October 6, 2013

Community-Based Analytics: Big Data and Decision Making for Community-Based Organizations

Michael P Johnson, Jr.

Community-Based Analytics: 'Big Data' and Decision Modeling for Community-Based Organizations

Michael P. Johnson
University of Massachusetts Boston

INFORMS Fall National Conference, October 6, 2013

Special thanks:

- Merritt Hughes, Leibiana Feliz, Sandeep Jani: Research assistants, Department of Public Policy and Public Affairs, University of Massachusetts Boston
- Holly St. Clair, Director of Data Services, Metropolitan Area Planning Council
- Peter Shane, Moritz School of Law, The Ohio State University and editor, *I/T: The Journal of Law and Policy for the Technology Society*
- Mark Warren, Professor, Department of Public Policy and Public Affairs, University of Massachusetts Boston and director, URBAN.Boston

What is the motivation for this project?

- The 'Big Data' and analytics revolutions have enabled many organizations to generate novel insights from large datasets that provide new opportunities for products and services that further their missions
- Government agencies and large nonprofit organizations have developed new data stores and applications that have the potential to revolutionize public service
- However, community-based organizations often do not and cannot avail themselves of these resources. Why is this the case? Can innovations in data, analytics and IT help them do so?

The ‘big data’ movement

- Definition: “high volume, velocity, and variety information assets that demand cost-effective, innovative forms of information process for enhanced insight and decision making” (Gartner 2013)
- Data sources: product and service transaction records, inventory management systems, IT system logs, forms, multimedia files, email, social media feeds, Web analytics, metadata, mobile devices
- Size: Estimated 2.7 zettabytes (2.7×10^{20}) of raw data generated from all sources (IDC 2012)

Analytics: a way to tame 'big data'?

- Analytics: "the extensive use of data, statistical and quantitative analysis, explanatory and predictive models, and fact-based management to drive decisions and actions" (Davenport and Harris 2007)
- Comprising descriptive, predictive and prescriptive analytics, this movement transforms data into action through analysis and insight (Liberatore and Luo 2010)
- Methods include data visualization, descriptive and exploratory statistics, operations research models and artificial intelligence

Big data and analytics are important to the public and nonprofit sectors

- CitiStat movements comprise large-scale municipal data collection and analysis performance management (Center for American Progress 2007)
- "Urban mechanics" represents government-supported innovation in data collection and analysis, and IT applications for constituent service (Crawford and Walters 2013)
- Initiatives such as the Boston Indicators Project (The Boston Foundation 2012) and MetroBoston DataCommon (Metropolitan Area Planning Council 2013) connect citizens and organizations to curated datasets and Web-accessible analytics applications

These benefits do not flow equally across the not-for-profit sector

- Nonprofits tend to lag behind others in exploiting big data (Boland 2012, 2013)
- Community-based nonprofit organizations face particular challenges in data-driven decision modeling (Johnson 2011)
- Smaller and/or less progressive governments may only be starting to embrace principles of performance management (Daniels 2006)
- Community-engaged initiatives such as URBAN focus more on community engagement and participatory research than investigator-driven inquiry and data-focused solutions (Community Innovators Lab 2012)

Pilot project: Understanding the ecosystem of data, analytics and CBOs

- How can community-based organizations create information and make decisions to better fulfill their missions?
  - How do CBOs access and use data for operations and strategy?
  - What challenges do CBOs face in making best use of data and analytics?
  - How can data and analytics enable CBOs to identify and solve novel decision problems?

Certain assumptions motivate study

- There is a mismatch between data, methods and IT infrastructure that CBOs have, and that which is available to them
- CBOs lack data & analytics resources to take advantage of existing planning, service and policy opportunities
- CBOs lack capacity to identify decision opportunities

How valid are these assumptions?

What support is there in the field for certain hypotheses?

- CBOs can effectively articulate their information needs
- CBOs lack knowledge of and access to expertise and technology to create the information that meets defined needs
- CBOs lack capacity to identify and solve decision problems that are aligned with their needs

Community-based organizations are different from larger nonprofits and government

- 'Grassroots organizations' (The Boston Foundation 2007):
  - Low expenses
  - Large share of public charity tax filers (and likely non-filers)
  - Small fraction of economic activity
- Especially likely to meet needs of low-income or underserved populations through community development, human services and advocacy

CBOs have special needs

- Data limitations:
  - Service populations under-represented in datasets
  - Non-standard service areas
  - Indicators of social impact not typically measured, not widely available
- Competing tasks and obligations:
  - Fund-raising
  - Strategic planning
  - Service delivery
  - Community organizing & advocacy
  - Organizational design & management

Answering the research questions

- Literature review
  - Practitioner resources
  - Scholarly resources
- Field data collection
  - Key informant interviews
  - Data training observations
  - Focus group
- Survey
  - "Data, Analytics and Not-for-Profit Organizations" (http://www.surveymonkey.com/s/ZHCPM3Y)

