Envisioning a Culture of Community Data Sharing

A vision for Kent County, Michigan
from KConnect and its stakeholders

Neil Carlson, PhD
Calvin College Center for Social Research

Presented to the KConnect
Accountability Partners Council
May 18, 2017

KCONNECT
Fulfilled Potential. Community Prosperity.

Table of Contents

Table of Figures

Table of Tables

1 Executive Summary and Business Case

1.1 Problem

Given a complex voluntary environment with disparate partnership requirements, **how can stakeholders in collective impact initiatives design and build interoperable community data systems** to achieve alignment of social and educational service systems, enabling a clearer, better-resourced path to economic prosperity for all children in Kent County?

1.2 Problem-solvers

This white paper is a product of the KConnect stakeholder network (k-connect.org). KConnect is the collective impact backbone organization for Kent County, Michigan. The staff and their consultants convene stakeholders, facilitate meetings, encourage alignment of efforts, and motivate sustained action. The KConnect network includes local representatives from school districts, social welfare agencies, health care providers, education and social service nonprofits, foundations, universities and colleges, and a variety of other interested parties, including related networking initiatives like Talent 2025 and the Kent County Essential Needs Task Force.

The Community Data Trust idea emerged primarily from KConnect's Data and Capacity Workgroup, which advises and supports KConnect on metric selection and data visualization. Lead author Neil Carlson, PhD, has been a co-chair of the Data and Capacity Workgroup since its inception in the fall of 2014.

We are grateful to many others for contributing content and feedback. A complete list of authors and contributors is found in **Appendix A: Contributors** on page 44.

1.3 Proposed solution

We propose that collective impact stakeholders **apply collective impact's logic to collective impact data systems**. We should apply thought and effort to the development of a Community Data Trust (CDT)--or a similarly named entity. The CDT would be **a permanent governance and support structure to coordinate and facilitate data sharing and shared measurement in Kent County**.

The Community Data Trust concept is about developing people and creating a sustainable culture of technical cooperation. Committing to a people-first, collaborative culture comes before governance, infrastructure, and technology. The goal is not to take over data-sharing duties for an entire county, but to inspire a data-sharing culture. The CDT should inspire trust because important matters of skill, reliability, accountability, and motivation are transparent and well understood.

1.4 Core CDT activities and their benefits

1.4.1 Core activities

The Community Data Trust would do some or all of the following (more details in 2.5 on page 16):

a. Provide **membership-based governance** for shared measurement and data sharing efforts.

b. Organize **professional development and peer support** for personnel of member organizations.

c. Adopt or create appropriate **standards for interoperability** among membership- and community-facing data systems.

d. Foster adoption of **standards-based "data feeds"** among members and affiliates.

e. Recruit **major computing infrastructures** for public benefit and collective impact.

f. Provide a legal and technical infrastructure for **lower-cost development of data sharing agreements**.

g. Develop software and support systems for **syndicating and auditing data sharing**.

h. Certify **third-party auditing processes** for data-sharing systems.

i. Provide **local, experienced advisory support for data-system development and transitions**.

j. Support **best practices in the selection of vendors and contractors**.

1.4.2 Benefits

We expect that these activities would produce the following benefits:

a. Through collaboration to promote professional development: **rapidly rising community capacity for technical work**, with positive spillover effects on the day-to-day operational performance of member organizations.

b. **Slowly falling per-project costs and rising quality and performance for data sharing initiatives**, through standard-setting and economies of scale for technical, logistical, and legal functions, including improved software selection and/or design.

c. **Slowly growing capacity for more granular and more comprehensive analysis of social and economic problems**, across both geography and time.

d. **Gradually falling risk of serious legal and financial liability** through implementation among members of certified training, best practices, and auditing for compliance.

1.5   Scope

The Trust is a big idea for solving big, complicated, expensive problems. However, it is not a proposal for a big, monolithic, vertical organization or massive software development project. Nor is it a new idea. Nor does it depend on the development of radical new technology.

Just the opposite: like many disruptive innovations (see sidebar about Christensen, 1995), the Trust concept is a simple but powerful combination of an array of well-known strategies, organizational forms, and information technologies. These ready-made strategies include:

a. **Standard-setting coalitions**, like the World Wide Web Consortium, which defines the HTML 5 standard that allows web browsers to interoperate.

b. **Training tracks**, like those used to develop and certify professionals in dozens of disciplines.

c. **Data feeds** with Application Programming Interfaces (APIs), like those that allow manufacturers to share real-time inventory data with suppliers, or those that allow app programmers to integrate with Google Maps or Weather.com. Robust, affordable examples for nonprofits include tools based on API-rich platforms like QuickBase and SalesForce.

d. **Semi-automated production of modular legal agreements**, like tax forms produced by H&R Block software or last wills and testaments produced by LegalZoom. (These are just glorified versions of the mail-merge function found in word processors since the 1980s.)

e. **Easy single-sign-on implementation of authentication with complex roles and permissions**, now provided by many platforms, from Google, Facebook, Microsoft, Salesforce, and many more.

f. **Enforced and audited compliance with agreements**, as is at least partly implemented by the Michigan Homeless Management Information System (HMIS).

Table 1.1 below outlines how we might initiate the big county-wide idea through startup, small-scale, simple, affordable actions.

Disruptive innovation

A disruptive innovation so thoroughly upsets the old way of doing things that the old way eventually all but vanishes.

In the classic business book *The Innovator's Dilemma,* Clayton Christensen documents how small, simple changes in products produced irresistible waves of disruption in major industries, wiping out organizations whose business models depended on making sophisticated "sustaining" refinements to the status quo technology.

For example, Christensen studied how low-tech companies making 3.5-inch disks wiped out high-tech companies making 5.5-inch hard disks. The essential business value appeared to come from the sophisticated technologies that allowed more storage per square millimeter. However, in the end, the organizations' real value was in their adaptive ability to alter the size of the low-tech plastic-and-metal box containing the fancy disk.

Collective impact is itself a disruptive innovation! More on collective impact below.

Table 1.1 Strategy and tactics for the Community Data Trust

Area of focus	Long-term, countywide strategic vision	Short-term, single-team tactical actions
Broad collaboration	A robust county-wide network of data-sharing relationships with powerful, efficient shared infrastructure	Identify initial small-scale teams and projects that will commit to develop CDT principles
Shared standards	Implementation of software development best practices, including versioning and documentation, especially for a system for matching and/or aggregating data records while fully protecting confidentiality and/or anonymity.	Build plans into future projects to test and publicly document software development, especially matching and aggregation processes.
Technical capacity in the public and social service sectors	A deep bench in every stakeholder organization of technically savvy leaders who not only consume evidence from data systems, but regularly cooperate freely to produce and maintain new analytical infrastructures.	Organize new teams and projects from the ground up to include CDT-inspired professional development and relationship-building roles for both senior and junior staff. Plan ahead for the time necessary to learn and gel together.
Smoothly syndicated access to data	Multiple sector- and team-specific portals listing available data feeds from partners and guiding users through automating machine-to-machine connections.	Design new projects to conform to the vision by preferring systems and vendors that use fully-documented APIs.
Modular, best-practice legal language for data sharing agreements	Fully automated "wizard" for creating, editing, and signing bi- and multi-lateral data sharing agreements.	Convene a small group of legal and technical talent from major stakeholders to support the next major community agreement, collecting and standardizing the language in a shared database.
Independently audited compliance	Fully independent third-party review of data access patterns, both scheduled and on-demand, with reliable logs and appropriate expertise.	Prototype the auditing function on a forthcoming local project by including it in the program evaluators' tasks and budget.

1.6   Risks

Risk assessment for the Community Data Trust concept is based on the premise that demand for data sharing and community-level analysis is not going to go away. When resources are scarce, data will be demanded to guide where to allocate them. When resources are plentiful, stakeholders will often turn to information systems as a top priority on the list of backlogged investments.

Given this premise, we believe the risks will grow that are associated with status quo bi-lateral and/or *ad hoc* data sharing systems. The Trust is intended not to eliminate risks, but to assist in their control by reducing transaction costs and increasing local capacity for risk management.

Table 1.2 below documents risks of maintaining the status quo and of implementing the CDT.

Area of focus	Risks of the *status quo*	Risks of Community Data Trust
Broad collaboration	**Loss of major investment due to a) misalignment of technology with projects and b) turnover that destroys organizational technical memory.** Due to misplaced but common concerns about the efficiency of frequent meetings and loss of hierarchical control, current collective impact efforts run risks while they remain limited to the top executive and program-management strata of stakeholder organizations.	**Loss of major investment due to CDT failure or inefficiency.** If the CDT project fails, as many standards-setting governance initiatives do, or is coopted by a stakeholder faction with a narrow agenda, then the time, effort, and investment necessary to create it would be wasted or inefficiently spent. However, as discussed in Section 2.8 on page 18, adaptability is more valuable than efficiency in solving complex problems.
Shared standards	**Missed opportunities** due to the high transaction costs of creating ad hoc per-project data matching and aggregation protocols. **Lack of best practices** due to the weak influence among low-profile organizations with strong, innovative practices.	**Some loss of organization-level flexibility** due to standardization. However, standards should be more about how data is handled than about specific codes and values. Training, communication, documentation, and governance-level voice for affected parties should reduce this risk further.
Technical capacity in the public and social service sectors	**Severe lack of capacity across the public and nonprofit sector** due to difficulty in competing for out-of-sector talent and lack of technical professional development emphasis for in-sector talent.	**Brain drain**: the risk that the CDT might elevate individuals' skills only to lose them to higher wages in other sectors and regions. This risk makes the development of a rewarding, developmental culture all the more important.
Smoothly syndicated access to data	**High transaction costs:** Many high-value collaborative data analysis projects are easy to imagine, but hard to imagine implementing due to the lack of economies of scale.	**Little to no risk**: it is hard to imagine a downside to the CDT strategy in this area. Data-feed networks and "mashups" have been transforming the Web for over two decades.
Modular, best-practice legal language for data sharing agreements	**High transaction costs, yet the risk of an expensive legal error remains**: The appropriate expertise is rare and specialized. Even large organizations' corporate counsels may spend weeks reviewing data sharing agreements, yet still have little confidence they have done well.	**Wasted legal and technical fees**: A community-wide expert legal advisory group composed of independent attorneys and major stakeholders' counsels would add significant value by reviewing data sharing agreements and building up a modular knowledge base through the CDT. But if the group were severely underutilized, investment would be wasted.

Area of focus	Risks of the *status quo*	Risks of Community Data Trust
Independently audited compliance	**Major legal and financial liabilities** in the event of a breach of confidentiality, security lapse, or other violation of an agreement. Stakeholder organizations are more concerned about these risks than any others, yet the status quo generally leaves partners to monitor their own behavior. The lack of any third-party verification of practices makes self-defense far less credible.	**Modest risk of intrusive oversight and red tape** imposed due to overreach by the CDT auditors. However, proper governance should give unhappy CDT members a quick route to relief.

1.7 Impacts

In theory, major stakeholders would supply the Community Data Trust with governance delegates. These delegates should possess an appropriate mix of technical skills and managerial authority. The membership should begin with those parties directly responsible for the data systems needed for current or near-future collaborative projects. The CDT would seek to become a presumptive first-stop context for discussing, organizing, and implementing the data-sharing components of projects and initiatives.

1.7.1 Prospective membership

Initial CDT membership could include one member from each of the following categories:

a. Network initiatives such as KConnect, ENTF, and Talent 2025.

b. Universities and colleges; key faculty of research centers and Data Science programs, with their computing and skilled personnel, might become the central sustaining partners of the Trust.

c. School districts and educational collaboratives.

d. Health systems and health-related agencies.

e. Governments, government agencies, and government contractors involved in service provision.

f. Large nonprofits and collaboratives of smaller nonprofits.

g. Major local employers with aggregate data relevant to public purposes.

h. Foundations and other major funders.

i. Local research and evaluation consultancies.

j. Local software vendors and software development firms interested in public data.

k. Corporate counsels and other local lawyers and law firms.

1.7.2 Prospective leadership and initial funding

We prefer to leave these matters largely to the judgment of early adopters of the CDT concept. However, from a completely abstract perspective, it could be effective to recruit startup executive leadership with strong technical skills and evidence of recent effectiveness as a negotiator serving a major standard-setting body in a technical or computing-related field. Since we are not, as of the time of writing, aware that a qualified person or group exists in our community, this recommendation is wholly unbiased. However, it seems unlikely that a highly effective and interested person or group would not exist in Grand Rapids, or at least in Michigan.

Initial funding might be accomplished through membership fees paid by the parties interested in an initial project or round of projects. Fees could be pro-rated to reflect the scope of work of each party in the proposed project. This funding model reflects the basic recognition that such projects routinely underinvest in converting ad hoc systems into sustained value. A key performance indicator for the CDT would be members' evaluations of whether the membership fee enables creation of sufficient sustained value.

1.7.3 Business impact

The business impact of the CDT is expected to be generally positive for all involved. Some organizations—perhaps especially software providers and evaluators—might see short-term fluctuations in demand as the CDT exposes their traditional clientele to previously unknown options. However, much as a stock market or free trade agreement does, the information and infrastructure facilitated by the CDT infrastructure should create the circumstances for organizations to shift focus to higher-value activities. Members and non-members alike should be able to thrive in previously inaccessible fields, markets, and niches.

1.8 List of prospective CDT actions

To achieve a working Community Data Trust, we need at least some progress on the following:

1. **Create a working group** to discuss, refine, and experiment with the concepts laid out in this section.

