

# DATA ECOSYSTEM PROJECT

## Appendices

bcANALYTICS  
MAY 2016

#### PREPARED BY

The buildingcommunityWORKSHOP ([bc]) is a Dallas based nonprofit community design center seeking to improve the livability and viability of communities through the practice of thoughtful design and making. We enrich the lives of citizens by bringing design thinking to areas of our city where resources are most scarce. To do so, [bc] recognizes that it must first understand the social, economic, and environmental issues facing a community before beginning work.

#### SUPPORTED BY

As the largest community foundation in Texas and one of the largest in the nation, Communities Foundation of Texas (CFT) works with families, companies and nonprofits to strengthen our community through a variety of charitable funds and strategic grantmaking initiatives. The foundation professionally manages more than 900 charitable funds and has awarded more than \$1.3 billion in grants since its founding in 1953. Increasing financial stability of working families is one of the two key focus areas of CFT's community impact funds. To support this area, CFT has launched the Data Driven Decision-Making (D3) Institute. The D3 Institute is designed to provide organizations that offer programs and services for low-income working families the power to accelerate their development of enduring solutions to the social and economic problems facing this population. [www.cftexas.org/D3](http://www.cftexas.org/D3)

## ABOUT bcANALYTICS

Overcoming the major economic and social issues that affect the working poor requires coordinated data gathering and analysis. To address these issues, buildingcommunityWORKSHOP has developed a data and analytics team, bcANALYTICS, focused on providing high quality services to nonprofit and community based organizations. Additionally, bcANALYTICS, has begun researching and discussing a variety of ways to improve data collection, analysis, and sharing in the region.

### *D3 Analytics*

Through Communities Foundation of Texas' Data-Driven Decision-Making Institute (D3), bcANALYTICS has offered graduating organizations an analytics package to answer pressing organizational questions - from appropriate locations for a new facility to demographic analysis of an organization's existing service area. To date, bcANALYTICS has completed reports for 10 organizations working across North Texas.

### *Data Ecosystem Project*

In 2014, bcANALYTICS began researching national best practices in community-focused data analysis, collection, and sharing in order to identify scalable solutions for Dallas. This work has been focused in three primary areas: a review of organizations nationwide, a survey of Dallas nonprofit data community, and conversations with stakeholders working across North Texas. The project seeks to identify practical approaches to improving the ecosystem of data and information in the region.

### *Fee-For-Service Analytics*

bcANALYTICS offers services to nonprofit and community based organizations. Services offered include data collection, analysis, cartography, and visualization. Additionally, bcANALYTICS provides support to projects housed in other units of buildingcommunityWORKSHOP.

# APPENDIX A - NATIONAL BEST PRACTICES

# COMMUNITY DATA BEST PRACTICES

**buildingcommunity**WORKSHOP  
MAY 2015

# BEST PRACTICES



COMMUNITY ENGAGEMENT includes any activity whereby the broader community is convened and engaged to discuss data. It is through community engagement that leaders convene, common narratives emerge, and networks expand to solve complex problems. Activities range from hack-a-thons which engage the civic tech community, to conferences and public events which increase data awareness, public participation, and best practice sharing partner organizations.



DATA ASSEMBLY, MANAGEMENT, and PROVISION is a common best practice whereby a trusted central organization is responsible for the collection and management of a shared database system for data requests, external or internal analysis, and open data sharing. This often serves as a one stop shop for many organizations looking for facts about the city. Data sharing agreements and other trust mechanisms protect sensitive data, while public data is assembled and transformed for maximum public use.



COMMUNITY BASED REPORTING - or INDICATORS is a very common best practice where an organization generates an annual report chronicling the health and well-being of the region. This includes collecting information considered relevant by the majority of the community and translating that data into a coherent narrative. The reports are meant to serve as an informative and a strategic mechanism for service delivery and fund generation.



PROJECT BASED REPORTING - many take on a consulting role and produce strategic materials or issue-specific reports for local organizations wanting to leverage internal or public data. These can be produced through sustained funding within a long-term partnership or through a one time fee-for-service arrangement.



WEB OR SOFTWARE DEVELOPMENT - some organizations with strong relationships with local technical talent often engage in web and software tool development and intensive capacity training. Live web maps, dashboards, and online queries allow users to visualize and learn from data. Tools are often issue specific but broad enough to add value to the larger community. Additionally hack-a-thons can leverage local talent to address community needs.



CROSS SECTOR ENGAGEMENT is any activity that brokers strategic relationships between government, non-profit, and commercial entities around a single project or sustained effort for aligned or strategic action. This can also take the form of issue-specific hack-a-thons, ongoing indicator development and evaluation, or networking events meant to incubate best practices for data locally.



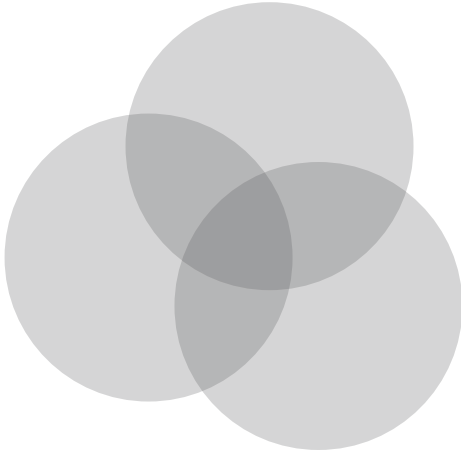
CAPACITY BUILDING is any activity where partner organizations engage in best practice sharing, staff training, and collaborative learning facilitated by the core model. Experts are paid or volunteer their time in events and meetings meant to engage and build internal staff capacity for partner organizations. Otherwise tools and materials are disseminated to build capacity passively.

# ORGANIZATION TYPES

While the structure of each organization is deeply dependent on local conditions, three general organization types begin to emerge.

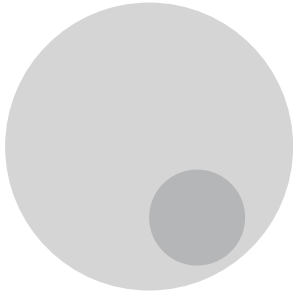
## COLLABORATIVE

This is often an early conception of an indicator or data-to-impact project. Multiple trusted partners invest through in-kind donation, staff time, or basic funding. Significant leadership is also usually provided by an advisory board to identify priorities, establish key deliverables, and evaluate community value over time.



## NESTED

Another common type of organization is one that acts independently within a trusted, larger, and more established parent organization. The parent organization often provides significant leadership, resources, and sustained funding. Similar to the collaborative model, significant leadership is also usually provided to identify priorities, establish key deliverables, and evaluate community value over time.



## INDEPENDENT

Independent organizations is the most mature manifestation of the data-to-impact projects and are typically established after an incubation period within a larger non-profit or collaborative framework. These function independently as an incorporated non-profit and are responsible for their own funding.



# OAKLAND | URBAN STRATEGIES COUNCIL



Organization Type: Independent Nonprofit  
City Governance: Mayor-Council  
Founded: 1989  
Website: [www.urbanstrategies.org](http://www.urbanstrategies.org)

Annual Budget: ~\$205,000  
FTE: 8  
PTE: n/a  
Director: Junious Williams

## MISSION

The focus of the Council's work is to support youth, family and community development as a means of building healthy, vibrant communities in order to reduce persistent poverty.

## DESCRIPTION

The focus of the Council's work is to support youth, family and community development as a means of building healthy, vibrant communities in order to reduce persistent poverty. The Research and Technology Program has 8 employees and works to increase availability and access to data regarding Oakland area residents. This team also works with organizations and individuals to build the local capacity for data use and literacy as well as technological skills in general. The organization as a whole has an operating budget over \$1 million, although only around \$205,000 is set aside for the Research and Technology team.

## PROJECTS & ACTIVITIES

**Research and Technology Program**

This program has four main goals: to increase the availability and access to public data by residents and CBOs; to increase capacity of data use by residents and CBOs, to increase access to technology in low-income communities of color, and to increase the capacity of low-income communities to use information technologies.

**Info Alameda**

Through the [infoAlameda.org](http://infoAlameda.org) platform, USC collects, analyzes, and shares a wide range of data for the local community to use. This data can be viewed online or downloaded for a user to use in their own software. This data is available free of charge, and needed data can be requested to [infoAlameda/USC](http://infoAlameda/USC).

**GIS Consulting**

USC engages in GIS consulting for a small fee to non-profits, community organizations, student organizations, public agencies, or neighborhood groups (for-profit/private companies are excluded).



# NEW HAVEN | DATA HAVEN

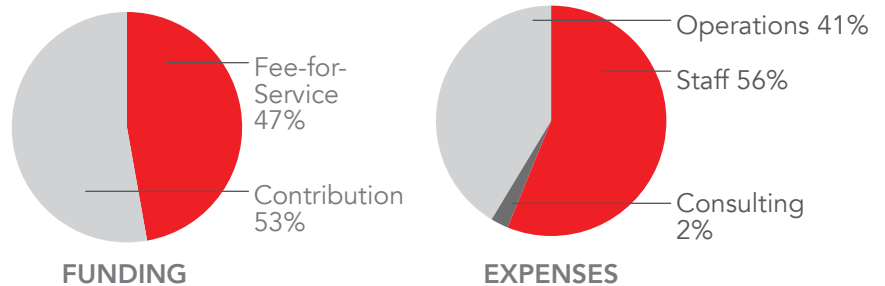


Organization Type: Independent Nonprofit  
City Governance: Mayor-Council  
Founded: 1992  
Website: [www.ctdatahaven.org](http://www.ctdatahaven.org)

Annual Budget: ~\$254,000  
FTE: 3  
PTE: NA  
Director: Mark Abraham

## MISSION

To improve quality of life by compiling, sharing, and interpreting public data for effective decision making.



## DESCRIPTION

Data haven is considered the “data steward” for the Connecticut Data Collaborative, a public-private partnership formed to change the way Connecticut uses data. DataHaven assembles and processes large quantities of information from various public and private entities to make it more useful to local communities. This is done through tools and resources available on its website and the creation of community indicator reports, web mapping, survey design services, grant writing assistance and reports. DataHaven has transitioned from a simple convening of cross-sector partners, to a community reporting service, into an independent organization with online interactive tools, and technical support services. Data-haven’s website is central their work. The website makes over 4,500 indicators about the region available for download at multiple scales, hosts a news blog and community knowledge forum to pool topic-specific resources, invite discussion and build local data capacity.

## PROJECTS & ACTIVITIES

**Indicators**     

DataHaven maintains over 10,000 indicators of community wellbeing, many at the neighborhood level, and are made available for download at multiple scales.

**Community Knowledge Center and Forum**     

In order to build local capacity and increase public knowledge, resources made available by contribution are compiled in one place. The site features different portals organized by geography, topic area, and contributing organization. As a result, links to other data efforts, national and local reports, and grant writing guides are organized in one place and allow a space for discussion.

**Greater New Haven Community Index**   

An annual assessment of well-being and quality of life is published to illustrate the opportunities and challenges faced as a community and comparisons between different geographies. Topics include demographics, economic growth, health, and civic life.

**Executes Wellbeing Survey**   

Nationally recognized program interviews randomly selected adults living in region on a variety of quality of life issues and public health. Originally only focused around New Haven in 2003, it has since and expanded include the entire state. Program is designed with and supported by 35+ government, academic health care, and community partners. Results are used to unify vision and strengthen community knowledge.

**Technical Assistance**     

Data Haven provides over \$20,000 worth of technical assistance to organizations looking to build capacity in grant writing, data development, and research support. More in-depth technical assistance and analyses are provided within a fee-for-service agreement.

# BALTIMORE | BALTIMORE NEIGHBORHOOD INDICATORS ALLIANCE



Organization Type: Nested Academic  
City Governance: Mayor-Council  
Founded: 2000  
Website: [www.bniajfi.org](http://www.bniajfi.org)

Annual Budget: ~\$300,000  
FTE: 3  
PTE: 1  
Director: Seema Iyer

## MISSION

The goal and mission of BNIA-JFI is to strengthen the practice of data-driven decisions to support stronger neighborhoods, improved quality of life, and a thriving city by empowering Baltimore residents, neighborhoods, and organizations with the data and analyses required to shape and impact policies that affect them and to support improvements within Baltimore.

## HISTORY & DESCRIPTION

The BNIA was formed in 2000 as way to enhance and synchronize efforts for three separate efforts in the city. The first was the Baltimore City Data Collaborative which was a collaborative effort to provide crucial datasets to local community-focused organizations. The city's internal data improvement plan CitiStat, and The Pratt Center's resource and capacity building program represented the second and third programs respectively. The Association of Baltimore Area Grantmakers convene the groups and formed BNIA, which now provides an annual community well-being report, technical expertise for cross-sector alignment efforts, consulting and capacity building services for nonprofits, and multi-source anonymized datasets for download on their website. BNIA-JFI continues to assist community-based organizations in strategic planning, grant writing, neighborhood advocacy, data requests, technical assistance and training.

## PROJECTS & ACTIVITIES

### *Vital Signs and Community Profiles*

The core product of BNIA-JFI, is a annual report that identifies data-points from a variety of sources that “take the pulse” of Community Statistical Areas (CSA). CSAs are neighborhoods defined socially by community planning organizations and physically a collections of census tracts. While this cannot reflect the unique fluidity of neighborhoods, it does allow the city to consistently track indicators over time. There are a total of 55 CSAs in Baltimore. Students, community organizers, researchers, nonprofits seeking grants, and grant funders all access the data.

### *Baltimore Data Days*

Since 2010, BNIA has hosted an annual gathering of community organizations, civic hackers, BNIA staff, and others to facilitate learning tools, and sharing best practices around the community. Sponsorships allow the conference to be FREE and OPEN to the public.

### *Custom Research*

BNIA performs research conducted on behalf of its parent organization(JFI), other nonprofit clients, and cross-sector efforts focused on Baltimore communities. BNIA also performs custom data searches, data analysis, benchmarking, GIS, market and policy analysis, and training. Web maps focused on visualizing key variables are also produced for cross-sector projects.

# AUSTIN | COMMUNITY ADVANCEMENT NETWORK



Organization Type: Collaborative  
City Governance: Council-Manager  
Founded: 2005  
Website: [www.canatx.org](http://www.canatx.org)

Annual Budget: ~\$300,000  
FTE: 4  
PTE: 1  
Director: Vanessa Saria

## MISSION

CAN is a partnership of governmental, non-profit, private, and faith-based organizations which leverage mutual resources to collectively improve social, health, educational and economic opportunities in our community.

