



The Global Leader in  
Enterprise Data Management

**WHITE  
PAPER**  
January 2018

# THE NEED FOR A CENTRALIZED ANALYTICAL PLATFORM

## 2 Introduction:

There is no shortage of data out there. The question is: how good and reliable is this data, and is your business making the best use of it? Decisions that drive businesses hinge on how accurate and timely this data is. Maximizing insight from an organization's data is therefore a fundamental business process in any industry. The ability to analyze both internally generated and external sources of data allows businesses to evaluate their current approaches, strategies, and risk mitigation methodologies with reference to external market prices, volumes and forecasts.

The first step in extracting value from data is to pool the data in a secure, transparent, and centralized platform. By creating a "data universe", that is, an ecosystem encompassing data inputs, outputs, and storage, an organization has a secure base to grow and expand. Building a scalable, reliable data universe is a major challenge every organization has to contend with.

Analysis is changing for many industries. Where once siloed corporate data governance structures existed, the drive for self-serve entitlement and access to data now prevails. What was once a single analytical application has now become a multitude, each with a variety of strengths, features, and abilities. Where once data was from very few sources, in a single format, location and type, the velocity, variety, veracity and volume of data has exploded. These changes are challenging, but they represent opportunities across all data-centric industries.

A crucial way to take advantage of this opportunity is by using a centralized analytical platform. This would allow data professionals, such as data scientists, engineers, analysts, architects or statisticians, to get the high quality, centralized data they need into their dedicated analysis tools. With this, the maximum amount of useful information can be wrung from an organization's data universe, maximizing the value and return on investment of a corporation's analytical tools and data assets.

### Analytical Challenges

The expansion of the overall data universe is growing rapidly, and if current trends are anything to go by, will increase by 40% per year well into the next decade <sup>(1)</sup>. This is just as true of a business's data assets, making a centralized repository of secure, high-quality, and normalized data a business need for all organizations. For more on data, data management and how ZEMA secures and adds value to organizational Intellectual Property, please refer to the white paper "The Challenges of an Expanding Data Universe and The ZEMA End-to-End Data Management Solution" <sup>(2)</sup>.

With all this data coming into an organization, the expectation would be that now, more than ever before, businesses would be utilizing data-driven analysis and predictive, descriptive and prescriptive methodologies to garner the maximum amount of insight from this data. However, for most organizations this is not the case. There are several hurdles that organizations must overcome to be able to get the maximum value out of their data through analysis:

- Automation of reproducible reports and analyses

- Transparency through the business intelligence and analytical tool black box
- The wider approach to analytics

the key features of an analytical platform that can meet these challenges will be discussed in the following chapters.

## Automation

Manual data preparation, human data validation and by-hand creation of reports is not needed for most reproducible and repetitive processes and reporting. However, analysts, scientists, engineers, data architects, and statisticians must often perform these low-level tasks instead of the high-value tasks they specialize in. This is where automation can help.

Automation simply allows businesses to do more with less. No longer should data professionals spend most of their time on data collection, preparation or normalization, as this can be automated by data management systems. They should not need to prepare and feed data manually into their chosen platforms or tools required to analyze their data sets. Additionally, they should no longer waste their time on replicating analysis or reporting on a scheduled basis when robust and mature process and workflow automation engines could automate this.

The return on investment from data management automation is based on several factors. Firstly, valued and experienced data professionals can spend their time investigating risks and opportunities found through self-service data discovery, thus increasing the value of an enterprise's data assets, accuracy of forecasts and mitigation of risks. Secondly, the risk of human error is reduced. Finally, increased transparency and security of intellectual property. Automation in analysis and reporting can be achieved with the correct overarching enterprise data management approach and strategy, and will create a more efficient and effective business.

## Transparency of Analytics and Business Intelligence

Analytics in the past have been considered a 'black box'; data assets go in, insight comes out. Understanding of the process that data goes through to create this actionable intelligence was always limited to the technical and mathematically inclined personnel of an organization. With the proliferation of Business Intelligence (BI) tools and mathematical engines, along with the relatively new professions of Business Analyst and Data Scientist, a lack of transparency around mathematical and logical manipulation of data, along with the formulation of models and methodologies, has been found.

This opaqueness around analytical processes has several consequences. The first is a lack of trust. Given data-driven processes and decisions are made using these analytics, how can decision-makers have confidence in the information from analysts if they have no understanding of what data has gone into this 'black box', what manipulation has taken place, or how accurate and robust the used models were<sup>(3)</sup>? Furthermore, how can effective data governance happen within an organization when data collection, validation, analysis and integration flows are not transparent, centralized, and administered through a single point of access?