2. **Develop a common agenda to grow our community's technical capacity** and deepen our personnel bench by becoming a "Deliberately Developmental Community." This label is a logical extrapolation from applying collective impact thinking to the concept of Deliberately Developmental Organizations (DDOs) from *An Everyone Culture* (Kegan, Lahey, Miller, Fleming, & Helsing, 2016). This task includes both promoting specific technical skills for individuals and promoting a culture of teamwork among individuals and across organizations. We discuss this concept in **Section 3**, "Become a Deliberately Developmental Community with a technology," beginning on page 24.

2. **Reduce data sharing transaction costs and liability risks**: assign expert legal and technical resources to modularize and automate the production of and compliance with Data Sharing Agreements (DSAs). See **Section 4**, "Organize expertise and automate DSA development and auditing" on page 32.

3. **Increase adaptability, shared value, and economies of scale by setting standards and promoting adherence to best practices** among CDT member organizations and their relevant vendors. See **Section 5**, "Inspire technical best practices among members and vendors" on page 39.

1.9 Conclusion: improvements welcome

We believe that movement in the direction of a Community Data Trust (CDT) would benefit Kent County tremendously. It would begin with the modest costs of convening and planning among early interested stakeholders, and be governed according to models offered by collective impact and standards-setting organizations.

This white paper, like the CDT, is a working, living, evolving entity. Feedback and additional content are welcome. We are interested in releasing repeated iterations of this document with multiple authors, including perhaps entire sections dedicated to specific problems.

2   A vision for the Community Data Trust

2.1   Purpose: set a common agenda for community data system development

This white paper seeks to propose a **common agenda** for the development of the community culture and infrastructure needed to sustain multiple **shared measurement** efforts across Kent County, thus achieving the economy of scale necessary to make our Kent County **collective impact** work attractive, reproducible, and durable.

2.2   Problem: how can we develop community data systems to achieve collective impact?

Given a complex voluntary environment with disparate partnership requirements, **how can stakeholders design and maintain interoperable community data systems** to achieve "collective impact" alignment of social and educational service systems, ensuring a clear path to economic prosperity for all children in Kent County?

Demand for evidence-based decision-making continues to rise throughout our society. While skepticism is warranted about some of the wilder claims of "big data" advocates, it is clear that huge gains can be made by making better use of data—especially longitudinal data about individual children—to tailor social services better to meet needs in health, education, and welfare. In addition to serving individuals, we need better demographic and economic radar to forecast the long-term sustainability of foundational social institutions like schools and hospitals.

For example, we still have little measurement of influential forces like local residential mobility, a force growing in power as urban housing prices rise. How much of the challenge of our educational system is attributable to mobility? We don't know.

In Kent County (and in many other communities around the world!), we have strong and capable organizations and many networks, yet we have largely failed to establish a stable center of gravity for the collection and analysis of community data. Instead, promising, collaborative projects are siloed, captured, protected, then intentionally terminated or gradually tapered off by the original sponsors. Project budgets and mandates do not include preparing data for general use. Talented leaders leave projects, which then die off, because we do not have a deep technical bench. The public and nonprofit sectors have paid too little attention to building "farm teams" to develop and retain talent. We also engage in too little succession planning to maintain complex technical systems that require substantial institutional memory to survive.

Instead of recognizing our systemic and structural deficits, we usually follow a specific project's starvation and death with a rationalizing narrative that blames a particular partner or vendor for the community's general loss of confidence and lack of investment. Some working systems are useful, but their technology is quickly superseded by attractive new opportunities, and momentum is abandoned for the next "shiny new toy."

Major community organizations take turns trying to attain the critical mass necessary to disseminate their solution to everyone else, but each such attempt ends up being perceived as a "power trip" that renews our cynicism and sours our collective interest in the next "big thing."

2.3   The wickedness of our problems

In confronting chronic poverty, inequity, and educational achievement gaps, Kent County faces an array of "wicked problems." Jon Kolko (2012) argues that wicked problems are impossible to reduce—let alone solve—through traditional transactional, task- and project-driven approaches. Pushing on one part of the problem accentuates other parts of the problem. Cooperation is not enough. Wicked problems are composed of social countercurrents on generational time scales. Their difficulty gives team members incentives to define success narrowly and declare victory prematurely.

Meanwhile, the underlying social syndrome goes untreated. Wicked problems require sustained, coordinated, strategic action, including vigilant attention to measures that may reveal diminishing returns and unintended consequences of actions.

A central proposition of this white paper is that **the creation and maintenance of community data systems itself should be regarded as a wicked problem**. Data systems are highly interconnected by technical capacity issues, even when they are not connected directly. Sharing affects the availability of skilled labor, executive attention, and capital computing resources. It is impossible to design a bilateral data exchange between community partners—a relatively simple solution on its surface—without unintended con-sequences for other organizations and data transfers.

Without significant attention to growing capacity, the partners in a bilateral exchange of data may exhaust their ability to respond to new requests. Lack of capacity forecloses receptivity to further sharing. First movers enjoy the power of *de facto* exclusive access to a data provider's data by monopolizing their capacity.

Fortunately, partners in our community are coming to recognize that we are approaching critical mass. We have growing opportunities to amortize the costs of data sharing procedures over many projects. If we sustain our focus on the wicked problem, we can reduce it into a merely complex problem that is addressed by systems with robust precedents. These precedents are found both in the public and non-profit sector and in business, especially among agency-to-agency and business-to-business standards-setting practices.

Wicked Problems

"A wicked problem is a social or cultural problem that is difficult or impossible to solve for as many as four reasons:

- incomplete or contradictory knowledge,
- the number of people and opinions involved,
- the large economic burden, and
- the interconnected nature of these problems with other problems.

"Poverty is linked with education, nutrition with poverty, the economy with nutrition, and so on. These problems are typically offloaded to policy makers, or are written off as being too cumbersome to handle en masse. Yet these are the problems—poverty, sustainability, equality, and health and wellness—that plague our cities and our world and that touch each and every one of us. These problems can be mitigated through the process of design, which is an intellectual approach that emphasizes empathy, abductive reasoning, and rapid prototyping."

From *Wicked Problems: Problems Worth Solving* by Jon Kolko, which can be read online for free: *www.wickedproblems.com/read.php*.

2.4 Designing for messiness

Kolko's book calls attention to the value of intentional design work in addressing wicked problems. Just as it is complex to plan a major public event, to manufacture a new jumbo jet, or to program a new operating system, solutions for wicked problems require careful, attentive, ongoing design effort. We are highly unlikely to solve wicked problems by accident.

Wicked problems are so named precisely because they are much more than merely complex. Wicked problems are not susceptible to the orderly, short-run, hierarchical, quarterly-report-driven tactics most organizations use to address problems or create solutions. On the contrary, wicked problems make it prohibitively expensive to organize hierarchical, predictable, centralized solution strategies. It is a routine fact of contemporary public life that traditional industrial-style approaches cause organizations to tire quickly and become frustrated with the high-risk, low-reward environment of collaborating to solve wicked problems.

In their article titled "The Upside of Messiness: Clumsy Solutions for Wicked Problems," management scholars Steven Ney and Marco Verweij write:

> *Cultural Theory postulates that four "ways of life" are the building blocks of social life: individualism, hierarchy, egalitarianism, and fatalism. Every social domain—from a kindergarten to an international regime—is said to consist of an ever-changing mix of these four ways of organizing, justifying, and perceiving human relations. Case study evidence from a wide range of domains suggests that one way of dealing with wicked problems is through forms of governance that creatively and flexibly combine these four ways of organizing social relations.*
>
> *Even though these ways of organizing and perceiving emerge in opposition to each other, Cultural Theory shows that they are also dependent upon one another. Furthermore, they all contain a "kernel of truth" as to how people can and would like to live, and as a result, social diversity and disagreement comes with the territory. As a consequence, any form of governance that attempts to impose a single way of organizing, perceiving, and justifying on a particular social domain is destined to fail.*
>
> *In contrast, more effective forms of governance nimbly mix all possible ways of organizing and thinking. Borrowing a term coined by law professor Michael Shapiro, we call these pluralist solutions "clumsy," because unlike their "elegant" counterparts, these approaches acknowledge that solutions need to be as multifaceted as the problems themselves.*
>
> *We refer to the organizational set-ups that are most likely to generate clumsy solutions as "messy institutions" because, unlike the sleek organigrams found on corporate websites, these types of organization embrace and engage with messy pluralism. Cultural Theory assumes that there are four ways of approaching wicked problems, each with its own strengths and weaknesses.*

It is commonplace for messy institutions and clumsy solutions to receive dismissive criticism from the stakeholders of sleekly elegant, transaction-driven organizations governed by orderly hierarchies. Yet the evidence is abundant that sleek, elegant organizations too often abandon the field quickly when faced with direct responsibility for solving wicked problems.

2.5   Proposed solution: create a Community Data Trust to apply collective impact strategies to community data system development

The central proposition of this proposal is that **community data systems should also be designed, governed, and maintained through collective impact infrastructure**. Technical solutions will be inadequate without a significant infrastructure of cultural, political, and economic supports. The "shared measurement" goal of collective impact itself presents a wicked problem, which needs to be addressed through messy institutions that can create the clumsy solutions we need.

Best practices in the business and technology sectors increasingly recognize the need for culture change, organizational reform, and transparency. We can develop the human and machine capacity we need through agile practices (Beck et al., 2001) and by cultivating talent and capacity through Deliberately Developmental Organizations (DDOs) throughout our community (Kegan & Lahey, 2009).

To these ends, we propose to **create a permanent governance and support structure to coordinate and facilitate data sharing and shared measurement in Kent County**. This structure, tentatively named the "Community Data Trust," would do the following:

1.   Provide membership-based governance for shared measurement and data sharing efforts, including both executive and technical representatives of municipal government, school districts, nonprofits, businesses, and universities. The new structure would be independent of, yet interlaced with, supportive to, and supported by local collaboratives and initiatives such as KConnect, ENTF, Talent 2025 or others.

2.   Mobilize and deploy aspects of general-purpose academic, government, nonprofit, and business computing infrastructures for public benefit purposes.

3.   Adopt or create appropriate standards for interoperability among membership- and community-facing data systems.

4.   Document and foster adoption of a network of standards-based data feeds among members, as opposed to centralizing sensitive data in a vulnerable central system.

5.   Provide a legal and technical infrastructure to reduce transaction costs for creating and implementing data sharing agreements through modular templates, document automation, and certification of expertise.

Additional possible functions or later phases of the CDT and its network might include:

6.   Develop and support an open-source software and support systems for role and permission management, syndicating services, and auditing data access.

7.   Provide certification for third-party auditing processes for data sharing systems.

8.   Organize and certify third-party training and professional development for the personnel of member and affiliate organizations.

9.   Ensure a sustainable community pipeline of personnel with the managerial and technical expertise to maintain momentum.

10.   Provide critical mass and best practices for the county community in negotiations with vendors, contractors, state and federal agencies, and major businesses.

2.6   Abundant collaborative and technical resources

In Kent County, we have adopted the collective impact strategy set (and its many conceptual cousins) as a promising solution for social change. We are aware of the following initiatives and working groups. Many of these overlap, and all of them engage too many stakeholders for them all to be mentioned equitably:[1]

1.   KConnect, fostering collective impact for Kent County and organizing a community dashboard (http://k-connect.org/data).

2.   The Kent County Essential Needs Task Force (ENTF, http://entfkent.org), aligning basic needs provision, functioning as the Home and Family Stability workgroup of KConnect, and contributing to that section of the KConnect dashboard.

3.   Talent 2025, workforce development for the West Michigan region, hosting a 13-county dashboard (http://talent2025.org/dashboard), and acting through its Kent County subsidiary as the KConnect workgroup on High School to Career.

4.   Kent County's municipal government, convening an array of partner organizations to work together for common ends through the Kent County Coordinating Council.

5.   The Great Start Collaborative, aligning early childhood education efforts.

6.   Kent Intermediate School District is collective of 300 schools, 20 public districts, 3 non-public districts and many public school academies and non-public schools. They serve the broader community by helping schools prepare nearly 120,000 students for school and life success.

7.   The Expanded Learning Opportunities (ELO) Network, administered by Our Community's Children, and sponsoring the Youth Community Data Center (YCDC) created by CRI.

8.   Believe 2 Become, an educational initiative of the Doug & Maria DeVos Foundation, which worked with the Grand Rapids Public Schools and the Community Research Institute of the Johnson Center for Philanthropy at Grand Valley State University to create a Master Data Sharing Agreement model (Carlson et al., 2011), with neighborhood engagement support from LINC UP.

9.   Challenge Scholars, an educational initiative of the Grand Rapids Community Foundation, in partnership with the Grand Rapids Public Schools and the Kent School Services Network, with neighborhood engagement support from the WestSide Collaborative.

10.   The Coalition to End Homelessness, a constituent group of the ENTF, makes regular use of the Michigan Statewide Homelessness Management Information System (MSHMIS), whose supporting vendor is Bowman Internet Systems. MSHMIS is notable for a robust feature that tracks which specific agreement grants users access to which data at any given time.

11.   StartGarden has recently been transformed into a for-profit/non-profit hybrid that encourages entrepreneurship in business and social services; among its partners is a

---

[1] Please note that mention of any organization here does not imply any endorsement or sponsorship of this white paper or the concepts proposed here.

Grand Rapids-based software firm called Collective Metrics which provides data-collection services to startups in West Michigan.