## HISTORY & DESCRIPTION

In 1995, CAN emerged from a desire to effectively invest in solutions to the problems of Travis County. To achieve its goal, CAN was founded on the belief that all organizations working in and with the community need to work together to affect change. A coalition of 25 organizations participate in CAN and provide various of levels of funding - giving the project the opportunity to work without the need to raise money. Although CAN has struggled at times to maintain a vision, bringing their annual Dashboard report online helped to unify and organize the partner organizations. This has allowed CAN to work as the intermediary to release data to the public, listen and act on input from partners, demonstrate where action is occurring, and convene conversations necessary to addressing the issues faced in Travis County. While some projects work to solve projects, CAN is a neutral player and is not involved in action or political debates. Rather, CAN focuses only on helping Travis County understand how it is performing on selected metrics every year.

## PROJECTS & ACTIVITIES

### *CAN Community Dashboard*

The Community Dashboard is both an annual report and an interactive website that shows how the Austin and Travis County region is performing along key indicators. While data is available for the specific indicators, the Community Dashboard goes further than some indicator projects in identifying organizations and agencies that are working to “move the needle” for certain indicators.

### *Community Forums*

Within the annual theme of the Community Dashboard, CAN hosts multiple forums each year to bring stakeholders and community members together to talk about issues important to solving the overall theme. These forums include both presentations from experts and breakout discussions on specific sub-issues.

# DETROIT | DATA DRIVEN DETROIT



Organization Type: Nested Nonprofit  
City Governance: Mayor-Council  
Founded: 2008  
Website: [www.datadrivendetroit.org](http://www.datadrivendetroit.org)

Annual Budget: ~\$1,600,000  
FTE: 7  
PTE: -  
Director: Erica Raleigh

## MISSION

Data Driven Detroit (D3), an affiliate of the Michigan Nonprofit Association (MNA), provides accessible, high-quality information and analysis to drive informed decision-making. Our vision is that essential and unbiased information is used by all.

## HISTORY & DESCRIPTION

From the Michigan Metropolitan Information Center at Wayne State University, to the Southeast Michigan Information Center at United Way, there have been many initiatives to collect and democratize data about Detroit and its neighborhoods. In 2008, The Skillman Foundation and The Kresge Foundation awarded City Connect Detroit a \$1.85 million grant to incubate Data Driven Detroit (originally named the Detroit-Area Community Information System). Within its first year, D3 was selected by the Urban Institute to join the National Neighborhood Indicators Partnership. In December 2012, after an extensive period of review, discussion and due diligence, Data Driven Detroit became an affiliated program of the Michigan Nonprofit Association.

## PROJECTS & ACTIVITIES

### *Open data portal*

D3 stores and offers data from a variety of sources through their open data portal through ESRI's open data platform.

### *Interactive Web Tools*

D3 offers a wide range of interactive web maps and tools to visualize demographic, socioeconomic, housing data. Data topics are focused on common issues known to affect the well-being of Detroit and are available for download through D3's data warehouse.

### *Consulting Services*

D3 provides a number of independent research services for cross-sector strategy and action and can range greatly in scale and scope. Services offered include data collection, visualization, analysis, and provision.

### *Motor City Mapping*

To provide a baseline dataset to strategize cross-sector action and planning against blight, D3 has completed a citywide survey of 380,000 parcels in partnership with the Detroit Blight Removal Task Force, the Michigan Nonprofit Association, LOVELAND Technologies, Quicken Loans, and the Detroit Employment Solutions Corporation. Project formulation was developed through a series of community workshops.

# MINNESOTA | MINNESOTA COMPASS



Organization Type: Nested Nonprofit  
City Governance: n/a  
Founded: 2008  
Website: [www.mncompass.org](http://www.mncompass.org)

Annual Budget: ~\$350,000  
FTE: 4  
PTE: n/a  
Senior Researcher: Craig Helmstetter

## MISSION

Minnesota Compass is a social indicators project that measures progress in our state, its seven regions, 87 counties and larger cities. Compass tracks trends in topic areas such as education, economy and workforce, health, housing, public safety, and a host of others.

## HISTORY & DESCRIPTION

Wilder Research, an arm of the Wilder Foundation, was doing community-based research and reporting well before founding the Minnesota Compass project and its predecessor, Twin Cities Compass. The demand for indicator data from within the Twin Cities and the rest of the state was the impetus in expanding the project statewide in 2008. To do so, strong advisory boards were convened to establish a wide variety of topic areas and the indicators needed to successfully measure the performance within the topic areas. These key measures show how the state, regions, and select cities compare to the US and each other through interactive charts and tables that may be broken down by various socioeconomic classifications. The underlying data pertaining to these indicators is made available for download in machine-readable formats, and the charts or tables are also available for use in grant applications or media articles.

## PROJECTS & ACTIVITIES

### *Minnesota Compass*

The primary project of the Minnesota Compass, is to compile ~70 indicators and key measures from multiple governmental and local sources within a number of focus areas. These include Aging, Children & Youth, Civic Engagement, Demographics, Disparities, Early Childhood, Economy, Education, Environment, Health, Housing, Immigration, Public Safety, Transportation, and Workforce. Key measures are selected by topic-focused advisory groups comprised of experts in the field, academic, and business leaders. These measures are also visualized on the website and made available for download in a variety of methods.

# CHICAGO | SMART CHICAGO COLLABORATIVE



Organization Type: Collaborative  
City Governance: Mayor-Council  
Founded: 2010  
Website: [www.smartchicagocollaborative.org](http://www.smartchicagocollaborative.org)

Annual Budget: ~\$1,300,000  
FTE: 3 + consultants  
PTE: n/a  
Director: Daniel O'Neill

## MISSION

Smart Chicago is a civic organization devoted to improving lives in Chicago through technology. The organizations work on increasing access to the data and Internet tools, improving skills for using technology, and developing meaningful products from data that measurably contribute to the quality of life of residents in our region and beyond.

## HISTORY & DESCRIPTION

In 2007, The mayor's Advisory Council on Closing the Digital Divide a commissioned a report titled "The City that NetWorks" highlighting a need for cross sector collaboration to increase capacity and data use within the city and solve complex urban problems. By leveraging this strategy for equitable technology access, the government's commitment to open data, and Chicago's growing technology sector, the SCC was formed in 2010. Federal grants from the American Recovery and ReInvestment Act have provided funding for public computer programs implemented by LISC, Lumity, and other organizations building capacity for communities with a digital disparity. Project-based collaborations are built around cross-sector issues identified by funding partners, civic leaders and developed by Chicago's civically engaged tech community. All projects are vetted through the SCC principles of internet technology, open data, and equitable access to information and tools, and a focus on Chicago. The organization was founded and supported by Chicago's Department of Innovation and Technology, is primarily funded by the MacArthur Foundation, secondarily funded by other, project-based funders, and housed within the Chicago Community Trust.

## PROJECTS & ACTIVITIES

### *Web tools*

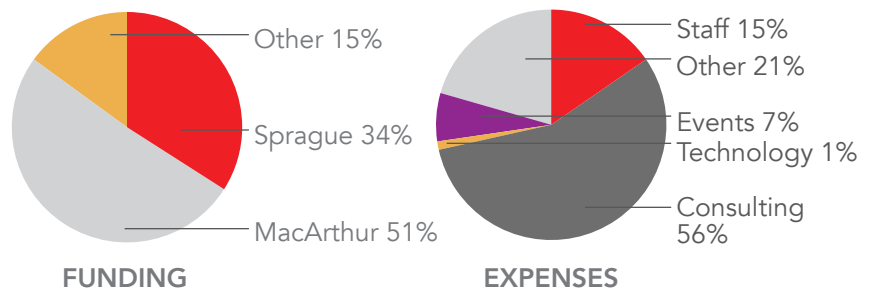
Central to the function of SCC's functioning, project managers administer funds and coordinate between developers, designers, civic users, and data providers to generate community wide and topic-specific web tools to leverage available community data. Tools are focused on Health, Education, Justice, and City services. Projects can be conceived in cross-sector hack-athons and vetted through SCC principles, experts, data providers, and potential users in testing groups.

### *Developer Resources*

To incubate developer participation, free web hosting, workspace, code access, and regular meetups and hackathons are provided within the offices of Smart Chicago. A paid civic user testing resource (CUTGroup) is also made available to developers to refine user-experience design and build skills for both tool builders and tool users.

### *Connect Chicago*

Funds are also administered through SCC, to a loose network of public programs where internet and computer access, digital skills training, and online learning resources are made available for free.





## APPENDIX B - DATA ECOSYSTEM SURVEY



# Dallas Data Ecosystem Survey

Thank you for taking the Dallas Data Ecosystem survey. Funders, city officials, and nonprofits in the Dallas area would like to better understand how data is generated, shared, and used to drive local action. Understanding existing connections, barrier, and finding possible linkages can help foster a culture of collaboration and build a better city. While your individual responses will be kept confidential, aggregated survey results will be made available to all survey takers in the near future.

\* Required

1. **First Name \***

.....

2. **Last Name \***

.....

3. **Title**

.....

4. **Organization/Company/Government Agency or Department \***

.....

5. **Email \***

.....

6. **Phone Number**  
(XXX-XXX-XXXX)

.....

7. **Organization Address**

.....

8. **Organization City**

.....

9. **Organization Zip Code \***

.....

## Organization Information

10. How would you categorize your organization? \*

Mark only one oval.

- Non-profit
- For-profit
- Academic
- Education
- Local Government
- Regional Government
- Religious Organization

11. What category below best describes the work your organization is primarily engaged in (direct activity)?

Select One

Mark only one oval.

- Education
- Community Building
- Economic Development
- Food
- Health
- Environment/Sustainability
- Law Enforcement/Civil Action
- Advocacy/Lobbying
- Design/Planning
- Arts
- Technology
- Faith-based
- Funding/Philanthropy
- Consulting/Strategy
- Collaborative
- Other: .....

12. **What categories below best describe the work your organization is secondarily engaged in (through funding or provisional services)?**

Select all that apply  
Check all that apply.

- Education
- Community Building
- Economic Development
- Food
- Health
- Environment/Sustainability
- Law Enforcement/Legal Services
- Advocacy/Lobbying
- Design/Planning
- Arts
- Technology
- Faith-based
- Funding/Philanthropy
- Consulting/Strategy
- Collaborative
- Other: .....

**Organizational Data Usage**

All organizations use data to inform decisions and drive outcomes. How frequently does your organization use the following data types to inform decisions? (0=never, 1=annually, 2=quarterly, 3=monthly, 4=weekly, 5=daily)

13. **Anecdotal accounts from practitioners, clientele, or donors**

Mark only one oval.

	0	1	2	3	4	5	
Never	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Daily

14. **Articles and reports from National and State Governments**

Mark only one oval.

	0	1	2	3	4	5	
Never	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Daily

**15. Articles and reports from academic institutions**

*Mark only one oval.*

	0	1	2	3	4	5	
Never	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Daily

**16. Articles and reports from private companies or organizations**

*Mark only one oval.*

	0	1	2	3	4	5	
Never	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Daily

**17. Articles and reports generated from within my own organization**

*Mark only one oval.*

	0	1	2	3	4	5	
Never	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Daily

**18. Contact information for clientele/donors collected by my organizations**

*Mark only one oval.*

	0	1	2	3	4	5	
Never	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Daily

**19. Contact information for clientele/donors collected by a third party**

*Mark only one oval.*

	0	1	2	3	4	5	
Never	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Daily

**20. Demographic or personal information on clientele directly collected by my organization**

via surveys, applications, program evaluations

*Mark only one oval.*

	0	1	2	3	4	5	
Never	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Daily

21. **Demographic or personal data on donors directly collected by my organization \***

via surveys, applications, program evaluations

Mark only one oval.

	0	1	2	3	4	5	
Never	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Daily

22. **Physical measurements from the field collected by my organization**

Mark only one oval.

	0	1	2	3	4	5	
Never	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Daily

23. **Data collected by national and state government agencies (ex. U.S. Census data, TxDOT data, etc.)**

Mark only one oval.

	0	1	2	3	4	5	
Never	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Daily

24. **Data collected by local and regional government agencies (ex. public city data)**

Mark only one oval.

	0	1	2	3	4	5	
Never	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Daily

25. **Data collected by a partner organization**

Mark only one oval.

	0	1	2	3	4	5	
Never	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Daily

26. **Data purchased from a data vendor**

Mark only one oval.

	0	1	2	3	4	5	
Never	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Daily

**27. If applicable, which of the following data types have you purchased from a data vendor?**

Select all that apply.  
Check all that apply.

- Agriculture
- Demographic
- Education
- Energy/Environment
- Economic/Finance
- Health
- Jobs/Skills
- Organization
- Public Safety
- Real Estate/Development
- Transportation

**28. Are there any datasets that you would like to be see made available or more accessible? If so, what are they?**

.....

.....

.....

.....

.....

**29. How does your organization use data?**

Select all that apply  
Check all that apply.

- Resource allocation (staffing, funding)
- Understanding needs of constituency
- Project or program development
- Project or program evaluation
- Measuring and reporting impact (reports to funding sources)
- Lobbying and policy formation
- Donor development
- Internal operations
- Research
- None of the above
- Other: .....

## Capacity Barriers

Organizations may face barriers in their capacity to access or use data for specific purposes. Consider these barriers in regards to your organization.

### 30. What are some of the greatest challenges your organization faces in using data?

Select all that apply

Check all that apply.

- Not sure how data projects can further our mission
- Not sure where to find, or how to collect, the data we need
- Unable to gain access to the data we need
- We are unaware of a software/website/tool that meets our data needs
- Unable to afford the tools to make use of the data
- Staff lack the necessary technical skills
- Unable to dedicate the time to work with data given other demands
- Unable to sustain long-term data projects and partnerships
- We have not faced these barriers
- Other: .....

### 31. What skills does your organization currently desire, but struggles to develop, retain, or hire?

Select all that apply

Check all that apply.

- We are satisfied with our current capacity
- Basic computer literacy
- Data literacy (understanding how data can inform decisions)
- Basic excel/spreadsheets
- Database use and development
- Data cleaning and preparation
- Basic data analysis
- Advanced data analysis
- Mapping
- Other: .....

## Data Collaborations

**32. Does your organization PROVIDE any of the following services WITHIN A PARTNERSHIP? If so, which?**

*Check all that apply.*

- My organization provides funding
- My organization provides raw data collected internally
- My organization provides data analysis and/or visualization services
- My organization provides strategic input
- My organization provides strategic action
- My organization provides data-training and coaching
- My organization provides data-based planning or design services
- My organization DOES NOT provide these services
- Other: .....

