

Data Driven Decision-Making in Utah Government

All across the country, businesses are finding tremendous value in data. New compute capabilities are creating opportunities far beyond anything seen in previous decades. Eighty-four percent of businesses believe that Big Data analytics will completely change the competitive landscape of their industries within the next 3 years. State and federal governments are launching promising initiatives based on these technologies.

Insurance fraud costs US taxpayers hundreds of billions of dollars every year. Studies indicate that each Utah family pays more than \$1,000 annually in increased insurance premiums due to insurance fraud. Meanwhile, organized crime is increasing its role in insurance fraud. Almost daily, large scale identity thefts are reported that increase our sense of vulnerability. The state of Utah has created an entire division of state government to address this growing problem with a team of attorneys and investigators. That talented team is only able to respond to about ten percent of the potential cases they are given. In today's world, attorneys and investigators are not enough. Government needs access to large amounts of data and the ability to effectively digest, monitor, and analyze the data dynamically in real-time. This problem is not going away and it is not going to get any easier. Combatting insurance fraud is one of many challenges facing government that are continuing to grow and that require more effective utilization of large data sets if they are to be responded to effectively.

Executive Summary

Utah state government deals with many of society's biggest problems. As the state's population continues to grow, these challenges become more complex and the public expects government to respond to these issues as effectively as possible. State agencies have turned to data and analytics as a way to reduce cost and improve service delivery. Big data analytics is the next evolution in technology that businesses and government are utilizing to accomplish these objectives. Just as Utah invested in PC's and networks in the 1980's, in the Internet and Online Services in the 90's and Digital and Open Government in the 2000's, and in mobility in the last 7 years, an investment in big data and analytics technologies is prerequisite to remaining competitive in the next five years. This business case identifies a few of the many areas of state government that could benefit from the use of advanced analytics. To be done right, this will require statewide coordination, in addition to investment in technology and skill development. With the proper coordination, Utah can reduce crime, including insurance fraud, improve educational outcomes, create performance-based budgeting processes, and even reduce recidivism through use of the technologies discussed briefly in this document.

What do we do today?

Although Utah has been working for years to become a data-driven state that makes decisions based on the best available data, it cannot reach its full potential using data management and analytics technologies that are even 5 years behind. Today, the state utilizes

8 Benefits of Big Data for State & Local Government

1. Make better decisions more quickly
2. Improve mission outcomes
3. Identify and reduce inefficiencies
4. Eliminate waste, fraud, and abuse
5. Improve productivity
6. Boost ROI, cut total cost of ownership (TCO)
7. Enhance transparency and service
8. Reduce security threats and crime

source: State Tech Magazine

<http://www.statetechmagazine.com/article/2013/05/8-benefits-big-data-state-and-local-governments>

primarily relational database technologies that have limits on their ability to process and analyze huge amounts of diverse data and high velocity data streams effectively. Some individual departments have their own data warehouses and may be using last generation business analytics such as IBM Cognos. Others have plans to use business intelligence tools in the near future, while some have hardly begun.

The good news is that new technologies, proven in business and government, provide the capabilities to do what is needed. In order to benefit from these technologies, the state will need to invest in an enterprise technology platform and in skill procurement or development of capabilities that are currently in short supply in Utah state government. By effectively combining new analytic skill sets and making better use of state data through a new technology platform, the state can make data-driven decisions that will improve the way that government services are managed and delivered in Utah.

Some agencies have already made progress in the use of analytics to drive government decisions. For example,

- **UDOT** uses in-pavement real-time traffic data sent to the UDOT Traffic Operations Center to create the congestion layers on the UDOT Traffic app and website, so travelers can know about delay and congestion information for their trips. They can also analyze traffic flows and make near real time adjustments to traffic signals, etc.
- **Corrections'** new BI dashboard provides goals, trends, actuals and ties financial costs to results.
- The Economics Analysis team of economists and labor market specialists at **Workforce Services** provides economic analyses and recommendations to local workforce employment centers, partners in education, economic development and the public.