12. The Council of Michigan Foundations (CMF) is working to provide common data systems to foundations and their grantees through an initiative called Impact Michigan..

13. "Data Bridges" is the working title for a broad working group on prenatal to kindergarten data sharing convened by Family Futures and the Michigan State University School of Medicine, with a focus on preserving and providing community benefits from the complex, expensive infrastructures often built for longitudinal academic research studies.

2.7   The magic of setting standards

Meanwhile, there are abundant collaborative, networked organizational precedents for the kind of standard setting we need to establish, especially within the technology sector itself.

1. The World Wide Web Consortium (the "W3C" to most) is a successful standards-setting organization that has defined the common global technological basis for the Internet, web browsers, HTML, networking, data exchange, and more (Pilgrim, 2011; World Wide Web Consortium, 2016).

2. The IEEE sets engineering standards that facilitate interoperability among networked devices (Institute of Electrical and Electronic Engineers, 2016).

3. Businesses have defined stable data exchange standards for manufacturing (MESA International, 2016; Nurmilaakso & Kotinurmi, 2004) and retail supply chains (Standards for Technology in Automotive Retail, 2016).

4. There are established standards for data exchange in health care (AHIMA Work Group, 2013), education ("Common Education Data Standards (CEDS)," n.d.), criminal justice (National Center for State Courts, 2016), and children's welfare services (US Department of Health and Human Services, 2016).

These efforts all require extensive cooperative organization, a culture of collaboration. Once achieved, such culture is highly productive and valuable.

What we can do now, on the more modest but significant scale of Kent County, Michigan, is to join our growing public capacity for collective impact work to the standards and resources of contemporary information technology. We need to design and build community data systems through collective impact processes, building culture, deploying standards, and developing software collaboratively, in alignment.

2.8   Adopting a "team-of-teams" approach

We face a common cultural barrier to acceptance of the organizational complexity and regular, extended face-to-face meetings required by collective impact. Busy executives are particularly likely to perceive these approaches as inefficient and unfocused.

A potent means to hurdle this cultural barrier is the book *Team of Teams* by retired Gen. Stanley McChrystal and several co-authors (2015). The general commanded the U.S. Joint Special Operations Task Force (JSOTF) during the extended anti-Al Qaeda and anti-insurgent wars in Iraq and Afghanistan from 2003 to 2009. For many people, the modern U.S. military embodies the

logistical efficiency of a command-and-control hierarchy. So how does it fare handling wicked problems?

The book vividly describes how Al Qaeda's highly adaptable network organization repeatedly frustrated the highly efficient, powerful, yet rigid practices of the U.S. military. The terror network's "operational tempo" could field an operation and vanish faster than the U.S. could react, despite abundant intelligence and operational resources and a tremendous capacity for rapid deployment. The pattern of failure persisted until McChrystal and his colleagues recognized that the U.S. would not win until it learned to mimic Al Qaeda by prioritizing adaptability over efficiency, organizing themselves as a network of networks.

Figure 2.1 *Command vs. network (McChrystal et al., 2015, p. 25)*

What we were designed for   What we were facing
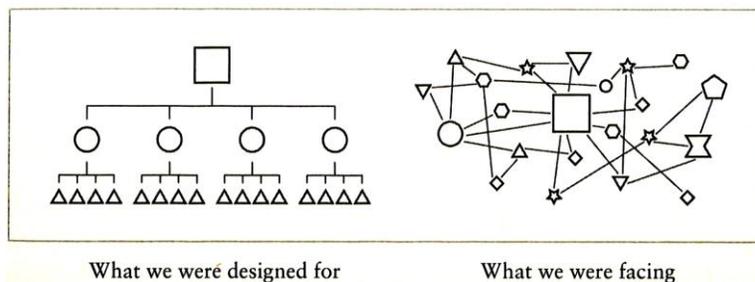
McChrystal and his colleagues found that it was not sufficient merely to reorganize the constituent commands into teams; this merely reinforced U.S. units' frustration with the vertical silos between operational, logistical, analytical, diplomatic, and executive teams.

> *Our circumstances were unique, but the problem is not. Though teams have proliferated across organizations from hospitals to airline crews, almost without exception this has happened within the confines of broader reductionist structures, and this has limited their adaptive potential. (p. 123)*

The clumsy solution to the wicked problem of Al Qaeda in Iraq was a team of teams: inviting and reinforcing horizontal connections between members of different units, growing trust through transparency, and encouraging autonomous unit action given growing awareness of the priorities and resources of the whole U.S. counterinsurgent network.

Figure 2.2 illustrates the team-of-teams model:

Command

Command of Teams
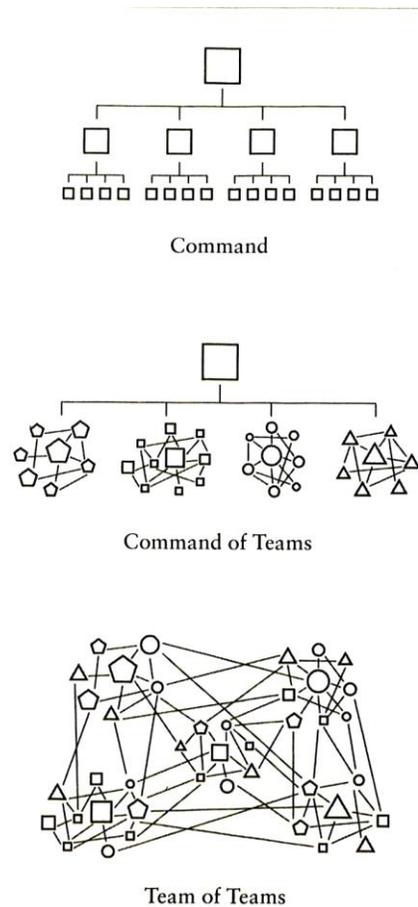
Team of Teams

The team-of-teams concept was incarnated in one of the more startling innovations of the JSOTF under McChrystal, an expanded, global, daily, two-hour videoconference called the "O&I," the Operations and Intelligence brief (2015, pp. 164–9):

> When I assumed command in 2003, the O&I was a relatively small video teleconference between our rear headquarters at Fort Bragg, a few D.C. offices, and our biggest bases in Iraq and Afghanistan. Quickly, though, that audience grew. We urged everyone from regional embassies to FBI field offices to install secure communications so that they could participate in our discussions. (p. 164)

> … However dysfunctional the internal competition within our command, it was dwarfed by that between our organization and the CIA, NSA, FBI, and other external agencies. Much as von Braun found with NASA contractors, we realized that no group could be useful if it did not understand the full context. We could not simply ship our intelligence requirements out to these agencies and expect them to realize all the intricacies of what we wanted and needed. (p. 166)

> … Many people sent to work with us found the environment distracting, or were uncomfortable in the participatory atmosphere. For bureaucrats who had built careers on discretion and never putting a toe out of line by oversharing, our way of

working was anathema. One partner agency offered the same response every day for the first year of our experiment: "Nothing new to report on our end." (p. 166)

… Our process began to develop its own gravitational pull as more and more groups recognized what the speed and transparency we had put in place could offer. Our forces were in daily contact with Al Qaeda, the nation's highest counterterrorism priority, and we were offering to share whatever we were learning. (p. 167)

*… In time, people came to appreciate the value of systemic understanding. O&I attendance grew as the quality of the information and interaction grew. **Eventually we had seven thousand people attending almost daily for up to two hours**. [emphasis added] To some management theorists, that sounds like a nightmare of inefficiency, but the information that was shared in the O&I was so rich, so timely, and so pertinent to the fight no one wanted to miss it. (p. 168)*

Seven thousand people on one conference call? A gross inefficiency indeed, right? And that on a scale only the U.S. government could muster. Yet results speak volumes, and the O&I process produced results—fast-paced, cooperative action against the enemy—where the previous model failed. Adaptivity-first produced success, where efficiency-first had reproduced deadly failure at lower cost. The appropriate role of efficiency was to be subordinate to adaptivity; for example, the government invested in telecommunications innovation to make the daily O&I conversation reliable and easy to connect to.

When people think of cutting-edge military hardware, they usually picture weaponry, not a bulked-up version of Skype, but that was our main technological hurdle and point of investment for several months. (p. 164)

When addressing wicked problems, technological and administrative investments should be driven by the quest for the adaptivity of the team of teams first, not the efficiency of single teams first. Placing priority on efficiency accentuates short time horizons and develops solutions that are too rigidly defined for the problem of the month or year, not the wicked problem of a generation or century.

2.9   Re-learning an old lesson

Is this "new" idea valid? Actually, the most trenchant criticism of *Team of Teams* is that it's far from a new lesson: German leaders learned it during and after World War I, and American troops demonstrated a mastery of it during World War II:

The real team-of-team fighters during World War II, however, were the Americans. Their success in the field resulted largely from a combination of effective small unit leadership and technical innovation at the tactical level by small independent groups. … If the American military really is the magnificent, formidable machine it claims to be, it should be able to deliver forces that can turn on a dime and adapt to the dilemmas they face. The fact that it took so long for some of the military's best, like McChrystal, to figure this out is not something admirable. And suggesting that they have discovered something new and important is equally dismaying. (Carafano, 2015)

Just as our community must recognize that our culture's path of least resistance leads not to "colorblind" equity but to racism and racial inequity, our managerial class needs to learn that it's not hard work to achieve autocratic hierarchy, organizational silos, and mistrustful workplace relationships built around suspicion and an expectation of insubordination and resistance. These are our culture's natural path of least resistance. We will have to work hard, continuously, to accept and promote the counterintuitive reality that apparently inefficient mass conversations and broad, open networks are the practices that produce solutions to wicked social problems.

How, then, can we apply the "team of teams" lesson to Kent County's wicked problems of poverty, racial inequity, and educational achievement gaps?

2.10 Organizing data systems initiatives

Through experience with at least a dozen similar initiatives to build shared data systems, both at the community level and within large, complex organizations, we observe a general pattern. One of our chief challenges is to learn how to develop community data systems for collective impact *through a collective impact process*. We need to apply the same "clumsy solution" to data systems development that we are using to mobilize alignment for programs.

Figure 2.3 is an idealized version of the contrast between a siloed, hierarchical network (left) and a "team-of-teams" network (right).

Figure 2.3 Moving to a team-of-teams model for community data systems
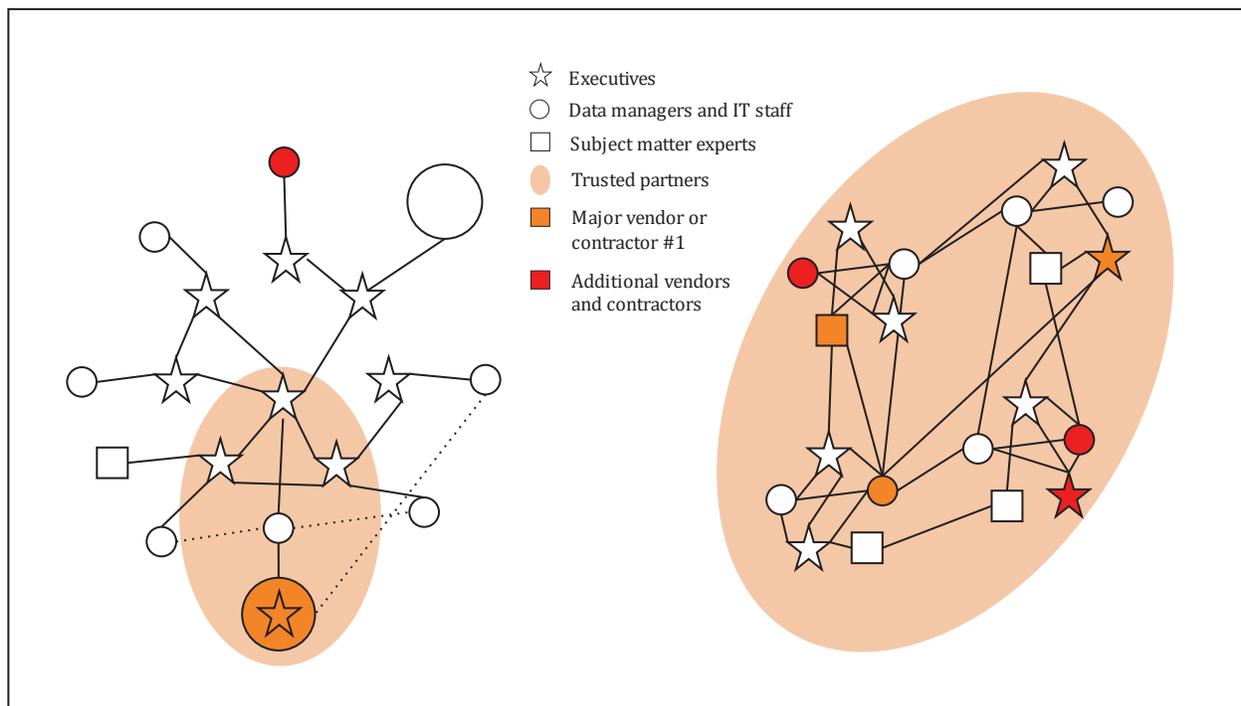
How we often organize
community IT projects

How we need to organize
community IT projects

On the left, executives are alone "at the table" or are accompanied by a relatively small representation of the technical expertise necessary to implement the desired work. Data managers and IT staff are relegated to the periphery, functioning primarily as advisors to the executives. Major vendors and contractors have contractual obligations that encourage self-

protective, unitary behavior. They prefer to provide a single point of contact to the network that is directly occupied by or strongly monitored by the vendor's executive lead. Indeed, the major vendor (dark orange) has incentives to protect its stake by working constantly to exclude other solution providers (red) from the network. The network is weak on social capital, to the point that most of the "stakeholders" can readily perceive that they are outside the circle of trusted major partners (pale orange). Transaction costs (such as information discovery and skill acquisition) are high for those outside the circle of trust.