**33. Does your organization RECEIVE any of the following services WITHIN A PARTNERSHIP? If so, which?**

*Check all that apply.*

- My organization receives funding
- My organization receives raw data collected by another organization
- My organization receives data analysis and/or visualization services
- My organization receives strategic input
- My organization receives strategic action
- My organization receives data-training and coaching
- My organization receives data-based planning or design services
- My organization DOES NOT receive these services
- Other: .....

**34. Does your organization PROVIDE any of the following services FOR A FEE? If so, which?**

*Check all that apply.*

- My organization provides funding
- My organization provides raw data collected internally
- My organization provides data analysis and/or visualization services
- My organization provides strategic input
- My organization provides strategic action
- My organization provides data-training and coaching
- My organization provides data-based planning or design services
- My organization DOES NOT provide these services
- Other: .....



35. Does your organization RECEIVE any of the following services FOR A FEE? If so, which?

Check all that apply.

- My organization receives funding
- My organization receives raw data collected internally
- My organization receives data analysis and/or visualization services
- My organization receives strategic input
- My organization receives strategic action
- My organization receives data-training and coaching
- My organization receives data-based planning or design services
- My organization DOES NOT receive these services
- Other: .....

36. Have you directly engaged with or partnered with any organizations, such as other non-profits, businesses, academic institutions, consulting firms, or government organizations for data-sharing, research, planning, or analysis? \*

Mark only one oval.

- Yes, currently
- Yes, in the past year
- In the past, but no longer
- No, but have tried
- No, never *After the last question in this section, stop filling out this form.*

37. Were any of these partnerships based on pro-bono work or reduced fees?

Mark only one oval.

- Yes, currently
- Yes, in the past year
- In the past, but no longer
- No, never

38. Please describe a notable data-based partnership occurring since 2008. What was the nature of the relationship? How did it form and what partners were involved?

.....

.....

.....

.....

.....

39. **If the partnership had a name, please share it with us.**

.....

40. **Thinking of the partnership described above, which of the following organization types were involved?**

*Check all that apply.*

- Non-Profit
- Business
- Academic Institution
- Government Agency (including public schools)
- Other: .....

41. **Please list one other organization with which your organization has strong informal connections.**

.....

### Collaboration Outcomes

42. **Thinking of the partnership described on the previous page, what benefits, if any, were experienced by the parties involved?**

*Check all that apply.*

- We learned more about other organizations
- Our organization built subsequent partnerships with organizations from WITHIN our sector
- Our organization built subsequent partnerships with organizations from OUTSIDE our sector
- Organizations aligned service delivery to target population
- Organizations received more funding through demonstration of need
- We sustained our collaboration through long-term investment
- Our organization developed greater capacity for data use
- Other: .....

43. **Thinking of the partnership described on the previous page, were there any unexpected outcomes of the partnership or unintended uses for the data gathered? This can include any lessons learned from the processes of data-collection, analysis, sharing and/or collaboration.**

*Mark only one oval.*

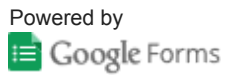
- Yes
- No

**44. Describe them.**

.....  
.....  
.....  
.....  
.....

**45. Thinking of the partnership described on the previous page, please also describe any barriers or obstacles to the partnership?**

.....  
.....  
.....  
.....  
.....



## APPENDIX B - DATA ECOSYSTEM RESPONSES

# DALLAS DATA ECOSYSTEM SURVEY

## INITIAL FINDINGS

bcWORKHSOP  
JANURARY 2015



# INTRODUCTION

The power of data has grown as quickly as the technological innovation that enables its analysis. This data revolution has been spearheaded by the work of high-tech companies, government agencies, academic institutions, and other organizations. However, change has been slow in some sectors. Nonprofit organizations, often under-funded and overwhelmed with need, do not have the funding or time to invest in the technology and staffing necessary for data analysis.

Many organizations across the United States have begun to fill this role, providing data and analysis to nonprofits, community organizations, or others that need information about their local communities. The role these organizations play within their local and regional “data ecosystems” vary. Organizations such as Data Driven Detroit, Data Haven, and the Baltimore Neighborhoods Indicators Alliance work to solve local problems with the use of data. They enable other organizations to easily access the data and information necessary to fulfill their missions. This network of community data organizations is so widespread that umbrella organizations such as the Community Indicators Consortium and the National Neighborhood Indicators Partnership have emerged to provide a place for these organizations to discuss areas of common interest.

The Smart Chicago Collaborative launched their Data Ecosystem Survey to understand and map Chicago’s complex and dynamic data ecosystem. This survey was meant to provide insight into the capacity, needs, and talents that can be connected to solve problems in Chicago and Cook County, IL. Much like Chicago, Dallas-Fort Worth has a large and complex data ecosystem.

Following the precedent set by Smart Chicago, buildingcommunityWORKSHOP, has launched a similar survey, in partnership with Communities Foundation of Texas, to gauge the needs and capacity of the DFW data community. We will continue to collect survey responses through January 2015. The following is a first look at responses through December 22, 2014

# ABOUT THE SURVEY

The survey gauges five main aspects of the data ecosystem. Each of the 5 sections focuses on particular areas of interest in understanding Dallas' data ecosystem.

- **Organization Information:** Organizations use data differently and have varying internal capacities for data usage. This group of questions provides a lens to not only understand the sample population but to explore other responses by organization type and roles within the community.
- **Organizational Data Usage:** The various types of data used in decision-making are likely to have various levels of use within organizations and types of organizations. This set of questions provides information on how often different types of data are used and what they are used for.
- **Capacity Barriers:** Some organizations cannot hire or retain the talent necessary to pursue advanced data analysis, however there may be a number of reasons organizations do not or cannot pursue these skillsets. Two questions attempt to gauge barriers to organizational capacity regarding data analysis.
- **Data Collaborations:** Organizations nationwide work through partnerships to advance data availability, data analysis, and community outcomes. Similar partnerships have formed in Dallas; these questions attempt to identify other partnerships that have occurred around data.
- **Collaboration Outcomes:** Partnerships and collaborations may not always work as intended. The two questions in this section are meant to provide insight into the outcomes of partnerships in Dallas-Fort Worth.

We distributed the survey through Google Forms to over 3,000 people. 2,700 organizations and individuals received the survey through Communities Foundation of Texas, with smaller distributions through the City of Dallas Information Services department, the Dallas Makerspace's Civic Hacking Committee, and other networks of local and regional data users.



# TAKEAWAYS

To date, we have had 229 total respondents to the Dallas Data Ecosystem Survey. Distribution through the CFT network accounts for more than 90% of possible responses – this is reflected in the dominance of responses from non-profit, religious, and educational organizations (91%). Information from outside the nonprofit community would also be useful for comparative purposes. We hypothesize, for example, that many nonprofits likely use more accessible data sources or anecdotal evidence to drive decisions more frequently than for-profit businesses or consultants. Ultimately, the dominance of nonprofit responders gives us a strong look at that sector uses data and where it faces barriers in improving analytical capabilities.

Additionally, executive staff members have submitted the majority of the responses (52%). These executive staff members, CEOs, Executive Directors, Directors, Presidents, and board members, provide important information about how data is used across an organization. More responses from the staff members that regularly use and make decisions from data would provide a wealth of new information for analysis and understanding.

We have been able to explore how data is being used across the ecosystem: the different types of barriers organizations face in using or expanding their data and analytical capacity, the types of data organizations want, and the way that organizations have worked with others for their data and analytic needs. Analysis of many questions is on-going, however myriad questions have emerged from the first month of survey data.

## who responded?

The largest number of responders were non-profit directors and CEO's (49%), followed by marketing /development staff (24%), operations/management staff (10%). Other responders included governmental employees, board members, technologists, reseachers, and consultants, and for-profit companies. 9 responders did not give a title.

Directors and CEOs 49%  
Communications/Development/Marketing 24%  
Operations/Management 10%  
Governmental employees 7%  
Board Members 3%  
Technology, Research, Consulting 3%  
No Response 9%

## what do they do?

The majority of responders worked within community/social services, education, and health fields. Other organization types represented include animal welfare groups, government agencies, and environmental organizations.

Community and Social Services 23%  
Education 21%  
Other 15%  
Health 14%  
Arts 12%  
Faith-based 7%  
Environment 7%  
Food 4%

# how is data used?

So how does Dallas use data? We're not trying to find out which specific analytical techniques are being used, but rather what the ultimate goal is when data is employed. For the most part, almost 70% of organizations identified that they use data in five specific ways:

- Project of Program Development 79%
- Understanding Needs of the constituency 73%
- Measuring and reporting impact to funding sources 72%
- Donor Development 71%
- Project or Program evaluation 69%

Four other response categories were provided which had far fewer responses overall. It seems that most organizations that use data do so to help provide more efficient services (Project or program development, Understanding needs of constituency, and Project or program evaluation) and to help report their impact to current and potential donors (Measuring and reporting impact and Donor development). Internal Operations (57%), Resource Allocation (52%), Research (32%), and Lobbying / Policy Formation (18%) have far fewer overall responses, although many organizations do still use data to help with resource allocation and internal operations.

# what are the challenges?

What are some of the greatest challenges an organization faces in using data? Organizations in Dallas that use data have many different barriers to achieving their goals with data. Many organizations have too much on their plate to really jump into data analysis or they simply cannot afford the appropriate tools they need, in addition to a whole subset of other challenges they must overcome.

Unable to dedicate the time to work with data given other demands  
127 responses (57%)

Unable to afford the tools to make use of the data  
107 responses (48%)

Staff lacks the necessary technical skills  
71 responses (32%)

Unable to gain access to the data we need  
65 responses (29%)

Not sure where to find, or how to collect, the data we need  
60 responses (27%)

We are unaware of a software, website, or tool that meets our data needs  
48 responses (21%)

Unable to sustain long-term data projects and partnerships  
43 responses (19%)

Not sure how data projects can further our mission  
31 responses (14%)

We have not faced these barriers  
21 responses (9%)

# what skills are lacking?

Expectedly, organizations face many different barriers in using data from a personnel standpoint. Organizations find it difficult to perform advanced data analysis, work with databases, and to an extent have difficulty in understanding how data can further their impact. Results for this section do not provide as much insight into the data needs of the Dallas – Fort Worth Metroplex, but they do show that there is likely no one method of helping organizations improve their capacity to use data strategically.

Advanced data analysis

111 responses (51%)

Database use and development

94 responses (43%)

Data Literacy

72 responses (33%)

Data cleaning and preparation

71 responses (33%)

Mapping

58 responses (27%)

Basic data analysis

57 responses (26%)

We are satisfied with our current capacity

39 responses (18%)

Basic Excel/spreadsheets

26 responses (12%)

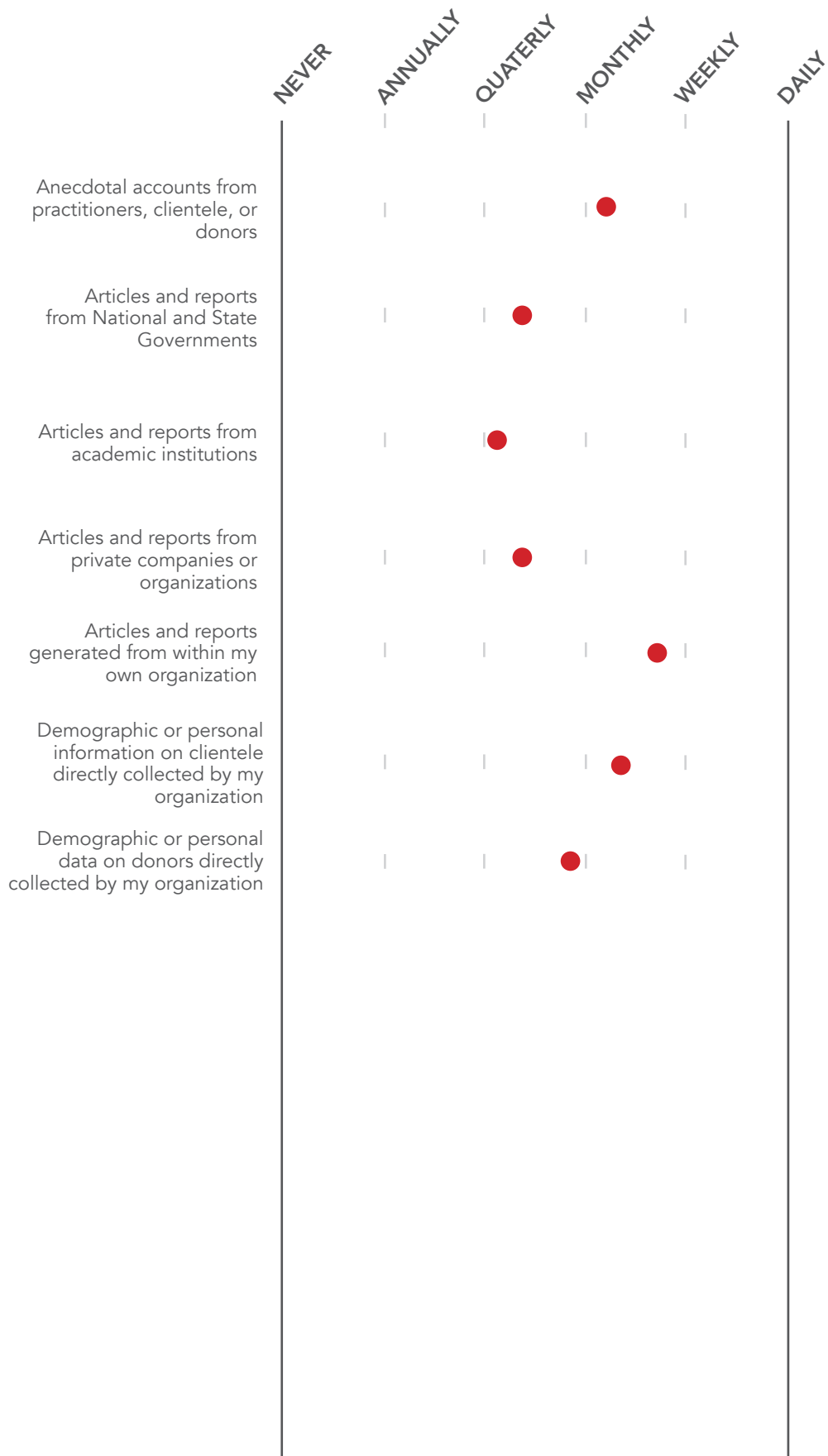
Basic computer literacy

15 responses (7%)

# how often is data used?

*Q: How frequently does your organization use the following data types to inform decisions?*

The most common data type used for strategy and practice are articles and reports generated from an organization internally, and are used almost weekly on average. This is followed by demographic and personal information on clientele directly collected by the organizations, and anecdotal accounts from practitioners, clientele, or donors. This indicates that many organizations are generally effective at using internal data to make decisions. Data types that are used far less frequently are public data sources, reports from academic institutions, data purchased from a data vendor, and data collected from a partner organization. This highlights an opportunity for organizations to utilize multi-source datasets. This is especially true for organizations that share a common agenda and work in similar service areas.



