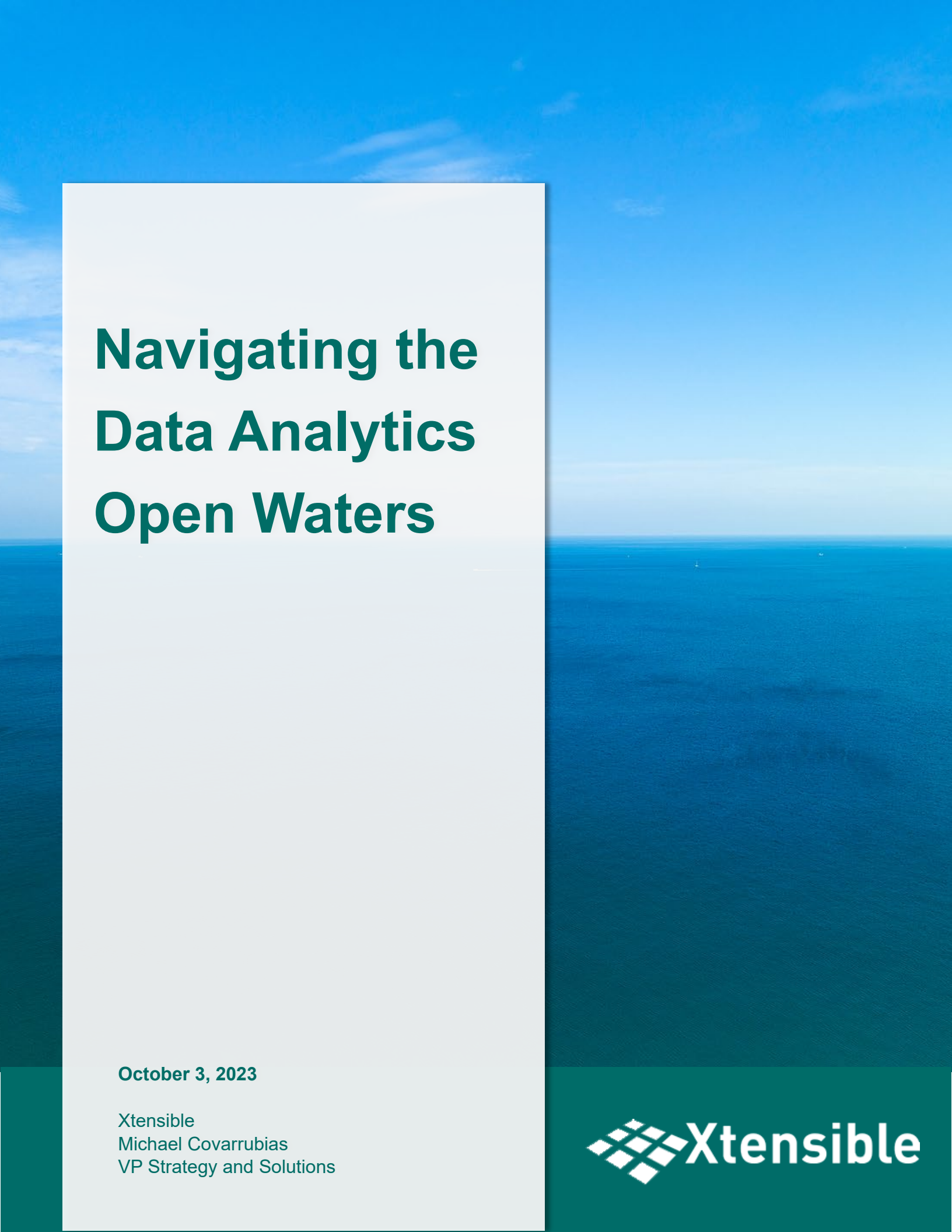


# Navigating the Data Analytics Open Waters

**October 3, 2023**

Xtensible
Michael Covarrubias
VP Strategy and Solutions

Xtensible

# Synopsis

**Navigating the waters of enterprise analytics can be seen as a daunting journey. If you're not careful you end up floating in the endless waters of data trying to figure out where you're trying to go and what to look for.**

## Why do you need to "plot" an analytics course?

Navigating the waters of enterprise analytics can be seen as a daunting journey. If you're not careful you end up floating in the endless waters of data trying to figure out where you're trying to go and what to look for. A CIO once remarked, at the height of the "big data" buzz cycle, that little data was the problem.

