What's Needed for a Community to be Good at Using Data?

As the San Antonio region works to build our local data capacity, what conditions should we be working to achieve? The following summary outlines some key human and technical capacities a data-driven community would ideally have in place.

1. **A widespread culture of valuing data**. People in all kinds of different settings value good data and recognize that using it can make them more effective. Organizationally that might translate into a commitment to collect and share data, to use data internally, to provide data-related services, to financially support local data infrastructure, etc.

2. **Widespread data literacy**. People who collect, generate, store, manage, analyze, interpret, disseminate, communicate, or review data have a basic working knowledge of how to handle and talk about it effectively and responsibly. People across organizations and issue areas are able to communicate effectively with each other about data, speaking a common language.

3. **Widespread processes and structures to support data use and communication**. Public and private organizations of all kinds, including collaboratives of organizations, have some systematic, routine way of engaging with data to inform their decisions and actions. There also needs to be some way to engage and support the general public in using data.

4. **Widespread good data management practices**. People and organizations who collect, generate, store, manage, analyze, interpret, disseminate, communicate, or review data understand data quality and data privacy and security and have appropriate policies, practices, and formal agreements in place.

5. **Widespread access to data**. The general public and public and private organizations of all kinds are able to get the data they need at low or no cost.
    - Data should be available for a wide range of issue areas and geographic areas across the entire San Antonio region
    - Data should be trustworthy, timely, and highly disaggregated demographically and geographically
    - Where appropriate, datasets should be integrated or linked across sources and issues to form a more complete picture
    - Access to non-private data should be quick and easy. It should not require personal relationships with data holders or any special hardware or software that isn't already in wide use
    - Data of any sensitive nature should be protected appropriately through a variety of tools like data aggregation, data suppression, usage agreements, required user training, privacy and security policies and practices, hardware and software, etc.

- Data should be available in a form that is useful to the person needing the data. That form will differ widely by user and purpose, e.g., a downloadable machine-readable data file, a paper handout tailored to a low-literacy audience, an online neighborhood profile, or a technical research report or community assessment. (The degree of time and effort required to prepare data files may be just as great as for highly-synthesized formats.)
- Data should come with information that helps people use it effectively and responsibly. Depending on the data or situation, that might mean metadata files, lay-friendly user notes, mouseover text in online visualizations, direct conversations with people very familiar with a particular dataset, etc.
- Proprietary datasets, analysis tools, storage solutions, etc. should be leveraged whenever possible. Resources should not have to be purchased separately by multiple local organizations if extending access in some way is possible without violating the terms of purchase/license/use agreement.
- Where appropriate, data and indicators should line up with the evidence base and be able to be compared against other geographies (e.g., other cities, state, nation)

6. **Widespread access to help in using data effectively**. People/organizations can get competent help:
   - determining exactly what data they need given the problem or need at hand, including developing appropriate indicators
   - finding or accessing data
   - understanding and applying data for a particular purpose
   - disseminating data and communicating about it
   - establishing informal or formal data-sharing relationships and agreements
   - making data available to other organizations or to the general public

7. **Collaborative and continual planning and action to build local capacity to use data**, e.g.,
   - Effective governance of and shared responsibility for data-sharing initiatives, central data repositories, integrated data systems, community dashboards, etc.
   - A process/structure to identify whatever is needed – data, integrations, tools, services, communication methods, etc. – and secure or develop it
   - Ways to develop or incubate datasets that don't currently exist or are not strong enough for use
   - Ways to establish agreements, protocols, common indicators, data standards, or other things that support community data use
   - Ways to establish and foster good working institutional and personal relationships at the local, state, and national level, especially among organizations with major roles in collecting/generating, managing, analyzing, or communicating data
   - Ways to learn about and potentially adopt innovations and effective practices from other cities/regions