Minneapolis
City of Lakes

IT Analytics Hub Consultation

City of Minneapolis Internal Audit Department

August 19, 2019

Table of Contents

Executive Summary

Data is one of the City of Minneapolis' largest assets. However, data at the City has historically been siloed in teams or departments, resulting in multiple versions of the truth, inability to find data from another department to aid in analysis, decentralized security, lack of understanding about what data a department contains means and how it should be used, and inability to effectively use data for analysis.

The IT – Data & Analytics Services team was formed with the mission to aid departments in understanding and using their data, making it accessible and usable across the enterprise. To meet this goal, the team created the IT data strategy, a holistic process that drives how data is collected, managed, stored, and used at the City. The team also created the Analytics Hub, a platform for data use and analysis that brings data from multiple departments together in a usable, repeatable method. Released in July 2017, the Analytics Hub is a tangible product of data strategy. It brings data from multiple systems and data classifications together in a way that is easy to use, repeatable, expandable, and secure.

As the Analytics Hub has matured over the last two years, departmental visibility to and trust in data has increased. Departments are placing more value on their data to help make decisions and demonstrate outcomes. As IT – Data & Analytics Services continues to build and increase support for the Analytics Hub, they requested Internal Audit conduct a review of processes used in development and maintenance of the Analytics Hub today, with the goal being to identify and recommend best practices needed to continue a sustainable model for data and analytics within the City.

Objective, Scope and Procedures Performed Approach

The City's Internal Audit Department is conducting a consultation of the Analytics Hub to review governance and controls around data collection and reporting of information. The scope for this engagement is to perform the following activities:

- **Analytics Hub Data Governance**: Provide compliance advice, counsel and best practice recommendations for the policies, standards and processes developed by the IT Data & Analytics team to support the Analytics Hub Data Governance initiative.

- **Analytics Hub Access Management**: Provide compliance advice, counsel and best practice recommendations for the policies, standards and processes related to logical access controls for provisioning/deprovisioning and user access review and management.

Because the IT Data Strategy is an adaptable, holistic process and certain data elements are utilized by other departments to drive the business, data quality risk index must be considered. The progression and success of each of these recommendations may rely heavily on the successful implementation of the others.

The City's Internal Audit Department gathered supporting documentation, facilitated meetings with key stakeholders and obtained feedback in the creation of the recommendations listed below.

Results

As a result of this consultation, the following observations were noted and presented to management:

1. The City should work to formally develop a Data Governance Policy to help communicate the value of data governance to both business users and leadership.
2. The City should work to formally develop a data classification policy. The creation of data classification policy should be executed through review or performance of the following:
    a. Ensure compliance/alignment with the Minnesota Government Data Practices Act (MGDPA)
    b. Perform a risk assessment of data/information collected by the Analytics Hub
    c. Creation of risk management plan based on the risk assessment
3. The City should work to gather enterprise support for the Analytics Hub through the following methods:
    a. Operationalize the Enterprise Data Strategy
    b. Identification of working groups and data stewards
    c. Creation of Champion Story
4. The City should work to formalize/refine the following logical access internal control areas associated with the Analytics Hub.
    a. Access Requests
    b. Transfers
    c. User Access Reviews

Conclusion

Internal Audit noted a limited number of non-critical opportunities for process and technology improvement to help reduce the number of inaccurate complaints received from residents. No substantial risk exposures or material control weaknesses were identified during the scope of this consultation.

The Internal Audit team would like to thank Information Technology and Data and Analytics Services staff for their cooperation, time and effort during this engagement.

Internal Audit Team on this Engagement
Sam Boterman, CISA, CCSFP, ISO27001 Lead Implementer, Manager, Baker Tilly
Justice Kanu, Consultant, Baker Tilly

IT Analytics Hub Primary Contacts
Eero Kilkson, IT Director, Data & Analytics Services
Pam Helf, Database Engineer, Data & Analytics Services