Secondly, this lack of transparency means increasing the risk of lost intellectual property. With models, validation techniques, and analytical methodologies locked away, how can this information be transferred to the next generation of analysts, scientists, architects or engineers? This is especially pertinent now, with intellectual property, skills and knowledge transfer from experienced professionals to the newer generation needed to continue operations within many organizations, as the population in these professions aging.

Lastly, with ‘self-serve data discovery’ on the rise, and business processes and workflows becoming more Data-centric, the lack of transparency in BI moves an organization in the opposite direction to current corporate culture. Cross-functional groups, inter-disciplinary jobs, collaborative teams, all lead to a transparent, efficient and effective enterprise. Transparency allows these teams to operate better, make more accurate, collaborative, data-driven decisions, and allow the management or responsible parties to have more confidence and trust in the outcomes.

## Centralized Analytical Platform Approach

It’s estimated that 80 % of the time and effort invested into any Business Intelligence (BI) project is spent on the data acquisition process<sup>(4)</sup>. Getting reliable and accurate data is an ongoing struggle for most data professionals. Data professionals (Data Scientists, Data Engineers, Data Architects, Data Analysts and Statisticians), utilizing an organization’s data assets are demanding a better, and faster, way to get the data they need on a daily basis. The focus should be on extracting useful information from an enterprise’s data, but currently too much effort is spent getting the data into a suitable form for analysis. Companies that are serious about mastering their data create their own “data universe” that is depended upon by data professionals within their organizations to use it to derive actionable insights.

These derived actionable insights require analysis tools to delve into an organization’s data assets and generate information that can be used in decision-making. Today, no single analytical tool has supremacy within the Business Intelligence sphere. Microsoft’s Excel is a single tool that has long dominated the market. Now, space for a wide array of purpose built tools, with different core competencies, attributes and new abilities, has opened up. There has never been a more exciting time, from a data professional’s perspective, given the expansion in the choice of analytical tools on the market. The challenge that comes with choice is the need for a rational approach to integration of these tools with the data universe and data management platforms.

Two key factors allow data professionals to optimize their performance by using tools that promote self-serve data-discovery:

- The ability to manage an organization’s data assets into a centralized, secure, accessible, normalized and validated data universe.
- The ability to make this data easily ingestible by any tools that these data professionals require. This challenge applies equally to BI tools, trade-risk management systems, ERPs, statistical or mathematic modelling platforms, along with many others.

Thusly, a wider approach to analytics is needed for enterprise level IT architecture. No longer will a single tool suffice; an organization's data assets must be able to feed a multitude of BI-tools, mathematical engines, in-house or proprietary systems, along with downstream applications such as trade and risk management systems, or feed market data to distributed file systems such as Hadoop. This way, an enterprise's data universe becomes available to the tool of choice, for the analytical methodology chosen, for any data at any time.

Enterprise data management solutions should enhance integration, automation and analytical flexibility, not place restrictions on data access, analysis, or tools used.

## References

1. EMC Digital Universe with Research and Analysis by IDC, "The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things", <https://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm>
2. ZE PowerGroup, "The Challenges of an Expanding Data Universe and the ZEMA End-to-End Data Management Solution", <http://www.ze.com/learning/resource-library>
3. Vaughan, Jack, "Data Trust Gap Confronts Analytics – Time to Open the 'Black Box'", [http://searchdatamanagement.techtarget.com/feature/Data-trust-gap-confronts-analytics-time-to-open-the-black-box?utm\\_medium=EM&asrc=EM\\_NLN\\_67749757&utm\\_campaign=20161110\\_Data%20scientists%20should%20open%20analytics%20%27black%20box%27%20to%20boost%20data](http://searchdatamanagement.techtarget.com/feature/Data-trust-gap-confronts-analytics-time-to-open-the-black-box?utm_medium=EM&asrc=EM_NLN_67749757&utm_campaign=20161110_Data%20scientists%20should%20open%20analytics%20%27black%20box%27%20to%20boost%20data)
4. Xu, Wei, et al., "Incremental data feed maintenance of a data warehouse system derived from multiple autonomous data sources", [https://www.researchgate.net/publication/4185411\\_Incremental\\_data\\_feed\\_maintenance\\_of\\_a\\_data\\_warehouse\\_system\\_derived\\_from\\_multiple\\_autonomous\\_data\\_sources](https://www.researchgate.net/publication/4185411_Incremental_data_feed_maintenance_of_a_data_warehouse_system_derived_from_multiple_autonomous_data_sources)



**The Global Leader in Enterprise Data Management**

Vancouver • Calgary • Toronto • Houston • Raleigh •  
London • Singapore

130-5920 No. Two Road, Richmond, BC V7C 4R9 Canada

Toll Free: 1-866-944-1469 [www.ze.com](http://www.ze.com)

© 2018 by ZE PowerGroup Inc.