Currently, the state of Utah maintains hundreds of databases and a small number of data warehouses. These include data warehouses in the Division of Finance, the Department of Workforce Services, the Department of Human Services, the Department of Health, and perhaps others. At a minimum, we should begin to ask if these large data stores can be better managed as an enterprise using newer models for data storage, management, and analysis.

Opportunities

Prison Recidivism

A recent article in The Economist suggests that better use of data when making parole decisions will reduce recidivism by 15%. Utah's prison population is growing at an annual rate of 6 percent, significantly higher than the national average of 4 percent.

Michigan claims that the use of computerized assessments have helped reduce the three year recidivism rate by 10 percent. Several states are sharing data in order to better understand these risk mitigation factors.

Reducing recidivism in Utah is expected to produce the following benefits:

- Lower crime rates
- Less victimization
- Decreased incarceration
- Improved communities and families

Former New Jersey Attorney General, Anne Milgram, used smart statistical analysis and data analytics to substantially reduce crime in her state.

“Technology could help us leverage data to identify offenders who will pose unacceptable risks to society if they are not behind bars and distinguish them from those defendants who will have lower recidivism rates if they are supervised in the community or given alternatives to incarceration before trial. Likewise, it could help us figure out which terms of imprisonment, alternatives to incarceration, and other interventions work best--and for whom.”

- Anne Milgram

Crime Lab New York is one of a growing number of analytics driven programs around the country seeking to fundamentally change the way key policy decisions are made, substituting data-driven objective analysis to determine the impact of various kinds of interventions. The expectation is that analytics can increase public safety while also reducing recidivism.

In 2014, the Utah Department of Corrections continued a nascent investment in business intelligence, working to implement emerging social intervention approaches contained in the “Evidence-Based Practices” model.

Expected Goal: Reduce recidivism by 6%.

Budget

The Governor’s Office of Management and Budget has identified 4 key questions for budget-making in Utah.

- How do we know if the programs and services the state provides really work?
- How much should we spend?
- Are we getting the outcomes we want?
- What will we need in the future?

Data-driven analytics can help answer all of these questions by providing a more complete picture of what is happening across the state, while also potentially speeding up the analytical process.

Harvard Professor, Stephen Goldsmith, recommends that local government should begin with a data driven performance program that may be based on transactional data collected by agencies and then introduce data mining and analytics to identify the root cause of specific service challenges. Next, they should *“advance to predictive data analytics to solve more complicated and systematic problems.”*

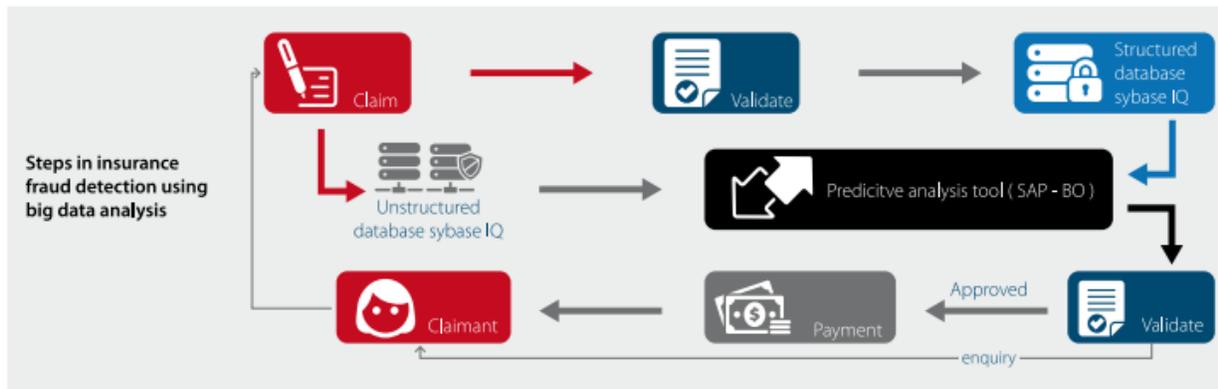
There may be a tendency through the budget process to propose additional “big data silos” as agencies seek to address their growing data needs at the agency level. NASCIO and a team of experts strongly recommends that states avoid this tendency.