> "I don't have a big data problem. I have a little data problem. I have little data here, little data there, little data spread across the entire organization. I need to integrate this data to gain intelligence from it. "
>
> **Utility CIO**

## Understanding the waters of enterprise analytics

Enterprise analytics offers a wide range of benefits to organizations across all industries. But often an organization has different views and definitions as to what analytic is, what it is not, and what it can bring. They understand analytics is needed but how to effectively deploy analytics and for what purposes is often subject to many discussions across stakeholders.

> "We need a common analytics language across our organization, we need to understand key decision point for deployment. We need education! "
>
> **Utility CEO**

## What are different types of analytics?

Ask the question within your organization as to what are different types of analytics? Can everyone answer the question, and will they answer the same way? Different types of analytics require different data sets and methodologies to solve a specific problem or reach a business objective. The choice of what type of analytics to use for what purpose depends on the nature of the available data, the goals of the analytics, and the skillset and availability of resources as well as technology. Analytics can be broken down into five major groups:


- **Descriptive Analytics**: What has happened?
- **Diagnostic Analytics**: Why did it happen?
- **Predictive Analytics**: What could happen?
- **Prescriptive Analytics**: What action to take?
- **Cognitive Analytics**: What is the best action?

It is important to note that the further down the list you go, the more complex, the more data and specific approaches or technologies may be required. If you are just starting out on your journey you might not want to start with cognitive analytics as it leverages artificial intelligence, machine learning, and natural language processing. Ask yourself, do we have the data to support it? Do we have the technology to support it? Do you "trust" the data you have such that you would "trust" the cognitive analytics result?

Each of these analytic types you can further be decomposed. For example, let's consider Diagnostic Analytics. Diagnostic Analytics is a type of data analysis that seeks to understand why certain events or patterns occurred in the data. The primary goal of diagnostic analytics is to answer questions about what happened and why it happened. Key characteristics or features are:

- **Causal Analysis:** Diagnostic analytics explores the relationships between different variables to determine cause-and-effect relationships. It seeks to uncover the factors that led to specific outcomes or events.
- **Root Cause Identification:** The focus here is on identifying the underlying reasons or root causes behind observed trends or patterns in the data.
- **Hypothesis Testing:** Diagnostic analytics often involves hypothesis testing to validate or refute possible explanations for certain events.

- **Comparative Analysis:** It compares different datasets or segments to understand variations in outcomes and identifies factors that contribute to those variations.

It's worth noting that diagnostic analytics is often followed by predictive analytics, which uses the knowledge gained from diagnostic analysis to forecast future outcomes and recommend proactive actions.

## What is metadata and how is it used?

Ask the question within your organization as to what metadata is and how it is used? The simplest definition is that "metadata" is "data about data". Metadata is data that provides information about other data. It describes various aspects of a dataset, database, document, file, or other information resources. It helps you to understand, manage, and interpret the underlying data. In essence, metadata provides context, structure, and meaning to the data it represents including:

- **Descriptive Metadata:** Descriptive metadata provides information that describes the characteristics, content, and context of the data.
- **Structural Metadata:** Structural metadata defines the organization and structure of the data.
- **Administrative Metadata:** Administrative metadata contains information related to the management and administration of the data.
- **Technical Metadata:** Technical metadata provides details about the technical characteristics and properties of the data.
- **Preservation Metadata:** Preservation metadata relates to the long-term preservation and maintenance of the data.
- **Usage Metadata:** Usage metadata captures information about the usage and behavior of the data.

> ***Take Away***
>
> *Ensure everyone is in the "know" and has a common understanding. Take the time to ensure alignment across your teams; understand the real world behind analytics, educate, and apply.*

# What are core components for defining a successful analytics course?

A plan for the organization's analytics course helps it stay relevant and compliant in a rapidly evolving data landscape. There is much more to enterprise analytics than just selecting a technology for implementation. While technology is a critical enabler, focusing solely on acquiring the latest tools and platforms can lead to neglecting the importance of data strategy, people, and processes in analytics programs.

