public affairs graduates are competing with students from other disciplines for these positions. Given the multi-faceted perspective that policy informatics imparts, graduates with policy informatics training can act as change agents or boundary spanners in their agencies. They are likely to be better communicators about and savvy consumers of data and evidence. Their higher levels of data and technology literacy, grounded in the public affairs context, can allow them to play leadership and facilitation roles that demonstrate the special value of hiring managers and analysts with a public affairs degree.

Advancing Policy Informatics in Public Affairs Education
The workshop emphasized the importance of positioning policy informatics within the public affairs curriculum in schools of various sizes, as well as the need to train all students to be savvy consumers of data and analysis and trusted stewards of government information.

To do so, public affairs programs will be challenged to answer some important questions: How do we get students to understand how complex systems operate so that they can think critically, analytically, and ask good questions? What technical skills are important for different career goals, and what is the right balance between preparing expert technicians and data-aware professionals?

Policy Informatics Competencies for Public Affairs Graduates
1. At least basic understanding of data as it intersects with technology, policy, management, and context
2. Critical appreciation for data, analytic tools, and models
3. Analytical agility or the ability to choose and apply different tools and data sources to different types of public needs

---

policy makers and managers? How do we train our students to be discriminating consumers of data? What and how do we teach students about the elements and tradeoffs in data quality? How do we think about mid-career versus new career core competencies? How do you incorporate policy informatics elements into programs of any size? How do you integrate its essential elements into the core public administration curriculum?

Next Steps
The workshop was a first step toward more discussions about how to incorporate policy informatics into public affairs education. This report is one vehicle for sharing the workshop discussions with a wider audience. However, in order for the discussion to expand, other more action-oriented steps were also identified:

LinkedIn Group (Policy 2.0). Invite all workshop participants and interested readers to join a LinkedIn group called Policy 2.0 which has been established by the eGovPoliNetProject and provides a way to participate in relevant discussions, announce events, and connect to a global community. To join visit: http://www.linkedin.com/groups/Policy-Making-20-4165795.

Policy Informatics Network listserv (PIN-L). All workshop participants were invited to subscribe to and invite others to join the PIN-L listerv, which was established at the University at Albany several years ago to share news about policy informatics developments, publications, and events. It currently has over 200 community members. To subscribe, send a blank email to: PIN-L-SUBSCRIBE-REQUEST@LISTSERV.ALBANY.EDU.

NAASPA conference panel. The findings from this workshop will be presented during a panel discussion at the 2014 National Association of Schools of Public Administration and Affairs (NASPAA) conference. The session entitled, #OpenData #BigData: Data, Big and Small in the Public Affairs Curriculum, will feature presentations by participants from this workshop and will also invite conference attendees to join the community of interested faculty.

Resource repository. As part of its National Science Foundation grant for policy informatics, the Center for Technology in Government at the University at Albany will explore ways to create a repository for interested faculty to share syllabi and other curricular material.

Academic publications. The workshop organizers plan to submit a manuscript to the Journal of Public Affairs Education based on the workshop results and offering an initial set of recommendations for incorporating policy informatics into mainstream public affairs education.

A Network of Networks. Finally, during the workshop several participants expressed interest in exploring the idea of an International Policy Informatics Network (iPIN). Envisioned as a network of networks connecting existing organizations and professional communities to promote activities and exchange among scholars working in policy informatics. As this idea develops further, information will be distributed on the PIN-L listserv and LinkedIn Group.