Contents lists available at ScienceDirect

Children and Youth Services Review

Using big data for evidence based governance in child welfare

Hye-Chung Kum [a,c,*], C. Joy Stewart [b], Roderick A. Rose [b], Dean F. Duncan [b]

[a] School of Public Health, Texas A&M Health Science Center, 1266 TAMU, College Station, TX 77843, United States
[b] School of Social Work, University of North Carolina at Chapel Hill, 325 Pittsboro St, Chapel Hill, NC 27599-3559, United States
[c] Department of Computer Science and Engineering, Texas A&M University, 3112 TAMU, College Station, TX 77843, United States

ARTICLE INFO

ABSTRACT

Numerous approaches are available for improving governance of the child welfare system, all of which require longitudinal data reporting on child welfare clients. A substantial amount of agency administrative information – big data – can be transformed into knowledge for policy and management actions through a rigorous information generation process. Important properties of the information generation process are that it must generate accurate, timely information while protecting the confidentiality of the clients. In addition, it must be extensible to serve an ever-changing policy and technology environment. Knowledge discovery and data mining (KDD), aka data science, is a method developed in the private sector to mine consumer data and can be used in public settings to support evidence based governance. KDD consists of a rigorous 5-step process that includes a Web-based end-user interface. The relationship between KDD and governance is a continuous feedback cycle that enables ongoing development of new information and knowledge as stakeholders identify emerging needs. In this paper, we synthesis the different frameworks for utilizing big data for public governance, introduce the KDD process, describe the nature of big data in child welfare, and then present an updated KDD architecture that can support these frameworks to utilize big data for governance. We also demonstrate the role KDD plays in child welfare management through 2 case studies. We conclude with a discussion on implications for agency–university partnerships and research-to-practice.

1. Introduction

An important goal of policy and management research in child welfare is to improve the governance of the child welfare system. As a term applied to public policy, *governance* refers to the act of governing (Emerson, Nabatchi, & Balogh, 2011), the social coordination of public services (Lowndes & Skelchner, 1998), or the "creation, execution, and implementation of activities backed by the shared goals of citizens" (Bingham, Nabatchi, & O'Leary, 2005, p. 548). A number of related approaches such as evidence-based management (Kovner, Elton, & Billings, 2000) or performance management (Heinrich, 2007) have the potential to improve child welfare system governance through the use of performance information to inform the design and selection of interventions by managers and caseworkers. These approaches all require measures of the experiences of child welfare clients over time and across a spectrum of desirable social outcomes. The child welfare system, like the social welfare system in broader perspective, benefits from a wealth of administrative data collected by agencies for the discharge of their duties to clients and the public.

Converting these data into performance information that can be used by agencies and other stakeholders to improve the lives of clients and the effectiveness and efficiency of the child welfare system is not a straightforward task. This task involves creating consistent measures (Usher, Locklin, Wildfire, & Harris, 2001) while serving the diverse information needs of various stakeholders, maintaining accuracy in an ever-changing policy and technology environment, and ensuring client confidentiality. Knowledge discovery and data mining (KDD) is a rigorous approach to the integration, processing, modeling and distribution of such data. In the KDD process, administrative data are transformed first into longitudinal data consisting of key social indicators at different units (e.g. county, judicial district) and then subsequently modeled into "data cubes" for retrieval through a dynamic on-demand web interface. Data cubes are multi-dimensional tables that store the various pre-calculated measures from all aspects to support easy use. In addition, the individual level longitudinal data are used to support in depth analysis for reports, research, and evaluation. The relationship between KDD and governance is not one way, but rather exists in a feedback cycle that enables the continued development of new data infrastructure and reporting tools as needs arise or are recognized by stakeholders.

In this article, we extend prior works by situating KDD as a concept important to public governance, discussing several contemporary approaches to management that public agencies can use to effectively serve their populations, introduce the basics of big data and the KDD

* Corresponding author at: School of Public Health, Texas A&M Health Science Center, 1266 TAMU, College Station, TX 77843, United States.
E-mail addresses: kum@tamhsc.edu (H.-C. Kum), jstewart@unc.edu (C. Joy Stewart), rarose@email.unc.edu (R.A. Rose), dfduncan@email.unc.edu (D.F. Duncan).

process, and then further develop the KDD architecture. Data and information technology needs are presented as central to these approaches or any approaches that purport to translate knowledge about the population served into action and monitor performance such as implementation sciences (Fixsen, Blase, Naoom, & Wallace, 2009). We describe the generic KDD process for processing big data effectively and then present our updated system architecture developed for public governance applications detailing the new elements including examples of more advanced work in analytics. We demonstrate its efficacy by describing two demonstrations of KDD system in use. We conclude with a discussion of the strong partnerships between government agencies and interdisciplinary teams at public universities that can lead to successful implementation of comprehensive KDD information systems for child welfare.