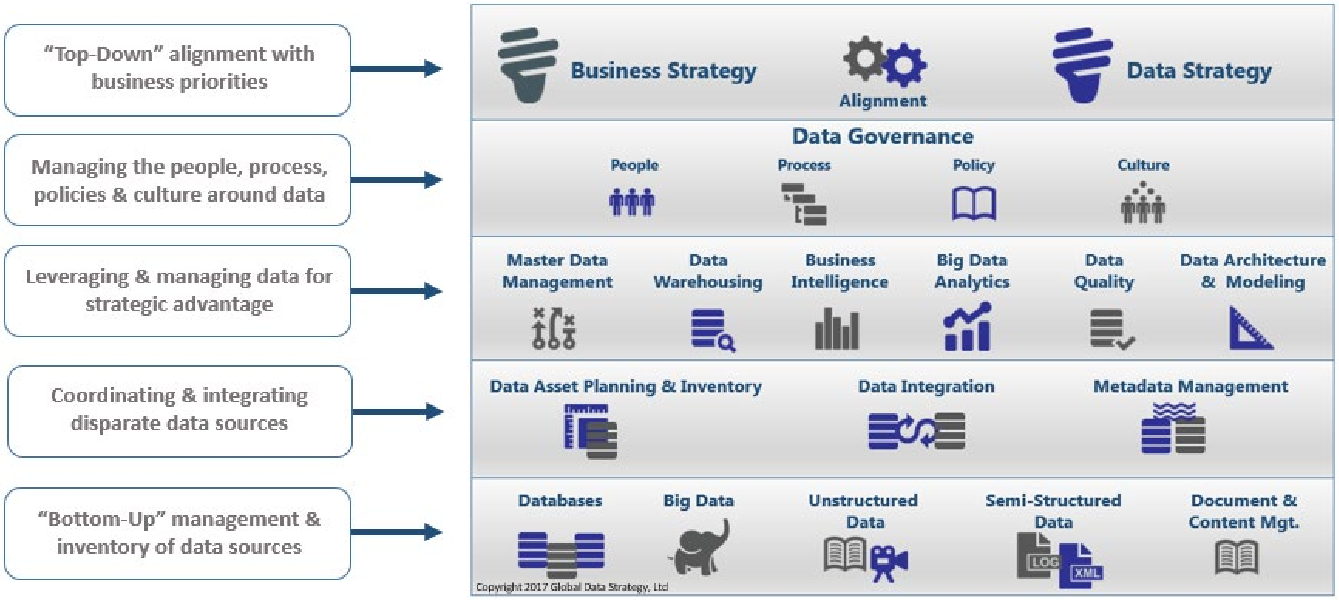
City of Minneapolis Office of Internal Audit
Ginger Bigbie, CFE, CPA, Internal Audit Director
Phone: (612) 673-5938

Analytics Hub Background

The City's 2018 State of Data report[1] states data is one of the City's greatest assets and liabilities. Managing data appropriately is critical in an age where it can be used both to inform and harm. Accordingly, IT – Data & Analytics Services was formed with the mission to help departments understand and use their data accurately and appropriately. The team designed the IT Data Strategy, which is built on top of four core tenets:

- Data should be *trustworthy*.
- Data should be used, stored, and managed *intentionally*.
- Data should be *available* to those who need it, and secured appropriately.
- Data should be *transparent*, enabling better conversations between the City and the public.

Data strategy is a complex, holistic process. The graphic below highlights all the components of an effective data strategy, which are all taken into account and implemented as part of the work IT-Data & Analytics Services performs.


"Top-Down" alignment with business priorities
Managing the people, process, policies & culture around data
Leveraging & managing data for strategic advantage
Coordinating & integrating disparate data sources
"Bottom-Up" management & inventory of data sources
Business Strategy
Alignment
Data Strategy
Data Governance
People
Process
Policy
Culture
Master Data Management
Data Warehousing
Business Intelligence
Big Data Analytics
Data Quality
Data Architecture & Modeling
Data Asset Planning & Inventory
Data Integration
Metadata Management
Databases
Big Data
Unstructured Data
Semi-Structured Data
LOG
XML
Document & Content Mgt.
Copyright 2017 Global Data Strategy, Ltd

Source: *https://www.dataversity.net/data-management-vs-data-strategy-a-framework-for-business-success/*

In July 2017, IT-Data & Analytics Services released the Analytics Hub, a reusable, sustainable data warehouse platform for data analysis and understanding. The Analytics Hub combines information from multiple systems and departments across the City in a meaningful, easy to use way. In releasing the Analytics Hub and empowering departments how to understand, use, and improve their data, the team enables data to be used as an asset.