Fraud Detection and Prevention

States have already achieved huge savings through the use of big data in reducing fraud.

“Big Data is saving the state of Michigan \$1 million each business day. Cynthia Edwards, director of Medicaid health information technology, said that the Medicaid savings have been measured by the Department of Community Health for the state of Michigan, which achieved the savings by billing Medicaid patients’ existing insurance plans when they had them, using death records to recover payments made to clients who had died, and reducing fraud.”

The federal government saved \$4.3 billion by reducing medical fraud using big data in 2013 and expects to save a total of \$19.2 billion over a five year period. A model for using data analytics in fraud detection might work like this:



Predictive analytics using real time data processes will be able to help identify potential fraud even before a transaction is complete. A recent example of where this might be applied is the recent discovery of as many as 8,000 fraudulent 2013 Utah tax returns.

Expected Goal: Reduce fraud in Utah by \$30 million

Transportation

UDOT has already demonstrated the value of data to improve transportation planning and management. With various sensor networks now providing real-time data feeds and new technologies such as self-driving vehicles on the horizon, the ability to process and analyze these data flows is critical. Big data technologies and predictive analytics are needed for this to happen.

According to George Leopold, “Transportation planners are embracing big data analytics as the nation’s highways become more congested and data science in some instances replaces costly road-construction projects.” Predictive analytics tools, combined with the data management and processing capabilities of Hadoop will enable agencies to better manage and utilize terabytes of previously disassociated data as a strategic planning asset, enhancing capabilities that already exist in a platform such as UPlan.

The state of Oregon is looking at big data analytics as a way to more holistically look at data about the public, demographics, work patterns, housing and employment trends, travel preferences, freight flows to feed into transportation models that serve as the basis for transportation plans. Big data will be a key component in future transportation applications such as the connected vehicle, electronic ticketing for transit, traveler information systems, and advanced traffic management systems.

Healthcare

The healthcare ecosystem is increasingly using electronic databases and tools to deliver services. Utah's all-payer database will soon have hundreds of millions of records and the ability to make rational decisions based on this data and support improved healthcare solutions is dependent on our ability to process and analyze the data. The Utah Department of Health has recognized the need to use analytics to understand, process, and manage.

The state of Indiana is using big data analytics to identify the core causes of its higher-than-average infant mortality. The state expects to reduce the death rate from 7.4 deaths per 1,000 births to 6.89 by 2016. The U.S. expects to use big data to reduce waste in the healthcare industry and lower costs. McKinsey Global estimates that up to \$300 billion in value can be captured by applying big data to the U.S. healthcare sector. With the rapid adoption of electronic health records and the growth (in Utah) of extremely large data sets like the All Payer Claims database (APCD), a big data and analytics infrastructure will be critical to provide a way to effectively use the data to provide better service to the public.

Other

Other areas where the state can benefit from improved use of big data and analytics include:

- Education
- Cyber threat detection and response
- Intergenerational poverty
- The management of intelligent sensor networks / smart cities
- Performance management
- Occupational planning and workforce development
- Environmental quality

Intangible Benefits

The use of advanced analytics using large, diverse data sets provides capabilities to do analysis at a large scale that cannot be done at a small one. This presents a host of new, albeit intangible opportunities. Bigger data is not better, but it is different and the results that you are able to achieve may be more significant. In order to unlock its potential, we will have to make changes to how we think, manage and operate.

By vastly increasing the data we use, we can incorporate lower quality sources and still be very accurate. We can continue to reevaluate, correct errors in initial assessments and make adjustments as facts from agency operational experience change. Some specific intangible benefits that may result from the effective use of analytics may include:

- Data driven decision-making. Over time, we will develop patterns of decision-making that are more effective and produce better results.
- Deep learning
- Ability to see complementary relationships between public sector and private sector data
- Increased understanding of data patterns outside of traditional data silos
- Ability to define and apply benchmarks from multiple sources of data with unanticipated results

What does good look like?