On the right, in the "team-of-teams" model, executives, technical personnel, and subject matter experts (who might include program staff, evaluators, and other internal and external consultants) are all interacting regularly around the development of the needed data systems. The web of connections is rich in social capital; cross-team and inter-disciplinary communication is profuse. Expertise is widely available throughout the network, and technical staff from partner organizations are in regular direct contact with each other. (A subteam or user group composed entirely of technical personnel would be a reasonable part of such a model.) The circle of trust is inclusive and diffuse; no one is left out, and transaction costs are reduced for all. The network's openness extends to dispersing major vendors' personnel loosely throughout the network. Other vendors and contractors are included in the network and encouraged to contribute to standard setting and maximizing interoperability.

In a team-of-teams approach, the CDT would convene the expertise of a wide range of stakeholders and facilitate standard-setting among large stakeholders and software vendors small and large. Organizations would continue to choose between vendors for specific purposes, but the CDT environment would coordinate the "sports league" whose "off-the-field" coordination of "the rules of the game" and, importantly, social relationship-building (read: post-game Beer City USA activities?) among friendly competitors. Such work would promote on-the-field competition that is more productive for the goals of the community.

Note that, for simplicity's sake, **Figure 2.3** takes individuals' roles and specialties as given. This assumption does not reflect the important role of professional skill development for individuals and teams in creating greater fluidity among roles, blurring the distinction between stars and circles in the figure. Investments in long-term, in-place professional development can create a common foundation of skills, conceptual vocabulary, shared conventional practices, and mutual knowledge. New and more productive forms of individual and team specialization can build on that new foundation. **Section 3** of this paper expands on this idea.

3 Become a Deliberately Developmental Community with a technology focus
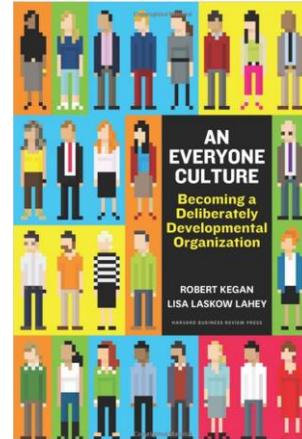
3.1 Foster Deliberately Developmental Organizations (DDOs)

Kent County will see little benefit from the pursuit of shared measurement if we don't develop a strong culture of learning and boost our capacity to produce, consume, analyze, and act on data. At present, the necessary technical skills are weak across the board and especially scarce in the social sector. But it is not necessary to recruit a few all-stars from Ann Arbor, Lansing, or Chicago, nor is it necessary to wait for the next generation of West Michiganders to assume the mantle, though planning for their arrival as leaders is critical. The people we already have can learn, and they can learn to love learning—to have a "growth mindset."

In *An Everyone Culture* (Kegan et al., 2016), the authors describe the benefits of "Deliberately Developmental Organizations" or DDOs (see sidebar). A DDO rejects the usual view that personal lives should be hidden and suppressed at work. Kegan and Lahey write that in all kinds of organizations, "most people are spending time and energy covering up weaknesses, managing other people's impressions of them, showing themselves to their best advantage, playing politics, hiding their inadequacies, hiding their uncertainties, hiding their limitations. Hiding."

Instead of hiding, a DDO's entire team is invited to bring life to work and to stop hiding worries, fears, health problems, and personality quirks. Far from an indulgent playground for spoiled children, a DDO is a challenging environment dedicated to fostering a growth-oriented mindset in its people. The goal isn't just individual growth—it's team growth. The culture isn't just a means to generate revenue; it is the reason for generating revenue at all: to maintain the growth culture. Instead of settling for stagnant practices, the DDO members work together constantly to improve their culture and its practices. In three case studies, the authors find that the DDO features aren't an expensive luxury for those enjoying rich revenue streams; they are actually the strategy that generates those rich revenue streams.

An Everyone Culture

From An Everyone Culture: Becoming a Deliberately Developmental Organization by Robert Kegan and Lisa Laskow Lahey, page 88:

"Think about your own organization. Is your company designed to support its employees' development as people (and not only their careers)? If we walked up to a random member—a leader, a manager, a support staff member—would she say yes to any of the following questions?

- Does your organization help you identify a personal challenge—meaningful to you and valuable for the company—that you can work on in order to grow?

- Are there others who are aware of this growing edge and who care that you transcend it?

- Are you given support to overcome your limitations? Can you name or describe this support?

- Do you experience yourself actively working on transcending this growing edge daily or at least weekly?

- When you do become a more capable version of yourself, is it recognized, is it celebrated, and—when you're ready—are you given the opportunity to keep growing?

3.1.1 Promote a growth mindset and reduce resistance to constant learning

The "growth mindset" is an important pillar of the Deliberately Developmental Organization (DDO). Decades of research by Carol Dweck and her colleagues and peers prove that our ability to learn is strongly affected by how much we *believe* we can learn. Carefully designed and validated studies show that people who believe they can get smarter, do, while those who believe intelligence is fixed don't improve at comparable rates. Now, Kegan and Lahey's *An Everyone Culture* calls attention to research showing that adults are not exempt from the mindset logic.

"Mental complexity" is a psychological measure of the degree to which we can comprehend and absorb a complex world. **Figure 3.1** differentiates the three adult plateaus (children and adolescents experience the first two).

Figure 3.1 The three adult plateaus (Table 2-1 on page 63 of *An Everyone Culture*)

The three adult plateaus

The socialized mind
- We are shaped by the definitions and expectations of our personal environment.
- Our self coheres by its alignment with, and loyalty to, that with which it identifies.
- This sense of self can express itself primarily in our relationships with people, with schools of thought (our ideas and beliefs), or both.

The self-authoring mind
- We are able to step back enough from the social environment to generate an internal seat of judgment, or personal authority, that evaluates and makes choices about external expectations.
- Our self coheres by its alignment with its own belief system, ideology, or personal code; by its ability to self-direct, take stands, set limits, and create and regulate its boundaries on behalf of its own voice.

The self-transforming mind
- We can step back from and reflect on the limits of our own ideology or personal authority; see that any one system or self-organization is in some way partial or incomplete; be friendlier toward contradiction and opposites; seek to hold on to multiple systems rather than project all except one onto the other.
- Our self coheres through its ability not to confuse internal consistency with wholeness or completeness, and through its alignment with the with the dialectic rather than either pole.

When researcher Keith Eigel compared performance evaluations of CEOs and middle managers across the mental complexity levels, those with greater complexity performed significantly better. "Taken together, the cumulative data supports the proposition that for those at a higher level of mental complexity, a complex world is more manageable." (Kegan et al., 2016, p. 73).

As shown in **Figure 3.2**, two major studies document the progression of mental complexity from "socialized" to "self-authoring" to "self-transforming." The small number of people at the topmost level in the figure is not a biological given; many more can attain such advancement through shared effort. A culture of personal and team development can help us.

Figure 3.2 Mental complexity in adults from two studies (Figure 2-6 on page 76 of *An Everyone Culture*)

FIGURE 2-6

Results from two large-scale studies of the distribution of levels of mental complexity among adults

Complexity

< 1%
< 1%
6%
7%
Stage 5
Self-transforming mind
34%
35%
32%
47%
Stage 4
Self-authoring mind
14%
8%
8%
2%
Stage 3
Socialized mind
5%
1%

Study A
n = 342

Study B
n = 497

Time

Sources: Study A: R. Kegan, *In Over Our Heads* (Cambridge, MA: Harvard University Press, 1994). Study B: W. Torbert, *Managing the Corporate Dream* (Homewood, IL: Dow-Jones, 1987).

3.1.2 Technology in an everyone culture

The developmental goals of a DDO are comprehensive. Self-understanding, confidence, health, teamwork, creativity, responsibility, strategic insight, aesthetics: all areas are of interest. So our focus here on technology skills should be understood as just an illustration of the broader workplace culture of a DDO.

One of the DDOs profiled in *An Everyone Culture*, Decurion, tracks talent development through "competency boards" on which employees' transition from trainee to master to coach is encouraged and celebrated.

> *"… [C]ompetency boards are literal poster boards hanging on the walls of every theater's back-of-house area. … Every crew member has the potential to gain certified competency in a series of about fifteen roles. … Standing backstage near a competency board feels a little like hanging out in a communal watering hole—a place where attention and conversation are focused as people stop to look at the board, seeing whether anyone has gained a new pin. The message is unmistakable: growing people into greater levels of skill, at the very least, is a focus of this community. Moreover, that growth is not a private set of goals between you and your manager but rather a public resource. Everyone knows what others are working on, and, as a result, everyone can step forward with support*

and feedback. Everyone also knows what kinds of capability, manifested in the roles, are required across the organization for the business to run effectively."
(Kegan et al., 2016, p. 147)

Among many roles, the Community Data Trust could support competency-board-like skill development across the community. The contemporary digital economy offers a vast field for skill development that too few people are challenged to consider, let alone master and then coach.

For example, consider the mundane matter of word processing. The classic book *The Mac is Not a Typewriter* was first published twenty-eight years ago (Wiliams, 1990; Williams, 2003), yet many young adults are still being taught to use word processors with typing and formatting habits inherited from the typewriter training of the 1920s. Too few professionals have had any exposure, let alone practice, with many of the basic features of Microsoft Word implemented in the creation of this very document:

- Styles can reformat multiple headings instantly, populate tables of contents, serve as targets for dynamic cross-references, and export as PDF bookmarks.

- Paragraph formatting ensures text that belongs together, stays together.

- A professional look is all about consistent grids. Invisible layout tables with nonbreaking rows of grid content for fast, flexible reformatting and prevention of unwanted page break locations.

- Special characters such as line breaks and nonbreaking spaces are important timesavers.

- The Review tools and Track Changes aren't just a last resort for shared editing; they are the best way to do individual work as well, to prevent inadvertent publication of embarrassing "notes to self" in the main text of printed or PDF documents.

- Almost any tedious, repetitive, time-wasting task can be avoided through automation.

These are just a few examples. Almost everyone in any office job has occasion to use Microsoft Word and its siblings. It only takes a few hours of training and a few weeks of practice to move from tedious drudgery to enjoyable mastery. Yet we meet people with decades of experience who've never been challenged to move beyond the small miracle of not typing Enter at the end of every line of a paragraph.

The above might seem trivial, until we consider that it's likely that hundreds of government, nonprofit, and philanthropic executives and staff people in Kent County spend dozens of hours per year being annoyed about doing tasks in Word that could be faster and more enjoyable. Supposing a modest average cost of $30 per hour for just 1,000 people countywide working 40 hours a year on tasks that could take 10, the cost is a whopping $900,000. And that's not to mention time spent managing files in the operating system, working in Excel, PowerPoint, Access, exporting content to PDFs and JPGs, posting the content to the web, and so on. **Figure 3.4** on page 30 lists a sampling of important "competency board" skills beyond Word.

In an everyone culture, we wouldn't need to haul 1,000 hypothetical trainees in for coursework. We'd just need to change the conversation, train someone in each network how to self-teach and to learn by teaching others, challenge everyone to learn more, and reward both learning and coaching of others. This is what DDOs do for all kinds of competencies, and it is transformative.

3.1.3 Beyond the basics

Supposing we agree that more learning about productivity software would be valuable, what does that have to do with the Community Data Trust concept? Surely the skills needed are not comparable?

Actually, these basic computer skills are quite comparable and transferable to data analytics and "big data" work:

a. At a purely **technical skill** level, the learning required to get past superficial usage of Microsoft Office opens conceptual windows into understanding enterprise data systems. Excel is a gateway to Tableau, Tableau is a gateway to SQL, and SQL is the gateway to fundamental insights into community data systems.

b. At the **recruitment and retention** level, the process of developing a learning culture will also reveal those who have a special aptitude and appetite for technical work. Have you been advertising for a database administrator for years, but are unable to pay enough to attract strong applicants? Have you ever offered your loyal, committed administrative assistant or your bookkeeper the chance to learn database administration?

c. At the all-important **culture** level, technology training is an effective icebreaker in the painful process of learning not to hide our weaknesses and ignorance from our teammates. The same "stop hiding" ethic of mutual aid and constructive critique that makes learning software together fun can spill over into conversations about organizational strategy and interpersonal relations with teammates.

The example of Microsoft Word skill development is highly relevant for another reason. While a long-term goal for the CDT is a full-featured web application to produce modular data sharing agreements (see section 4 on page 32), the project to develop modular, rapidly customizable, general-purpose data sharing agreements can begin as an automation-light collection of Word documents, along with Excel spreadsheets or Access databases.

3.2　Lay a shared community-wide foundation of core concepts and skills

Figure 3.3 is a hypothetical illustration[2] of the distribution of technology skills by role in a typical small- to midsize nonprofit. In the status quo (red bars), the skill distribution is strongly skewed to the right, where the more technical roles are found, and the overall level of advancement is low. In our recommended future, newly acquired skills (orange bars) would raise the overall skill level and reduce the rightward skew somewhat.

