Roadmap to Capacity Building in Analytics

White Paper

By the APHSA National Collaborative for Integration of Health and Human Services Analytics Committee

Executive Summary

State and local health and human service (H/HS) agencies are at a crossroads today relative to their need to modernize how they operate and finance enterprise-wide projects while implementing and sustaining change, demonstrating accountability, and adopting modern technologies to achieve multiple program objectives and benefit clients with cross-agency needs.

Analytical tools can be invaluable in assisting such agencies with their transformation toward enterprise-wide, client-centric services. But the task is not a simple one and will require considerable leadership on the part of the agencies that believe in the value of analytics.

This guidance provides information on how to effectively implement successful analytic initiatives and the organizational capacities that will be needed to support it. Specifically, it addresses the following questions:

- How does an agency define its data and business needs to achieve its goals?
- What types of assessments and planning should occur prior to undertaking an analytic project?
- What are the needed skill sets and types of staff (internal and external) to implement an analytical strategy?
- What data governance structure and processes should be in place to facilitate collaboration and management cross-programmatic strategies?
- How and why is change management critical to embedding data sharing and analytics throughout an agency?

In addition, this guidance provides a number of real-world examples of how H/HS agencies have actually achieved cross-enterprise solutions by using analytics to effectively address their business needs.

Acknowledgments

Table of Contents

Section 1. Introduction

State and local H/HS agencies increasingly appreciate the value of analytics as they seek to balance their program participants' needs with taxpayers' insistence upon efficiency and effectiveness.

Multiple challenges arise, however, when one seeks to implement such powerful tools that are inherent in any approach rooted in the use of analytics. Historically siloed data systems and programs originally created to address specific, mandated business objectives and output measures[1] required by statute or regulation have made cross-boundary integration extremely challenging. Out-of-date technologies that lack analytical capabilities sufficient to evaluate multi-variant factors associated with enterprise-wide effectiveness and shared outcomes are rampant throughout the H/HS public in spite of billions invested in eligibility and enrollment systems.[2] And the skill sets required of such an ambitious initiative are in scarce supply. Yet, there are countless examples of where states and localities, regardless of the limitations of their budgets or political direction, have succeeded in making superb use of analytics to monitor the present and plan for the future.

How did they do that?

APHSA's National Collaborative for Integration of H/HS (the National Collaborative) established the Analytics Committee in 2014 to extract the lessons learned from the most successful practitioners of H/HS analytics—state and county leaders, faculty members from leading institutions, and industry leaders. The result is this paper, which we call *Roadmap to Capacity Building.*

Initially, this committee developed the *Analytic Capability Roadmap 1.0 for Human Service Agencies*[3] as a first step to help human service organizations assess the extent to which data analytics was a viable option for their programs. *Roadmap 1.0* described the Analytic Capability Curve,[4] together with a Capability Assessment Tool, to help agencies define where they were along the evolutionary pathway described on the *Curve,* as well the components of an analytic strategy appropriate to their particular situation.[5]

Building upon that foundational document, this guidance provides additional insights on:

- The analytic capacities that underpin successful analytical efforts
- The range of skills needed
- What the appropriate governance structure and change management processes look like to ensure the decision-making authority and processes are clear
- The importance of communicating, early and often, with all staff regarding their specific roles and making sure to seek and include their input, together with regular communications about project milestones and expectations
- Finally, we emphasize the need for agencies to dedicate efforts to ensure cross-programmatic data cleanliness, data integration, extraction, analyzing, and information delivery.

---

1. *Outputs* are defined here as process measurements of time, quantity, or quality relative to a particular reporting requirement (e.g., how many individuals were served, how much time did it take to process an application).
2. *Outcomes* are defined as changes that have taken place over time in knowledge, skills, behaviors and conditions as a result of an organization's work.
   Footnotes 1 & 2 Source: APHSA's Health and Human Services Integration Maturity Model v2 (2015).
   http://www.aphsa.org/content/dam/aphsa/pdfs/NWI/APHSA%20Maturity%20Model_2%200.pdf
3. APHSA. *Analytic Capability Roadmap 1.0 for Human Service Agencies. (*2014).
   http://www.aphsa.org/content/dam/aphsa/pdfs/NWI/FINAL_NWI%20Analytics%20Capability%20Roadmap_4.17.14.pdf.
4. Capability Curve is adapted from Davenport, Thomas H. and Harris, Jeanne G. *Competing Analytics: the New Science of Winning* (2007).
5. The Analytic Capability Curve and the Capability Assessment Tool are located in Appendixes A and B of this document.

Data Stratification

Additionally, and as all H/HS departments are deeply committed to gathering data, using data for decision-making and increasing competencies in analytics, we must also acknowledge the levels of data and the related activities that occur at each level. This shall determine the scope, capacity, and data-sharing policies related to data production and consumption.

- *Population Level Data*—data that are aggregated and de-identified and largely responsible for assessment of trends in investments, program design, and policy at the population level. These kinds of data can largely be modeled off of a public health template. They include data sets such as the American Communities Survey, the Youth Risk Behavior Survey or the Behavior Risk Factor Surveillance System, the Adverse Childhood Experiences risk data, and so on.

- *Program Level Data*—data that are focused on the service menu offered by a program, i.e., the evidence that supports the program and the outcomes expected at the program level for consumers using the program. These data can be aggregate but are more nuanced than population level data and, depending on the size of the universe of consumers, could be de-identified data or may require higher levels of confidentiality protection. These data are often used to build evidence on the efficacy of programs funded, to build the case to assess the return on investment and social return on investment for the program, and to justify program investments and level of evidence for the effectiveness of a program.

- *Individual Client and Household Level Data*—these data need the greatest level of privacy protection and clear policies around data sharing and safeguards against breach. This is the level of data production and use that requires a well-trained workforce and compliance with federal, state, and local laws and regulations. These data are most useful to evaluate alleviation and mitigation of presenting issues of safety and risk, health and well-being, and self-sufficiency for individual clients and households receiving services from the public or private agency. They measure impact of service plans and integration and interoperability across multiple service providers and across the H/HS enterprise.