The big data and analytics platform of tomorrow will ensure privacy and security while also creating opportunities to analyze vast amounts of data in a wide variety of formats, including streaming data. Creating this platform can be done in house or augmented through the

purchase of cloud-based services. This platform should be enterprise in nature and facilitate a wide range of analytical methods.

The technological platform to deliver advanced big data and analytical technologies is not enough however, if the state wants to fully benefit from the potential of these technologies. The state will need a team of well-trained data scientists that understand advanced analytics and managers that are willing to explore and act on new information. The enterprise analytics team will be applied to a wide variety of tasks and improve the state's ability to manage complex challenges based on data.

Overall, the data analytics solution should include

- Cloud-based or on-premise big data platform with NoSQL database structure to support real-time querying and analytics
- Advanced analytics services
- Enterprise team of data scientists
- Data flow capabilities to ensure the ability to analyze and respond to high speed data flows such as network traffic analysis and sensor networks.
- Effective Governance: One of the biggest potential obstacles to the effective use of big data is that of data access.

The data analytics solution requires problems to which data can be applied, with management willing to make data-driven decisions. Various types of analysis can be applied to problems already facing the state as shown in the following table. Expected impacts (results) may vary depending on the ability of management to apply data-driven decisions.

Solution Workflow

The analytics team uses their expertise in statistics, data mining, machine learning, and visualization to answer questions that agency and state leaders pose. They draw on data from enterprise sources (e.g., economic data, financial data, or performance data) called business information, sometimes in combination with data from public sources.

So, if Natural Resources had a problem with defining the best solution for managing an endangered species, for example, they could present the problem to the analytics team who would have access to relevant data sets from across the enterprise (transportation data, health data, construction and population growth information, public sentiment data through social media, etc.). The analytics team will be able to run a wide variety of analytical processes to support the agency request, providing predictive analysis for use in a variety of scenarios. DTS would provide and support an underlying platform as well as the capabilities for securing and managing the data.

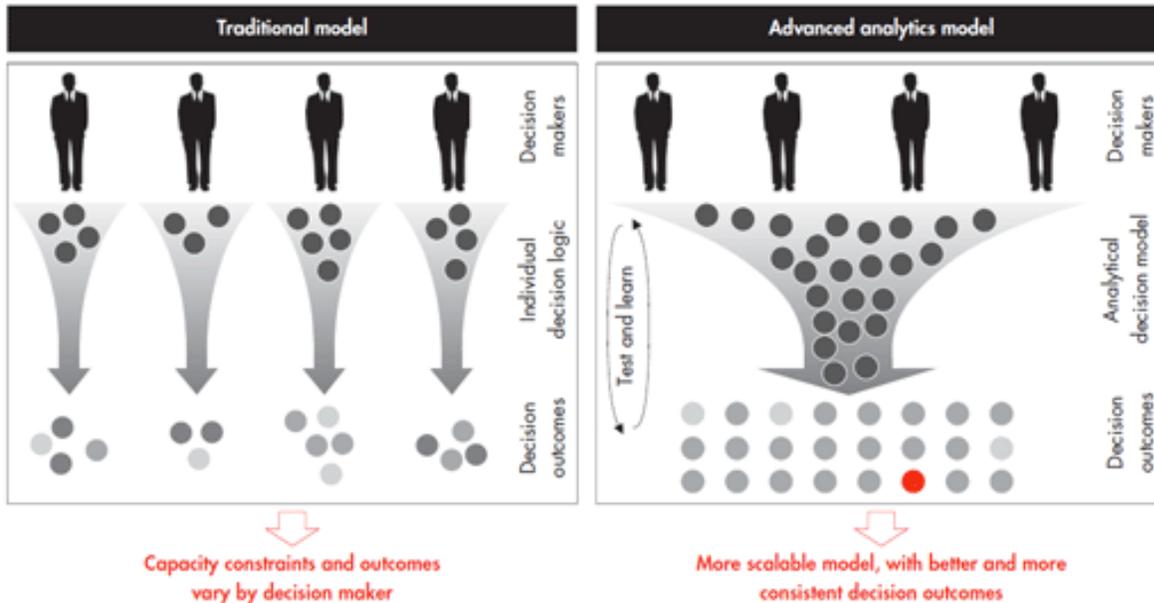
Necessary characteristics of the technology solution

1. Can it handle high data velocity?
2. Can it handle data variety?
3. Can it handle large data volumes - Terabytes to Petabytes?
4. Can it handle complexity?
5. Can it scale cost effectively without becoming overly complex to support operationally?