Figure 3.3 Software skill ranges by role

Software skills for prospective CDT members

← Increasing role fluidity →

Legend:
- Desireable developer skills
- Recommended advanced competency
- Existing advanced competency
- Existing base competency

Roles: Executive Director, Program Staff, Admin. Staff, Designer, Communications, Business Manager, IT Staff
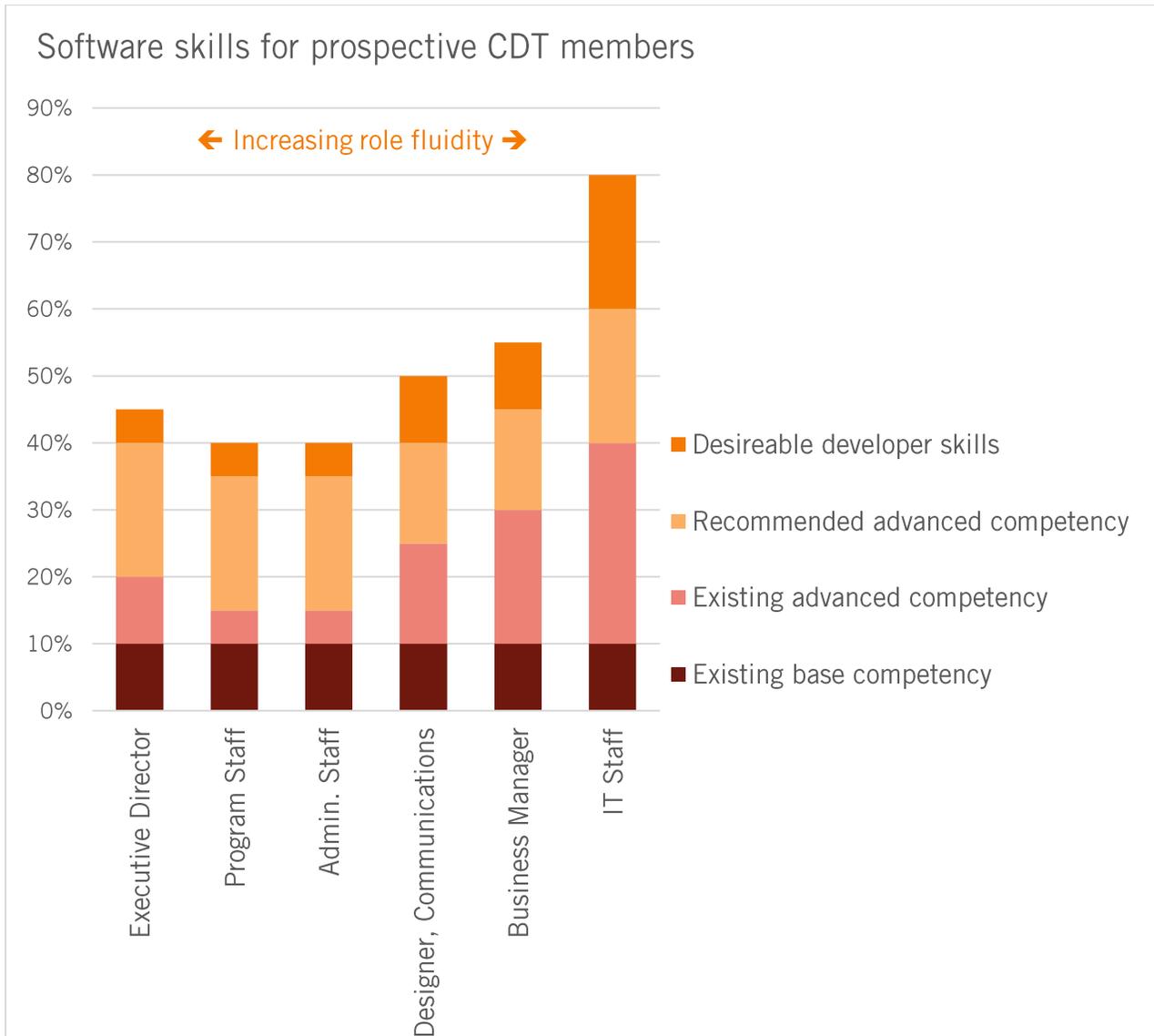
Figure 3.4 breaks the hypothetical categories into a list of specific skills, several of which we'll describe in greater detail below.

---

[2] It's not real data—but a personal technology skills survey could meet a community need and produce real data on this topic.

Figure 3.4 Expanded core of technical skills useful for stronger community data systems

Role *bottom = basic; top = advanced*	Executive Director	Program Staff	Admin. Staff	Graphic Designer, Comm. Director	Business Manager	IT Staff
Coding new applications from scratch						○
Predictive analytics (statistics)	○	○			○	○
IT platform management			○	○	○	◉
Web service calls (API exchanges)			○	○	○	◉
Data warehouse for analytics					◉	◉
Relational database modeling	○		○		◉	●
"Citizen developer" app customization[3]	○	○	◉	○	◉	●
Content management system (CMS)[4]	○	○	◉	●	○	◉
Online forms and survey data collection		◉	◉	○	◉	●
Skilled client for professional developer	◉	◉	◉	◉	◉	◉
Data visualization & business intelligence	◉	◉	◉	◉	◉	◉
Accounting transaction management[5]	◉	◉	◉	◉	●	◉
MS Word styles and automation	◉	◉	◉	●	◉	◉
Complex URL formation (key-value pairs)	◉	◉	◉	◉	◉	●
Password vault / account proliferation	◉	◉	◉	◉	◉	●
Cloud file syncing tools	◉	◉	◉	◉	●	●
MS Excel PivotTables	◉	◉	◉	◉	●	●
MS Excel spreadsheets as databases	◉	◉	◉	◉	●	●
Basic HTML use for email and web pages	◉	◉	◉	●	◉	●
Financial spreadsheets (basic calculations)	●	◉	◉	◉	●	●
Word processing	●	●	●	●	●	●
Email and electronic calendaring	●	●	●	●	●	●
Operating system / file management	●	●	●	●	●	●

LEGEND

● Currently common skills
◉ Recommended skill expansion
○ Valuable additional skills

---

[3] For online database applications such as Salesforce and QuickBase.
[4] Common examples are systems like WordPress and Drupal.
[5] E.g. direct entry of personal expenses, including mobile receipt storage

3.3 Do succession planning for a data systems personnel pipeline

If it is hard to imagine an all-volunteer team or modest stipends for staffing a Community Data Trust with existing personnel, it is only because we have very limited insight into the breadth and depth of our community team. Along with certifying compliance and promoting training and skill-building, the CDT should encourage and document 2-deep succession planning for important technical roles across the community.

Strong professional associations generally have a past president, a current president, and a president-elect, so that the current occupant of the role has formal access to and responsibility for both her predecessor and her successor. A similar model should be considered both within local DDOs and across our DDO landscape as a Deliberately Developmental Community.

Some possible resources for succession planning:

1. A CDT-maintained list of desirable skills and experiences for technical and executive leadership roles in CDT member organizations.

2. A list of all local personnel who demonstrably have these skills and experiences.

3. An apprenticeship or "bench strength" program to link each important data feed and related services with the names and contact information of multiple people who have been fully oriented about how to support and maintain the service.

4   Organize expertise and automate DSA development and auditing

Data sharing agreements are tedious, complex legal contracts that can require a great deal of effort to specify and monitor, especially when they are the data provider's first foray into sharing data. Federal and state legislation is generally designed to help facilitate useful data sharing and limit liability, but risk-aversion and high transaction costs generally limit organizations' willingness to share data.

This problem is not wicked, only complex, and the solution is not novel; it's the same solution of economies of scale, automation, and division of labor that we have exploited across multiple sectors and industries.

4.1   Picturing the Community Data Trust as a network of data feeds

The typical status quo for shared measurement work is shown in **Figure 4.1**. Each data provider maintains independent relationships with data requesters, and vice versa; few economies of scale are achieved, data sharing agreement (DSA) development is a redundant, time-consuming process, and data provision and permissions management is highly manual (such as emailing spreadsheet files).

Figure 4.1 Typical status-quo data sharing structures

Typical data sharing
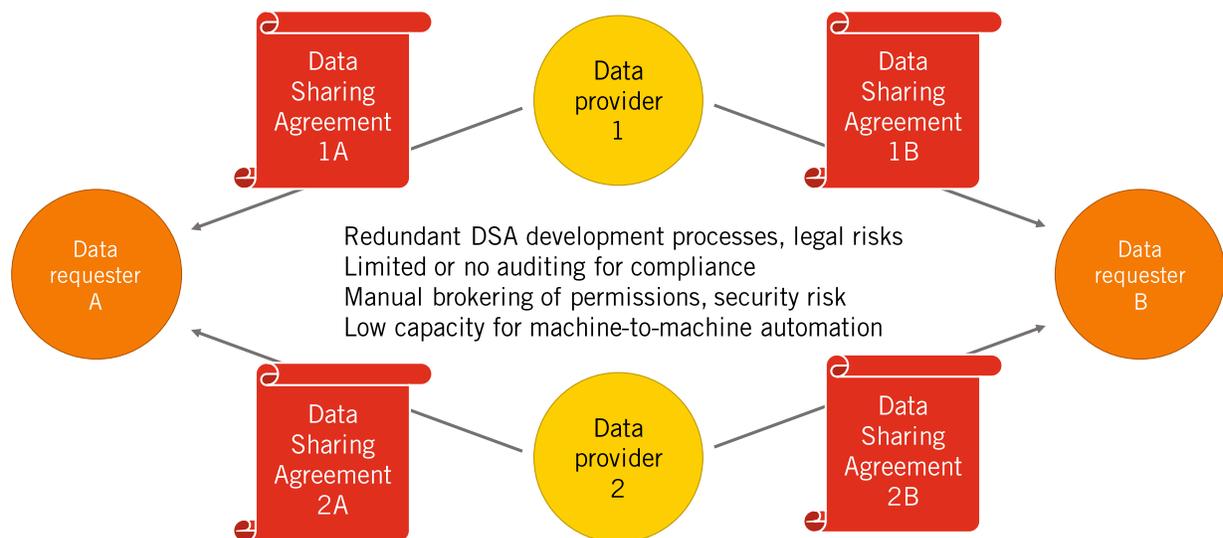
As shown in **Figure 4.2**, by pooling legal and technical expertise and resources, the Community Data Trust modularizes, centralizes, and streamlines data sharing agreements (DSAs) *without storing copies of sensitive data*.

Figure 4.2 Improved data sharing adaptability and efficiency through the Community Data Trust

Improved data sharing
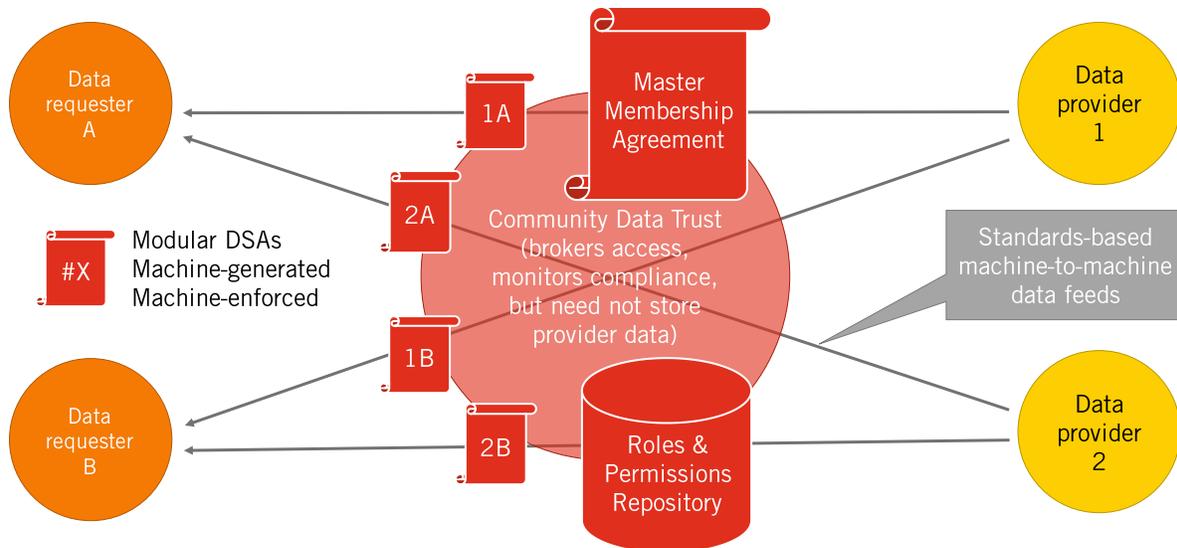
Efficient, modular DSAs with superior legal design
Professional auditing and logging for compliance
Machine-brokered permissions, better security
High capacity for machine-to-machine automation

Data requester A

Data provider 1

1A

Master Membership Agreement

2A

Community Data Trust (brokers access, monitors compliance, but need not store provider data)

#X | Modular DSAs Machine-generated Machine-enforced

Standards-based machine-to-machine data feeds

1B

Data requester B

2B

Roles & Permissions Repository

Data provider 2

As members of the CDT, data providers and requesters certify their compliance with the Master Membership Agreement (MMA, red scroll at top center), which also specifies the array of standard data-sharing protocols agreed to by CDT members. Data providers and requesters then use a form-based process (which might be fully automated in later phases of the CDT) to negotiate and complete modular DSAs (small red scrolls with "#X" designations, such as "1A"). Each DSA is a document created by combining language previously vetted and agreed to by members' legal counsel as part of the CDT membership process. Section 4 suggests a system for producing modular DSAs to reduce transaction costs.