## Alignment with Business Goals

A plan ensures that analytics efforts are closely aligned with overarching business goals and objectives. Without a plan, analytics initiatives may lack direction and focus, leading to wasted resources and efforts. As part of the plan, it is essential

to identify the measurable outcomes. A plan defines clear key performance indicators and success metrics, enabling the organization to measure the impact of its analytics initiatives and make data-driven decisions.

## Resource Allocation

A plan helps in the effective allocation of resources, including budget, personnel, and technology. This ensures that the organization invests in the right tools and people to achieve its analytics objectives.

## Data Management and Governance

A plan also needs to include data governance policies and procedures, for data collection, storage, security, and privacy. This is especially important as data becomes increasingly valuable and is subject to regulatory requirements. For large organizations, data often resides in multiple databases and systems. The data management and governance plan contain the strategy for integrating data from disparate sources to create a unified and comprehensive view of the business.

## Plotted Course

Just as a ship is going to leave the dock, they must first plot a course, taking into consideration external impacts so must you with your analytics journey.

## Scalability

As an enterprise grows, its analytics needs evolve. A plan takes into consideration scalability and ensures that the infrastructure and processes can accommodate the increasing volume and complexity of data and analysis.

## Risk Management

It helps identify and mitigate risks associated with data analytics, such as data breaches, data quality issues, and compliance violations.

## Communication and Collaboration

A well-documented plan facilitates communication and collaboration among different departments and teams within the organization. It is the common roadmap and language for discussing and prioritizing analytics initiatives.

## Continuous Improvement

A plan is not static; it should include provisions for ongoing assessment and improvement. Regular reviews and updates ensure that the analytics strategy remains relevant and effective.

### *Take Away*

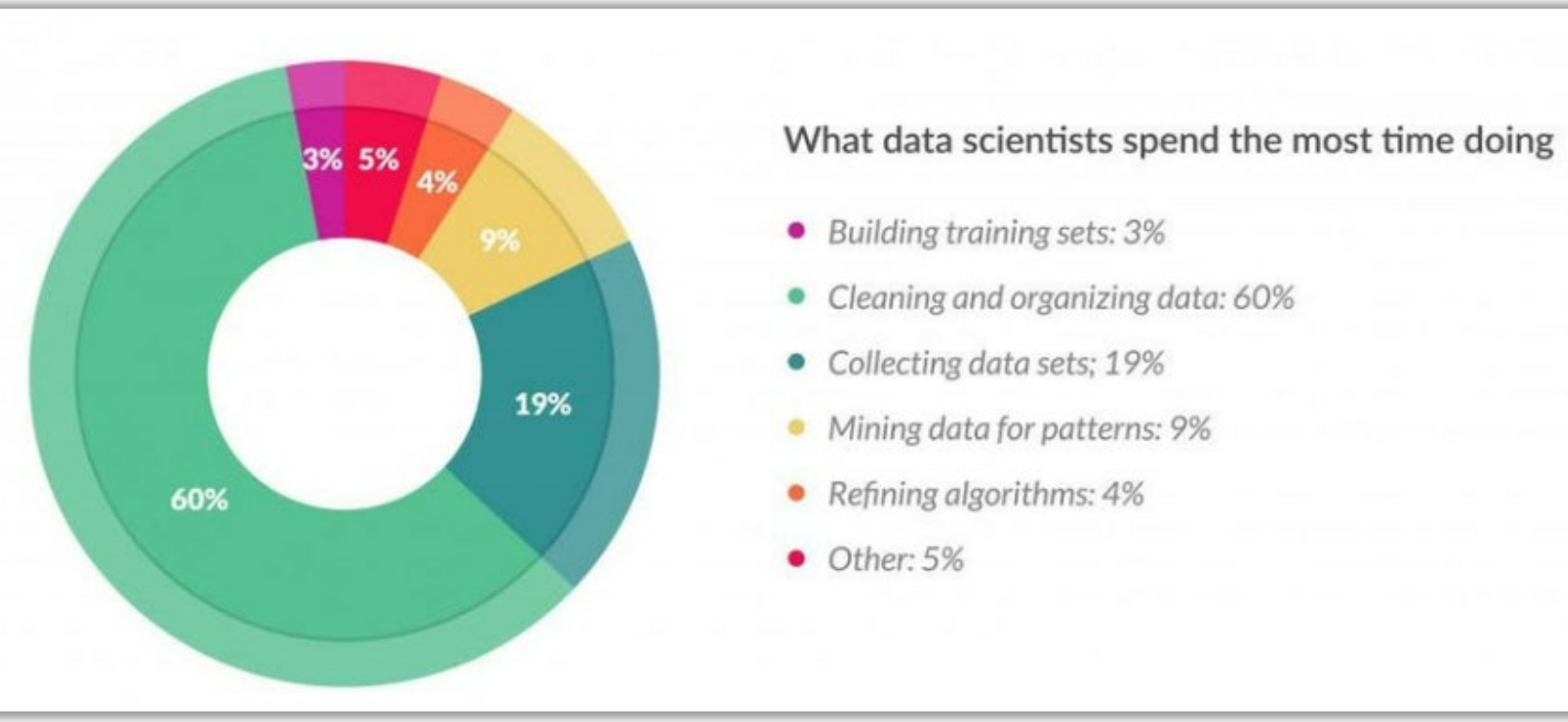
*From Xtensible' observed experience, organizations starting without having a "plotted course" will inevitably need to regroup and redirect the ship towards the correct organizational objectives. To ensure the success of enterprise analytics programs, organizations should focus on detailed planning, robust data governance, adequate resource allocation, user engagement, ongoing training, and a commitment to aligning analytics efforts with business goals and strategies.*