# what data is needed?

*Q: Are there any datasets that you would like to see made available or more accessible? If so, what are they?*

Data users in the region would like to see a wide variety! Some of the 51 responses to this question asked for datasets that do exist – Given the small response may be an indication that building capacity for data analysis among non-profits would be increasing the awareness of existing resources and tools.

**So what data is needed? Here are a few of the responses:**

- Detailed aerial imagery
- County level versus city-wide Middle School and High School statistics
- Service providers at campuses (resource/asset mapping)
- Email addresses for school administrators
- Regional public transit network shapefiles
- Where people live related to their jobs and how they commute
- Number of children in DISD below poverty level and schools they attend
- Student data from DISD and surrounding school districts
- Datasets on success of at-risk students in earning a college degree vs. access to college
- Nonprofit lists with contact names and e-mails

The availability of many of these datasets is dependent on money, time, relationships, and effort. While some datasets do exist as downloads on a website or through a governments open data portal, many datasets might only be available through existing relationships or formal data agreements. Otherwise, for datasets that do not exist, compiling and creation of the dataset may be required to help organizations gain access to the needed data. Steps can be taken to identify which data sets are most crucial to the most organizations.



# who's partnering?

*Q: Have you directly engaged with or partnered with any organizations, such as other nonprofits, businesses, academic institutions, consulting firms, or government organizations for data sharing, research, planning, or analysis?*

A majority of survey responses have reported current or past engagement in data-centered partnerships (59%). By contrast, roughly a third of the responding organizations have never engaged in this way.

In the past but no longer

18 responses (8%)

No, but have tried

15 responses (7%)

No, never

79 responses (34%)

Yes, currently

75 responses (33%)

Yes, in the past year

42 responses (18%)

## APPENDIX C - FOCUS GROUP 1 SUMMARY



# bcDATA Focus Group 1

MARCH 2015

#### PREPARED BY

The buildingcommunityWORKSHOP is a Dallas based nonprofit community design center seeking to improve the livability and viability of communities through the practice of thoughtful design and making. We enrich the lives of citizens by bringing design thinking to areas of our city where resources are most scarce. To do so, the [bc] recognizes that it must first understand the social, economic, and environmental issues facing a community before beginning work.

#### SUPPORTED BY

As the largest community foundation in Texas and one of the largest in the nation, Communities Foundation of Texas (CFT) works with families, companies and nonprofits to strengthen our community through a variety of charitable funds and strategic grantmaking initiatives. The foundation professionally manages more than 900 charitable funds and has awarded more than \$1.3 billion in grants since its founding in 1953. Increasing financial stability of working families is one of the two key focus areas of CFT's community impact funds. To support this area, CFT as launched the Data Driven Decision-Making (D3) Institute. The D3 Institute is designed to provide organizations that offer programs and services for low-income working families the power to accelerate their development of enduring solutions to the social and economic problems facing this population. [www.cftexas.org/D3](http://www.cftexas.org/D3)

# EXECUTIVE SUMMARY

buildingcommunityWORKSHOP has been funded by Communities Foundation of Texas to research national best practices that enhance access to reliable, intelligible data and analysis to drive strategic decision-making across the community; to investigate the current status of data use across Dallas; and to develop a model to address the needs of Dallas. As a part of this work, [bc] has convened a small group of stakeholders from Dallas' nonprofit data community to discuss challenges to data use in the nonprofit community and to strategize how national best practices might be leveraged in Dallas.

This report documents the first of these discussion, Focus Group 1, held at CFT in February 2015. This session was primarily focused on identifying the biggest and most pressing challenges to data use in Dallas' nonprofit community as seen by the 11 stakeholders who attended the session. As the first of two focus groups, the discussion is a primary driver of the activities and discussions for the second gathering in March 2015.



# INTRODUCTION

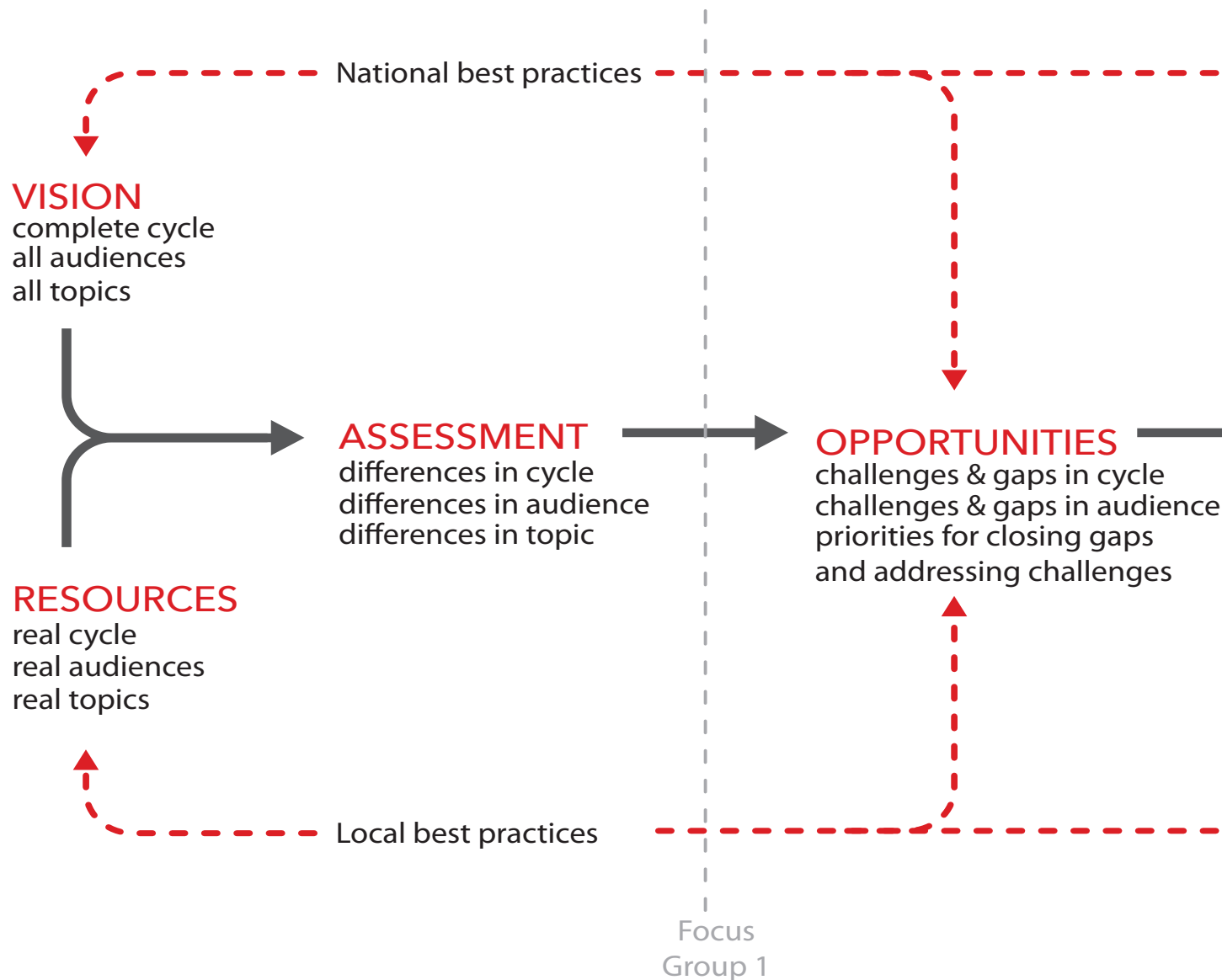
The City of Dallas faces a myriad of broad challenges as it moves through the second decade of the 21st Century. Affordable housing, rapid growth, aging infrastructure, poverty alleviation, children's health, and urban blight are all just some of the many complex interconnected problems that require an unprecedented amount of long-term motivation, cross-sector coordination, and common agenda development to address. Cities across the world have recognized that data is central to both understanding and acting on these complex issues and that the importance of making data accessible to a greater population is key to reaching these goals. A number of different strategies for achieving these goals have been adopted across the country, with each offering unique strengths and benefits.

[bc] has prepared several case studies of organizations across the country to understand how different collaboratives and data-centered organizations function. This includes a documentation of their products, funding structures, governance structures, and audience engagement strategies. However, identifying the approaches and methods necessary and plausible in Dallas requires input from a local audience of stakeholders working within Dallas. To do so, [bc] has convened a group of nonprofit data users to discuss the role of data in Dallas, challenges these stakeholders know to exist, and ways these problems can be addressed that make sense in Dallas.

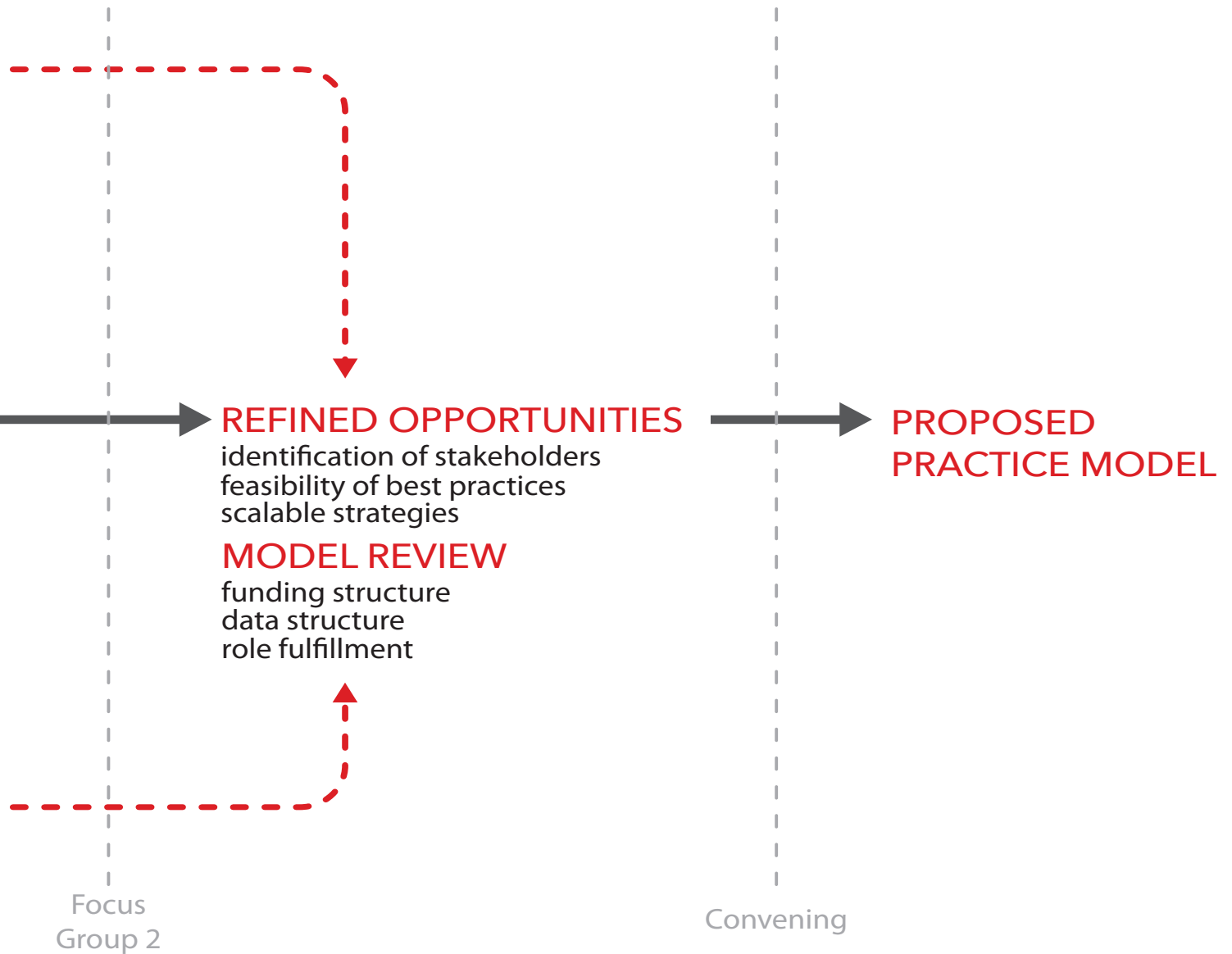
The following page lays out the framework of [bc]'s work plan, and where this first Focus Group fits within that framework.

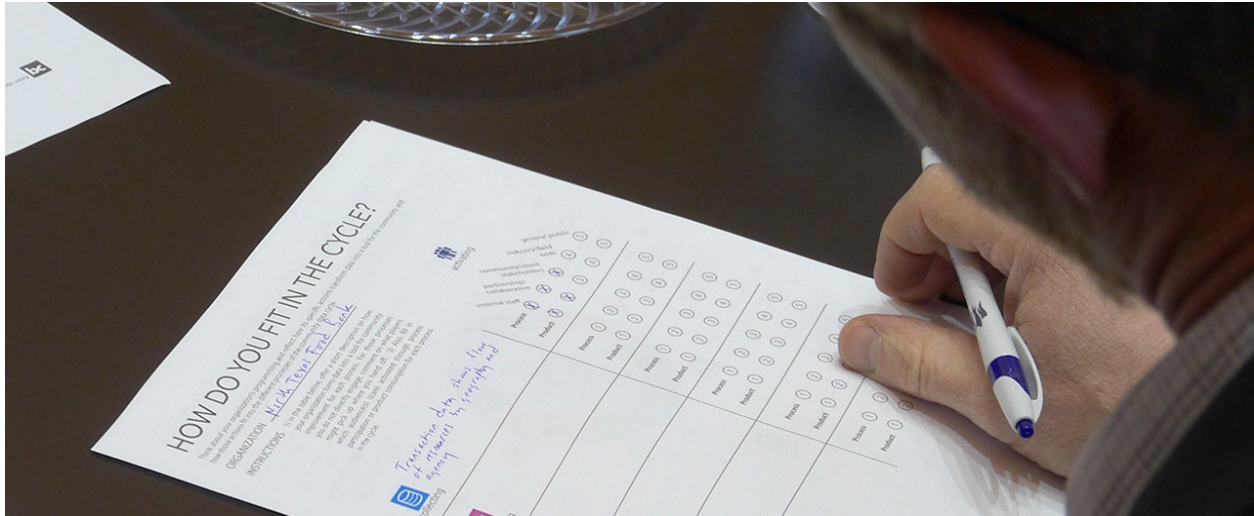
# A FRAMEWORK FOR BUILDING STRATEGY

The first step in the framework for building strategy is defining the VISION for ideal or effective community data practice. Comparing this ideal vision to the reality of existing local RESOURCES is the first step to identifying OPPORTUNITIES. After this initial comparison, feedback from the local data-driven community refines this understanding of needs and opportunities in Dallas and allows the prioritization of best practices to implement immediately and the development of a long term strategy for comprehensive, equitable, and efficient community data practice.