Ultimately, analytics provides a powerful tool to take existing data that is clean, reliable, and meaningful and use it to model future behavior, trends, and utilization modeling to support programs, investments, and outcomes for identified populations. At the same time, if the data are not modeled to address existing inequalities and lack of access and poor outcomes resulting from these existing disparities, the predictions and assessments made with the data will be flawed. This could, in turn, cause harm to communities and populations who are already disadvantaged and lead to further inequities in program, policies, and investments. More work needs to be done to determine how predictive models can be built to create greater equity and support program and investment goals aimed at reducing and eliminating disparities when analyzing the data. Primarily, can questions related to access to care, fairness of policy and distribution of resources, and disparities in outcomes be evaluated to add greater nuance to the predictions being made with the use of data?

By building out these capacities, identifying the types or tiers of data to implement strategies making use of advanced analytics, and taking into account considerations on how predictive models are used, agencies will be able to make truly informed decisions with cross-programmatic, real-time information to impact individual, population, and organizational outcomes.

A. Project Description

APHSA supports H/HS agencies in adopting Continuous Quality Improvement[6] (CQI) practices that involve cycles of evidence-based decision-making (e.g., Define, Assess, Plan, Implement, and Monitor or DAPIM), using advanced analytics to generate evidence about performance or answer business questions, and staff use of that evidence to inform their decision-making about future policy and practice investments.

In order for agencies to keep pace with the private sector and other public-sector programs moving to a data-driven culture, as well as ensuring that CQI be informed through the most current and cross-representative evidence, the strategies and capabilities for that effort must be considered and established first.

The 2014 National Collaborative's Roadmap 1.0 supports CQI practices and initiates this effort through the Analytic Capability Assessment. In this Roadmap 2.0, the committee expands upon those foundational and assessment tools with the goal of seeking a universal business strategy to operationalize analytics regardless of agency maturity, with the following areas of focus:

- Support the adoption of continuous improvement efforts and visioning;
- Assess and develop agency capacities; and
- Provide practical examples of existing business solutions.

Taken together, these tools should assist H/HS agencies in assessing their existing capabilities to use analytics, define how they envision using analytics, and develop an analytics strategy to help meet the organization's business objectives, including the structures and processes by which to get there.

---

6. APHSA. *A Guide to Building Capacity for Child Welfare Using the CQI Process.* 2014.
   http://www.aphsa.org/content/dam/NAPCWA/PDF%20DOC/Home%20Page/A%20Guide%20to%20Build%20Capacity%20for%20Child%20Welfare%20Using%20the%20CQI%20Process%201.23.15.pdf

Section 2. Capacity Building

Capacity Building is a conceptual approach to development that focuses on understanding the obstacles that could inhibit an organization from realizing its developmental goals while enhancing the abilities that will allow it to achieve measurable and sustainable results. The key is to have a guide or framework for your organizational analytic development and activities prior to launching the analytics journey to assure a sound return on your time and resource investments.

In the following capacity-building diagram, the four quadrants of planning, team building, governance and change management, and data processes represent an essential interrelated framework that operates independently and interdependently of the other. As an organization prepares to journey down the analytics path, understanding and identifying critical components described within this capacity building framework should significantly aid in delivering useful and impactful analytics products.

Capacity Building

- Problem Statement
- Types of Data
- Assemble & Structure
- Model & Algorithm
- Transform
- Process & Analyze
- Visualization
- Quality Reviews

Data Processes

Planning
- Purpose & Audience
- Short Term & Long Term Goals-Plan
- Business & Resource Assessment
- Budget/Cost Allocation
- Sustainability/ Reusability

Governance & Change Management
- Governance Structure
- Methodology & Process Documentation
- Privacy, Safeguards & Security
- Change Management Activities

Team Building
- Statisticians
- Researchers
- Business & Financial Analysts
- Data Architect
- Technical
- SMEs
- Stakeholders

APHSA

A. Capacity 1: Planning

Clearly, anyone who has worked in the fields of H/HS policies and practice appreciates the value of understanding the context they work in, the need to leverage existing investments, and to build upon strategies articulated by stakeholders with shared goals. In addition, a clear approach to the use of analytics that is similarly focused on goals, audiences, available resources, financing, and sustainability is critical to the success of any new initiative or extension of an existing one.

We have also provided a simple table in the following section that will help users of this guide articulate each of these key elements and, just as important, why they are needed. While at first glance this may seem so rudimentary that it is unnecessary, many analytic projects get off on the wrong foot because no one clearly articulated what these terms mean to the leadership. And, as a result, team members pursue different objectives. While the Capacity 1 table provides definitions for each element listed above, we recommend that states and localities take the time to work with their teams to arrive at a consensus view on each of these elements, their definitions, and why they are needed. You may find it useful to add additional ones for the sake of clarity and confluence of your team's objectives.

Capacity 1 Example: **Community Connections NY, Inc.'s EQUIP Policy and Practice tool with human service workforces in Erie County, NY.**

EQUIP P&P is a cascading business intelligence tool to inform management of programs, teams and caseloads. The tool is being used in Erie County's Department of Social Services to overcome data access and information synthesis barriers that human service workers were having when trying to make data-driven decisions for case planning and utilization management.

B. Capacity 2: Team Building

When an H/HS agency proceeds with incorporating the use of data analytics in the organization, the team that is assembled should be composed of individuals with a variety of skill sets in order to bring together the right expertise and experience. No one model of organizational structure is standard; rather it is unique to each agency. The use of an interdisciplinary team focused on analytics provides the best opportunity for success by allowing collaboration to identify the correct questions and to build models that provide results the business can use. The team composition should be flexible enough to allow for the addition or subtraction of the requisite skills as needed.

The Team Building table that follows in Section 3 of this paper further defines these job titles, descriptions, and functional relevance to the development of the appropriate strategy that should be considered. While this list is not exhaustive, it includes the members of a team that will be essential to the development and implementation of any analytics strategy: statisticians, researchers, fiscal analysts, business analysts, data architect, technical staff, program or subject matter experts, executive or manager, and other stakeholders.