Literature review findings

- GIS is frequently the basis for community engagement, community planning and participatory decision making (Elwood 2002, Elwood and Leitner 2003, Jankowski and Nyerges 2008)
- 'Community informatics' can provide theoretical frameworks and assessments of practice (Stillman and Linger 2009); 'knowledge transfer' describes process of transforming research into action (Wilson *et al*. 2010)
- Decision models developed in collaboration with CBOs have the potential significantly improve strategy design (Johnson *et al*. 2012)

Limited literature related to our research questions

Field research: Key informant interviews

- Description:

  Six key informant interviews with CDCs, advocacy organizations, service providers, and a data trainer

- Findings:
  - Data visualization is desirable but unavailable
  - Prefer outcome metrics based on values and mission rather than output measures mandated by administrative rules:
    - How to define?
    - How to measure?
  - CBOs seek specialized data usually to respond to funder reporting requirements

Key informant interview findings, continued

- Lack of knowledge of software resources intended to benefit organizations like them
- Do not specify analysis needs beyond descriptive statistics and maps
- Connections between multiple required software packages may entail wasteful double-entry
- Limited IT and analytic skills among staff members
- Enthusiasm for decision modeling applications
- Do not see ‘big data’ as relevant to their needs

Field research: Data trainings

- Description:

  Attended one training of neighborhood specialists for U.S. Census products and another to train users on proprietary software for public-service applications

- Findings & observations:
  - Such applications, intended for use by ordinary practitioners, are difficult to master
  - If used correctly and customized appropriately, could substantially improve of data-analytic skills and quality of data for decision-making

Field research: Focus group

- Description:

  Six directors of neighborhood branches of Federally-funded economic development enterprise

- Findings and observations:
  - Strong dissatisfaction with available IT applications for knowledge transfer and sharing
  - Required output measures do not capture neighborhood impacts
  - Want to quantify desired outcome measures
  - Strong interest in data sharing between sites, and site-specific 'dashboards' to measure performance
  - Some potential solutions are low-tech and inexpensive; other solutions require training; none require advanced degrees

CBO survey on data and analytics

1. How would you characterize your organization's need for data to do its daily work?
2. How would you characterize the purpose for which you most often retrieve data collected by sources?
3. How would you describe the quality of the data your organization has currently?
4. In what form does your organization usually access the data it needs?
5. How does your organization usually analyze the data it needs for routine activities?
6. What are the primary challenges your organization faces in acquiring the data it needs to do its work?
7. How would you characterize the problems your organization seeks to solve with the data it collects?

Survey results

- N = 10 (so far)
- Type: All are CBOs
- Size: 80% report 5 or fewer employees; 2 report 11 or more employees
- Service area: All report serving a defined geographic region

Need for data:

- Light (10%)
- Moderate (80%)
- Heavy (10%)

Primary data purpose:

- Research (40%)
- Funding (10%)
- Performance management (20%)
- Advocacy (10%)

Primary analysis methods:

- Office productivity software (70%)
- Databases or GIS (10%)
- No analysis (20%)

Problems sought to solve (% who choose 'very' or 'most' important):

- Service delivery (50%)
- Development (40%)
- Strategy design (90%)
- Analysis and research (60%)

CBOs have identified IT & decision modeling opportunities

- Information technology:
  - Geographic information systems training
  - Shared IT support among multiple organizations
  - 'Wiki'-style collaboration applications
- Decision modeling:
  - What metrics are most-closely linked to organizations' needs and values?
  - How can community economic development managers design a portfolio of services to maximize beneficial neighborhood outcomes?
  - What neighborhoods outside of a CBO's service region should be targeted for new initiatives?

Results analysis

- Propositions:
  - There is a mismatch between data, methods and IT of CBOs and that available from other sources – YES
  - CBOs miss planning, service and policy opportunities due to lack of data expertise – YES
  - CBOs often lack capacity to apply decision science – NO
- Hypotheses:
  - CBOs can effectively articulate their information needs – Supported
  - CBOs lack knowledge of and access to expertise and technology to create the information that meets defined needs – Supported
  - CBOs lack capacity to identify (Not supported) and solve (Supported) decision problems that are aligned with their missions

What have we learned about our research questions?

- CBOs access data primarily from required software and field data collection; limited use of customized data
- Data access highly constrained by administrative supports, competing priorities and available training
- Limited opportunities to articulate 'values' for data acquisition, communication and decision-making
- CBOs have identified multiple novel data-analytic and decision modeling applications

So what about 'big data' and 'analytics'?

- For CBOs, problem is not 'big data' and 'analytics' but small, customized flexible datasets whose variables reflect mission and values
- Case for analytics has not been made

Next steps

- Collect additional administrative, field & survey data to more rigorously validate propositions & test hypotheses
- Develop a theory of data & analytics usage for CBOs
- Develop solutions for 1 – 2 CBOs that have expressed an interest in data design, IT solutions and decision modeling

Questions?