The Solution

The Utah Center for Advanced Data Analytics can become a model for how governments can leverage data to tackle the most critical issues facing society. With access to data from across the state enterprise, the Center for Advanced Analytics will have access to capabilities that provide quicker, more accurate, and more complete results. The Center will tackle the most challenging, complex issues facing state government.

Traditional Decision Making vs. Decision Making with Advance Analytics



Source: Bain & Company

In order to achieve the results that we envision for the Center for Advanced Data Analytics, the State will have to invest in human resources. For buildout, we expect that the team will need the following personnel with their associated skillsets, some of which are new to state government:

- 4 Data Scientists
- Big Data Hub Administrator
- Enterprise Data Security Administrator

The Utah Center for Advanced Analytics will coordinate closely with the Automated Geographic Reference Center to integrate location-based analysis, as well as agency analysts and other resources. The Center should also coordinate with higher education resources to expand Utah's big data ecosystem.

Governance

In order to address issues that will arise with more effective use of data, the State needs a way to coordinate involvement, resolve issues of data security and access, determine data ownership and approve data sharing issues. A big data oversight board with executive participation from multiple agencies would also need to coordinate decisions, set priorities, and activity with the Security Advisory Board and the Transparency Board, both of which also have an interest in how data in the state is accessed and utilized.

Cost to get started

The cost of implementing a statewide big data and analytics solution will most likely be weighted more towards skill development, personnel costs, and solutions implementation than in the technology itself. The technologies became attractive to industry because they were, in fact, a lower cost way of delivering better results. Most of the original technologies, other than the storage itself, were developed as open source software. Since then, many of these technologies have been commercialized, providing even more alternatives, particularly if the state should require external support.

Hadoop proponents claim that fairly standard Big Data platforms cost between 1/5 to 1/20 the cost of other data management technologies and are at least ten times more scalable.

An enterprise big data and analytics solution should be implemented over a period of 3-5 years.

Year One - Setting Up a Big Data Lab

The Big Data Lab is a dedicated development environment, within the current technology infrastructure, that can be created explicitly for experimentation with emerging technologies and approaches to big data and analytics.

Key Activities

- Assemble a selected set of technologies to be evaluated during first 2-3 months
- Test permutations against high value use cases
- Develop recommendations from the testing scenarios to drive future architecture and usage

Objectives

- Deliver 2-3 “Quick Wins” to demonstrate the value of these technologies from both an IT and business perspective
- Create a “Proof-of-Concept” that shows how these technologies can be integrated into the State’s enterprise existing architecture
- Develop future-state AI architecture recommendations
- Deliver low-cost, high-performance, agile BI and data discovery, with a focus on big data technologies
- Pilot new analytical capabilities and use cases to prove business value and inform a long-term roadmap to compete on analytics
- Establish a permanent “Innovation Hub” within the state architecture and center for big data and analytics skill-building

Personnel

1 Data Scientist

Big Data Hub Administrator

2 Technical Support to Manage Data, Connectivity, Automation, etc.

Existing Personnel: CTO, Chief Architect, Data Manager, Other IT Staff

External Consultants

Technology & Infrastructure

The Big Data Lab's research mission is to identify, engineer and evaluate innovative technologies that address current and future data-intensive challenges. In order to unlock the potential of big data, we need to overcome a significant number of research challenges including: managing diverse sources of unstructured data with no common schema, removing the complexity of writing auto-scaling algorithms, real-time analytics, suitable visualization techniques for petabyte-scale data sets, etc. The Big Data Lab will provide the platform to test the hypothesis and integrate big data efforts across the organization.