# FOCUS GROUP 1

## GOALS & AGENDA

1. Share and discuss the vision for how different roles are filled within the system or cycle of informing audiences and decisionmakers with data locally.
2. Review Ecosystem Survey results as a tool to identify potential gaps and challenges to the broader community.
3. Identify challenges and gaps in the current data ecosystem to moving data toward action and collaborative decision-making.

## FACILITATORS

Owen Wilson-Chavez, Research Analyst, [bc]  
Ryan Williams, Research Analyst, [bc]

# PARTICIPANTS

NAME	POSITION	ORGANIZATION
<b>Anthony Galvan</b>	Senior Research Associate	Institute for Urban Policy Research - UTD
<b>Ashwina Kirplani</b>	Director of Analytics	Commit!
<b>Bill Finch</b>	Chief Information Officer	City of Dallas
<b>Corron Saunders</b>	Population Health Strategies Researcher	Children's Health
<b>Galen Smith</b>	Director of Community Financial Stability	United Way of Metropolitan Dallas
<b>Girish Ramachadran</b>	Chief IT Architect	City of Dallas
<b>Jane Massey</b>	Director of Neighborhood Research and Revitalization	Habitat for Humanity
<b>Junior Ortiz</b>	Executive Director	Economic Partners Investing in Communities
<b>Katie McConnell</b>	Chief Operations Officer and Director for Social Measurement and Evaluation	Children at Risk
<b>Richard Amory</b>	Director of Research	North Texas Food Bank
<b>Robert Munding</b>	Data and Analytics Strategy Associate	Commit!

# THE DATA CYCLE



## GOAL 1

Share and discuss the vision for how different roles are filled within the system or cycle of informing audiences and decision-makers with data locally.

Share and discuss the vision for how different roles are filled within the system or cycle of informing audiences and decision-makers with data locally.

[bc] introduced a “data cycle” model as tool or framework to discuss data work in the community. The cycle was synthesized from a number of models to understand how data moves through the community decision-making process. The cycle distilled the work of data-driven decision making into themes of work including collection, analysis, visualizing, informing, storytelling, and acting. This attempt was meant to prompt a critical conversation at how each of those work themes shared data with different audiences. Participants gave their reaction to the data cycle as a tool and discussed and contextualized their data work within the cycle. \*

The data cycle model did not capture all aspects that focus group participants felt had a place in the discussion of data use. However, the cycle did facilitate the development of a common language around data use. [bc] is reconfiguring the model to increase its usefulness as a discussion tool and as an instrument to represent the many roles that data use may take on within an organization and broader data ecosystem.

Key conversations about present and missing aspects of the data cycle are featured on the following pages.

\*Participants also engaged in an activity piloting a data collection tool designed by [bc] to measure data use.

# How should data inform the community?

Critically examining the data cycle as a group prompted a number of open-ended responses and follow-up questions. Here are the primary discussion points:

---

## CLEARLY DEFINING THE INTENTION AND PURPOSE OF DATA

There was a clear concern that all data collection and sharing have a very clear purpose with either a research question or tool in mind. It is necessary to document why datasets are necessary (or necessary in the future) for a project and make clear and be aware of the original intention of the data.

---

## ENGAGING WIDE AUDIENCES

Focus group participants believed that data users should be very aware of what audiences they were activating through storytelling with data and putting a great emphasis on making sure the resultant stories around data be compelling, relevant, and directly related to individual citizen experiences.

---

## COLLABORATION AND SHARING

There was also an interest in the concepts of data sharing, data use limitations, and cost-sharing with regards to data initiatives for community improvement. Especially the questions: "What can be shared?," "What should we focus on?" and "How should the burden of costs be allocated?"

*"...Common questions should be included in our model...the data has to support some research question or some type of reason why we are collecting it. It should be supportive of some type of purpose."*

*"I think we also have to be very cognizant.... to identify the questions for the future. Trying to figure out not only what the research questions are, but what are they going to be five, ten years for now."*

*"Why is data collected?... there's a lot of...process data, transaction data, that's substantially different from primary research data"*

*"Before you [implement a data initiative]. Pilot it. Ask the question: 'Is this even going to benefit you?...and if it's not tell us what is so we can focus our energy."*

*"Making sure we are identifying the audience to drive our questions... engaging [everyday citizens] along the way in our practice"*

*"...Looking at it from the [individual citizen] standpoint. What's the impact to them?"*

*"One of the things we can start doing is speaking in terms of how data-sharing can benefit not just the city but the individual child, the individual community center..."*

*"We have to put [data] in a language that [non-experts] understand."*

*"How do you engage the community and let them tell the story? What kind of tools can we put in the [citizen] community?"*

*"We haven't made an effort or found a way to share information or communicate information directly to the end recipients of the services we are a part of..."*

*"...If we look at it as a [data] community, how do we share the cost?"*

*"The reality is even though we're focusing on gaps today, there's also a lot of duplication happening. This [data] cycle is happening in parallel with a lot of organizations collecting the exact same data. Because it's difficult at times to get organizations focused on the same things, to come together and collect this exact same data."*

*"We have a very protective culture around data limits and sharing."*

*"How do we share data?...It can be a very closed system. There is data flowing around, some of it is closely guarded and private, some of it is shared to some degree, and some of it is openly shared.... That's what it comes down to for me.... what can you share?"*

## GOAL 2

# Review Data Ecosystem Survey results as a tool to identify potential gaps and challenges to the broader community.

[bc], in partnership with CFT, launched a community-wide survey to gauge how data is used in Dallas, and where data use faces barriers and challenges within organizations. Preliminary results were shared with the focus group, preceding a discussion about the results and their implication on data use in the nonprofit community. Focus group members were asked how the results compare to their personal experiences within Dallas' data ecosystem.

A few common themes emerged in discussions around the survey results:

---

### NEED FOR A PROSPECTIVE CATALOGUE

Many focus groups members described the clear need for a comprehensive catalogue of public information and regional data. This way it can serve as the one-stop reference for organizations seeking relevant data and information about its use.

---

### LACK OF COLLABORATION AND SHARING

Participants voiced their frustration of the duplication of data work and the barriers to collaboration may be highlighted in the survey results.

---

### SKILLS GAP

Participants voiced their concern in organizations reporting a greater capacity for data use and analysis than they actually have. Some also commented on how evaluation and impact measurement was very limited locally.



*"Sometimes we don't know the data exists... I would love to have a catalogue... [that shows where] demographic data is available and free of charge... if people are purchasing data that's free, sometimes it's because they are not aware that it's free."*

*"We need a common place for [local] information."*

*"The number one response to challenges, was people not having time, it takes time to go find these things, it takes time to go send an email to a colleague or a co worker or peer at another organization and say hey do you know where I can find that data... they're busy, they might not respond."*

*"[the survey] replicates a lot of what we experience. As we're talking and we're saying we need to work together... not have that kind of competitive edge, I think [sharing this survey] is a really good solution to overcome some of those hurdles."*

*"The reality is nonprofits require funding to keep their doors open, and provide those services, and that requires they submit applications to funders to say all of these great things are happening because of our work and that's just the reality right now until we find some sort of collaborative funding models that allow for... partnerships to more organically occur without any sort of penalty..."*

*"A lot of [local data] is free...if you have technical skill you don't need any money. I don't like the mindset where we say we don't have enough money or we can't afford this."*

*"93% have the skills to use the \$5,000 solution but don't have the skills to use the free solution."*

## GOAL 3a

Identify challenges and gaps in the current data ecosystem to moving data toward action and collaborative decision-making.

The third session was an open brainstorm of challenges and gaps to the existing data ecosystem. The concerns of the focus group centered around 9 general themes:

---

### TRUST

Focus Group members were concerned that there is a culture of mistrust between producers, consumers, and audiences of data. Data producers fear misuse and misrepresentation of the data, data users fear competition for funds and service delivery, the public fears invasions of privacy, and funders fear waste or duplication of effort.

---

### COMMUNITY ENGAGEMENT

Focus groups members expressed concern that data work be directly related to improving lives of the public and that communities had a stronger grasp on what information sharing can do for public benefit.

---

### FUNDING/COST SHARING

Participants cited that silos and redundancies of data work may be rooted in fundraising demands.

---

### STORYTELLING

Focus group members expressed the importance of accurate and engaging storytelling, which requires different skill sets apart from data management and analysis.

---

### PLANNING

Nearly all focus group members felt it was important that work done in response to data and analysis be meaningful and constructive and actively address questions or needs posed at the beginning of an initiative or innovative program.

---

### COLLECTION

The focus group considered standardized definitions, coordinated collection activities, and relevance of data as weaknesses in Dallas' data practice. This may be a result of gaps in collective impact initiatives, or a general lack of collaboration.

---

### RESOURCE AWARENESS

There was a near constant reference to the need for a centralized data catalogue where resources were made available to all audiences and fully explained in their entirety. This may reduce time and money wasted seeking data and misrepresenting information.

---

### TECHNICAL CONSIDERATIONS

There were also a number of concerns around the standardization, scale, history, purpose, privacy, and confidentiality of data. Not all data can be shared, and data that is shared should be useful and have very clear purposes and limitations for use.

---

### SKILL CONSIDERATIONS

Many focus groups members highlighted the general need for a greater capacity among the non-profit community in Dallas and the necessity for issue and business specific expertise in the use and proliferation of related data.

# challenges and gaps

raw responses

"Lack of trust among perceived competitors"

"Trust between producers and consumers of data"

"Willingness for law enforcement and other data producers to engage"

"Data impact silos caused by fundraising demands"

"Cannabilizing research staff for program needs"

"Centering data-use around needs of clients and consumers of social services"

"Engaging the general population throughout process activities"

"Data work should be aware of audiences... Aware audiences should understand datawork"

"General (lack of) understanding of power of data"

"Many organizations trying to change behavior - community gets overwhelmed"

"visualizing for different audiences"

"visualizing - different skill set needed"

"graphic support to tell stories"

"telling accurate stories"

"storytelling & informing audiences"

"storytelling is a key way to share & needs unique skill sets"

"Translating data into action. Knowing what to act on... the 'so what?'"

"Motivation, Coordination  
Reacting decisions need to have plan of action after story telling"

"Reacting should be based on questions and goals established prior to collecting"

"All parts beyond visualization due to gaps in collective impact infrastructure."

"Lack of shared client data systems limits potential for service coordination"

"Inconsistent definitions of indicators, services, clients"

"Demand for collective impact work needed to align data collection efforts"

"Collecting relevant data"

"Centralized data catalogue"

"lack of knowledge about available resources"

"Duplication and redundancy of data"

"Resource availability (breadth of skills)"

"Business knowledge"

"Capacity"

"...evaluation activities (esp. in nonprofits) are still pretty limited/ and rudimentary"

"issue sector specific expertise"

"What can be shared?"

"Not all data can be shared"

"Data confidentiality"

"Accessing past - 'old' data sets"

"use of 'unique' geographies, zip codes, attendance zones, etc."

## GOAL 3b

# How should the challenges and gaps be prioritized?

After brainstorming general challenges and gaps to the data ecosystem, focus group participants were asked to come to a consensus around which of the challenges to prioritize for any future plans for data initiatives and rank them in terms of urgency and importance. Here are the top challenges in order:

- 1 DATA AWARENESS & ACCESS**

Often non-profits /individuals can't find the data they are looking for, don't have the time it takes to find it, or the technical skills to put it in the form they need it in. Building on strategies to make data more accessible through awareness, outreach, and technical support will allow organizations to more easily access what they need to drive decisions, write grants, and track progress in a community.
- 2 DATA PRIORITIZATION**

Making sure that the data used in the non-profit sector is the right data is key. Strategies should be pursued to ensure that stakeholders are aware of the purpose for the data collection/assembly, the questions they are seeking to answer, and the correct interpretation of the data used.
- 3 STORYTELLING & COMMUNITY ENGAGEMENT**

Organizations that do use data work need more support for telling stories with data and building trust and capacity within the broader community around data work. Strategies that employ multiple sectors in using graphical representation, narratives, and accurate storytelling should build the both the usefulness and impact of data initiatives.
- 4 FUNDING/PARTNERSHIPS**

