unifi

WHITE PAPER

Governance in a Self-Service Data & Analytics Environment

Introduction

Today, a modern enterprise is awash with both structured and unstructured data, legacy data warehouses, and private and public clouds. Data volume is rapidly growing and evolving; and in some instances, in real-time. While organizations place more importance on data, enterprises also face growing challenges of protecting, governing, and tracking data that is now beyond traditional on- premise networks, in motion, in the cloud, and shared outside the organization. The amount of data being generated is astounding. And it's impacting analytics; in fact, the International Data Corporation (IDC), estimates global revenues for Big Data and business analytics will grow from $130.1 billion in 2016 to more than $203 billion in 2020.

But there are also questions pertaining to IT organizations. What challenges do IT orgs face in the changing data landscape? How will a self-service data analytics environment impact IT? What is the impact on data governance?

As a leader in Data as a Service, Unifi is uniquely positioned to share strategic and actionable insights on the dynamic analytics landscape. The goal of this Unifi white paper is to deliver insights on Data Governance, debunk a few misconceptions, review challenges for IT, make the case for implementing a true self-service analytics environment, examine considerations when selecting a solution, and discuss how some regulations (General Data Protection Regulation) may impact your organization.

Data Governance Defined Myths Debunked

Before delving into the challenges and making the case for a self-service data & analytics environment, it's imperative we establish a clear and concise definition of Data Governance. Simply defined, Data Governance refers to the overall management of the availability, usability, integrity, and security of the data employed in an organization.

Current misconceptions of Data Governance include:

Myth 1 "IT owns it all."—False. Organizations should stress a company-wide adherence to data governance and put the plan into practice at every level. After all, analysts continue to use Microsoft Excel as the most common form of data interchange. Data Governance is not just limited to leadership or IT; it is and should be a company-wide concern and followed by all members of an organization.

Myth 2 "More governance means less collaboration." —False. Proper Data Governance will augment the value of the data. By securing the data and operating a governance policy you are enabling communication and collaboration.

Myth 3 "Data Governance is only for large enterprises."—False. Small and medium-sized companies have growing data and analytic needs, thus, Data Governance is vital to ensure the availability, usability, integrity, and security of the data. Regulations like the General Data Protection Regulation (GDPR) in the

European Union does not discriminate on the size of company causing a GDPR breach; in fact, the fine for a violation is still a minimum of €20M.

Who develops the Data Governance program?

Data Stewards are often the developers and keepers of Data Governance programs. They are the docents of the management of data elements and are seen as experts on the content of the datasets under their management. Many organizations establish a Data Governance Board comprised of a cross-section of key stakeholders that may be led by the Chief Data Officer, Chief Compliance Officer, Chief Privacy Officer or Chief Information Security Officer. This group defines the rules, forms policies, standards, requirements, guidelines, etc.

Current Challenges for IT Orgs

Challenge 1 Fragmented Tools, Approach & Data Many IT organizations face the obstacles of using different analytics tools for various tasks and requests for data insights. Data is also fragmented and thanks to legacy infrastructure, stored in data silos, with various sources, and in different formats.

Furthermore, IT is often bombarded by massive amounts of data requests from business users. However, IT can't stop everything to answer every single question. In addition, IT must figure out what is a priority, who gets data access, how much data access, etc.

Overall, these challenges represent a disjointed approach that is also highly disruptive to already over-burdened IT organizations and creates friction among teams.

Challenge 2 No Scalability Data volume and speed are affecting almost every operational element for business today. As an example, organizations looking at their customer 360 models realize that more data sources are available today than in the past; this includes publicly accessible data and proprietary data. However, combining all this data for a holistic view of how customers interact with their brand is vital and requires a scalable solution.