C. Capacity 3: Governance and Change Management

With the average age of many state IT systems hovering around 20 years, it's quite likely several generations of technology will be leapfrogged during an analytics modernization effort. This degree of IT transformation requires accompanying organization and process change management to empower the affected staff and business operations within agencies. Measures identified should minimize project risk, maximize agency adoption, and secure an appropriate return on investment.

As further defined in the Capacity 3 table in the following section, elements of an effective change management plan will define the needed agency roles and responsibilities to enable analytics based decision-making. This includes governance to mentor agencies to share analytic resources through the establishment of a sound (data) governance structure providing clear decision-making and communication plans, policies, and procedures; and motivation to promote a secure data exchange culture defined for specific purposes and in accordance to applicable state and federal law and regulations and incentive to undertake business operations innovation triggered by information technology modernization. Knowing that change is typically not easy for people to assimilate, the Organizational Change Management (OCM) plan needs to proactively identify the business based need for analytics change, how the project might affect each person's responsibilities and how they can engage assistance to evolve to the road ahead.

Capacity 3 Example: **Global Memorandum of Understanding (MOU) between the California Department of Social Services (CDSS), the California Department of Health Care Services (DHCS), California Counties, and Title IV-E Tribes**

This Global MOU sets forth the terms and conditions for the exchange of confidential data, collected and retained by CDSS and DHCS, for the purpose of matching the confidential data, hereinafter referred to as "matched data," to administer and implement the applicable federal and/or state health and public social service programs.

For more information, see Appendix C.

Capacity 3 Example: **Commonwealth of Virginia's (CoV) Health and Human Resources Agency—Enhanced Memorandum of Understanding (eMOU)**

HHRA developed an eMOU which is a reusable data governance framework able to be extended or modified as different data assets for different purposes/projects are identified. The agency used the federal Data Use and Reciprocal Support Agreement (DURSA) as the model for the framework and modified to accommodate the state-specific, agency-wide goals and structures.

HHR is comprised of 11 divisions, each with their own security officer. The initiative to get all division officers on the same page was undertaken by a central dedicated staff within the Secretariat's office to keep the momentum and direction focused on how data sharing would be accomplished rather than why it could not. Upon development of the eMOU, the Attorney General had to weigh in and endorse the use of the new data governance framework as well as its replacement of the point-to-point MOUs between the divisions. A significant benefit of this process was the demystification of the actual development of the eMOU as it engaged those from the beginning and then received official legal approval. As of July 8, 2014, all 11 directors had agreed to use the eMOU.

For more information, see Appendix C.

D. Capacity 4: Data Processes

Having clean, quality data is a critical foundation for data analytics. Obtaining such high-quality data needs to start at the beginning of the process and build upon the capacities discussed earlier. There should be a plan for what data are needed from the people who know the data both from a technical and a business perspective. A data governance process for how the data are to be shared amongst the data owners must be established. To ensure the integrity of the data, a process for obtaining the right data sets must be in place; as well as how data will be extracted, transformed, and loaded into an integrated analytical environment such as a data warehouse.

- Problem Statement
- Types of Data
- Assemble & Structure
- Model & Algorithm
- Transform
- Process & Analyze
- Visualization
- Quality Reviews
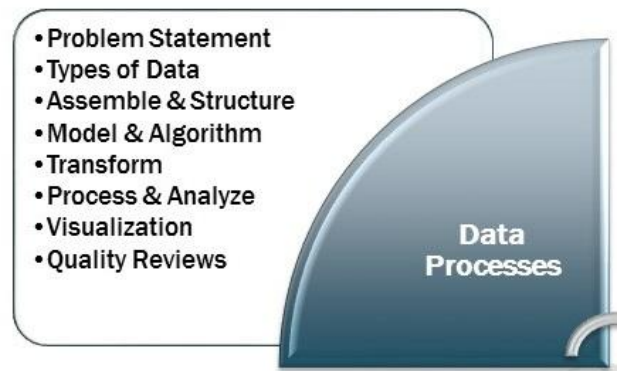
Data Processes

Data processes must support the main objectives of improved decision-making, improving program outcomes, and better service delivery. Agencies often say they are drowning in data, but they don't have enough usable and actionable information. Employing sound data processes facilitates the achievement of having actionable information to better inform policymakers and all levels of program staff to make decisions. The Capacity 4 table in the next section details the key components in a sound data process environment.

Capacity 4 Example: **Improving Child Welfare Systems of Care Using Data Analytics within the Florida State Family Network (FSFN)**

Mindshare Technologies and the FSFN are using the "Immediate Care Accountability and Case Escalation" (ICARE), an analytic and case management solution, to rebuild quality management practices and enable improvement in direct care.

ICARE can take data from any source, match/compile/organize/analyze within their system, then export to dashboards, mobile apps, etc.

See Appendix C.

Section 3. Capacity Matrices

This section provides further detail on each of the capacities laid out in the previous sections of this paper. The elements or components of Capacities 1–4 are identified and defined. In addition, their value toward an agency-wide analytic strategy or within an analytic project is provided.

Capacity 1: Planning		
Topic	**Definition**	**Why Needed**
Identify Purpose and Audience	Establishes the project scope, main objectives, and intended recipients of the outcome.	Summarizes and communicates a strategic vision and core purpose of the project to stakeholders and others, and most likely remains unchanged over time.
Short- and Long-Term Goal Setting	The act of identifying, prioritizing, and assigning tasks and activities with timeframes for completion. This operationalizes your project's strategy to achieve objectives.	In essence, it provides a detailed agreement of activities to be accomplished. By delineating short- and long-term goals, results are achieved throughout the project's lifecycle and it retains stakeholder interest.
Business and Resource Assessment	A comprehensive review of your agency's assets and business needs to assist in decision-making and planning for the project start and maintenance support. Assets and business needs reviewed typically include: staff, RFP/contracts, procurement, technical equipment, software, architecture, office space, and business operations that are affected.	Making decisions that impact an organization involves careful decision planning. Projects may be dependent on external factors (contracts/approvals) that should be planned for in advance of convening all stakeholders, convening too early may lose stakeholder interest.
Budget/Cost Allocation	Determining the funding stream(s) to support the planned effort.	Projects often have delays, or unintended costs not foreseen at the start of the project. Allocating a reserve budget will keep the project's lifecycle moving.
Sustainability/Reusability	A foresight and planning methodology, the ability to endure or continue over time. Plan purchases and create business practices that include the capability to be expanded upon, enhanced, or operationalized with minimal maintenance or cost.	Agencies may make large dollar investments in assets and agencies should procure assets that project future growth of its organization.