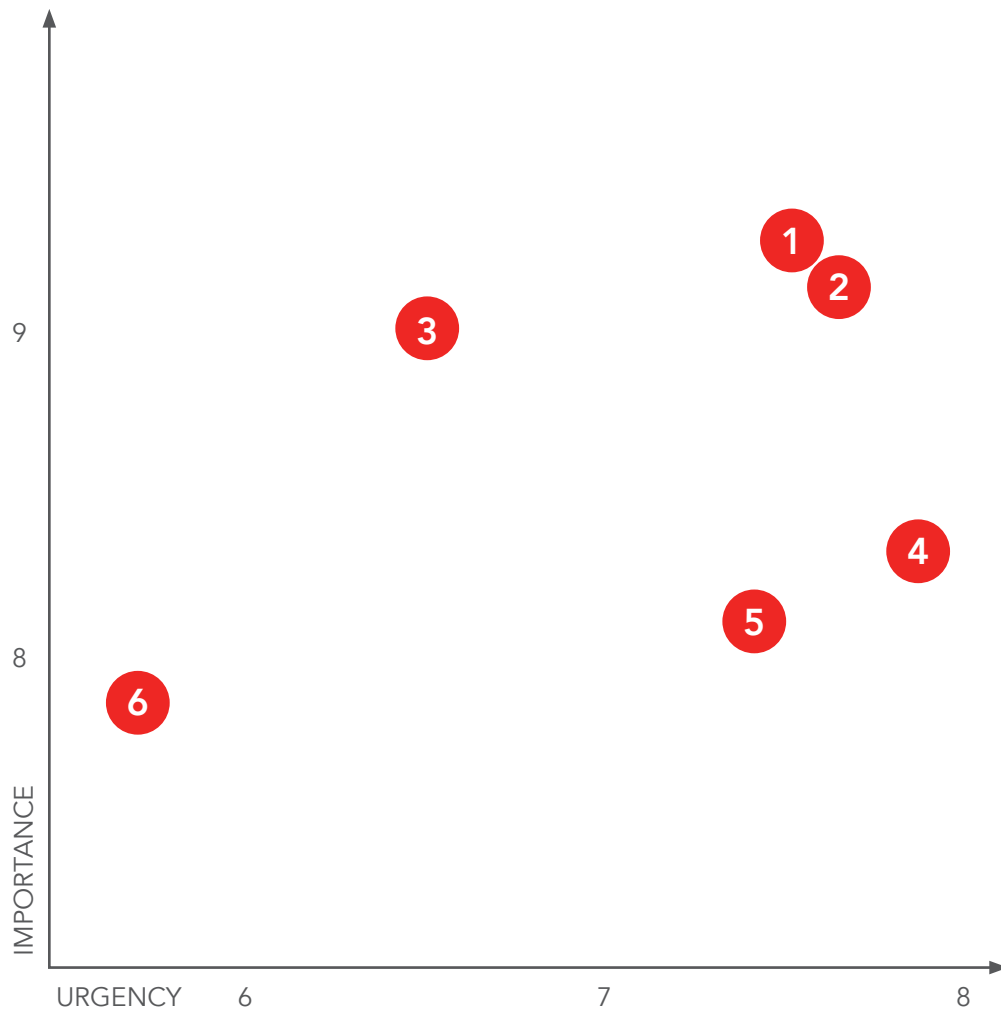
Focus group participants were very aware of the reality of how data initiatives and impact measurements are often duplicated due to the systems of funding and a general lack of collaboration and coordination. Funding strategies that allow for partnerships and collaborative to develop organically are necessary to support the work.
- 5 DATA STANDARDIZATION**

Related to the prioritization of data, is the need for data and definitions around data employment to be clearly defined and standardized for aggregation and multiple-inquiries.
- 6 "SO WHAT?"/ACTING STRATEGICALLY**

Related to the funding piece is the general challenge of building capacity within the non-profit sector to understand data's usefulness and employ data initiatives that help organizations strategize with data fluently.

# PRIORITIZATION MATRIX

Participants were asked to rank priorities on a scale of 1 to 10 with 1 being “less important and less urgent” and 10 being “more important and more urgent.” above On average respondents ranked all priorities above 5 for both categories.



## Dallas must:

There is clear indication the the data community in Dallas needs to create order out of the chaos in both data-use and capacity. The issue of trust, which should be isolated to the protection of personal case/transactional data, has expanded to affect the city's basic needs assessment and interpretation of neighborhood level aggregate data. This is due in part to the city's "protective culture" and general lack of trust within the research community and among non-profits competing for the same funds to acheive similar missions.

### **The Practice model in response must:**

Make public data easily available, accessible, and the public aware of its existance.

Have an infrastructure that drives collaboration without fear of competetive penalties from direct performance comparison.

Develop a research community that shares data, best practices and mutual goals within and across sectors.

Build a culture of excellence, trust, and inclusiveness through transparent disclosure of funding and activities for public good.

Be responsive to the needs of the communities by engaging place and issue specific expertise in the development of data initiatives and use of sensitive information.

Have clear strategy that effectively communicates the resources available and local best use cases to improve community confiditions using data.

Improve skills and capacities of organizations and individuals to use and share available data.







## CONCLUSIONS AND NEXT STEPS

Dallas faces a number of real world challenges that can only be solved by smart, coordinated, and efficient action. Effective data sharing, collaboration, and sharing of resources is the only way this can occur. As costs rise and access to resources become more competitive, challenges to foster a collective culture in Dallas grow. Like many nonprofit sectors, the data community suffers from a over protective culture fraught with misperceived competition, duplication of effort, and general lack of trust. The discussion around the data ecosystem survey revealed the need for a local shared public data resource center, the frustration of work duplication due to fund-raising cycles, and a misrepresentation of skill and capacity.

The discussion and activities of Focus Group 1 were aimed at pin-pointing these challenges and gaps within Dallas' data ecosystem - the places where improvements or new ideas may be needed to help organizations more readily utilize data to enhance strategic decision-making. The 6 challenges identified on the previous page will provide the framework for identifying different means of addressing those problems, either by leveraging work already undertaken in Dallas or by replicating work done in cities across the country.

Various strategies and programs identified by [bc] will be shared with Focus Group 2 participants prior to the second gathering and will form the backbone of a more strategic discussion centered on understanding the feasibility and viability of national best practices in the context of Dallas. The discussions held in Focus Group 2 will play an important role in developing a model for enhanced community-driven data work in Dallas.

## APPENDIX D - FOCUS GROUP 2 SUMMARY

### *Overview*

In early 2015, a small group of stakeholders in Dallas' data community was assembled by buildingcommunityWORKSHOP to identify barriers in advancing data use within the city. The first meeting, Focus Group 1, was primarily focused on assessing the types of barriers and constraints identified in [bc]'s Dallas Data Ecosystem Survey, however at the end of the meeting participants were engaged in an activity to identify and rank the primary challenges and gaps within the data ecosystem. The six challenges identified by the group formed the foundation of Focus Group 2, where [bc] presented a number of solutions to Dallas' primary challenges employed by organizations across the United States.

However, Focus Group 2 was primarily aimed at identifying three types of solutions to Dallas' data problems: 1) solutions entirely new to Dallas; 2) solutions that exist in Dallas but may need some adjustment to become more effective; and 3) successful solutions that exist in Dallas but could be scaled up or expanded to cover additional topics or geographies. The group was tasked with brainstorming solutions, however wild or seemingly unrealistic, in order to launch a discussion and understanding of ways that organizations working in the city can build a more robust community of data users and practitioners. Participants were also asked to identify which challenges, identified in Focus Group 1, each proposed solution could address.

Out of this exercise, a number of key takeaways emerged:

- Dallas could benefit from data-focused "user groups" as a way to highlight problems and concerns within the data ecosystem, accelerate data use and sharing, and to pressure decision-makers and data holders to develop more robust data-specific policies and strategies. These user groups could increase trust across organizations, and help drive awareness to problems faced by a particular sector.
- Identifying, building, and engaging the social innovation/civic technology community is a vital step in addressing a variety of problems in new and creative ways.
- Many participants identified the successes of collective impact organizations, such as Commit!, in rallying organizations and activity around specific topics. This approach could create incentives for more uniform collection and sharing of data on any number of topic areas.
- Finally, the idea of using data and analysis to link back to specific policy changes was a common theme. This focus on policy also directly relates to the needs of the data community – a more informed data policy effort is necessary in order to increase the amount of data collected and shared by a variety of organizations.

### ***Outcomes***

From this Focus Group, [bc] began to distill the overall research process into a statement of how Dallas' data ecosystem should function. The following principles have been used to further explore the various ways that Dallas' data community can bring change to current status quo of data collection, provision, and analysis occurring in organizations across the city.

Dallas' data ecosystem should:

- Organize public data to be easily available, accessible, and the public aware of its existence.
- Have an infrastructure that drives collaboration without fear of competitive penalties from direct performance comparison.
- Build a culture of excellence, trust, and inclusiveness through open sharing of best practices, data, and strategies for public good.
- Engage place and issue specific expertise in the development of research, data initiatives, and use of sensitive information to be fully responsive to the specific needs of the communities.
- Improve skills and capacities of organizations and individuals to use and share available data.
- Engage and participate in the local the technology sector to accelerate and test civic innovations to keep pace with the rapid clip of technological innovation.
- Engage the policy sector to be fully informed and saturated by data and research derived from it.

### ***Next Steps***

Following Focus Groups 1 and 2, [bc] has begun working on a more organized model for bringing about changes within Dallas' data ecosystem. This work has tentatively taken the form of a business plan for a solitary organization, however this is not the desired end result. A number of existing organizations are currently forging their own projects and services to help bring about change within the data ecosystem. [bc]'s approach has been to identify the primary functions a solitary organization would have, following the model of organizations in cities across the country.

[bc] plans to present this approach to focus group participants and other stakeholders in Fall 2015, in an attempt to refine the approach and identify organizations or individuals interested in advancing each of the functions identified. With more engaged stakeholder participation, [bc] seeks to unveil the group's work at a convening early in 2016.

## APPENDIX E - LAB 1 SUMMARY



## DATA ECOSYSTEM PROJECT LAB 1 OVERVIEW AND NEXT STEPS

bcANALYTICS  
FEBRUARY 2016

#### PREPARED BY

The buildingcommunityWORKSHOP ([bc]) is a Dallas based nonprofit community design center seeking to improve the livability and viability of communities through the practice of thoughtful design and making. We enrich the lives of citizens by bringing design thinking to areas of our city where resources are most scarce. To do so, [bc] recognizes that it must first understand the social, economic, and environmental issues facing a community before beginning work.

#### SUPPORTED BY

As the largest community foundation in Texas and one of the largest in the nation, Communities Foundation of Texas (CFT) works with families, companies and nonprofits to strengthen our community through a variety of charitable funds and strategic grantmaking initiatives. The foundation professionally manages more than 900 charitable funds and has awarded more than \$1.3 billion in grants since its founding in 1953. Increasing financial stability of working families is one of the two key focus areas of CFT's community impact funds. To support this area, CFT has launched the Data Driven Decision-Making (D3) Institute. The D3 Institute is designed to provide organizations that offer programs and services for low-income working families the power to accelerate their development of enduring solutions to the social and economic problems facing this population. [www.cftexas.org/D3](http://www.cftexas.org/D3)

# DATA ECOSYSTEM PROJECT

Access to information has changed dramatically over the last two decades. Data about our country, states, cities, and communities is being collected at an unprecedented scale. These datasets, and the tools built from them, are becoming more and more accessible to a wide range of audiences. Today, organizations, government jurisdictions, and policy or community advocates seek to make informed decisions based on the wealth of data and information that flows out of our communities.

Despite this increased demand for local information, data in Dallas is often not easily accessible, collected infrequently or not at all, in usable formats, outdated, or exceedingly costly. Meeting today's data needs will require an overhaul of the status quo.

The Data Ecosystem Project, initiated by Communities Foundation of Texas and supported by buildingcommunityWORKSHOP ([bc]), has begun a process to learn, understand, and assess the gaps in the current system. Working with local organizations and professionals, the Data Ecosystem Project seeks to identify strategies that build a more complete data ecosystem. The outcome of which should be a more robust open data environment that serves the needs of the city, local organizations, communities, and residents working to build a better region.

To date, The Data Ecosystem Project has involved three key learning and engagement activities:

- **National Best Practices** - A research and interview process identifying best practices in organizations across the country, including short case studies of select organizations.
- **Dallas Data Ecosystem Survey** - A survey gauging the data and analytics capacity of organizations working across North Texas.
- **Focus Groups** - Two stakeholder-driven focus groups aimed at identifying challenges within and solutions for Dallas' data ecosystem.



## GOALS

On January 25, 2016, [bc] and CFT convened a number of key stakeholders from Dallas' Data Ecosystem to discuss the current state of the local data ecosystem and brainstorm ideas to transform it for the better. Lab 1 identified solutions that can be built upon to transform Dallas' data ecosystem. These activities will form the basis of a report issued by [bc] and can be used to help demonstrate their usefulness/urgency to funders, local governments, philanthropists, and the business community.

- Discuss the current state of Dallas' data ecosystem
- Evaluate [bc]'s 6 Key Functions of the Data Ecosystem
- Assess missing elements from the 6 Key Functions
- Brainstorm ideas to positively transform Dallas' data ecosystem
- Identify missing voices to include in Lab 2



# DATA ECOSYSTEM PROJECT LAB 1

Through a series of focus groups, one-on-one conversations, discussions with peer organizations, and a survey administered primarily to nonprofit organizations, [bc] began mapping out how organizations are interacting with data. A vibrant data ecosystem engages a collection of decision-makers, researchers, analysts, civic technologists, service providers, and students who rely on data to serve their communities, drive program development, inform constituents, or conduct cutting-edge research. In the ideal setting, [bc] argues that a data ecosystem should:

- Organize public data to be easily available and accessible to all audiences.
- Create an infrastructure that drives collaboration, and minimizes the need for direct competition.
- Build a culture of excellence, trust, and inclusiveness through open sharing of best practices, data, and strategies for public good.
- Engage place and issue specific expertise in the development of research, data initiatives, and use of sensitive information to be fully responsive to the specific needs of the communities.
- Improve the skills and capacities of organizations and individuals to both use and share data.
- Engage and participate in the local technology sector to accelerate and test civic innovations.
- Engage the policy sector to be fully informed by data and research and to be responsive to the needs of the data community.

From a parcel-by-parcel mapping project in Detroit to Code For America's fellowship program, initiatives across the country are working to transform how governments interact and serve their constituents through the use of data and technology. Lab 1 is a step towards making these changes in Dallas.



## 6 FUNCTIONS OF THE DATA ECOSYSTEM

To begin the day, [bc] shared the origins of the project with the group as a preface to a presentation focused on the 6 Functions of the Data Ecosystem. Through previous work, the Data Ecosystem Project has identified 6 Functions key to a healthy data ecosystem. Each of these functions is present, at varying capacities, in cities across the country. [bc] feels that improving the data ecosystem requires renewed attention in these 6 functions, which represent a framework for discussing a variety of related activities.

As defined by [bc], each function either feeds into or is fed by another function. For example, an Open Data Catalog provides data for consultants and researchers who in turn provide evidence to drive policy or produce new data to share through the catalog. Examples from each function were shared with the group to help provide additional context for the days discussions.

The 6 Functions are listed below, and defined in Appendix A.

- Open Data Catalog
- Data Policy Advocacy
- Civic Technology Engagement
- Data Cohorts
- Small-Scale Consulting
- Citywide Reporting



## DOT VOTING & DISCUSSION

After [bc]'s short presentation on the 6 Functions of the Data Ecosystem, lab attendees were asked to participate in a dot voting activity to develop an understanding of the group's attitudes towards each function. Overall, the group consensus leaned towards the need for slight definitional adjustments to each function - with the exception of the Open Data Catalog which was more or less agreed upon by the group. Participants were given the following definitions for voting with green, yellow, and red dots.