[1] *http://lims.minneapolismn.gov/Download/File/1804/2018%20Open%20Data%20Portal.pdf*

The Analytics Hub is dynamic, and since its initial release, the IT-Data & Analytics Services has continued to work with departments to add additional datasets to the platform and to empower them to understand their data, such as through the creation of dashboards to inform management in decision-making processes. Each dataset is secured appropriately according to the business use, leading to a trustworthy, intentional, available, and transparent platform that continues to grow, and on which City leadership and analysts can rely.

Recommendations

Recommendation #1

Creation of a Data Governance Charter

The City should work to formally develop a Data Governance Policy to help communicate the value of data governance to both business users and leadership. Data governance helps ensure the following:

- Stakeholder needs, conditions and options are evaluated to determine balanced, mutually agreed enterprise objectives to be achieved through the acquisition and management of data and information resources
- Direction is set for data/information management capabilities through prioritization and decision making
- Performance and compliance of data and information resources are monitored and evaluated relative to mutually agreed upon directions and objectives

Data governance reflects the practice of evaluation requirements and brings direction and control over data and information so that users have access to that data and can trust and rely on it. Data governance involves monitoring and performance citywide, through the Records and Information Management group, and IT operations, specifically those areas that relate to data and its availability, integrity, and confidentiality. The key elements of a data governance policy should include but are not limited to:

- Statement of Purpose/Goals – Establish and communicate overarching goals for each area below to achieve success. This should align to the core IT Data Strategy principles the IT - Data & Analytics Services group has identified surrounding trustworthiness, intentionality, availability and transparency.
- People - Define key data-related roles throughout the organization. For every data system, identify the data stewards who manage data as an asset and focus on data quality; the data owners who have decision making authority and define data quality standards and the IT staff who provide technical support and monitor compliance. The creation of a responsibility assignment matrix (RACI chart) would help to formally define these roles.
- Data Access – Formally define permissions and who has access to what data.
- Data Inventory – Inventory and document current data sources using the Analytics Hub. Regularly review and update the inventory to include new sources and remove any legacy sources.
- Data Usage/Content Management – Identify purpose for which data is collected and used for as part of the Analytics Hub. Communicate the purpose to business users, leadership, and the public as defined.

- Data Records Management – Develop and adhere to policies that define how records should be created, maintained and deleted.
- Data Quality – Assign responsibility for data quality to appropriate staff. A data steward should perform regular audits to ensure quality.
- Data Security – Define policies around data security, sharing of data and access to data through logical access controls. A data security risk assessment could be performed to indicate the risk and probability of risk acceptance.

Management Response – Recommendation 1

Management agrees with this recommendation. Over the past year the IT – Data & Analytics Services group has drafted an IT Data Strategy that it has used to develop solutions such as the Analytics Hub and analytics engagements, such as the Public Works mobility analysis. Within the Data Strategy consists a data platform strategy, design principles, procedural framework, and a higher-level organization chart to data design, storage, use, and governance. One of the elements of the implementation of the data strategy is the Analytics Hub which in its design, population and maintenance embodies many of the core elements of the data strategy. The Data Strategy also identifies the roles of City staff in the stewardship of data, the access granting process, and data design, both in the Analytics Hub and the source systems. Elements listed above in the recommendations are covered in the Data Strategy but should be fully documented as part of Data Governance formalization.

Recommendation #2

Creation of a Data Classification Policy and Risk Assessment

The City should work to formally develop a data classification policy. A data classification policy's primary purpose is to ensure that information and data is properly managed and handled with the respect to the threat it poses to an organization. Data and information being exported and received by the City should be maintained in an accurate, secure, and consistent manner and be readily available for personnel with authorized access. The creation of data classification policy should be executed through review or performance of the following:

- Ensure compliance/alignment with the Minnesota Government Data Practices Act (MGDPA)
- Perform a risk assessment of data/information collected by the Analytics Hub
- Creation of risk management plan based on the risk assessment

Ensure compliance/alignment with the Minnesota Government Data Practices Act (MGDPA)

The MGPDA is the state law that controls how government data is collected, created, stored (maintained), used and released (disseminated). The MGDPA outlines requirements relating to the right of the public to access government data and the rights of individuals who are the subjects of government data. Alignment to the MGPDA helps to ensure that all data housed within the Analytics Hub is maintained properly and has internal controls to safeguard the data. The data classification policy would integrate MGPDA and associated risks surrounding data in additional to the already enforced classifications administered by the City Clerk's office and Chief Information Security Officer (CISO).

