

Large-Scale Data Management in San Antonio

Working Paper

Draft May 1, 2016

History

In the late 1990s, a Community Network (SATNET) and the SA Community Information System (SACIS) were working separately to achieve a common goal: easy public access to information related to healthcare, education, and social services. After planning meetings with community leaders and over 30 public, private, and non-profit organizations, it was decided in 2000 that merging of SATNET and SACIS to become the Alamo Area Community Information System (AACIS) was the best strategy for making the proposed information system a reality.

In 2008, AACIS expanded the collaboration to include the San Antonio Regional Campus of the University of Texas Health Science Center at Houston School of Public Health in order to update and expand data holdings and create a community portal with which the community could access data and other resources. In 2009, AACIS developed a visionary project known as NOWCastSA, a community news and public service journalism organization.

In January of 2010, the need to revise the mission and vision of AACIS became clear, and brainstorming and visioning meetings resulted in the rebranding of AACIS as Community Information Now, or CI:Now. Then, in October 2010, CI:Now was accepted as a local partner in the National Neighborhood Indicators Partnership (NNIP). CI:Now continues to an active participant and leader in NNIP but local challenges remain.

Challenges

Currently, gathering information on targeted variables is necessary but can be frustrating. In some cases it is still necessary to justify the expense and time required for this work: relevant data sets allows us to identify needs, establish baselines, develop new policies and programs, monitor trends, evaluate outcomes, and assess productivity and efficiency.

The challenges associated with data management in San Antonio include:

- lack of common definitions of even basic terms
- variation in metrics and standards
- lack of funding/cost of maintaining data sharing operations
- difficulty tracking person-level outcomes through a 3rd party (i.e. workforce agency)
- privacy concerns (FERPA) vs. facilitating services
- lack of automation for common information requests
- creating a culture of willingness to share data

Large-Scale Data Management in San Antonio
Working Paper
Draft May 1, 2016

With regard to the data itself, problems include who, how, and what is available vs. what is needed; whether existing data is current, complete and accurate; technical compatibility between data sets; the availability and use of local data vs. state or national data, and the linkage of data sets lead to aggregated assumptions.

What Would “Data Nirvana” Look Like?

Data Nirvana is envisioned as a “one stop shop,” a centralized repository for all local data that will be available both through a web-based data system and a physical location where highly trained employees manage, share, and report data. This system would bring together several groups that focus on data reporting, measurements, and analytics to reduce duplication and to create standardized definitions and methodologies that can connect to national outcome metrics.

It is ideal that data be transparent for residents and community groups while protecting the confidentiality of citizens by adhering to FERPA’s data privacy law. Information sharing for non-local organizations would be available through data sharing agreements.

Data would be comprehensive local information--accurate, current, and useful to reliably report history, measure current status, and predict future outcomes. Data gaps will be systematically identified and subsequently addressed. The web-based data system should be an easy-to-use interactive instrument that displays visually represented data, cause and effect, and related factors to provide users the opportunity to draw accurate conclusions and determine the appropriate response. Lastly, in “Data Nirvana” the effectiveness of implemented activities/services, programs, and policies will be frequently assessed and both measures and policies subject to continual improvement based on data results.

Building a More Robust Data Management Culture: Next Steps

In February, 2016, the Office of Mayor Ivy R. Taylor, the San Antonio Area Foundation and CI:Now convened a group of 40-50 data professionals from various fields to develop an overview of the current state of data management in our community, and to suggest next steps. This working paper, provisionally titled “Large Scale Data Management in San Antonio,” is an outgrowth of that meeting (item 1., below). The action items agreed upon during the convening are:

1. Develop Data Management Summary
 - a. Vision
 - b. Challenges, Vision, Next Steps

Large-Scale Data Management in San Antonio

Working Paper

Draft May 1, 2016

2. Compile data management history in San Antonio and comparison cities
 - a. Timeline and description
 - b. Include data success stories
3. Begin to build a network of “data geeks”
 - a. Create a directory providing contact information
 - b. Develop a discussion forum/platform
4. Standardize terminology and metrics
 - a. Develop lexicon
 - b. Develop common metrics

With regard to item 4., we have identified research and data lexicons that could be used as a basis for a local product:

- <http://www.aisp.upenn.edu/integrated-data-systems/establishing-an-ids/>
- <http://files.eric.ed.gov/fulltext/ED544710.pdf>
- http://www.aisp.upenn.edu/wp-content/uploads/2015/03/DataIntegration_paper_Kumar.pdf
- https://books.google.com/books?id=C7x_BAAAQBAJ&pg=PA72&lpg=PA72&dq=integrated+data+warehouse+city&source=bl&ots=WoleJ1iUlo&sig=dOktQniP4m_YhMtGicHpftK60h0&hl=en&sa=X&ved=0ahUKEwiN7s-c-bfMAhWEKCYKHfoWBkU4ChDoAQggMAE#v=onepage&q=integrated%20data%20warehouse%20city&f=false
- <https://books.google.com/books?id=1bpEifVEi2MC&pg=PA1356&lpg=PA1356&dq=integrated+data+warehouse+city&source=bl&ots=WyH6-kwhJK&sig=9qLCWlkSjOyOe-f9s1HuvSRNgtU&hl=en&sa=X&ved=0ahUKEwiN7s-c-bfMAhWEKCYKHfoWBkU4ChDoAQg3MAU#v=onepage&q=integrated%20data%20warehouse%20city&f=false>
- <http://www.aisp.upenn.edu/resources/aisp-best-practice-papers/>
- <http://uar.sagepub.com/content/50/4/577.full.pdf>
- <https://www.data.gov/impact/#local> (Can establish using another website)
 - <http://www.data.gov/about#who>

Examples and Ideas

IDS allow government agencies to integrate various databases and bridge the gaps that have traditionally formed among them and between government agencies and community providers.

Large-Scale Data Management in San Antonio

Working Paper

Draft May 1, 2016

- http://www.chapinhall.org/sites/default/files/old_reports/69.pdf
- http://www.aisp.upenn.edu/wp-content/uploads/2015/08/SouthCarolina_CaseStudy.pdf

Researchers can compare data across various agencies and ask probing questions about issues that have stymied public agencies' past efforts to improve policies.

- <http://www.aisp.upenn.edu/network-site/alleggheny-county-pa/>
- <http://www.aisp.upenn.edu/network/multi-site-research-projects/>

Integrated data reveal patterns of risk and resilience, which, in turn, more readily links cause and effect, allowing executive leaders, policy makers, and policy analysts to test more targeted interventions and higher-impact policies.

- <http://www.aisp.upenn.edu/network-site/washington-state/>
 - http://www.aisp.upenn.edu/wp-content/uploads/2015/08/WAState_CaseStudy.pdf