# Focus on navigating the data

A recent survey from Forbes asked Data Scientists what they spend their time on with respect to data management. The result supports what Xtensible has been referring to as the "analytic iceberg effect". Over the past few years, basically 80 percent[1] of time is spent on what we refer to as "Data Janitoring". The surface of the iceberg seems small, but it is what lies beneath the surface that is big, namely data janitoring, and this can sink your enterprise analytic ship.

Time spent on:

- Finding and sourcing the data
- Understanding the data
- Cleaning the data
- Mapping the data




Ask yourself, are you managing your data and analytics program in a manner that starts to reduce the level of effort when implementing additional analytics? This is essentially "chipping" away the bottom of the iceberg, resulting in less data management and more actual analytics value. If effective data modeling and governance is not put in place, each analytics

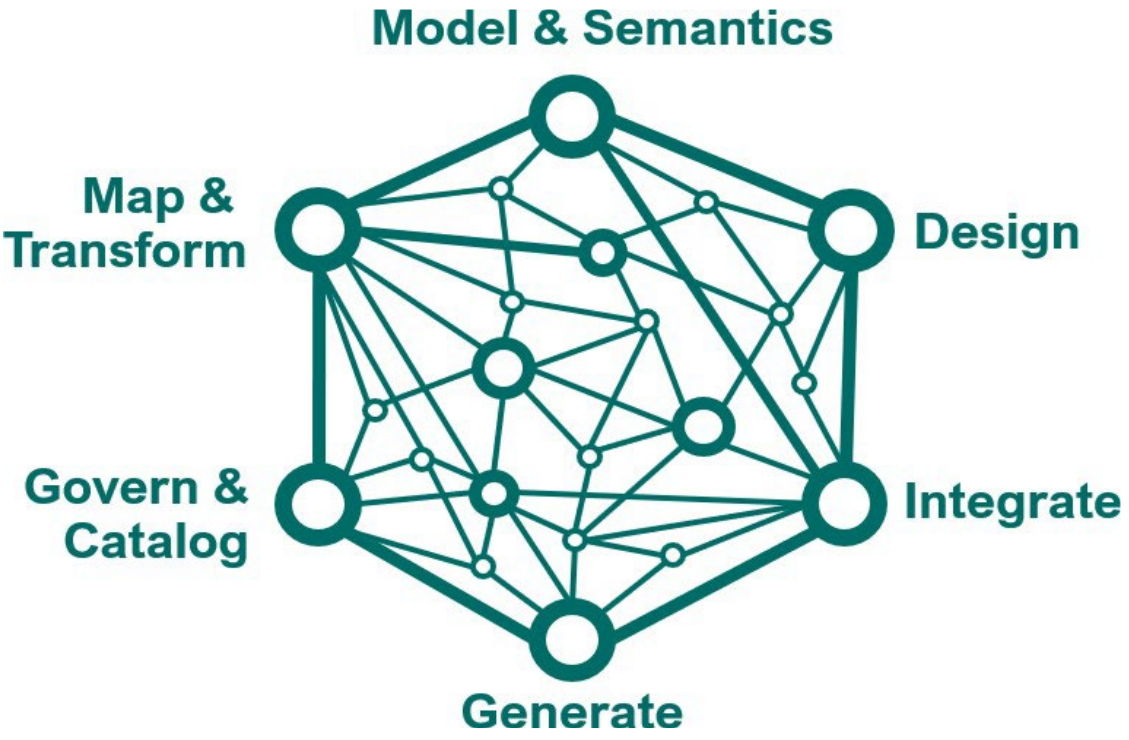
[1] Managing The Data For The AI Lifecycle (forbes.com) (including figure)