Estimated Year One Cost: \$2,350,000

Table 1

Year 1	Units	Cost / Unit	Total
Data Scientist	1	175000	\$175,000
Big Data Admin	1	150000	\$150,000
Technical Support	2	135000	\$270,000
Terabytes of Storage	500	1000	\$500,000
Networking	1	30000	\$30,000
Servers	20	20000	\$400,000
Security Administrator	1	125000	\$175,000
Analytics Software	2	25000	\$50,000
Visualization / Presentation Software	5	20000	\$100,000
Consultant Support			\$500,000
Total			\$2,350,000

Year Two - Creating the State Big Data Hub

By the second year, the State should have an established governance model and process for identifying and loading data to the big data hub. Parameters and technology for the hub should be in place. Statewide, we should have addressed the following questions:

- Alignment: What is the alignment to the big data effort across the enterprise?
- Architecture: How advanced is the big data architecture, and to what degree do groups adhere to architectural standards?
- Data: To what degree does the data provided by the big data environment meet business requirements?
- Data governance: Does the organization have a data governance model, and how effective is data governance with the big data program?
- Delivery: How aligned are reporting and analysis capabilities with line-of-business user requirements, and what is the extent of the usage?

- Development: How effective is the big data team’s approach to managing projects and developing solutions?
- Measurement: What are the different strategies for measuring success? Who defines these measurements and their associated success criteria?
- Program governance: Does the organization have a program governance model? How effective is the program governance with the big data program?
- Scope: To what extent does the big data program support all parts of the organization and all potential users?
- Skills: Does the organization have the skills needed to support the big data initiative?
- Sponsorship: To what degree are big data sponsors engaged and committed to the program?
- Statistical model: What are the requirements for measuring and monitoring the performance of the enterprise? What are the different techniques used today and expected from big data technologies and solutions?
- Technology: What are the key capabilities available within the organization for big data technologies?
- Value: How effectively does the big data solution meet business needs and expectations?
- Visualization: What are the key requirements for data visualization and analytics delivery?

Once we have assessed these questions at a reduced scale, the state should be able to expand the Lab concept to a statewide big data hub implementation. It will be particularly important that all security protocols are in place and that access procedures have been fully defined and tested.

Personnel

3 Data Scientists

Big Data Hub Administrator

2 Technical Support Personnel

Full-time big data security specialist

Estimated Year 2 Cost: \$3,110,000

Year 2	Units	Cost / Unit	Total
Data Scientist	3	175000	\$525,000
Big Data Admin	1	150000	\$150,000
Technical Support	2	135000	\$270,000
Terabytes of Storage	600	1000	\$600,000
Networking	1	50000	\$50,000
Servers	25	20000	\$500,000
Security Administrator	1	125000	\$525,000
Analytics Software	3	25000	\$75,000
Visualization / Presentation Software	5	20000	\$100,000
Consultant Support	1		\$315,000
Total			\$3,110,000

Year Three - Maturing the Statewide Big Data Environment

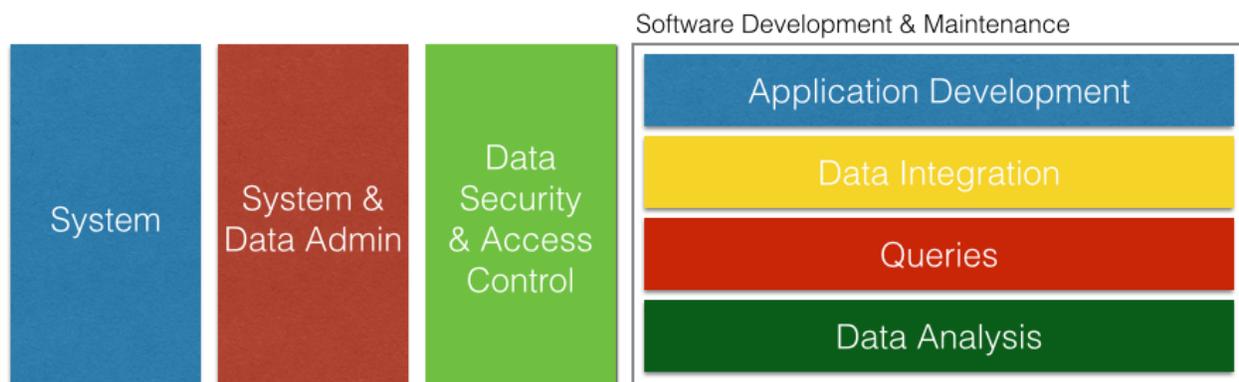
In year three, the expanded data hub should be available for use across the enterprise. If the first two years are successful and we have seasoned data scientists along with a well-oiled governance structure, the Utah Center for Advanced Analytics should be ready to accept diverse workloads from any state agency.