Capacity 2: Team Building		
Topic	**Definition**	**Why Needed**
Statistician	Trained and experienced in applied statistics, mathematics, or both, the statistician uses statistical methods for collecting and analyzing data to help solve real-world problems. This is accomplished by the processes associated with producing reliable data, analyzing data to understand the environment, and drawing practical conclusions from the data. In addition, statisticians often have domain-specific expertise that provides additional insight to the analysis activity, including understanding pertinent questions and familiarity with potential outliers or anomalies. Statisticians will apply algorithms on top of the data sets to draw out richer information, which goes beyond simple counts or other operational reporting.	A statistician brings quantifiable and scientifically backed approaches to the decision-making process. With the increased use of computing power, statistical activities have become faster and even more accurate. Quantifiable justification is often necessary for planning and implementing federally funded projects. The expertise of statisticians can also serve to help organizations by cleansing their data and making it useable for analysis.
Researcher	Researchers conduct comprehensive research in a given industry or domain area. They identify, analyze, and synthesize data sets (sometimes in collaboration with a statistician). The researcher understands the broader landscape of a domain, is informed of significant academic research outcomes, is knowledgeable about evidenced-based practices, and is often well connected to academia.	Qualitative and qualitative research can be used to help hone and craft areas of focus by providing an objective and informed perspective of objectivity, identifying metrics and measuring results, establishing standardized methods for approaching research and analysis, offering explanations for complex situations with a depth that goes beyond surface observations.
Fiscal Analyst	A fiscal analyst provides budgeting and accounting functions using an organization's financial data to spot trends, develop forecasts, and provide cost benefit and risk analysis, guiding leaders in making cost-effective solutions that meet an organization's strategic goals. Staff contribute knowledge and understanding of federal, state, and local revenue sources, potential avenues of change, asset management and resource level of effort, and apply data to key performance indicators and outcome-based measures.	Financial strategies support an organization's vision and assist with maximizing performance outcomes for strategic planning. The fiscal analyst assists by navigating procurement practices to ensure effective solutions are met.
Business Analyst	A business analyst enables change in an organization by using analytical principles and practices to define needs and recommend solutions that deliver value to stakeholders. Best practices dictate that the business analyst should also be in tune with current organizational processes. According to the International Institute of Business Analysis, the business analyst performs tasks in "knowledge	Business analytics can help an organization to realize identified project or change benefits, cost avoidance, identify new opportunities for automation and improvement and an understanding of required capabilities. The business analyst can also create a bridge between business and IT in order to find or use correct data attributes for effective meaningful

Capacity 2: Team Building		
Topic	**Definition**	**Why Needed**
	areas" including Elicitation, Requirements Management and Communication, Enterprise Analysis, Requirements Analysis, Solution Assessment and Validation, and Business Analysis Planning and Monitoring.	reporting.
Data Architect	A data architect works closely with business analysts, systems designers, and project teams to design database solutions to support business needs. The data architect will also plan for issues associated with data migration, validation, clean-up, mapping, and data dictionaries. The data architect is knowledgeable about industry data models, enterprise planning, and regulatory structures such as HIPAA, FERPA, Sarbanes Oxley, and Basel II.	A good data architect understands the vision of the organization and the close alliance between data architecture and business architecture. In a world of complex business environments and technology environments utilizing data integration or interoperability of disparate data sources, the data architect position is critical to the success of any data analytics effort.
Technical Staff	Technical staff, such as report writers and programmers, supports changes or implementations of technical solutions that result from the data analytics efforts. Technical staff is crucial to conduct data integration because they can help take complex data systems and make it useable in the organization's analysis.	The most successful outcome of a data analytics effort is the institutionalization of data-driven operations and decisions. The technical staff will incorporate the findings into visualization tools, web applications, and reports that allow for broader adoption of these information tools and empowerment of the stakeholder community. Technical staff can add value by transferring knowledge of report writing and data understanding to business analysts and other staff. They also can bring context to data that may reside outside the organization's operational system.
Program Subject Matter Expert	A program subject-matter expert is an authority in performing a job, task, or skill within a domain or topic within the organization. Depending upon the data analytics initiative, it may be required to have cross-functional representation of multiple subject matter experts. Examples of high-level domains include Integrated Eligibility, Child Welfare, and Child Support. Within these domains, however, there may be sub-categories such as IV-E funding, adoption, hotline, front line or intake, case management, and collective bargaining (using Child Welfare as an example) that should also be included.	Key program subject matter experts will inform and validate the findings of the statistician and researcher, as well as the solution or interventions that may be implemented as a result of the analytics effort. Engaging program subject matter experts helps nurture a culture of inclusivity that will support the institutionalization and operationalization of data-driven functions.
Executive Manager	An executive manager is responsible for overseeing the data analytics operation with strategic vision and the required level of authority to manage budget, resources, tools,	A visionary and influential executive manager is critical to the start-up and institutionalization of data analytics in an organization. Given the variety of resources

Capacity 2: Team Building		
Topic	**Definition**	**Why Needed**
	and other necessary requirements of the operation.	needed to manage a data analytics effort, the executive manager must champion the cause in a matrix approach throughout the breadth of the organization and beyond.
Other Stakeholders	Other stakeholders are internal or external interested parties that can affect or are affected by an organization. Examples may include staff from another office, division, or department; contractors; advocates; or consumers.	Including stakeholders adds value to the process, achieving buy-in of a particular process or use of product.