- **Green:** No substantial change to the concept or definition is needed.
- **Yellow:** The concept is right, but the definition needs minor adjustments.
- **Red:** This Function does not play an important role in the data ecosystem.



With no “Red” votes, the Open Data Catalog function was confirmed as a key component of a healthy data ecosystem. While there were some concerns with the specific definition, a centralized location for accessing a variety of data is needed to help improve activities across a data ecosystem.

In the discussion following the dot voting activity, participants identified the need for a strong institutional home for an ideal data catalog, such as at a university or foundation, but did not recommend a specific entity. Wherever a data catalog is hosted, it will need to address issues related to data quality, documentation, standardization of metadata, and stewardship of catalog creation and maintenance.

Additionally, there is a need to identify a data schema to handle a variety of data types to be stored in the data catalog.

### Open Data Catalog

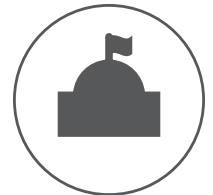


Further clarification and refinement of the Data Policy Advocacy function is needed, a task that the affiliated work group will address - primarily focused on identifying exactly what the relationship between data and policy could be.

However, several important points were raised by participants. In many organizations, working to refine and adopt policy that relates to collection, dissemination, or use of data might require an assessment of the return on investment (ROI) that can be expected from any change.

Finally, the group identified a need to further understand the scope of existing data policies across North Texas and how those relate to broader Federal and State policies and directives. In order to enhance this groups discussion of the Data Policy Advocacy function, policymakers and legal representatives from various data holders will be necessary.

### Data Policy Advocacy



While attendees were familiar with many examples of civic technology, the group voiced a number of concerns related to this function. Primarily, do tools developed by civic technologists serve the needs of a those most in need? Beyond that, the group voiced a desire for identifying mechanisms to enable citizens to gain skills and expertise that can let them participate in this function.

Without participants working in the technology world in attendance, much of this conversation needs to be informed in conversations following Lab 1 and during Lab 2. To achieve this, bringing a number of civic technologists to the discussion will be vital for further defining this function.

### Civic Technology Engagement





### Data Cohorts



A function similar to the Data Cohorts is something needed in order to provide training, oversight, and management of other functions in the data ecosystem. The existing definition for the function needs to be refined in order to fit the needs of the broader data community, which this group identified it was not diverse enough to do itself. Community members, business leaders, and government officials are a major part of the ecosystem, but are not in the room.

Additionally, the possibility for aligning any work towards the Data Cohort function could further enhance the data catalog, support data policies, and establish training to build ecosystem wide capacity. An idea was promoted that a group-defined vision for the data ecosystem might allow the Data Cohort function to further solidify its purpose and activities.



### Small-Scale Consulting



Voting revealed a common understanding for the need of consultants within the data ecosystem. The group discussed the need for consultants in the ideal data ecosystem, where organizations have the internal capacity to perform needed functions. However, most organizations do not need access to certain services frequently enough to justify a staff position.

To address these concerns, an idea for developing tools to help organizations access the types of information consultants provide more easily was discussed democratizing the power of data and analytics services. In addition to these tools, consultants within the data ecosystem could begin to take on a role more akin to coaching or partnership to provide services while building capacity for organizations to tackle tough problems on their own.



### Citywide Reporting



While the group mostly agreed with the function's definition, a need for some refinement was identified, including a name that more accurately represents the type of work that could be included under this function - from city and region wide indicators or reports, to smaller neighborhood level dashboards or scorecards. This activity, much more a product of the data ecosystem, is integral in advancing policies and programs aimed at improving life in the community.



## TAKEAWAYS & WORK GROUPS

Overall, the dot voting activity and following discussion raised a number of key points which work groups were instructed to address in more detail later in the day. There were a number of common themes mentioned throughout this discussion:

- A need for improve trust within the data community and between the data community and the public and private sectors, institutions of higher education, and the general public.
- In order to reform the data ecosystem, incremental steps must be identified and developed to 1) catch up with other cities across the US and 2) meet the needs for the future that can't be anticipated at present.
- A collective vision and understanding for data quality standards, uniform metadata and documentation, data collection efforts, and priorities within the ecosystem should be developed to address the needs of the entire data ecosystem.

## WORK GROUPS

Following the discussions around the 6 Functions in the first part of the day, participants were sorted into four work groups. Work groups were combined for Data Policy Advocacy & Civic Technology Engagement and Small-Scale Consulting and Citywide Reporting. Facilitators from [bc] joined each group to help participants move towards ideas that can transform the way the local data ecosystem functions and document each group's conversation.

Each work group was asked to focus their discussion on four primary areas:

**Clarity:** In cases where the previous discussion and dot voting identified a need for a more refined definition or re-focused function, groups were tasked to begin those conversations and begin to identify what changes are needed to reflect the group's understanding of the data ecosystem.

**Opportunities:** Lab 1 was designed for a group of stakeholders to identify workable solutions for improving the data ecosystem, through the working groups opportunities within each function were explored in order to assess the different ways an Open Data Catalog or Data Cohort program could be successfully developed.

**Barriers:** Work group members were asked to discuss the barriers they know might stand in the way of their idea being implemented, driven by past experiences in Dallas or knowledge of challenges faced by similar approaches in other cities. Work group members were also asked to brainstorm any challenges that might be too complex to address in this setting.

**Missing Players:** As this process is meant to be stakeholder driven, work group members were asked to identify individuals and organizations not present at Lab 1 who need to be invited to Lab 2.

## WORK GROUP 1



### Open Data Catalog

Facilitator  
Thomas Simpson

Work Group  
Thomas Brindle  
Sweety Badihya  
Kristin Kuhne

While other work groups had to identify an activity that could enhance work within their function(s), the Open Data Catalog team prescribed the following solution: creation of a robust, inclusive data catalog of free, publicly available local, regional, state, and federal datasets. However, definitional changes were discussed by the work group, and members will be pushed to further refine their definition prior to Lab 2. Concepts discussed included the needed to expand targeted data providers, identification of realistic standards for inclusion, and the explicit need for uniform metadata and documentation.

However, the group recognized the need to include a more extensive group of stakeholders from organizations that collect, maintain, and disseminate data in order to better understand challenges to implementing the ideal data catalog. While an institutional home would be ideal, due to their inherent trust and sustainability, a number of opportunities exist for establishing a temporary catalog through open-source technologies. At a minimum, this could function purely as a catalog for available datasets that can direct users to the source for raw data.





## Data Policy Advocacy & Civic Technology Engagement

Data Policy was seen as foundational to a healthy Data Ecosystem. Developing standards and practices for data collection, sharing, and hosting is vital to the productive use of open data resources, and sharing analyzed data for the use of Policy advocates. The group agreed that this is not an activity that is currently active in Dallas, and gaining the support of key local leaders would be necessary to move this forward.

Current challenges facing data policy, and particularly a robust open data catalog are: the legal barriers to sharing or accessing organizational collected data, duplication of data collection efforts, gathering data from smaller groups or crowdsourcing data.

The conversation around civic tech was limited, in large part to the absence of a person participating or driving that effort in the room. Identifying a person to participate in future effort will be critical to understanding the role of this function. One of the key questions surrounding civic tech, and its role in the data ecosystem is “Who is the end user”, and what role it plays in supporting community efforts for change and growth.

Facilitator  
*Lisa Neergaard*

Work Group  
*Gary Hufstedler*  
*Richard Amory*  
*Jenny Eyer*

## WORK GROUP 3



### Data Cohorts

Facilitator  
Bernardo Salazar

Work Group  
*Michael Veale*  
*Mark Draz*  
*Corron Sanders*

The Data Cohort team recognized the potential for this function to establish itself as the framework for organizing and streamlining activities within the data ecosystem. The recommended development of a number of groups, or domains, operating at various scales within the data ecosystem on discrete topics, geographies, or disciplines. A tiered structure would allow for the creation of an overarching advisory cohort comprised of key stakeholders and members from sub-domains (health, education, housing, etc). These groups could inform the development of, and provide stewardship for, the Open Data Catalog and identify opportunities for citywide reporting, changes in data policy, and connect consultants with organizations or entities in need.

The team was challenged to identify ways for including access to a wider variety of audiences at different levels of participation. Capacity building activities for professionals are needed, but so are opportunities for broader community wide capacity building around data collection, visualization, storytelling, and decision-making.



## Small-Scale Consulting & Citywide Reporting

The work group felt that the distinct definitions for both Small-Scale Consulting and Citywide Reporting could be rolled into one function that focuses primarily on their fundamental similarity: reporting on a topic, using data, to inform various audiences. The work group identified the need for a comprehensive research catalog devoted to research conducted in/on the city/region and a scholarship program aimed at graduate students at area universities for research interested in conducting their research in North Texas.

Facilitator  
Owen Wilson-Chavez

Work Group  
Dylan Farmer  
Duane Danksreiter  
Jane Massey

Before any effort is made to establish annual or semi-annual reports on topics important to the City or Region, understanding the body of research focused on North Texas is key. Members of this work group felt that this effort should be undertaken by a reliable institutional body with an advisory group able to maintain consistency and continuity to the work. Commitments to sharing future reports, developing collaborations to drive needed research, and establishing a protocol for peer review to enhance quality were all discussed.



## MISSING PLAYERS

A variety of individuals and organizations needed to advance these conversations was compiled by each work group.

- City of Dallas
- Dallas County
- Dallas County Community College District
- Dallas County Appraisal District
- Dallas-Fort Worth Hospital Council
- Dallas Independent School District
- North Central Texas Council of Governments
- Parkland Center for Clinical Innovation (PCCI)
- Texas Education Agency - Region 10
- United Way of Metropolitan Dallas
- University of North Texas, University of North Texas - Dallas
- University of Texas at Arlington
- Members from the business community, including private consultants
- Neighborhood Groups

## LAB 1 TAKEAWAYS

Ideas developed by each of the four Work Groups presented a number of ways for connection, interdependence, and accountability shared across the 6 Functions. While no group fully laid out a roadmap for successfully implementing their idea, Lab 1 was intended to be the launching point for those roadmaps. To implement each group's ideas, there is a clear need for building trust between different actors in the data ecosystem, from large data providers (e.g. cities, counties, hospitals) to individual citizens.

Additionally, the group expressed a desire to better define the collective vision for Dallas' data ecosystem. This vision could help provide additional answers and context for other recurring themes in the day's discussion, such as:

**Audience:** Identifying the end-users of the data ecosystem is an important step for each work group to take moving into Lab 2.

**Prioritization:** How are different elements identified by the group prioritized?

**Relationship between Functions:** For the most part, work groups developed ideas that are closely related to the needs of other functions, indicating the degree to which the group has already developed a common vision for transforming the data ecosystem.

**Institutional Partner:** Throughout the day, different participants echoed the need for an institutional entity to take the lead in as a way of helping sustain the body of work over time.



## UPCOMING EVENTS

March 2, 2016  
8:30 AM - 2:30 PM

Data Ecosystem Lab #2  
*Communities Foundation of Texas*  
*Mabel Peters Caruth Center - Bluebonnet Room*  
*5500 Caruth Haven Lane at Central Expressway*  
*Dallas, TX 75225-8146*

March 23, 2016  
8:30 AM - 2:30 PM

Data Social Hour #5  
*buildingcommunityWORKSHOP*  
*416 S. Ervay Street*  
*Dallas, TX 75201*

## NEXT STEPS

The discussions of Lab 1 are vital in developing a consensus vision for Dallas' data ecosystem. In order to meet the challenges and needs identified by lab participants, [bc] will critically evaluate each of the 6 Functions, the relationships between functions, and relay this information to each of the four work groups. As this work moves into Lab 2, [bc] will work with stakeholders to determine the types of procedures, roles, and structures needed to support a healthy data ecosystem in North Texas.

As these elements are defined, discussed, and settled on by Lab participants, [bc] will distill the information and establish a roadmap for achieving the group's vision. With this roadmap, [bc] will work with stakeholders to bring the vision to fruition. How this emerges will depend on the vision that unfolds, the participation of key stakeholders, and investment from the funding community.







## APPENDIX F - LAB 2 SUMMARY

# DATA ECOSYSTEM PROJECT LAB 2

## OVERVIEW AND NEXT STEPS

bcANALYTICS  
MARCH 2016

#### PREPARED BY

The buildingcommunityWORKSHOP ([bc]) is a Dallas based nonprofit community design center seeking to improve the livability and viability of communities through the practice of thoughtful design and making. We enrich the lives of citizens by bringing design thinking to areas of our city where resources are most scarce. To do so, [bc] recognizes that it must first understand the social, economic, and environmental issues facing a community before beginning work.

#### SUPPORTED BY

As the largest community foundation in Texas and one of the largest in the nation, Communities Foundation of Texas (CFT) works with families, companies and nonprofits to strengthen our community through a variety of charitable funds and strategic grantmaking initiatives. The foundation professionally manages more than 900 charitable funds and has awarded more than \$1.3 billion in grants since its founding in 1953. Increasing financial stability of working families is one of the two key focus areas of CFT's community impact funds. To support this area, CFT has launched the Data Driven Decision-Making (D3) Institute. The D3 Institute is designed to provide organizations that offer programs and services for low-income working families the power to accelerate their development of enduring solutions to the social and economic problems facing this population. [www.cftexas.org/D3](http://www.cftexas.org/D3)

# DATA ECOSYSTEM PROJECT

Access to information has changed dramatically over the last two decades. Data about our country, states, cities, and communities is being collected at an unprecedented scale. These datasets, and the tools built from them, are becoming more and more accessible to a wide range of audiences. Today, organizations, government jurisdictions, and policy or community advocates seek to make informed decisions based on the wealth of data and information that flows out of our communities.

Despite this increased demand for local information, data in Dallas is often not easily accessible, collected infrequently or not at all, in usable formats, outdated, or exceedingly costly. Meeting today's data needs will require an overhaul of the status quo.

The Data Ecosystem Project, initiated by Communities Foundation of Texas and supported by buildingcommunityWORKSHOP ([bc]), has begun a process to learn, understand, and assess the gaps in the current system. Working with local organizations and professionals, the Data Ecosystem Project seeks to identify strategies that build a more complete data ecosystem. The outcome of which should be a more robust open data environment that serves the needs of the city, local organizations, communities, and residents working to build a better region.