Estimated Year 3 Cost: \$3,760,000

Year 3	Units	Cost / Unit	Total
Data Scientist	4	175000	\$700,000
Big Data Admin	1	150000	\$150,000
Technical Support	2	135000	\$270,000
Terabytes of Storage	800	1000	\$800,000
Networking	1	50000	\$50,000
Servers	30	20000	\$600,000
Security Administrator	1	125000	\$700,000
Analytics Software	5	25000	\$125,000
Visualization / Presentation Software	5	20000	\$100,000
Consultant Support			\$265,000
Total			\$3,760,000

Over time, the state of Utah should be able to eliminate some existing infrastructure costs as current architectures are replaced. About 2 petabytes of data are currently stored in the data center supported by the Department of Technology Services. This data is stored in a diverse array of data warehouses, relational databases, etc. Implementation of an advanced big data hub will not only enable us to leverage these valuable data stores to better use today, but it is critical for the state to prepare for new and emerging technologies that are already developed and will make their way to state government in the next 3-5 years.

Total Cost of Data



Perhaps the biggest cost to the state may be that of not getting started on this project in the near future. The result will continue to be lost opportunities and less-than-adequate analysis while the data continues to grow at an ever increasing pace and the state falls further behind in its ability to deal effectively with its data. Fortunately, several local universities have already implemented big data analytics into their curricula and have programs to develop data scientists and other professionals that understand how to work in these environments.

Cost Factors

	Cost Factor	Value	Comment
1	Hadoop cluster acquisition cost	\$1000	per TB - open source software is free
2	Hadoop cluster maintenance and support cost		per TB per year
3	Total # of TB managed in cluster	500	TB of data in year 1
4	Cost per year for a Hadoop administrator	\$150,000	Salary+Benefits / Year
5	Expected rate of annual growth	20%	
6	Lines of Java/Map Reduce Code for data refinement	300,000	
8	Cost for Data Scientist	\$175,000	Salary+Benefits / Year
9	Number of Data Scientists		
10	Data Security Administrator	\$125,000	Salary+Benefits / Year

Success Stories

In the Los Angeles and Santa Cruz police departments, a team of researchers have taken an algorithm used to predict earthquakes, modified it and started feeding it crime data. The software can predict where crimes are likely to occur down to 500 square feet. In LA, there's been a 33% reduction in burglaries and 21% reduction in violent crimes in areas where the software is being used.

KTH Royal Institute of Technology in Stockholm, which uses streaming data in their congestion management system, is already reducing traffic in the Swedish capital by 20 percent, lowering average travel times by almost 50 percent and decreasing the amount of emissions by 10 percent

The Delaware River Commission used analytics to increase trout and shad habitat by about 200% and provide a modest increase in flood-mitigating reservoir voids with a minimal impact on the risk of drought in New York City. The states of New York, New Jersey, Pennsylvania, and Delaware, along with New York City, unanimously approved the solution. The Delaware River Basin Commission and the U.S. Geological Survey Delaware River Master administer it.

Economic benefits include an estimated \$163 million annual increase in fishing and boating income in the Upper Delaware River Valley and potential flood mitigation.