Capacity 3: Governance and Change Management		
Topic	**Definition**	**Why Needed**
Governance Structure	A team or teams of people that are convened for a defined purpose, or for a limited term, that institute a system of roles, responsibilities, and decision-making rights to teams that manage the organization to achieve a defined purpose. Governance structures with more than one team are typically hierarchical in nature, which establishes a process of interactions with final authoritative decision-making and includes executive, management, and working-group levels of cross-functional stakeholders of policy, legal, program, fiscal, and technical interests. The following artifacts may be used to establish governance: • Project Charter or Work Group Charter • Stakeholder Analysis (including financial, political, and programmatic stakeholders) • Plans, Policies, and Procedures • Governance Maturity Roadmap or Plan	Through the authority and activities of governance bodies, entities can collaborate and manage efforts to achieve mutual goals independent of their daily agency business operations. Sound data governance will guide all other analytics activities, offer cross-program analytics opportunities, fiscal savings, and programmatic efficiencies. It will also show where the decision-making authority lies relative to any activity as well as establish the parameters by which to select entities or groups with vested interests as stakeholders. Sound data governance will also provide a sense of boundary or ownership relative to the various data and systems included within a given project or activity.
Methodology and Process Documentation	A chosen system of methods, rules, guiding principles, processes, and procedures associated with a branch of knowledge (data analytics) based upon industry standards and best practices designed to be applied to carry out the particular research, project or study purpose. Choosing a particular paradigm, model or technique to be applied is an effort that takes time and commitment from all stakeholders. The following may be used to document this	Provides a constructive framework or a foundation of principles that justifies decision-making. A defined methodology and standard processes ensure consistency across the enterprise and expedite task time through creation of reusable work artifacts and allow continuous refinement and improvement efforts.

Capacity 3: Governance and Change Management		
Topic	**Definition**	**Why Needed**
	effort: • Research or Project Analytics Methodology • Data Collection, Inventory and Data Dictionary • Information Management Lifecycle • Master Data Management Plan • Data Governance Implementation Plan • Quality Management Plan • Data Change Management Plan • Validation and Utilization Plan	
Privacy, Safeguards, and Security Tools	Signed statements, documents or agreements used to govern the privacy and protection of the physical, administrative, legal and technical use of data through its access, collection, transmission, and storage. The purpose of data activities shall be clearly defined for specified purposes and must be in accordance with applicable state and federal law and regulations. All documents shall have a signatory. Separate documents or statements may be created independently or combined, and are generally established at the outset of a project, study or initiative when governance is being established: • Memorandum of Understanding (MOU) or Agreement (MOA) • Data Sharing Agreement • Security Agreement/SLA • Data Access and Confidentiality Statements • Enterprise Memorandum of Understanding or Agreements	Entities are required by law to protect and govern data to minimize risk of compromising secure information of the people they serves. More important, the credibility of the analytic activities are placed in serious jeopardy relative to users and the public in general to the extent that privacy and security protocols are breached by those without a right to have access to the information.
Organizational Change Management and Institutionalization	An organization's engagement in an effort or initiative with clear and timely planning, communicating, training, educating, and reporting for that initiative or effort. Incorporating the use of data analytics into daily business practices may bring about new approaches to current business efforts by highlighting early successes and is one strategy for securing champions and early adopters to support the data governance enterprise initiatives. Examples of institutionalization activities include: • Change Management Assessment	It operationalizes the analytic effort into standard business practices. Through formal planning and execution of an institutionalization movement, the data governance efforts will be achieved. The result is realization of all the benefits of solid data governance strategy. Risks to realizing the benefits of the data governance strategy are addressed proactively by the OCM effort: • Effective communication to educate and inform stakeholders on the reason for analytics-driven change. Future vision? Why? How does it apply to me?

Capacity 3: Governance and Change Management		
Topic	**Definition**	**Why Needed**
	• Communications and Organizational Change Management Plans • Standardize Reporting • Training Plans, including Data Usage, Data Security, and Data Analytics training curriculum	When? Where? • Engagement and continuous inclusion of business owners to understand and support business process reengineering that may be triggered with the culture change. • Education, training or skills upgrading for the organization. • Identification of counter resistance, either passive or overt within the organization. Align at-risk staff with the overall strategic direction through mentoring, education, communication, and administrative services.

Capacity 4: Data Processes Integration, extraction, analysis, and information delivery		
Topic	**Definition**	**Why needed**
Problem Statement and Requirements	The process by which to specify business needs and identify data needed	To achieve purpose identified in Capacity 1
Types of Data	*Data Dictionary:* The data dictionary is a log of all data elements in the data analysis environment. For each data element the log should include information such as: • name (standard across the environment) • description • format(s) • source(s) • source of record • source(s) owner • name in source(s) • active date • closed date	The data dictionary enables the data stewards to understand: • data elements available for analysis • potential for bringing other data elements into analysis • integrity of the data elements • comprehensiveness of the data elements The data dictionary is not a mechanical component of the technology infrastructure but an essential part of its content description. The data in the dictionary are not just the content data held in the environment, e.g., the case data, the provider data, the claim data but also metadata, audit/access data as well as historical data (see definitions below). The data dictionary should be compiled and maintained as part of the governance process, see Capacity 2 above.

		Demographic Data (examples): • Name • Address—various types • Phone—various types • E-mail • Gender • Race	The most basic form of information contained in human service case management systems is demographic data relative to the individual being served, household composition, caseworkers, providers, and so on. This type of information is essential to analytics at all maturity levels and is often the starting point for foundational analytics efforts.
		Unstructured data—text and comments	Case notes and comment fields often have valuable content that, once analyzed, yields incredib*le information regarding the case itself* and practice. This information is critical to the analytics process to help guide decision-making at all levels of the organization.
		Historical/Audit (Transactional) data: Data that are continually stored as the data are accessed and updated. The tables created by the Audit feature of a database provide a historical view of the versions of the data elements through time. • Audit/access data, which are data continually stored as the data are accessed and updated, e.g., service/username updating data, date of updating data. • Historical data, which are previous version(s) of data elements, e.g., prior and now paid claims.	Audit logs are required for a number of reasons. Most commonly they are needed to provide security and fraud detection. Depending upon the contents of the database, audit logs are required by law, such as Sarbanes-Oxley,[7] Payment Card Industry Compliance,[8] HIPAA,[9] FERPA,[10] 42 CFR Part 2,[11] and FOIA.[12] *Note:* The Audit data feature in a database is a feature of all database products and generally does not play a role in what this paper is discussing and should just be considered informational. Data analytics only uses this data for internal use for fraud and security. Audit features of a database allow for quality control and forensic review of what transactions took place, by user, by data element with a date and time stamp, providing oversight in the administration of a database. An audit tracks data changes made in the data that become the history.