Each data feed (gray arrows) has its own micro-DSA, which specifies which roles have access to what data. Access is then brokered and audited by the CDT system, with reference to a central repository of roles and permissions (red cylinder at bottom center). But shared data itself is held only by the members, as per the terms of each DSA.

A third illustration, **Figure 4.3**, demonstrates how later enrichments of the CDT concept might provide members with the means to achieve "on-the-fly" integration of public, de-identified or aggregate data sets into analytics models.

Figure 4.3 More CDT value added through integration of public data feeds and repositories and data-handling tools

More CDT value added

Infrastructure supports integration of public and team feeds
Automated data handling service reduces burden on providers and requesters
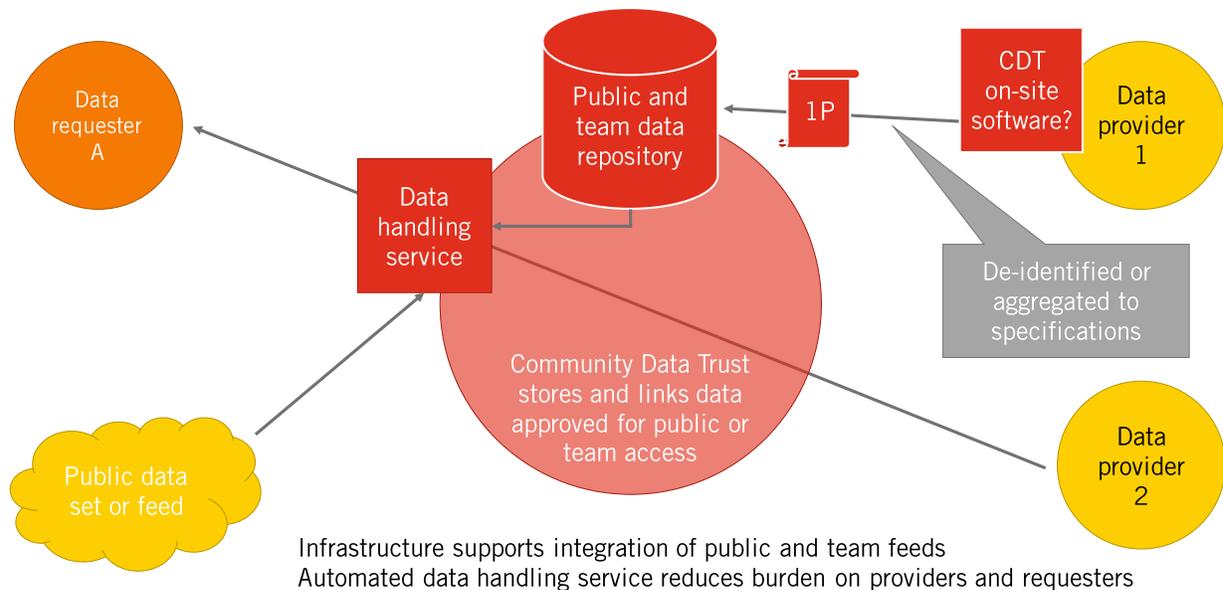
This approach builds new public-access capacity on the existence of the non-public CDT structure; for example, the creation of public data streams can be piggybacked on non-public projects. It also enables public-interest governance of the public repository independent of any other single partner in the network.

4.2   Understand legal standards for data security

The following is an expansion and commentary on "How secure does it need to be? A quick guide to compliance," a presentation to the inaugural "Big Data Ignite" conference at DeVos Place in Grand Rapids by Elliott Church, J.D., of the local law offices of Kreis Enderle (Church, 2016).

Data security standards are affected by the following laws and legal agreements:

1.   **Education:** FERPA ("Wikipedia: Family Educational Rights and Privacy Act of 1974," 2016). Quoting from Wikipedia:

> *"FERPA gives parents access to their child's education records, an opportunity to seek to have the records amended, and some control over the disclosure of information from the records. With several exceptions, schools must have a student's consent prior to the disclosure of education records after that student is 18 years old. The law applies only to educational agencies and institutions that receive funding under a program administered by the U.S. Department of Education. Other regulations under this act, effective starting January 3, 2012, allow for*

greater disclosures of personal and directory student identifying
information and regulate student IDs and e-mail addresses.

"Examples of situations affected by FERPA include school employees
divulging information to anyone other than the student about the student's
grades or behavior, and school work posted on a bulletin board with a
grade. Generally, schools must have written permission from the parent or
eligible student in order to release any information from a student's
education record."

FERPA's rules require detailed attention to students' and families' confidentiality, generally by requiring explicit informed parental consent. However, the law's intent is not to hamstring educational benefits to parents and students (Petrila, 2011). The useful "Forum Guide to Protecting the Privacy of Student Information" documents how to share information outside an education agency (National Center for Education Statistics, 2004). FERPA explicitly allows sharing of identifiable data with school officials and partners for a "legitimate educational interest" (Section 4.B.) and release of de-identified data to researchers (Section 6.E.) Section 6.D describes the conditions for release without prior consent:

Within the agency or school, education records may be released and used
by personnel who are considered to have a legitimate educational interest
or need-to-know without prior written consent of the parent. Section 4
contains guidelines regarding this type of release. Examples of personnel
who may have authorized access to the student records include research
and evaluation directors and service providers or coordinators of special
programs in which the students participate.

Explicit policies and practices on data release can provide a context to reinforce "legitimate educational purposes." Among these is regular submission of research projects to university or public Institutional Review Boards (IRBs), which can weigh the risks and benefits of data release and approve release of low-risk, high-benefit data without placing an unnecessary added burden on parents and guardians of reading and interpreting dense, complex statements of consent.

2. **Health:** HIPAA ("Wikipedia: Health Insurance Portability and Accountability Act of 1996," 2016). HIPAA's Privacy Rule requires "covered entities" and their "business associates" (contractors) to protect the confidentiality of Protected Health Information (PHI), which includes any part of an individual's medical record or payment history. The Privacy Rule also grants patients the right to request access to and corrections of PHI. Meanwhile, the Security Rule requires administrative, physical, and technical safeguards to the security of Electronic PHI (EPHI). HIPAA documentation is well-developed and will not be further elaborated here. Note that PHI is not the relevant legal category when health data is held by an educational institution in educational records, even if the educational institution provides health services; HIPAA must be complied with in the context of health service provision, but FERPA controls data release issues (Petrila, 2011; U.S. Department of Health and Human Services & U.S. Department of Education, 2008).

3. **Finance:** GLBA ("Wikipedia: Gramm–Leach–Bliley Act of 1999," 2016). This law includes a Privacy Rule that requires financial institutions to protect customer financial information and to send customers an annual notice of the financial institution's privacy policy. It also includes a Safeguards Rule that requires designated personnel to monitor safeguards, risk analysis, security audits, and attentive updates to safeguards.

4. **Personally identifiable information (PII):**

   The Office of Management and Budget (OMB) defines personally identifiable information [PII] as "information which can be used to distinguish or trace an individual's identity, such as their name, social security number, biometric records, etc. alone, or when combined with other personal or identifying information which is linked or linkable to a specific individual, such as date and place of birth, mother's maiden name, etc." (General Services Administration, 2015)

   Though the language above is federal, state laws govern the definition of security breaches exposing PII. In Michigan, the relevant law is MI 445.72 ("Michigan Legislature - Section 445.72," 2004). California tends to set the pace for changes in PII protection. Michigan does not require notification if lost data was encrypted, but as of 2017, California now requires notification for the loss even of encrypted data (Lazzarotti, 2016).

5. **Trade secrets:** organizations may be required to protect trade secrets through legal vehicles other than those specified in direct inter-organizational contracts. For example, individuals may have employment contracts that require them to protect an organization's trade secrets even after their employment is terminated.

6. **Contractual language:** Any contractual language between parties must be honored.

Church (2016) also provided helpful insight into the legal standard for judging the amount of investment required to show appropriate care. The precedent is the "Learned Hand Formula," named for its author Judge Learned Hand; it's also referred to as the "calculus of negligence" ("Wikipedia: Calculus of negligence," 2016). Paraphrasing Church:

If the potential harm is of size X, and probability it happens is Y, then the level of care I need to take is X*Y. For example, if the potential harm would cost $100,000 and the probability harm occurs is 1%, then I'd better exercise at least $1,000 in appropriate care to satisfy a court that I've paid attention.

According to Church, particular areas of concern for data security include:

- Failure to take reasonable steps to protect PII
- Failure to encrypt information in transit
- Allowing anonymous access to PII
- Failure to adapt readily-available security measures
- Insufficient monitoring for unauthorized access
- Preserving information longer than needed

4.3   Compare risks of isolated partnerships to those of a shared community system

This proposal is based partly on the strong expectation that pooled legal and technical resources can reduce liability for CDT participants, by both:

- Bolstering the best-practice rationale for organizational data sharing policies and procedures, and

- Providing external certifications and auditing services to detect and reduce risk on a regular basis.

However, this expected risk-reduction should be subjected to evaluation by prospective CDT members and compared with actual experiences of past data security and confidentiality breaches or emergencies at prospective CDT members and in other settings.

4.4   Convene or retain legal expertise charged with serving community goals

Federal law is intended to provide a framework to enable responsible data sharing, not to punish it (Petrila, 2011). But organizational leaders are naturally and reasonably risk averse, and an organization's legal counsel may find it less risky to recommend refusal to share over uncertainty.

The role that's lacking is community counsel: a legal mind or minds tasked with representing the community at large and counting the opportunity cost of not achieving responsible data sharing. Such counsel should not be responsible to any organization, government, or association; the counsel's job should be to provide legal advice, document templates, and other outputs to improve the quality and reduce general liability around community data systems.

4.5   Build an automated system for DSA development

A thriving community might need literally hundreds of bilateral and multilateral data sharing agreements (DSAs). Like any output, their production can be improved by professionalization, converted into the combination of modular components, an assembled by automated systems.

This proposal may sound complicated, but technically, it's just a moderately glorified version of the standard word-processor-based mail-merge process that offices have used since the 1980s. The LegalZoom service has been operating a document-generation service on a similar architecture for many years; we propose a supervised application of such technology.

Here are the steps:

1.   Dissect existing data sharing agreements into named components; a previously developed Master Data Sharing Agreement (Carlson et al., 2011) is comprehensive enough to provide a "backbone" with which other modules can be compared and integrated.

2.   Input modules as data objects into a database, perhaps a content management system (CMS). Each module includes placeholder field codes for variable text, such as the names of the parties to the DSA. Modules also include appropriate paragraph styles to ensure readable headings and page breaks in the final product.

3.   Further populate the database with names of standard variables, such as roles, permissions, and names and originating URLs of the relevant data feed(s).

4.   Create a form that collects the names of parties to the agreement (data-contributing and data-requesting organizations, DCOs and DROs)

    a.   The DCO selects applicable modules

    b.   The DRO approves or suggests revisions until the DCO agrees to them

5.   Create a data form to connect the DCO's data feed to the CDT system and allow the DCO representative to select which roles have which access to which fields of the feed. The simplest DSA's will give one role all access to all fields. This information is stored in the database and populates the roles-and-permissions-by-data-element module, known as "Attachment C" in the Master DSA (Carlson et al., 2011).

6.   Create a script that merges the variable text into the selected modules and then outputs the modules to a master PDF.

7.   Circulate the PDF for review and signatures.

8.   Integrate any lessons learned into the data system and repeat steps 4 to 7 until successful.

Fully developed, step 7 might further include an authentication system to allow electronic approvals by authorized executive or legal counsel users at DCOs and DROs.

5 Inspire technical best practices among members and vendors

The general practice of data sharing comes with a collection of high-pressure, high-temperature, risky conditions that rightly give executives pause in signing on the Data Sharing Agreement's dotted line. Potential legal liability affected by the practice of the staff of partner organizations; technical security issues; the high cost of getting locked in to a predatory or incompetent software vendor; the risk that all these will work, yet yield no actionable insights.

A major goal of the Community Data Trust concept is to reduce transaction costs and lower the figurative temperature and pressure in each of these areas. We can do so by encouraging and validating the diffusion of best practices among partner organizations, their staff, and their contractors and vendors.

5.1 Provide certified training and user groups for CDT member personnel

5.1.1 Certify user training in HIPAA-, FERPA-, and CDT-compliant data handling

Federal privacy legislation is strict and detailed, but it is not intended to hamper legitimate data sharing. Training and support procedures are sufficient to reduce risks and enable compliance. However, large health care and educational organizations are better equipped to provide such training and support.

The CDT could set standards, invite and communicate shared training opportunities, and provide organizations—especially smaller nonprofits—with confirmation that their staff have been trained in ways that meet a defensible standard.

5.1.2 Schedule and promote user group meetings for a range of software and data systems

One of the best means to promote efficient, skilled use of software is to foster the creation of in-person and virtual user groups. Peer-to-peer support with personal contact is a much more effective means of support than impersonal Internet searches.

5.1.3 Provide wiki-style group-authored knowledge base for all CDT resources

"Crowdsourcing" is a cost-effective and largely accurate means of sharing information. The CDT could host or organize sponsorship for hosting a wiki or similar system for all CDT member organization staff to discover solutions to technical problems.

5.2 Provide shared compliance auditing capability

Most organizations are depending on internal resources to monitor their compliance with the terms of DSAs. A robust CDT implementation would provide better risk management and greater system trustworthiness by pooling auditing resources across the CDT membership. If successful, CDT members could also benefit from further reduction in liability by following standardized crisis-management protocols that include outside parties to advise and witness organizational response.