2. Public governance

Generally, governance involves the management of activities and services informed by and contributing to the public interest with such management typically undertaken by multiple stakeholders (e.g., Emerson et al., 2011; Hill & Lynn, 2004). These stakeholders can include practitioners from government agencies (e.g., public managers), non-profit organization and private for-profit firms (Salamon, 2001). In social services, the stakeholders typically include county and state departments or divisions of social service managers and caseworkers, and private non-profit organizations.

Evidence and theory suggest that governance characterized by horizontally organized and collaborative partnerships between stakeholders,

such as between a government agency and a private organization, may be more effective in achieving the conditions necessary to promote socially desirable outcomes than top-down hierarchical relationships between government agencies and their subordinates (Emerson et al., 2011). Reorganizing public services into partnerships between public and private organizations, and a change in perspective from counting inputs or adhering to rules to a performance or accountability orientation are therefore central parts of public service effectiveness efforts (Heinrich, 2002). These new structures and performance orientations also have the potential to improve efficiency, offering these services at lower overall cost than their hierarchical predecessors (Benish, 2010; Head, 2008). Public policy scholars have long recognized that the effectiveness and efficiency of governance can be improved using social and behavioral science methods (Hill & Lynn, 2004). Three inter- related threads in the policy and management literature, evidence-based policy (EBP), performance management (PM), and evidence-based management (EBM), organize effectiveness efforts into cohesive frameworks that support public governance. We compare the three frameworks in Table 1, and demonstrate its relationship to KDD in Fig. 1.

3. Role of information systems in public governance

Scholars across fields such as policy (e.g., Bardach, 2003) and medicine (e.g., Sim et al., 2001) recognize the need for information systems that support the use of evidence-based best practices. Because child welfare agency managers (using, for example, PM or EBM) must make plans on the basis of performance and have ways to monitor the

Table 1
Comparison of public governance frameworks.

	Evidence-based policy (EBP)	Performance management (PM)	Evidence-based management (EBM)
Description	Policy decisions implemented on the basis of scientific research & evaluation.	Practices, technology, and information agencies use to monitor performance; decisions informed by monitoring in the context of goals; and actions taken to achieve goals[5,8]	"The conscientious, explicit, and judicious use of current best reasoning and experience in making decisions about strategic interventions"[11]
Goal	Identify and institutional effective policy interventions through the use of experiments[1]	Generate timely institutional knowledge through the day-to-day activities of governance[8]	Improve organizational performance through trial-and-error experimentation and continuous evaluation of staff activities and program outcomes[12]
Practice	Agencies can use extant literature to choose the most effective and efficient practices for addressing the unique challenges identified by data on agency performance[2]	Agencies make the most informed decisions possible that support their objectives, even in the absence of clear scientific evidence[9]; agencies conduct performance monitoring[8]	Multi-step processes: Scientific management, continuous quality improvement, and organizational learning; adapt, combine, and discontinue programs in response to changing agency conditions and client information[12]
Requirements	Information that identifies areas of need; access to evidence connecting these needs to appropriate policy interventions	Timely and accurate information technology structures[10]; organizational capacity to interpret data	Accumulation of long term data trends and information technology structures that handle these data[10]; organizational capacity and partnerships with research organizations to interpret data and change structures accordingly[10]
Challenges	Evidence and theory cannot keep pace with the needs of practitioners; agencies must perform certain tasks without evidence of their effectiveness[3,4]; Practices are partly informed by values of practitioners[1,5]; goals change over time[6,7]	Evidence may not be scientifically valid without controls or comparisons to a hypothesized counterfactual; agencies may not be staffed to handle the data visualization and analysis needs	Evidence may not be scientifically valid without controls or comparisons to a hypothesized counterfactual; agencies may not be staffed or have the partnerships needed to handle the data visualization and analysis needs
Relationship to Other Approaches	None inherent	Complementary to evidence-based policy[8]	The methods and processes of PM subjected to scientific scrutiny[13]
Role of KDD	Identify needs	Identify needs and monitor and evaluate program implementation and outcomes	Identify needs, monitor and evaluate program implementation and outcomes

[1] Jennings & Hall, 2004.
[2] Cannon & Kilburn, 2003.
[3] Bingham et al., 2005.
[4] Bardach, 2003.
[5] Head, 2008.
[6] Benish, 2010.
[7] Maynard, 2006.
[8] Heinrich, 2007.
[9] Schorr & Auspos, 2003.
[10] Meier & O'Toole, 2009.
[11] Kovner et al., 2000, p. 10.
[12] McBeath, Briggs, & Aisenberg, 2009.
[13] Briggs & McBeath, 2009.