To date, The Data Ecosystem Project has involved three key learning and engagement activities:

- **National Best Practices** - A research and interview process identifying best practices in organizations across the country, including short case studies of select organizations.
- **Dallas Data Ecosystem Survey** - A survey gauging the data and analytics capacity of organizations working across North Texas.
- **Focus Groups** - Two stakeholder-driven focus groups aimed at identifying challenges within and solutions for Dallas' data ecosystem.



## DATA ECOSYSTEM PROJECT LAB 1

On January 25, 2016, [bc] and CFT convened a number of key stakeholders from Dallas' data ecosystem to discuss the current state of the local data ecosystem and brainstorm ideas to transform it for the better. This event, Lab 1, identified a number of solutions that can be built upon to transform Dallas' data ecosystem to better support and enable data collection, sharing, analysis, and use in decision-making processes. Following the previous phases of the Data Ecosystem Project (DEP), Lab 1 focused activity on six key functions identified through a review of national best practices, the Dallas Data Ecosystem Survey, and focus groups held in 2015. These six functions, broad collections of activities that occur within a data ecosystem, were used as a way to organize solutions, identify key players, and draw connections between each function.

Following a broader group discussion, participants were split into 4 work groups to dive deeper in the structure of each function and to identify steps needed to carry these functions to fruition. As a whole, the group expressed a desire to better define the collective vision for Dallas' data ecosystem. This vision could help provide additional answers and context for other recurring themes in the day's discussion, and a document was circulated to participants for input in this manner following Lab 1.

Prior to any major changes in the data ecosystem, those who attended the lab felt that a number of issues needed to be addressed or defined, such as: establishing an audience of end-users of the data ecosystem; addressing the need for prioritization of activity and effort in the ecosystem; fully defining each function and their relationships; and establishing what kind of entity or organization can realistically carry out the needed work.

## PRE- LAB 2 WORK

Following Lab 1, [bc] reconnected with work groups formed in Lab 1 to further refine ideas generated at Lab 1 and brainstorm better ways of organizing elements of the data ecosystem. Through work group calls and internal discussions, a distinction arose between the 6 Data Ecosystem functions presented at Lab 1. We recognized that functions fell in one of two groups: primary functions and secondary functions. The primary functions are needed to create and support a vibrant data ecosystem, while others are byproducts of different organizational activities. These *primary* functions allow for better collection, sharing, and use of data and can create additional opportunities for organizations and individuals to create products through the use of data (*secondary* functions). [bc] chose to focus more heavily on the primary functions in Lab 2 in order to better focus on steps needed for improving the Dallas' data ecosystem.

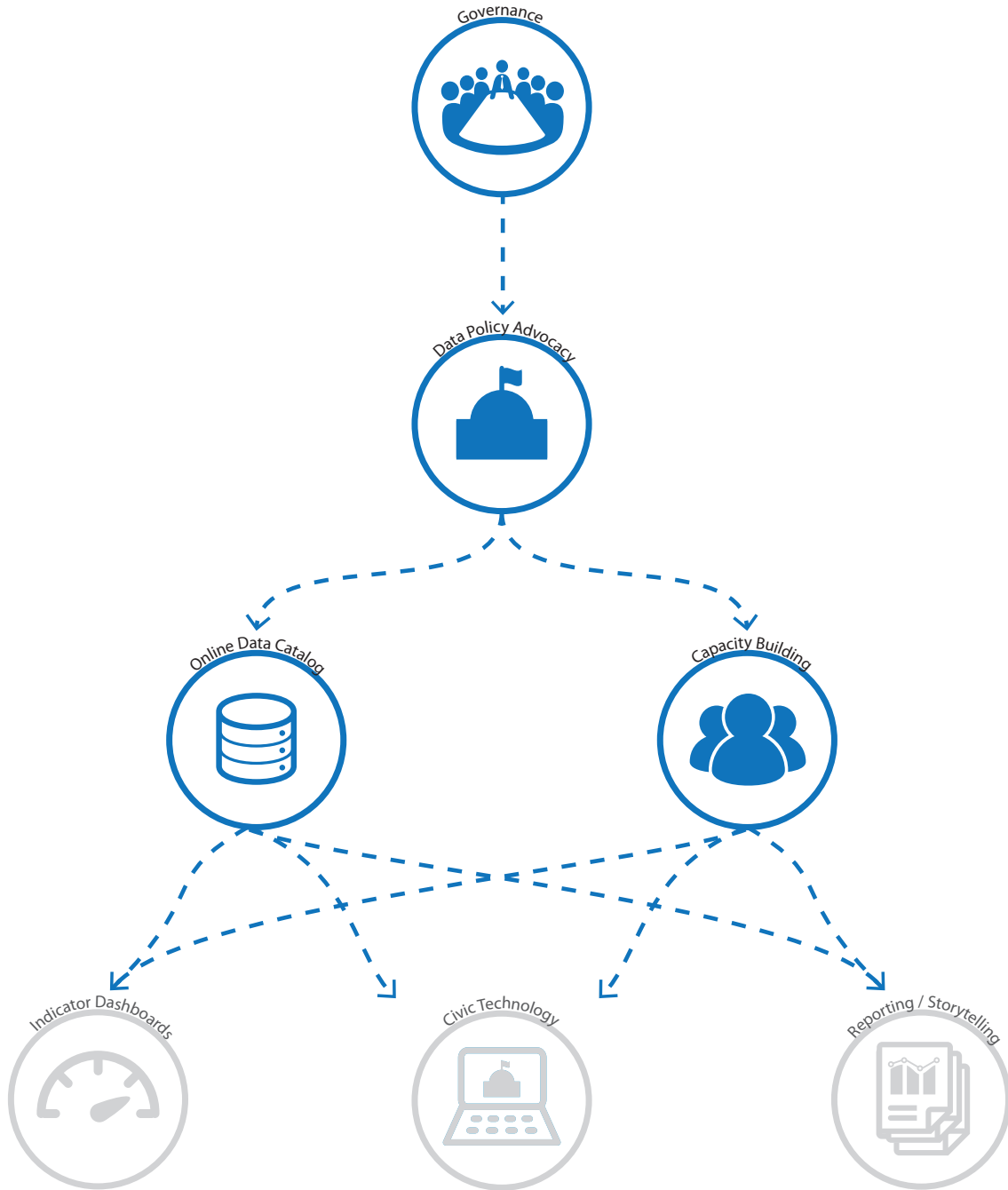
To do so, Lab 2 focused heavily on the roles of each primary function within the broader data ecosystem and the relationships between activities that occur under each function. An overwhelming theme from Lab 1 focused on the need for a cohesive strategy for prioritization across the ecosystem. Based on this, [bc] felt that a crucial element was missing from the Data Ecosystem model, which relates to governance. Establishing the approach to this governance mechanism became a major focus for Lab 2. The remaining primary functions - Open Data Advocacy, Open Data Catalog, and Data Cohorts - are under the domain of governance but still driven by their own core activities.

The wealth of data and analytics activity happening in organizations across North Texas is testament to the talents found across the region. Research still occurs, and data is an increasingly important, if not a major, part of the decision-making process in nonprofit, for profit, and public sector organizations. Those invited to Lab 2 come from a variety of organizations that engage in activities that fall within the scope of the Data Ecosystem Project. Their voices, therefore, were critical in identifying the mechanisms and approaches necessary to improve the ecosystem.

Major concepts that emerged from the discussion around each primary function's key roles and activities are detailed in the following pages and actionable next steps, identified by the group, are listed at the end of the document. Much of the conversations revolved around ways to re-conceptualize the relationships between the primary functions and secondary functions, as opposed to the data ecosystem relationship diagram presented at Lab 2.



# Primary & Secondary Functions of the Data Ecosystem



## GOVERNANCE

Governance took a central role in the discussions of Lab 2. Those in attendance felt the need for a strategic planning process and creation of a strong value proposition to guide and shape the development of the data ecosystem. While those in the room identified the DEP as a starting point, effort is needed to engage a wider number of organizations in any governance structure. Major generators of data in the region - cities, school districts, hospitals - would need to be engaged and included in the governing body.

This governing body would oversee and help direct work occurring in other functions by identifying priorities, establishing an agenda for moving forward, establishing data quality standards for the Open Data Catalog, and forming the preliminary Data Cohorts. In addition, the governing body would have an important role in fundraising. While the specific mechanisms for funding work in the data ecosystem have not been established, participants identified a number of possibilities each with their own pros and cons.

Overall, any governing body established through the DEP should be inclusive and have representation from a wide variety of organizations and the community. This board or committee needs to include members from a variety of sectors, backgrounds, and topic-areas to ensure issues are heard from across the data ecosystem.

## OPEN DATA ADVOCACY

While participants see the Labs and DEP as the beginning of the Governance function, it can potentially serve as a starting point for the Open Data Advocacy element. An essential role of the Open Data Advocacy team would be working with local governments to advocate for better ways to share data - either through open data policies or MOU's to provide raw data to the Open Data Catalog or other parties.

In addition to working with governmental or quasi-governmental organizations, the Open Data Advocacy team must also work with nonprofit organizations, foundations, for-profit businesses, and others to provide better access to data and information that may be beneficial to the ecosystem. In many cases, organizations collect a wide variety of data, for a specific purpose, that may have other possible applications in other fields. These types of arrangements are not new in North Texas. Crossroads Community Services and the North Texas Food Bank have partnered with UT Southwestern and UTD to collect a variety of client information as part of the Hunger Center Longitudinal Database. While this data is initially collected for administrative purposes, CCS and NTFB recognize the potential for this information to be used to better our understanding of issues affecting food security, poverty, and financial stability. The data is made available to interested researchers, upon request, and is used by UTD and UTSW in academic research.

Working to identify similar opportunities for expanding access to different types of data is a key role of the Open Data Advocacy element. To do this, a dedicated team would be needed to focus solely on this body of work. The Open Data Advocacy team would be an independent team funded to provide specific services to organizations wishing to participate in the Open Data Catalog, improve the ways their data is shared, or address barriers to their work that stem from issues of data access.

## OPEN DATA CATALOG

At the least, the group agrees that an Open Data Catalog is necessary for improving the health of Dallas' data ecosystem. The Catalog will provide researchers, individuals, and organizations an opportunity to quickly identify relevant datasets and reports to inform their work. This catalog would provide short descriptions of each entry, example uses, any necessary documentation, and information on where or how to access the data (either links to its location online or contact information for individuals in charge of specific data). While this approach would begin to improve how organizations and individuals are able to access data, broader changes would still be needed. The catalog approach was discussed by many as an easy win, and something that could be implemented quickly through prototyping and iteration.

To really improve access to data in North Texas, a strong online data portal is necessary. This would provide all of the same information contained in the catalog, but would enable direct downloads of data from the same website. This would put Dallas more squarely in line with other cities across the US, and would ideally enable users to access information in machine-readable formats and through API's for use in web and mobile applications.

In order to fully develop each scenario, a dedicated team would need to focus on building the resources and maintaining each option. As new data is collected and shared it will need to be documented, and updates to existing data need to be added. The team would primarily be tasked with identifying and documenting the wealth of data resources available at the federal, state, and local levels; providing value added elements to high priority datasets; and maintaining the catalog or portal over time.

## CAPACITY BUILDING

Capacity building activities can be targeted at a wide range of potential users - from community members to professional data analysts. Activities could focus on how to use tools for collecting, visualizing, and analyzing data available on the data catalog/portal or on more advanced techniques needed within organizations. Ultimately, the group agreed that the initial push of capacity building activity needs to focus on the small business and nonprofit communities. Targeting trainings and networking events to these groups will ensure sustained interest to grow the function for a wider audience.

Some in attendance felt that it was important to leverage existing community resources in order to better provide capacity building activities to community members. Facilities such as libraries, recreation centers, and schools might be able to offer a safe, trusted space to communities wishing to learn how they might use a new tool, access data relevant to their lives, or gain new skills.

Additionally, the experience Commit! and United Way had with their Data Cohort program - more in line with the capacity building function - was incredibly positive. One of the most productive aspects of that program was pairing individuals from the program with a particular staff member and allowing them access to "office hours" to take advantage of one-on-one training and coaching. The data coach approach used by CFT's D3 Institute allowed for a similar level of capacity building on-top of the structured activities of the programs. Adopting a similar practice within the data ecosystem might allow for a different level of capacity building activity for nonprofit organizations that have immediate needs that might not be addressed in more structured trainings or workshops.

## DATA COHORTS

Cohorts are seen by the group, overall, as a place to encourage discussions around a specific topic, method, or problem within the data ecosystem. In some cases a topic-specific cohort might focus on issues within that topic area that relate to data - such as changes in state or federal policy on collection mechanisms, quality issues from a data provider, new analytical approaches, etc. - or identify new ways for collaboration between organizations in the topic area.

While this broader vision for the Data Cohorts function was accepted, there was much debate on the nature of the cohorts over time. Are cohorts short-lived or static? There can be benefits under both scenarios. For one, short-lived, adaptive cohorts might be able to generate sustained interest for the duration and lay the groundwork for future cohorts. As different cohorts enter and exit the ecosystem, however, it might be challenging to develop long-term trust within the community. Static cohorts can build trust over time, but experience from participants other work highlights a challenge in maintaining interest and participation for several years.

Ultimately, a mix between the two approaches seems the most productive. The short-lived cohort offers interested individuals, with limited time commitments, an opportunity to play a meaningful role in the ecosystem, while the static cohorts can provide a continuing point of interaction for those working across the ecosystem. Cohorts would be established by the governing body but would emerge from the needs of the ecosystem as a whole and would communicate information and issues discussed in cohort meetings to the governing body.

## NEXT STEPS

At the end of Lab 2, participants identified activities that must proceed in order to stimulate the larger changes needed in the data ecosystem. These activities conform to the larger Functions established by the group, and represent work that can occur prior to a more robust push for changes to the data ecosystem. Major steps or work items include:

- Creation of a value proposition or statement that demonstrates the need for changes to North Texas' existing data ecosystem.
- Begin to inventory existing datasets maintained or collected by public sector, nonprofit, for profit, or institutional organizations in North Texas.
- Creation of a bare-bones data catalog documenting existing datasets available, organized by sector or topic area.
- Work with stakeholders involved in the process to-date to gauge interest in continuing with the work and shaping the governance element.

Additionally, the DEP project team will work to compile a full report of the projects work spanning July 2014 through April 2016. This report will reflect the input of each focus group and lab participant, as best as possible, and recommend specific actions needed to make transformative changes across the data ecosystem.