The form or medium used to create or store government data is irrelevant and does not impact the data classification; the law applies equally to data that exists in “hard copy,” such as paper, or in electronic or digital format. In Minnesota, data must first be classified user the law to be government data. This is a complex task made more difficult by the fact that classifications change over time as determined by the State Legislature. The current data classification structure is shown in the chart below.

Classification of Government Data in the State of Minnesota

Data Classification	Data Category	Meaning	Example
Public	Data on Individuals Data Not on Individuals	This data is available to anyone for any reason	-Names of government employees -Minutes of a public meeting
Private Nonpublic	Data on Individuals Data Not on Individuals	This data is available to: -Entities authorized by law to have access -Those whose work requires them to have access -The data subject -Those authorized by the data subject	-Social Security Numbers -Most education data -Public health data
Confidential	Data on Individuals	Available to: -Entities authorized by law to have access -Those whose work requires them to have access This data is *NOT* available to the data subject	-Active civil investigative data -Active criminal investigative data
Protected Nonpublic	Data Not on Individuals		

Source: *http://lims.minneapolismn.gov/Download/File/1804/2018%20Open%20Data%20Portal.pdf*

Perform a risk assessment of data/information collected by the Analytics Hub

The City should perform an internal assessment of associated risk related to the collection of data/information by the Analytics Hub. An initial assessment should be performed for each data set and an associated level of risk (low, medium, high) should be assigned. Additionally, core enterprise data strategy principles such as trustworthiness, intentionality, availability, and transparency, as defined by the IT - Data & Analytics Services team should be taken into consideration during the risk assessment process. By classifying data based on associated risk, the City should be able to:

- More efficiently identify areas of risk in data within the Analytics Hub
- Manage the risk and impact of access that may be unauthorized or prohibited
- More efficiently comply with regulatory mandates related to the governance of data

Creation of risk management plan based on the risk assessment

The City should work to develop a risk management plan based on the results of the risk assessment. Based upon associated risk, each data set should have an appropriate response in how risk will be managed depending on the classification. These responses should take into consideration a solution that will help to ensure the data's confidentiality, integrity, and availability.

As a result of the review of the MGDPA, the risk assessment and associated risk management plan, data should be able to be classified and categorized appropriately to mitigate risk and ensure compliance through the establishment of a formalized Data Classification Policy.

Management Response – Recommendation 2

Management agrees with this recommendation. The formalization of Data Classification, which extends to all data repositories, will be the culmination of three inflight efforts that are designed to examine and classify City data from the perspectives of the Minnesota Data Practices Act, information security, and data use and analysis. Each approach is based on national standard frameworks that can be merged into a single repository.

Recommendation #3

Establish Enterprise Support through Working Groups and Data Strategy

The City should work to gather enterprise support for the Analytics Hub through the following methods:

- Operationalize the IT Data Strategy within the Enterprise
- Identification of working groups and data stewards
- Creation of Champion Story

Operationalize the IT Data Strategy within the Enterprise

The City should work to operationalize the IT Data Strategy by leveraging the Analytics Hub as a current, tangible, implementation of elements of the strategy. An enterprise data strategy statement helps to create a roadmap and plan for obtaining, handling, managing, and storing information used by the City. The IT Data Strategy currently defines data strategy through the following key pillars:

- Trustworthiness – Standards & Procedures
- Intentionality – Resources & Choices
- Availability – Scope & Access
- Transparency – Understanding & Methodology

By defining enterprise data strategy in association with the Analytics Hub this will help to guide departments and users of data stored within the Analytics Hub to create, share and understand the information and reporting. Additionally, this data strategy will help in the creation of associated working groups and data stewards.

Identification of working groups and data stewards

Enterprise support is required to help in the aid and establishment of associated working groups and data stewards for the Analytics Hub. By clearly defining the enterprise data strategy, associated working groups and data stewards can effectively leverage a documented approach and common playbook for how to collect, store, use, and enable data at the City.