7. United States Code, Sarbanes-Oxley Act of 2002, PL 107-204, 116 Stat 745, Codified in Sections 11, 15, 18, 28, and 29 USC, July 2002.
8. Payment Card Industry Data Security Standard (PCI DSS) is a set of requirements for the secure exchange of credit, debit and cash card transactions and protection of cardholders. For more information, please visit *www.pcisecuritystandards.org*.
9. The Health Insurance Portability and Accountability Act of 1996 (HIPAA), P.L. 104-191, 110 Stat.1938 (1996).
10. Family Educational Rights and Privacy Act (FERPA), 20 USC §1232g; 35 CFR Part 99, as amended.
11. Confidentiality of Alcohol and Drug Abuse Patient Records, 42 CFR Part 2 (1987).
12. Freedom of Information Act (FOIA), 5 USC §552 (1966), as amended.

	Metadata: There are two "metadata types"—structural metadata, about the design and specification of data structures or "data about the containers of data"—and descriptive metadata about individual instances of application data or the data content. Typical information contained in metadata includes: • Clear and definitive business description and the intended meaning of the data fields contained • Specific rules or logic used for organizing or categorizing data elements • Language or local dialect details (e.g., http or HL7 version) • Code set assumptions applied to specific fields (date format, monetary base, ICD-10/LOINC/SNOMED codes) Standards for metadata include ISO/IEC 11179, Dublin core, NIEM, and NISTs' administrative standards for security.	The main purpose of metadata is to facilitate the discovery of relevant information about data being exchanged with a consumer or subscriber. Metadata assists the subscriber with information on how to understand the exchanged data without laborious research. By describing the contents and context of data files, the usefulness of the original data/files is greatly increased, administrative support costs are reduced and system independence is improved. Metadata fosters improved data exchange between partners through semantic interoperability. Semantic interoperability is concerned not just with the packaging of data (syntax), but the simultaneous transmission of the meaning with the data (semantics). This is accomplished by adding descriptive information or can be states as "data about the data" (metadata), linking each data element to a controlled, shared vocabulary. The meaning of the data is transmitted with the data itself, in one self-describing "information package" that is independent of any information system. It is this shared vocabulary, and its associated links to an ontology, which provides the foundation and capability of machine interpretation, inferencing, and logic.
Assembly and Structure	Organize the data in a computer so it can be used	This step is essential to the analytics activity. The data structures provide a means to manage large amounts of data efficiently for uses such as indexing, designing algorithms, and application processing.
Model and Algorithms	Process of selecting the right data sets, algorithms, and variables and the right techniques to format data for a particular business problem	The identification of the model and algorithm necessary to perform the appropriate analyses produces the intended results for the identified problem. It is akin to using the correct mathematical formula to achieve the correct result.

|		
Transform Data	Replacing variables so data can be used for an intended result	This function is usually applied so that the data more closely meet the defined assumptions, or to improve the interpretability or appearance of graphs.
Process and Analyze	Processing and analyses of data is the collection and manipulation of data to produce meaningful results.	This activity provides insight into business and operations to allow decision makers to take action (or not, as it may be) to intervene in a given scenario to achieve an intended outcome or result.
Visualization	Techniques used to communicate data or information by encoding it as visual objects or the visual representation of data	A primary goal of data visualization is to communicate information clearly and efficiently to users via the statistical graphics, plots, information graphics, tables, and charts selected
Quality Reviews	Data quality reviews are conducted to determine the data's fit for use to serve the purposes of a given context.	The outcome of data analyses is highly dependent upon the quality of the data analyzed. Given the intended purpose of using data analytics to inform business decisions, it is essential and expected that the data quality and, therefore, analytics results are accurately representative.

Section 4. Summary

One final consideration for which agency initiatives or specific projects will want to be cognizant is the available technology in the current marketplace required to implement the last phase of bringing analytic capability into the organization, program(s), or project. Much information is already published and available on the different types of automated tools to assist with advanced analytics like modeling and forecasting. As we did not wish to use this paper as a place to duplicate that information, we encourage readers to investigate existing resources like Gartner's Magic Quadrant.[13]

In conclusion, we encourage states and localities to use this Capacity Roadmap as a guide to assist in planning and executing successful use of advanced analytics at every level of the organization and for a variety of projects or when thinking about how to incorporate data into agency- or program-wide operations. Deciding what questions you want answered and for what purpose, assessing your current and projected resources including financial, technological and staffing assets, developing the structures and processes required to enable decision-making and change, and the types of data to be considered and harmonized, are critical steps to take no matter the envisioned use or where you may be starting from.