5.3 Select vendors and contractors that employ best practices

Even given decades of professionalization, there is a great deal of variation in the degree to which software developers and major software vendors adhere to best practices. CDT members should be encouraged and enabled to insist on software that provides most or all of the following:

- **Versioning:** Software version numbers and release dates should be explicit and easy to verify. (Some web-based software may use release dates only.)

- **Release notes:** Each software release should come with detailed, dated release notes on new features enabled, bugs fixed, and known issues remaining.

- **Knowledge base and support forums:** Known issues and major questions should be answerable by searching a knowledge base with stable URL links and by asking questions of support staff and fellow users in community forums in which links to knowledge base articles can be shared and discussed.

- **Application Programming Interface (API):** Any software—but especially web-based software—should have a thoroughly documented API available to the public or at least to authenticated developers. The API should be versioned.

In addition to the above, there are some "wish list" items that can distinguish excellent contractors and vendors from the crowd:

- **Participatory beta testing:** Ideally, advanced staff at CDT member organizations would be able to participate in beta-testing of upcoming versions of systems used by CDT members.

- **Road map:** Few vendors provide transparent road maps for future features and platform changes. Those that do are distinctively disciplined.

- **Shared or open source code:** Proprietary software is often the best solution, but only when supported by a robust organization. Smaller firms and contractors with significant risk of failing can boost confidence by posting their codebase in an online code repository like GitHub, either publicly or with private permissions for select customer representatives.

- **Collaborative development:** Contract developers and some smaller vendors may use collaborative systems (such as PivotalTracker from PivotalLabs) to allow CDT member staff to closely guide and approve development through direct proposal of user stories, evaluation of test code, and approval to release production code.

Table 5.1 Sample online resources for major software vendors

Product	Knowledge Base	Release Notes	API documentation
Salesforce CRM	Support	Spring '17	Docs (v 39.0)
Qualtrics survey research	Support	2/1/2017 *lacks known issues!*	API v3
QuickBase database	Help	March 2017	API reference *unversioned!*
Tableau data visualization	Knowledge Base	Release Notes	REST API 10.2 *Server product*

5.4   Promote the use of API-based software

5.4.1 What is an API?

According to Wikipedia today:

*In computer programming, an **application programming interface** (**API**) is a set of subroutine definitions, protocols, and tools for building application software. In general terms, it's a set of clearly defined methods of communication between various software components. A good API makes it easier to develop a computer program by providing all the building blocks, which are then put together by the programmer. An API may be for a web-based system, operating system, database system, computer hardware, or software library. An API specification can take many forms, but often includes specifications for routines, data structures, object classes, variables, or remote calls.*
("Application programming interface," 2017)

Microsoft Windows and the Apple MacOS had APIs all through the 1990s, but Google set the precedent for the modern web-based API with the release of the Google Maps API in June 2005 ("Google Maps APIs for Web," 2017). Google transformed web application development by allowing any programmer—no matter how small or large—to access instructions for incorporating Google Maps into any web application.

The web address (Uniform Resource Locator, URL) for each Google Map is itself a simple example of an API. To get a simple map, just follow this pattern, replacing the italicized values in brackets:

> https://www.google.com/maps/place/
> *[address here, substitute + for spaces]*/
> @*[latitude of map center]*,
> *[longitude of map center]*,
> *[zoom level 1-20]*z/

For example, the KConnect offices at the KISD Conference Center are here:

> https://www.google.com/maps/place/
> 1633+E+Beltline+Ave+NE,+Grand+Rapids,+MI+49525/
> @42.9917934,-85.5938342,
> 17z

A similar fill-in-the-blanks strategy can be used to assemble a URL that pulls data from a Qualtrics survey or that inserts a new record into a QuickBase database table.

The result is scalable, fluid, creativity-inducing interoperability among systems. For example, the Qualtrics survey system is a powerful form-based data collection system, but doesn't do custom reporting well. QuickBase is a powerful data storage and reporting system, but doesn't do long, complex data entry forms well. Tied together with their APIs, the two tools can solve a variety of problems. An employment application in Qualtrics can automatically insert records into a QuickBase table of job applicants, while references from past supervisors are inserted into a related table, including ratings. A custom report in QuickBase shows the top-rated applicants at the top of the list for each position.

5.4.2 Good and bad practice with APIs

The mere existence of an API is not enough to ensure useful software; the relevant system's developers have to respect it. This means using documented API calls for all internal functions of the front end, as shown in **Figure 5.1**. When the API is used consistently, the output from a
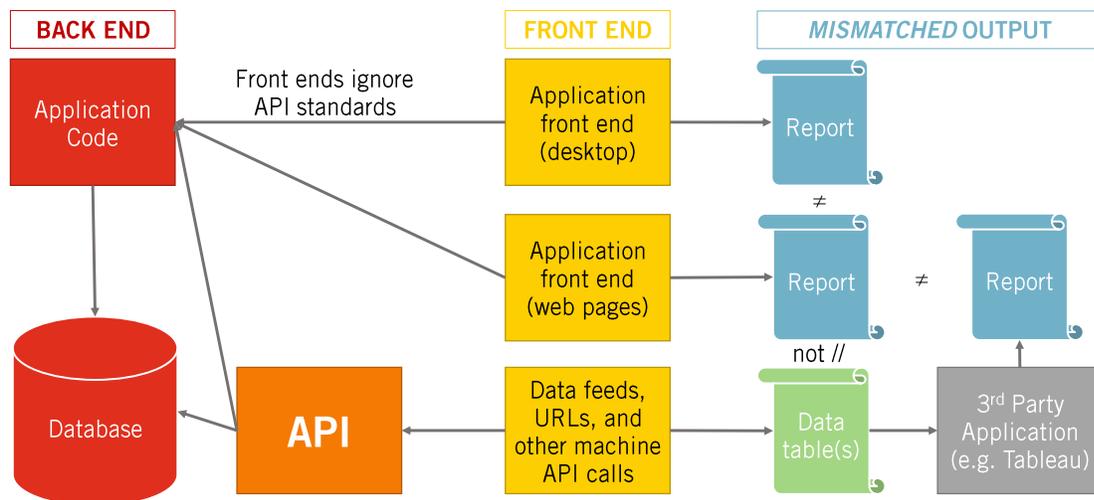
machine-to-machine data feed (green) can allow a third party application (such as Tableau) to reproduce the same results as a built-in report from the developer's designs.

Figure 5.1 Good API practice: respecting the API produces matching output due to matching API calls by application front ends
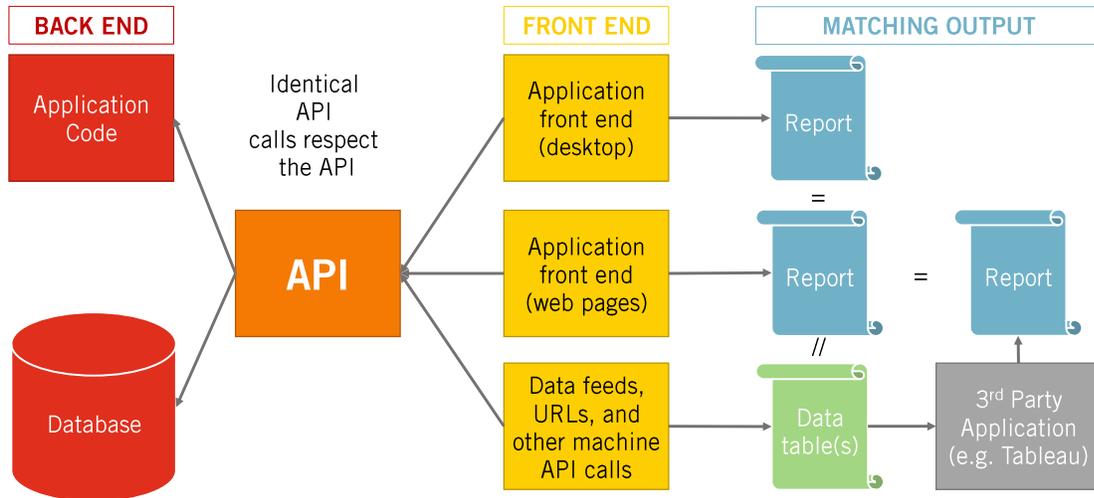
On the other hand, as **Figure 5.2** shows, failure to respect and use the API can result in built-in reports that can't be replicated, frustrating collaboration and limiting performance.

Figure 5.2 Bad API practice: disrespect the API, get mismatched output

Glossary of terms

API	Application Programming Interface, a documented set of commands or methods that can be used to control a software system.
Cloud	"The cloud" of the 2010s differs significantly from the "online hosting" of the 2000s. Cloud resources are not just "online," they are provided by computing-as-a-utility systems that migrate among servers as demand rises and falls. Cloud systems offer reduced expense and greater efficiency on a subscription basis, but the subscriber's data does not necessarily exist on a specific, predictable physical system. Whether this fact enhances or undermines data security is a matter of highly application-specific considerations.
DCO	"Data Contributing Organization."
DRO	"Data Requesting Organization."
DSA	"Data Sharing Agreement."
FERPA	Family Educational Rights and Privacy Act of 1974; governs data sharing by educational institutions.
HIPAA	Health Insurance Portability and Accountability Act of 1996; governs data sharing by health care providers.
KConnect	Kent County's collective impact network backbone organization, responsible for convening workgroups and inspiring experimentation with new strategies to help Kent County children achieve prosperity.
Mashup	The combination of multiple online applications, usually by linking them through API calls.

Appendices

Appendix A: Contributors

5.4.3 Data and Capacity Workgroup members

Mel Atkins, Grand Rapids Public Schools (Member, 2014-)
Erika Bolig, Michigan Department of Education (Member, 2014-)
Neil Carlson, Calvin College Center for Social Research (Co-Chair, 2014-)
Paul Isely, Grand Valley State University (Member, 2014-)
Julie Klausing, Great Lakes Health Connect (Member, 2014-2016; Co-Chair, 2017-)
Doug Kyle, Grand Rapids Public Library (Member, 2016-)
Luis Lugo, Doug & Maria DeVos Foundation (Co-Chair, 2014-2016)
Mark Maynard, Kent Intermediate School District (Member, 2014-)
David Riley, The Right Place, Inc. (Member, 2016-)
Rebeca Velasquez-Publes, Health Net (Member, 2014-)

5.4.4 Other contributors

Elliott Church, Kreis Enderle
Jodi Petersen, Grand Valley State University, Johnson Center for Philanthropy

Appendix B: A brief introduction to the Community Data Trust concept (February 2016)

The following document was originally prepared by:

> *Neil Carlson (Calvin College Center for Social Research),*
> *Jodi Petersen (GVSU Community Research Institute), and*
> *Mark Woltman (KConnect)*

for discussion by the KConnect Data & Capacity Workgroup in February 2016

Common Problem

Data sharing is an oft-desired and oft-abandoned pursuit. Data sharing can allow for better outcomes for community members by allowing:

1. organizations to better assess the impact of their programming,
2. policy makers to better understand community needs, and
3. funders to increase the impact of their investments.

Due to the ethical and legal concerns of consent and data security, along with the technological needs required for data matching, analysis, and reporting, data sharing is often complicated and costly. Data-sharing agreements are often project-based, requiring separate consent for each project, sharing only for specific analyses, and sharing only with parties defined a priori. Systems are usually dismantled at project completion, despite great sunk costs.

Proposed Solution

As data sharing efforts are time intensive, complex and costly, there is great potential value in achieving economies of scale, reducing the transaction costs of data sharing while maintaining the required legal, technical, and methodological rigor. One potential solution is a Community Data Trust (CDT), designed to better meet the data sharing needs of the community. The CDT would exist as a data-sharing entity, allowing data-sharing agreements to be made between data-contributing organizations (DCOs, or "contributors") and the Trust rather than with a specific project-related data-receiving organizations (DROs, or "receivers"). The Trust would be responsible, through human and software protocols, for controlling and auditing the release of data to receivers in accordance with the standards prescribed by contributors.

Data-Sharing Specifications

Data-sharing agreements would indicate the nature of the data shared and the attributes of permissible data-receiving organizations (in project type, organization type, level of aggregation, and so forth). Multiple methods of data contribution would need to be developed. Each data-contributing organization (DCO) would be responsible for obtaining consent to share data, verifying that consent, and only sharing consented data.

Each DRO would receive data only as approved in advance by the relevant DCOs. Data could be:

1. fed into various dashboard fronts with varying permissions (private, collective, and public),
2. provided as downloadable datasets to DROs:
   a. in deidentified individual level datasets,
   b. in aggregations to approved levels, or
   c. in individual level identifiable datasets.

Planning and Governance

Projects of this sort require both policies set in advance and interpretation throughout implementation, so a multi-stakeholder planning process and governance committee would be required. This committee should include program participants (both adult community members and parents of students), program front-line provider staff, nonprofit leaders, school data specialists, community data advocates, researchers, technology infrastructure specialists, and legal consultants.