The success of North Carolina's Criminal Justice Law Enforcement Automated Data Service led to the creation of the North Carolina Government Data Analytics Center, which allows state agencies to share data. One of its goals has been better data analysis to identify fraud, waste and compliance in areas such as unemployment insurance and workers' compensation.

"One of the first projects we tackled was unemployment insurance, as that program had nearly \$3 billion in debt," says North Carolina Senator Ralph Hise (R), chair of the Senate Health Committee and co-chair of the Senate Appropriations on Health and Human Services. "IT systems at the Employment Security Commission, labor department and tax office couldn't talk to one another, which made it difficult to identify fraudulent claims. As a result of data retrieved from the center and other changes to governing rules and regulations, we reduced that debt to under \$1 billion in 16 months."

Conclusion

Almost \$10 million dollars over a three year period sounds like a lot to spend on a new data management and analytics system. It is a significant investment. Results from the public and private sector indicate that it is an investment that merits strong consideration. A single "big win" in any of the areas discussed above would cover the cost of the investment.

Making an Impact with Big Data and Actionable Analytics

	Prison Recidivism	Budget	Fraud Detection	Healthcare	Transportation
Applications	<ul style="list-style-type: none"> • Risk Mitigation • Parolee Support Services • In-prison training and release preparation 	<ul style="list-style-type: none"> • Expense Analysis • TOS • Budget Effectiveness • Expenditure Forecasting • Revenue vs. Expense 	<ul style="list-style-type: none"> • Medicaid • Insurance • Welfare • Tax 	<ul style="list-style-type: none"> • Patient Community Analysis • Healthcare decision support 	<ul style="list-style-type: none"> • Long Term Transportation Planning • Traffic Management • Road Maintenance
Data	<ul style="list-style-type: none"> • Prison records • Arrest Records • Criminal Records • Crime Maps • Human services 	<ul style="list-style-type: none"> • FINET Expense Data • Performance Data 	<ul style="list-style-type: none"> • Medicaid Expenditures & Filings • Unemployment Records • eREP data • Tax Records 	<ul style="list-style-type: none"> • All Payer Database • Electronic Health Records • Immunization Data • Laboratory Results • 	<ul style="list-style-type: none"> • Environmental Data • LIDAR Data • UPlan • Local • Sensor Data • Traffic • Project expense records • Road maintenance, crash data
	<p>Characteristics: Fragmented information sources and legacy systems, rich textual content, unstructured informal citizen conversations</p>	<p>Characteristics: Structured financial data</p>	<p>Characteristics: Personal identity information, incomplete and deceptive content, rich group and network information</p>	<p>Characteristics: Disparate but highly linked content, person-specific content, HIPAA, IRB and ethics issues</p>	<p>Characteristics: Sensor data, streaming content, real-time, GIS,</p>
Analytics	<ul style="list-style-type: none"> • Association rule mining • Database segmentation and clustering • Anomaly detection • Graph mining • Social network analysis • Text and web analytics • Sentiment and affect analysis 	<ul style="list-style-type: none"> • Information integration • Privacy-preserving data mining • Content and text analytics • Government information semantic services and ontologies • Sentiment and affect analysis 	<ul style="list-style-type: none"> • Outlier analysis • Repetitive or continuous analysis • Benford's Law 	<ul style="list-style-type: none"> • Genomics and sequence analysis and visualization • EHR association mining and clustering • Health social media monitoring and analysis • Health text analytics • Health ontologies • Patient network analysis • Adverse drug side-effect analysis • Privacy-preserving data mining 	<ul style="list-style-type: none"> • Road condition analysis • Traffic pattern analysis
Impacts	<p>Reduced recidivism, prison population reduction, improved parolee re-entry and integration into society, reduced inconsistencies between judges</p>	<p>Improved budgeting process, better correlation of spending and performance, focus on results</p>	<ul style="list-style-type: none"> • Trust in government • Cost savings • Set expectations 	<p>Improved healthcare quality, improved long-term care, patient empowerment</p>	<ul style="list-style-type: none"> • Improved traffic flow • Optimized maintenance costs and project management • Air quality improvements