

### Looking at the Future: Data & Anticipatory Intelligence

*The Board will be asked to look at the possible future impacts on ALA and on ALA members of big data and anticipatory intelligence.*

Big data refers to data sets beyond those housed in single data repositories (databases). Big data encompasses data from transactional information, social media, enterprise content (documents, web content, search, records, digital assets, etc.), sensors, and mobile devices.

#### Seven Vs of Big Data<sup>1</sup>

- Volume: the amount of data generated and collected
- Velocity: the speed at which data are analyzed
- Variety: the diversity of the types of data that are collected
- Viscosity: the resistance to flow of data
- Variability: the unpredictable rate of flow and types
- Veracity: the biases, noise, abnormality, and reliability of datasets
- Volatility: how long data are valid and should be stored

Variety may be the most challenging element for nonprofits and associations as they build systems to collect data elements in very specific areas (membership, conference and event attendance). Related to collecting in very specific areas, nonprofits may not produce enough Volume to create big data insights.

Big data is used in both predictive modeling, where past data is used to predict what might happen in the future, and anticipatory modeling, where past data is combined with current (and ongoing) data collection to anticipate changes (“surge pricing” or Amazon’s potential “anticipatory shipping”).

#### Considerations

*Mobile devices:* 95% of Americans own a cell phone of some kind and 77% own a smartphone – and 51% of American own a table computer.<sup>2</sup> Many big data insights can be gathered at the consumer level, allowing individuals to conveniently and intentionally (i.e. responsibly adhering to their privacy preferences) provide information about their interests for the improvement of services.

---

<sup>1</sup> “Big data for social innovation.” Kevin C. Desouza and Kendra L. Smith. *Stanford Social Innovation Review*. Summer 2014.

<sup>2</sup> “Mobile fact sheet.” Pew Research Center. January 12, 2017. Available from <http://www.pewinternet.org/fact-sheet/mobile/>

*Sharing platforms:* Google, Facebook, Twitter, and the wealth of sharing economy services (Airbnb, Uber, Lyft, etc.) provide opportunities for users to share information and receive benefits (information, social connection, service) that can then be used by companies to target, refine, repurpose, and personalize to users. Users knowingly use products and services that gather data or even require the sharing of large amounts of personal information having determined that the benefits are worth the sharing.

*Privacy:* Data sets can be combined to infer lifestyles, consumer habits, personal networks, and more. Privacy safeguards built into data collection can help address concerns, including data minimization so that no personally identifiable information is collected without defined purpose; de-identification that strips all information that could identify an individual, either directly or through links to other datasets; and user access controls that allow users to grant or deny specific requests to obtain information.<sup>3</sup>

*Unauthorized access:* As data becomes more valuable, it will need to be protected from disclosure, loss, or data theft. In addition to the damage done to individuals' privacy, issues of reputational harm, regulatory sanctions, weakened customer loyalty, and risks to public safety.

*Inexactitude:* Big data is able to work at least partly because it embraces inexact forms of classification and structure. Without pre-determined hierarchies and systems of organization, big data can be sorted by emerging systems generated by consumers or developed through the process of data collection.<sup>4</sup> This inexactitude, however, might also provide opportunities for greater insight into users' behaviors and new opportunities for correlation across data sets.

*Prediction and profiling:* The sources and pre-requisites for big data (government response, technology availability, available income and consumption, cultural orientation) can bias the validity of big data, which becomes particularly concerning as big data is used for prediction and profiling. Big data predictions can empower the powerful (e.g. urban communities, younger and technologically inclined populations) and perpetuate existing systems by profiling certain communities.

---

<sup>3</sup> Have it all: Protecting privacy in the age of analytics." Ann Cavoukian, David Stewart, and Beth Dewitt. Deloitte. 2014.

<sup>4</sup> *Big Data: A Revolution That Will Transform How We Live, Work, and Think.* Viktor Mayer Schonberger and Kenneth Cukier. Houghton Mifflin, 2013.