Benefits

- Decreased transaction costs for DROs and DCOs
- Decreased consent burden for participants
- Improved data standards
- Increased access to research for resource-scarce organizations
- Increased access to community data
- Faster turnaround times through self-serve systems and automation
- Reinforced culture of and capacity for evidence-based standards for decisions and investments

Costs

- High non-project-specific infrastructure costs for set up
- High levels of community buy-in
- Diverse, sustained stakeholder engagement for planning and governance

Key Decision Points

1. System Scan – has anyone else done this? University or City?

2. Initial conceptualization – data sharing services provided

3. Staging/timeline (project management Gantt chart)

4. Consent/Legalities

   a. Modality – online vs paper

   b. Point of administration (at data contributing organization?)

   c. FERPA/HIPAA/CRIA

   d. Juvenile Assent

   e. Frequency

   f. Individual vs Aggregate Dissemination language

   g. Open use (not project specific)

5. Public Trust-Building

   a. Programmatic

   b. Individual

   c. Funder

   d. Academic

6. Data Sharing

   a. Agreements (in and out)

   b. Aggregate vs Individual data sharing (in and out)

   c. Aggregate vs Individual dissemination of results

   d. Methods of Data Sharing (APIs, Templates, Custom set ups, Paper Entry, CSV uploader wizard)

7. Data Security

   a. Hardware

   b. Infrastructure

   c. Access Roles

8. Hardware Infrastructure (Servers)

9. Software/Programming Infrastructure

   a. Individual matching algorithms

   b. External auditing

10. Staffing Requirements

11. Governance

   a. Legal counsel

   b. "Grey Area" decision making

12. Sustainability

   a. Ongoing hardware and staffing costs

   b. Business model for sustainable cost coverage

   c. Ongoing data changes, programmatic staff turnover

   d. Ongoing onboarding

13. Organizational capacity building (DCO/DRO needs for training for contributing data, interpreting findings, etc.)

KCONNECT
Fulfilled Potential. Community Prosperity.

Evaluation criteria for collective impact data sharing systems

KConnect Data and Capacity Workgroup
August 15, 2016

Task	What collective impact needs	Questions (Yes/positive answers preferred)
Governance	Persistent standards-setting body/ies that negotiate long-term common ground	1) Is intellectual property produced by the work of the community owned by the community? 2) Do customers own their data stored in the system? 3) Does the system design assist in decision support? 4) Does the vendor have a vision for setting standards? 5) Are there procedures in place to work with stakeholders to set common standards?
Reliability	Data is backed up, servers have failover provisions, and contracts provide for code and/or data retrieval in the event of vendor failure.	6) Do contracts and licenses protect customers' investments in the event the vendor fails or is acquired? 7) Does the vendor document its failover site (alternative system if the main servers fail)? 8) Is the cost structure sustainable over five years or more?
Security	Verify and trust: Permissions repository/ies with distributed, automated account setup based on organizational and role-based authentication: if your org and role are defined to have access and you can authenticate yourself, you have access	9) Can administrative roles be distributed without vendor intervention? 10) Can organizational administrators troubleshoot permissions issues easily without vendor support? 11) How are new users admitted to the system? Is initial access smooth and prompt?
Legal compliance	Systems that certify compliance with data sharing agreements and legislation (HIPAA and FERPA) through documentation of human users' certification and audit trails of potential and actual access by whom to what data.	12) Are data sharing agreements documented and accessible within the system? 13) Are data sharing agreements represented as business rules? 14) Is the system certified for HIPAA and FERPA compliance? 15) Are need-to-know restrictions enforced for identifiable data?

Task	What collective impact needs	Questions (Yes/positive answers preferred)
Accountability	Active audit-trail logs, with real-time records of access and instant reports from standard red-flag algorithms. Combined with standard-setting, this approach builds learning from past problems into business rules, and it helps protect innocent users and organizations from undue suspicion.	16) Are user access logs detailed and accessible? 17) Can business rules be defined and adjusted by customer administrators? [For example, if logs show a frequent user error, can administrators react by writing a data validation rule to prevent the error or warn users?]
Transparency	Regardless of ownership, partner organizations should have some insight into the platform(s) and tools used by vendors' developers.	18) Does the vendor disclose the vendor(s), language(s), platform(s) and toolkit(s) in use?
Matching and de-duplication	Systems that facilitate matching on standard key fields, provide code mappings for attributes, and provide deduplication protocols.	19) Does the system document and expose its data model in a way that ensures accurate matching of units on key fields? 20) Are there productivity tools for efficient handling of unmatched cases and for consolidating duplicate entries?
Intersystem interoperability	Open interoperability, based on publicly documented Application Programming Interfaces (APIs) for web services and similar tools. Public API documentation is compatible and desirable for proprietary systems as well as open-source platforms.	21) Is there a public API? 22) Is the API fully documented? 23) Is the API widely used and publicly evaluated by users? 24) Is the system designed for reliable machine-to-machine communication? 25) Can API calls be encrypted? 26) Are API calls logged and analyzed, including IP address and user authentication details? 27) Do vendor-provided front end systems (web sites and apps) use the same API calls as users? 28) Do API calls produce the same results as the front end(s)?
Data extract/ transform/ load (ETL)	Standardization and automation of ETL protocols protect data providers from inundation with requests, allowing staff to focus on quick online review of proposals for access to specific fields, geographies, and levels of aggregation.	29) Are ETL processes documented, logged, and reproducible? 30) Can ETL processes be implemented through the API? 31) Does the system include communications for users to request custom approvals from data providers?
Training & Support	Commit to constant learning at every level of the organization; choose products (or product platforms) with large, robust online communities of developers and users, with abundant evidence of idea-sharing and mutual aid.	32) Is the vendor's team accustomed to training and supporting users in organizations of the size and type expected by the collective impact initiative? 33) Do the vendor's software engineers have authority to answer questions and deliver solutions? 34) Is there a robust, visible user community with a knowledge base and discussion boards?

Task	What collective impact needs	Questions (Yes/positive answers preferred)
Data library quality	Access to appropriate public-record data sets	35) Does the vendor provide easy integration with US Census, Department of Education, and other important data sets? 36) Does the data library include appropriately fine-grained data? That is, is the level of data they have the level of data the collective impact initiative needs?
Public access	Collective impact systems mobilize public support, requiring public-facing systems	37) Does the vendor or product support public-facing dashboards or reports? 38) Can anonymous public users interact with data with analytics similar to those available to authenticated users?
Development	Provide professional-grade engineering systems, bug and issue documentation, versioning, release notes, feature requests, and product "road map."	39) Are release notes available for each version? 40) Does the vendor use an internal versioning system (e.g., Git)? 41) Are management and versioning integrated (that is, code approval automatically triggers production deployment?)
New small-scale development	Build internal "user developer" talent and generate capacity for customization.	42) Does the vendor encourage developer work by power users?
New large-scale development	Select and fund new large-scale development priorities based on distributed "user voice" systems.	43) Is there a user voice system? 44) Do engineers field test new internal ideas with user groups?
Ecosystem	Even when one or two vendors are prioritized, the system recognizes a need to interoperate with an ecosystem of a wide array of systems, both within the sector and through "mashups" with external best-of-breed tools.	45) Does the vendor articulate a vision for its wider ecosystem? 46) Are third-party vendors and ecosystem partners positive about their relationship with the primary vendor or the primary vendor's base platform(s)?

References

AHIMA Work Group. (2013). Data Standards, Data Quality, and Interoperability (2013 update). *Journal of AHIMA*, *84*(11), 64–69 [expanded web version].

Application programming interface. (2017, March 5). In *Wikipedia*. Retrieved from https://en.wikipedia.org/w/index.php?title=Application_programming_interface&oldid=768757981

Beck, K., Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., … Thomas, D. (2001). The Agile Manifesto. Retrieved December 28, 2016, from http://www.agilemanifesto.org/

Carafano, J. J. (2015, May 27). Lessons in Lousy Leadership. *The National Interest*. Retrieved from http://nationalinterest.org/feature/lessons-lousy-leadership-12975

Carlson, N., Hernández, E., Edmond-Verley, C., Rotondaro, G., Feliz-Santana, E., & Heynig, S. (2011). Developing a Master Data Sharing Agreement: Seeking Student-Level Evidence to Support a Collaborative Community Effort in Education. *The Foundation Review*, *3*(4). https://doi.org/10.4087/FOUNDATIONREVIEW-D-11-00024

Christensen, C. M. (1995). *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail* (Reprint edition). Harvard Business Review Press.

Church, E. (2016, September). *How Secure Does It Need to Be? A Quick Guide to Data-Security Compliance.* Presented at the Big Data Ignite Conference, DeVos Place, Grand Rapids, MI. Retrieved from http://www.bigdataignite.com/session/data-security-compliance/

Common Education Data Standards (CEDS). (n.d.). Retrieved December 27, 2016, from https://ceds.ed.gov/

General Services Administration. (2015, November 18). Rules and Policies - Protecting PII - Privacy Act. Retrieved December 29, 2016, from http://www.gsa.gov/portal/content/104256

Google Maps APIs for Web. (2017). Retrieved March 6, 2017, from https://developers.google.com/maps/web/

Institute of Electrical and Electronic Engineers. (2016). IEEE Standards. Retrieved January 1, 2017, from https://www.ieee.org/education_careers/education/standards/standards_index.html

Kegan, R., & Lahey, L. L. (2009). *Immunity to Change: How to Overcome it and Unlock Potential in Yourself and Your Organization*. Harvard Business Press.

Kegan, R., Lahey, L. L., Miller, M. L., Fleming, A., & Helsing, D. (2016). *An Everyone Culture: Becoming a Deliberately Developmental Organization*. Boston, Massachusetts: Harvard Business Review Press.

Kolko, J. (2012). Wicked Problems. In *Wicked Problems: Problems Worth Solving*. Retrieved from https://www.wickedproblems.com/1_wicked_problems.php

Lazzarotti, J. J. (2016, November 10). California Amends Its Data Breach Notification Law…Again. Retrieved December 30, 2016, from http://www.workplaceprivacyreport.com/2016/11/articles/data-security/california-amends-its-data-breach-notification-law-again/

McChrystal, G. S., Collins, T., Silverman, D., & Fussell, C. (2015). *Team of Teams: New Rules of Engagement for a Complex World* (1 edition). Portfolio.

MESA International. (2016, February 25). Technical Committee. Retrieved January 1, 2017, from http://www.mesa.org/en/committeeactivities/technicalcommittee.asp

Michigan Legislature - Section 445.72. (2004). Retrieved December 30, 2016, from http://www.legislature.mi.gov/mileg.aspx?page=GetObject&objectname=mcl-445-72

National Center for Education Statistics. (2004, March). Forum Guide to Protecting the Privacy of Student Information: State and Local Education Agencies. Retrieved December 30, 2016, from https://nces.ed.gov/pubs2004/privacy/section_4b.asp

National Center for State Courts. (2016). National standards. Retrieved December 27, 2016, from http://www.ncsc.org/Services-and-Experts/Technology-tools/National-standards.aspx

Ney, S., & Verweij, M. (2015, Winter). The Upside of Messiness: Clumsy Solutions for Wicked Problems. *Rotman Management*, (Winter 2015), 32–38.

Nurmilaakso, J.-M., & Kotinurmi, P. (2004). A review of XML-based supply-chain integration. Retrieved December 27, 2016, from https://www.researchgate.net/publication/245310235_A_review_of_XML-based_supply-chain_integration

Petrila, J. (2011, November). *Legal Issues Paper Presentation and Presentation of FERPA Notice of Proposed Rulemaking*. Presented at the Actionable Intelligence for Social Policy (AISP) Conference: Integrating Health and Human

Services and Education Data to Inform Public Policy, Washington, D.C. Retrieved from http://www.aisp.upenn.edu/resources/aisp-events/november-2011-conference/

Pilgrim, M. (2011). *Dive Into HTML5*. Retrieved from http://diveinto.html5doctor.com/index.html

Standards for Technology in Automotive Retail. (2016). The Standards. Retrieved January 1, 2017, from http://www.starstandard.org/index.php/star-standards

US Department of Health and Human Services. (2016, November 16). National Information Exchange Model | Children's Bureau. Retrieved December 27, 2016, from https://www.acf.hhs.gov/cb/niem

U.S. Department of Health and Human Services, & U.S. Department of Education. (2008, November). Joint Guidance on the Application of the Family Educational Rights and Privacy Act (FERPA) and the Health Insurance Portability and Accountability Act of 1996 (HIPAA) to Student Health Records. Retrieved from https://www2.ed.gov/policy/gen/guid/fpco/doc/ferpa-hipaa-guidance.pdf

Wikipedia: Calculus of negligence. (2016, October 26). In *Wikipedia*. Retrieved from https://en.wikipedia.org/w/index.php?title=Calculus_of_negligence&oldid=746332854

Wikipedia: Family Educational Rights and Privacy Act of 1974. (2016, October 19). In *Wikipedia*. Retrieved from https://en.wikipedia.org/w/index.php?title=Family_Educational_Rights_and_Privacy_Act&oldid=745097456

Wikipedia: Gramm–Leach–Bliley Act of 1999. (2016, November 29). In *Wikipedia*. Retrieved from https://en.wikipedia.org/w/index.php?title=Gramm%E2%80%93Leach%E2%80%93Bliley_Act&oldid=752098217

Wikipedia: Health Insurance Portability and Accountability Act of 1996. (2016, December 14). In *Wikipedia*. Retrieved from https://en.wikipedia.org/w/index.php?title=Health_Insurance_Portability_and_Accountability_Act&oldid=754715976

Wiliams, R. (1990). *The Mac Is Not A Typewriter* (First edition). Retrieved from http://www.powells.com/book/mac-is-not-a-typewriter-1st-edition-9780938151319

Williams, R. (2003). *The Mac is Not a Typewriter, 2nd Edition* (2 edition). Berkeley, Calif: Peachpit Press.

World Wide Web Consortium. (2016). Standards. Retrieved January 1, 2017, from https://www.w3.org/standards/