Michael P. Johnson

Department of Public Policy and Public Affairs

University of Massachusetts Boston

michael.johnson@umb.edu

http://works.bepress.com/michael_johnson/

URBAN RESEARCH • ACTION

Connecting Online with URBAN.Boston

Website:

http://www.urban-boston.org/

Email:

urban.boston@umb.edu

Facebook Page:

https://www.facebook.com/UrbanBoston?ref=ts&fref=ts

LinkedIn Group:

http://www.linkedin.com/groups?gid=4804225&trk=myg_ugrp_ovr

References

Boland, S. 2013. "Avoid Getting 'Stunned' by Big Data." *Nonprofit Quarterly*, May 7, 2013. Web: http://nonprofitquarterly.org/management/22245-avoid-getting-stunned.

Boland, S. 2012. "Big Data for Little Nonprofits". *Nonprofit Quarterly*, November 28, 2012. Web: http://nonprofitquarterly.org/management/21410-big-data-for-little-nonpr.

Center for American Progress. 2007. *The CitiStat Model: How Data-Driven Government Can Increase Efficiency and Effectiveness*. Prepared by Teresita Perez and Reece Rushing. Washington, D.C. Web: http://www.americanprogress.org/wp-content/uploads/issues/2007/04/pdf/citistat_report.pdf.

Crawford, S. and D. Walters. "Citizen-Centered Governance: The Mayor's Office of New Urban Mechanics and the Evolution of CRM in Boston." Boston: The Berkman Center for Internet and Society at Harvard University, Research Publication 2013-17. Web: http://cyber.law.harvard.edu/publications/2013/citizen_centered_governance.

Daniels, A.C. 2006. *Performance Management: Creating Behavior that Drives Organizational Effectiveness, 4th Edition*. Performance Management Publications.

Davenport, T. H. and J. G. Harris. 2007. *Competing on Analytics: The New Science of Winning*. Boston: Harvard Business School Press.

Elwood, S. A. 2002. GIS Use in Community Planning: A Multidimensional Analysis of Empowerment. *Environment & Planning A* **34**: 905 – 922.

Elwood, S. and H. Leitner, 2003. GIS and Spatial Knowledge Production for Neighborhood Revitalization: Negotiating State Priorities and Neighborhood Visions. *Journal of Urban Affairs* **25**(2): 139 – 157.

Gartner. 2013. *IT Glossary*. Web: http://www.gartner.com/it-glossary/big-data/.

IDC. 2012. "IDC Predictions 2012: Competing for 2020." Prepared by Frank Gens. Web: http://cdn.idc.com/research/Predictions12/Main/downloads/IDCTOP10Predictions2012.pdf.

Jankowski, P., T.L. Nyerges. 2008. "Geographic Information Systems and Participatory Decision Making", in *The Handbook of Geographic Information Science*, (John P. Wilson, A. Stewart Fotheringham, Eds). New York: Blackwell Publishing Ltd., pp. 481-493.

References, continued

Johnson, M.P. 2011. "Community-Based Operations Research: Introduction, Theory and Applications", in (M.P. Johnson, Ed.) *Community-Based Operations Research: Decision Modeling for Local Impact and Diverse Populations*. New York: Springer, p. 3 – 36.

Johnson, M. P., Drew, R. B., Keisler, J. M. and D.A. Turcotte. 2012. What is a Strategic Acquisition? Decision Modeling in Support of Foreclosed Housing Redevelopment. *Socio-Economic Planning Sciences*, **46**(3): 194 – 204.

Liberatore, M.J. and W. Luo. 2010. The Analytics Movement: Implications for Operations Research. *Interfaces* **40**(4): 313– 324.

The Boston Foundation. 2012. *City of Ideas: Reinventing Boston's Innovation Economy*. Prepared by Charlotte Kahn, Jessica K. Martin and Aditi Mehta. Boston. Web: http://www.bostonindicators.org/our-reports/latest-report.

Community Innovators Lab. 2012. "URBAN Proposal 2". Boston: Massachusetts Institute of Technology.

Metropolitan Area Planning Council. 2013. *MetroBoston DataCommon: A Partnership between the Metropolitan Area Planning Council and The Boston Indicators Project at the Boston Foundation*. Web: http://metrobostondatacommon.org/.

Pease, G . And B. Beresford. 2013. "Big Data Analysis Drives Not-for-Profit Performance". *T + D Magazine*, june 8, 2013.

Stillman, L. and H. Linger. 2009. Community Informatics and Information Systems: Can They Be Better Connected? *Information Society* **25**(4): 255-264

The Boston Foundation. 2007. *Passion & Purpose: Raising the Fiscal Fitness Bar for Massachusetts Nonprofits*. Prepared by Elizabeth Keating, Geeta Pradhan, Gregory Wassall and Douglas DeNatale. Boston. Web: http://www.tbf.org/reports/2008/june/passion-and-purpose.

Wilson, M. G., Lavis, J. N., Travers, R. and S.B. Rourke. 2010. Community-Based Knowledge Transfer and Exchange: Helping Community-Based Organizations Link Research to Action. *Implementation Science*, **5**: 33-46.