Working groups should be established for any high priority or high-risk data related issue as needed. These working groups would focus on the core principle of intentionality and would leverage defined data governance, classification and enterprise data strategy to help provide recommendations and insight around information contained within the Analytics Hub in association with the Information Governance Policy Committee to determine and maintain accountability.

Additionally, data stewards should be defined across business lines to help ensure the quality and accuracy of the associated data elements that are utilized by the Analytics Hub. Data stewardship roles require development of skillsets in order to understand how to evaluate quality and accuracy associated with the data being passed to the Analytics Hub. Development of these skillsets requires time allocation, research and knowledge of key data elements, and authority from the business line to successfully perform these duties.

Creation of Champion Story

The City should work to gather enterprise support for the Analytics Hub by creating a “Champion Story” that provides immediate insight into a problem that was encountered, how the Analytics Hub was utilized and the successful result. A “Champion Story” will help to increase stakeholder engagement and involvement throughout the City around the Analytics Hub. Additionally, IT executive leadership should be engaged in the Analytics Hub by understanding the value of the Analytics Hub brings and requirement of collaboration with the Analytics hub during the Systems Development Life Cycle (SDLC) for new projects.

Management Response - Recommendation 3

Management agrees with this recommendation. IT – Data & Analytics Services designed and implemented the Analytics Hub as a tool to function as a physical implementation of the broader IT Data Strategy. Through the Analytics Hub, City staff can easily see the benefits of a consolidated data source of truth that is designed and maintained with the City user base in mind. The Analytics Hub also serves to stress the value of good data quality and the inefficiencies that poor data quality introduce. As departments have increased their usage of the Analytics Hub they have identified issues in their respective data and have made process changes to improve the quality. Many departments have identified informal data stewards to monitor data quality and make recommendations for new data entry procedures that will improve the data. This has been the start of the creation of the data steward community that is part of the IT Data Strategy. Using the base tenants outlined in the Data Strategy (Trustworthiness, Intentionality, Availability, and Transparency) as a guideline, IT - Data & Analytics Services has begun to engage departments in pilot implementation of data stewardship and departmental implementation of the data strategy.

IT – Data & Analytics Services has worked with many City departments to bring data into the Analytics Hub since its inception. In particular, three departments have had demonstrated success stories based on their use of the Analytics Hub and the underlying data strategy, including beginning to build out their data steward roles, defining governance processes, prioritizing good data, and adopting data strategy principles. We have documented these examples and can use them as case studies as we talk with new departments to gain support for the data strategy, establishment of data stewards, and the Analytics Hub.

Recommendation #4

Strengthened Logical Access Controls

The City should work to formalize/refine the following logical access internal control areas associated with the Analytics Hub.

- Access Requests
  - Current State: Ad hoc requests may occur either verbally or through an informal method which does not provide adequate controls surrounding the provisioning process.
  - Recommendation: Requests for access to associated roles/security groups as part of the Analytics Hub should be formally documented through an associated service ticket. This will ensure that data access request and associated approvals are retained for security purposes.
- Transfers
  - Current State: Transfer notifications are provided to the Analytics Hub on an ad hoc basis and these requests are typically provided in an untimely manner.
  - Recommendation: Employee transfers between departments or groups within the City should result in a review of associated roles/security groups as part of the Analytics Hub and access should be modified as necessary to ensure least privileged access.
- User Access Review
  - Current State: A user access review is not performed at this time on the Analytics Hub.
  - Recommendation: A formalized user access review for the Analytics Hub should be performed on at least an annual basis to ensure user access data is limited to associated roles and responsibilities. The City should work to develop a user access review process to ensure that least privileged access is maintained in conjunction with information/data that end users have access to.

Management Response - Recommendation 4

Management agrees with the recommendations related to Logical Access Controls. Currently, the IT- Data & Analytics Services group together with the IT – Security Services group are engaged in a project to audit and document the security and access process of reporting and analytic databases. A goal of this project is to develop standards, processes, and documentation templates for the audit of existing databases and the implementation of future databases. Two sets of artifacts have been created as part of the development of standards in support the IT Data Strategy: 1) database creation standards, recommendations, and checklists, and 2) checklists and standards for application and administrator access. Both sets of artifacts will be integrated in the database security process review project.