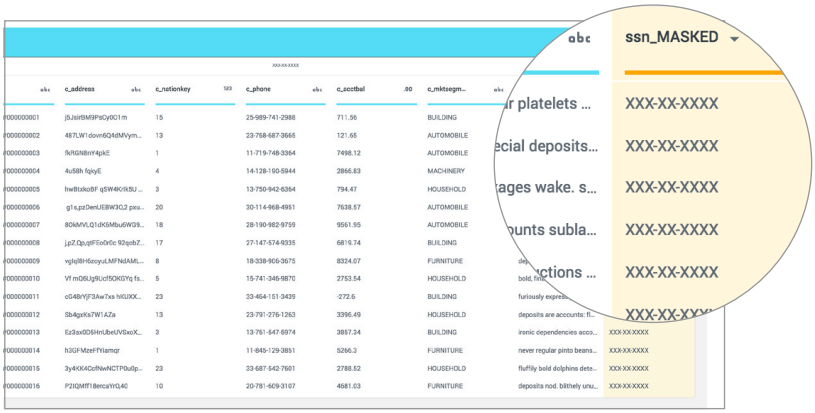
A similar scaling obstacle exists with organizations working on e-commerce shopping cart abandonment reduction.

Unfortunately, traditional analytics are unable to scale to massive data amounts, variations and real-time analytics needs of business users. Is IT meant to handle all this? No. IT organizations are erroneously seen as a one-stop-shop for data access. While IT is usually the steward(s) of Data Governance, they should not be the sole provider of data analytics. Unfortunately, some in the industry purport that so-called self-service BI tools are the key but to be self-service you must have access to all your data. You need to have have the platform to catalog, discover, cleanse, duplicate, enrich, normalize, transform, filter and format data for visualization.

Overall, as the analytics needs of business users grow, any organization chained to outdated data analytics tools and approaches will not scale adequately.

Challenge 3 Lack of Visibility & Control Data is everywhere today: On-premise, on mobile devices, in the cloud, outside the organization, and in motion. In terms of data analytics, data is flowing through sometimes multiple pipelines and creating value at different stages by users.

IT faces the challenge of controlling who is accessing the data. Often, when IT grants data access to business users,



A platform that provides a view into data lineage and enables row and column level security can address the challenge of lacking visibly & control.

GOVERNMENT DATA PROTECTION REGULATION (GDPR)

GDPR probably impacts your organization—are you ready?

What is it?

The EU General Data Protection Regulation (GDPR) is a new regulation, effective May 2018, that strengthens and unifies data protection laws for individuals within the European Union (EU). The regulation gives consumers a bill of rights which include:

- The right to personal data protection
- The right of notification in the event of a data breach
- The right to data portability
- The right to be forgotten/erasure
- The right for processing to be restricted

Who does it impact?

Consumers and enterprises. Any organization with any users in the EU needs to be compliant. This includes companies outside the EU that delivers goods, services or monitors EU user behavior.

Why you should be aware of the GDPR?

There are heavy penalties for violating GDPR with fines up to 4% of an organization's global revenue.

How does Unifi help prepare organizations for this regulation?

Unifi RegAlert! helps prevent GDPR violations. With RegAlert!, data stewards are automatically and immediately notified when a user tries to access GDPR data. Data stewards can grant unrestricted or restricted access to the requester. Even with restricted access, where row or column level data containing consumer PII is masked, an analyst using Unifi can use the data to derive quantifiable visualizations.

Unifi is helping organizations around the globe prepare for the upcoming GDPR requirements, its impact on data governance and consumer data privacy, while providing innovative and granular control and visibility to customers via RegAlert!

Unifi RegAlert! prevents unauthorized GDPR data access and alerts the assigned Data Steward via the policy engine when an analyst requests access. The Data Steward sets the level of access allowed and the analyst is notified so users can proceed with their analysis.

IT loses control and visibility of what happens to the data. They are unable to track and follow the lifecycle of the data. Furthermore, so many requests for data access are pouring into IT, they have difficulty managing and prioritizing the requests.

Self-service Analytics: Evaluation Considerations

A self-service analytics environment can overcome some of these obstacles while assuaging concerns.