---

13. *http://www.gartner.com/technology/research/methodologies/research_mq.jsp*

Appendices

Appendix A—Analytic Capability Curve[14]
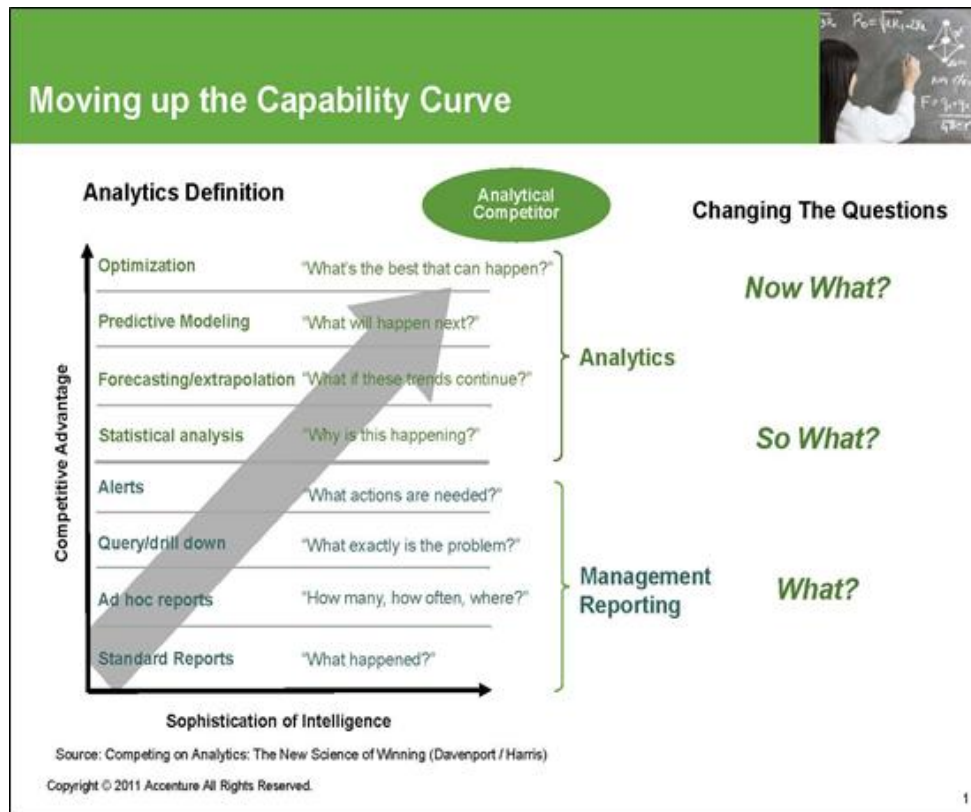
14. APHSA. *Analytic Capability Roadmap 1.0 for Human Service Agencies. (*2014).
http://www.aphsa.org/content/dam/aphsa/pdfs/NWI/FINAL_NWI%20Analytics%20Capability%20Roadmap_4.17.14.pdf
*Capability Curve is adapted from Davenport, Thomas H. and Harris, Jeanne G. *Competing Analytics: the New Science of Winning* (2007).
*Please note that for our purposes, the Y-axis will measure time and the X-axis will measure outcomes.

Appendix B—Capability Assessment Tool[15]

Domain	Basic	Advanced	Leading
Standard Reporting	• Formats suitable for federal reporting • Delivered quarterly	• Formats for federal reporting AND Key Program Indicators (KPI) • Delivered via the web at least monthly	• Formats for federal reporting, KPI AND geographic segmentation • Delivered via the web at least weekly
Random Reporting	• Basic query	• Complex query • Use multiple sources of data	• Joined, multiple complex queries • Use multiple sources of data
Focused Reporting	• Specific users are licensed to conduct analysis	• Key supervisory decision-makers are licensed to conduct analysis	• Staff administering program are licensed to conduct analysis
Automated Alerts	• Batch runs • Manual review and application	• Batch runs or real time • Automated review and application	• Multiple, real time alerts • Automated review and application
Statistical Analysis	• Structured data • Comparative analysis • Clearly defined sources	• Structured data • Comparative and trend analysis • Clearly defined sources • Results reviewed by statistician	• Structured and unstructured data • Comparative and trend analysis • Results reviewed/verified by statistician or data scientist
Forecasting and Extrapolation	• Linear extrapolation • Based on recent available historic data	• Linear and possibly multi-variant models • Based on recent, available historic data	• Multi-variant models • Based on recent, available historic and current data • Extrapolated forecast
Predictive Modeling	• Not used	• Uni-variant or multi-variant model • Linear extrapolation (e.g., predicting behavior for one question)	• Uni-variant or multi-variant models • Iterative regression analysis used to develop statistical probabilities (aka predictive analytic models)

15. APHSA. *Analytic Capability Roadmap 1.0 for Human Service Agencies. (*2014).
http://www.aphsa.org/content/dam/aphsa/pdfs/NWI/FINAL_NWI%20Analytics%20Capability%20Roadmap_4.17.14.pdf
*Abbreviated version of the Capability Assessment Tool. The full tool is available in the Roadmap 1.0.

Appendix C—Examples of Data Analytic Projects

1. Community Connections NY, Inc. (CCNYI) and Erie County, NY *(continued from p. 8)*

Erie County, NY implemented CCNYI's business intelligence tool (EQUIP) to assist staff at all levels of the agency incorporate data into their daily decision-making. That could translate to operations for managers or to service planning and provision for direct line workers. While EQUIP is the tool they use, CCNY also worked with Erie County toward developing what they termed as "data fluency" among workers from the early phases of planning, design, and then ultimately implementation. Having the right people from all levels of the organization at each level of the strategy was critical to the success of enculturating the use of advanced analytics into the work of the agency as a whole.

2. CA Department of Social Services, Department of Health Care Services and Counties Global Memorandum of Understanding (MOU) for Child Welfare Services *(continued from p. 9)*

This Global Memorandum of Understanding (MOU) between the California Department of Social Services (CDSS), the California Department of Health Care Services (DHCS), California Counties and Title IV-E Tribes sets forth the terms and conditions for the exchange of confidential data, collected and retained by CDSS and DHCS, for the purpose of matching the confidential data, referred to as 'matched data,' to administer and implement the applicable federal and/or state health and public social service programs. The Global MOU is available in full at http://www.cdss.ca.gov/pdf/GlobalDataSharingAgreement.pdf.

3. VA Health and Human Resources Agency enhanced Memorandum of Understanding (eMOU) *(continued from p. 10)*

The eMOU is an inter-Departmental document used to standardize data exchange agreements with governance process for additions and approvals.

Each and all division security officers meet at least once a month to provide an overview of what data assets are proposed and thus approved to be included as allowable data sets/elements within the construct of the eMOU. This block of time also is being used by the HHRA as a teaching forum for staff. Some of the topics have included providing a landscape of the different privacy rules and how one navigates privacy concerns within those.