use case continues to repeat itself, addressing the bottom of the "iceberg" repeatedly without measurable data management improvements.

## Changing to a holistic data perspective

We have tended to think about the data journey in a linear and largely segmented fashion. We utilize different technologies during different stages. Very often these technologies are loosely connected requiring manual intervention, or worse, they may not interact at all, reducing the ability to gain insights and increasing complexity.

There is a need to unify our thinking, approach, and delivery methodologies to bring different groups together in the organization to reduce complexity, for faster time to market, when realizing analytics to deliver business value.




Xtensible refers to this unified thinking as connected and managed data nodes described in a recent blog post from Sr. Business Consultant, Erika Ferguson, of Xtensible, entitled Unleashing the power of the 6 data nodes. These data nodes are:

- Govern and Catalog
- Model and Semantics
- Design
- Map and Transform
- Integrate
- Generate

These data nodes must be unified and connected to ensure alignment, reusability, and consistency of data AND analytics across the organization. From a stakeholder perspective consider a consumer of data such as a data scientist, data analyst, citizen data scientist, and others. To ensure that they get the results they require they need clear and unambiguous data semantics resulting in:

- Common terms and definitions, agreed by both business and information technology
- Relationships between entities are well-defined, easy to follow, and utilized
- Business rules are well-defined, easy to follow, and utilized
- Trust & Proof
- "Single version of the (metadata) truth"

It has become clear that master data management can be challenging – now throw in the added effort of having the associated meta data also be correct – if you are attempting to do this manually or with spreadsheets, it becomes an unmanageable problem.

Imagine what would happen if the organization spoke the same language, i.e., the semantic model. Imagine if this model was the basis for your designs, leading to mappings and transformations, for your integrations and ultimately for generating your data store as an example. How many spreadsheets or word documents with designs, and mappings etc. could you eliminate? How much time would your analytics team eliminate "data janitoring" efforts to combat the analytics iceberg effect?

# It is a journey

It is a journey. Organizations must be patient – but they must also have a sense of urgency – the problem with utilizing and gaining value from data will only get worse if proactive steps are not taken to build data management maturity. Patience is called for, but not procrastination.

Organizations need to begin to embrace some guiding principles for how they approach both data management and enterprise analytics. They need to plan for gaining data management maturity over time. They also need to think ahead and consider different analytics technologies. Maybe the addressing business objectives requires advanced analytics such as cognitive analytics which then necessitates machine learning and artificial intelligence.

Xtensible applies over a decade of experience with standards, best practices, and leading platform technologies to expand utilities real-world analytics capabilities through our deployment models.

To learn more about how Xtensible can help your organization in gaining value out of your data and managing it as a valuable asset visit us at: https://www.xtensible.net/about/insights/

To learn more about the data management platform that your organization visit us at https://affirma.xtensible.com/capabilities/

**ABOUT XTENSIBLE**

Xtensible is a leading US provider of semantic-based products and consulting services for the utility industry with a global client base. We provide strategy and associated architecture consulting, technology consulting, enterprise information management, integration, business intelligence and data analytics services, as well as utility standards training. Our products manage the data lifecycle, from design to run-time. We are a driving force behind the development and implementation of international open standards, that underpin current and future interoperability and technology initiatives. We deliver and empower our clients with sustainable end-to-end solutions in a repeatable and cost-effective manner and strive to become a trusted, reliable partner, putting our clients first.

**Sales and Business Development**

sales@xtensible.net

**Support**

support@xtensible.net

**Headquarter Office**

Xtensible
6900 Tavistock Lakes Blvd
Suite 400
Orlando, FL 32827
Phone: +1.321.558.6950