However, not all self-service analytics solutions are the same. Some analytic vendors only have limited features and lack robustness in other aspects; for example, an analytic solution may purport data discovery capabilities but have no features/technology to support data preparation.

Before selecting a self-service analytics solution, make sure the solution meets your organization's specific needs and challenges. Unifi has created some best practices for considering and evaluating a self-service analytics solution.

THE WORLD OF COMPLIANCE IS GROWING AND CONSTANTLY CHANGING

The role of the compliance analyst is complicated and getting worse as their areas of responsibility expand to include monitoring and analysis. IT can avoid an increase in requests for data sets and insights if they enable the compliance analyst to be self-sufficient.

Key Solution Evaluation Considerations

Single Platform—An analytics platform that delivers an end-to-end data tool set that eliminates the need for disparate and outdated analytics tools. Data Discovery and Data Preparation should exist within the same platform.

Improved Data Governance—Look for a solution that allows IT to grant data access at row and column-level granularity. Be sure that Data Governance is built into the fabric of the technology and not an add-on piece.

Data Lineage—Understanding how derived data sets were created is essential to help analysts and data scientists understand the provenance of the data. T his is especially important in compliance applications, breaches of compliance are most likely to occur with derived data sets rather than source data.

Scalability—Make sure your self-service analytics solution can handle the growing and evolving analytics needs of concurrent users and growing data volumes and velocity. The solution should be cloud-optimized and take advantage of the "elastic" scalability that leading environments deliver.

Higher Control and Visibility—IT must be able to track the data as it traverses through the lineage pipeline. Access to data may happen at different points and create value for different users, but IT should be aware of the data lifecycle to mitigate any risks to sensitive data. IT must know not only who is using the data, but how and when.

Greater Collaboration and Communication—Consider a solution that promotes communication not only between IT orgs and business users, but throughout the organization among key stakeholders. Does the solution allow users to comment and rate the quality of data? Collaboration should be seamless for users.

Operational Simplification—Data catalog and discovery features should display all data sources and derived data sets of interest to the business user, even if they are not authorized to access them. A simple access request operation should alert the designated data steward to evaluate the request and simply grant restricted or unrestricted access to the data sources.

Innovation—Does the vendor have a roadmap for new features or is the technology static? Is artificial intelligence or machine learning applied to automate functions that traditionally were hand coded or required a high degree of technical competency? Is innovation for customers and the industry a drive of the solution? Be sure to look for a self-service data & analytics environment that is putting innovation in motion with new features.

The above attributes are some examples of criteria to consider when evaluating a vendor. Another key capability is having a holistic, end-to-end solution on a single platform. As mentioned earlier, using disparate tools for various analytics functions only augments complexity instead of simplifying the analytics process. Using multiple, disparate vendor solutions leads to challenges such as:

- Increased complexity in terms of implementation and management
- Multiple connections to data sources
- Steep learning curve for business users and IT data governors
- Risk that if one of your vendors is acquired, the solution will no longer be supported

These are only a few of the obstacles organizations face when using various vendor solutions. If Data Governance is integral to a self-service analytics solution, it will create an environment where IT is the hero and enabler, giving power to business users to access the data themselves.

Conclusion

Data is fueling modern business organizations with strategic insights and driving incredible business value. However, as data constantly evolves, business users require faster data access and analytic insights. While IT orgs face challenges regarding Data Governance, a self-service data & analytics environment can overcome these challenges only if the platform embraces the principals and best practices described in this white paper. Fortunately, Unifi combines the technology, leadership and expertise—all within a powerful, single Data as a Service platform that benefits both IT and business users.

About Unifi Unifi provides Data as a Service in an integrated suite of self-service data tools that include Governance & Security, Cataloging & Discovery, AI-Assisted Data Preparation, Community Collaboration and is Cloud-Optimized. Governed by IT and operated by business users, Unifi alleviates data bottlenecks and delivers faster business insights.

1810 Gateway Drive, Suite 380, San Mateo, CA 94404 | 844-TO-UNIFI | info@unifisoftware.com | unifisoftware.com