By taking this type of iterative process toward consistent and intentional modifications to the eMOU, the agency has now begun to compile a data inventory. HHR is also receiving requests from external agencies across the CoV (e.g., DOJ) to use the eMOU as well. One of the goals moving forward is get to a point where the eMOU is being used by private partners too-which would further the data sharing efforts across the CoV.

An overview of the eMOU is available at
www.vita.virginia.gov/uploadedFiles/VITA_Main_Public/ITAC/HITSAC/2014/EMOUexecsummary.pdf. For more information, please contact Russ Sarbora *(Russell.Sarbora@governor.virginia.gov).*

4. Commonwealth of Pennsylvania Department of Human Services (formerly the Department of Public Welfare) and Department of Labor & Industry Data Sharing Agreement

The Department of Labor & Industry (L&I) was awarded a Workforce Data Quality Initiative Grant (WDQI Grant). One of the goals of the project is to obtain a better understanding of the efficacy of existing work training programs conducted within and external to L&I, whereby inclusive of the Department of Human Services' (DHS) SNAP & TANF programmatic (de-identified personally identifiable information) data for purposes of the research and analysis.

The framework used to develop this Data Sharing Agreement (DSA) is similar to existing DSA frameworks: it outlines which agency is responsible for doing what (e.g., who reports the data to whom); which agency will house and manipulate the data; and who will pay for the storage. One critical characteristic of this DSA is the amount of flexibility provided to incorporate other agencies and data as the project evolves. For example, L&I and DHS would like to be able to include data from the Department of Education at some point as they move forward so L&I could evaluate the supply and demand of the three departments in order to coordinate education and training and work engagement.

One of the key success factors related to the development of this DSA was the "why not" approach taken by the Commonwealth's Assistant Counsel. With the legal authority and directive coming from the Governor's office, it encouraged the departmental attorneys to negotiate and rely upon one another's knowledge to figure out how they would integrate the data and work collaboratively on meeting the goals and requirements of the WDQI grant.

For more information, please contact Megan Lape *(mlape@aphsa.org)*.

5. Florida State Family Network & Mindshare Technologies Use of Data Analytics to Improve Child Welfare Systems of Care *(continued from p. 11)*

Through ICARE, the Florida State Family Network's (FSFN) use of predictive analytics moved from standard business intelligence to applying data to vulnerable children in order to provide stakeholders with information to protect children:

- With the safety model, able to look at which agencies are underperforming; percentage improvement in specific cases
- Also able to use attribute ranking (with factors such as number of missed visits, living situation, etc.) to predict which children are at the highest risk of harm
    - Includes charts and graphs for visualization
- Models are specific to jurisdiction and must be retrained monthly to reflect activity, primarily within system of care
- MOUs were developed between county child welfare agencies and school districts

In Florida, the state used ICARE in to connect data between their SACWIS systems and education systems-using MindShare as the Hub to connect these systems. In addition, Mindshare had to go agency by agency to introduce the model in order to get child welfare and education departments to share data. The MOU developed with the initial participating county agencies was replicated and used as a template with other county child welfare and education agencies within the Florida State Family Network. The technology has had proven results, including reduced child deaths, improved housing permanence and overall wellness, enhanced efficiency in case management services, and reduced case manager turnover.

6. NYC ACS and KMPG's Predictive Foster Care Model for Capacity Management

KPMG worked with New York City's (NYC) Administration for Children Services (ACS) to develop an analytic model allowing for the agency to forecast the characteristics of future foster care populations, as well as what types of facilities and services would be required of these children and families.

NYC ACS is responsible for providing outsourced foster care facilities for children but the agency had no method for forecasting the future population and corresponding need. At the time, the agency had no way to figure out how best to allocate resources and build capacity-based data projecting future demand. This model enables executives, policy planners, and contract negotiators to forecast the facilities' needs.

Once the NYY ACS and KPMG defined the problem, they pulled together a team, including an executive sponsor, subject matter experts, and data modelers and designers. Upon assessing their resources, ACS decided to utilize

their existing Data Governance workgroup and staff that skilled in Software as a Service (SaaS). The model was funded by KPMG and support resources were provided by ACS to finance the work. The agency also realized they did not have the technological infrastructure (e.g., desktop server) available to run the new analytic model so they used KPMG's data center facility for the initial data storage and processing infrastructure. SaaS was the primary tool used for data transformation and modeling.

7. Chapin Hall's Multistate Foster Care Data Archive Web Tool

This web-based analytic application enables end-users to ask and answer business and research questions about trajectories and outcomes for children in foster care. The software is primarily intended for strategic planning and CQI purposes, enabling state child welfare agencies to observe variation within the system, identify sub-populations for whom outcomes can be improved, and develop evidence to support improvement efforts. An interface runs queries off of the Chapin Hall Multistate Foster Care Data Archive (FCDA), a longitudinal database containing foster care records from participating states.

The FCDA web tool was designed to give child welfare agencies a user-friendly analytic tool that they could use to answer mission-critical questions about system performance. The longitudinal approach enables agencies to ask and answer questions about changes in performance over time and develop evidence that links policy and practice innovation to changes in outcomes. Program managers and staff tasked with performance monitoring are the primary staff benefitting from this tool. Yet there are also clear linkages for fiscal staff—fiscal data applied to program performance outcomes produced by the web tool provide information about return on investment.

The web tool was developed by Chapin Hall and required the participation of researchers, data management experts, web development experts, and member states contributing data. Establishing the web tool for a particular state requires partnership with a state liaison, the state's data management team, and a leadership that is willing to champion the tool and put it to use.

Chapin Hall uses a Data Sharing agreement and Services Agreement to establish the relationship between Chapin Hall and the participating agency and to lay out rules regarding confidentiality, use of data, intellectual property, and so on.

The Multistate FCDA is created using member states' own electronic data, transferred to Chapin Hall by Secure File Transfer Protocol (SFTP). Data dictionaries are developed in partnership with each member state and clarify both the processes by which state data are transformed into FCDA data and the definitions of variables in the resulting database. States determine how they will fund membership to the Data Center, sometimes receiving support from foundations to do so.