ACCELERATING TO THE NEW

A CHIEF DATA OFFICER ROUNDTABLE DISCUSSION
JAN 17 - 18, 2018
INTRODUCTION

In January 2018, Accenture Applied Intelligence invited Chief Data Officers (CDOs) and the most senior data executives from leading financial institutions to participate in a roundtable discussion. The objective was to delve into perspectives on data, transformation, change and migration—each critical to organizational success and business impact. During the course of this roundtable discussion, the participating CDOs engaged in healthy discussion and debate concluding with a set of key take-aways. This report summarizes the discussions and observations.

ARE WE GETTING IT RIGHT?

1. What’s Different This Time?
2. Do We Have the Right Architecture?
3. What Happened to Big Data?
4. What is the CDO?
ARE WE GETTING IT RIGHT?

The session began with the critical question: Are we getting the data-play right this time? Enterprises have been involved with the data game for over three decades—bringing together multiple sources of company’s data assets into a useful and usable form—serving it up to business users at the right time and place, and in the right format. This data is imbued with intelligence from advanced statistical and mathematical analytics. These are the fundamentals that are key to making better business decisions, even as the technological ecosystem continues to transform.

What’s different this time?

The data movement as we now know it—the sourcing and integration of multi-source data—started as early as the 1980s, when the consumer products industry was faced with the explosion of data directly linked to supermarket scanners. Then the 90s saw a massive boom within the telecom and healthcare industry—a transformation which came about from the call-detail and managed care data explosions. The beginning of the twenty-first century had the web-click wave, followed closely by the mobile wave. We had hardly begun to deal with the latter when the initial swell from the Internet of Things and wearables began to rise—a data movement which continues to wash over us today. Data history has presented itself in cycles—data explosions, computer and storage power becoming more affordable, increasing demand from business for intelligence-enriched data to aid decision making. So, the question is, what’s different this time?

Massive changes in the power of hardware and its capabilities have led to huge technological leaps. Data has transitioned into an integral role in core business processes, across companies big and small, initially exemplified by data-driven companies such as Google. Increased types of data required more in-depth data skills, leading to today’s growth in opportunities, and availability of talent dedicated to data. And the digital revolution has exponentially exploded the volume and variety of data available to anyone who wants to analyze it.
However, what’s different this time is in data consumption, the group argued. The more we use it, the more capable we are of making better decisions and differentiating ourselves. This produces opportunities not only to explore analytics, but also focus on operational transformation. Algorithms invented in the 1950s and 1960s are still being applied today. Running an algorithm on modern systems has no limitation now. What has really changed is the underlying technology—storing, organizing and analyzing massive quantities of data—and the value this has created.

The group agreed that, in addition to the areas described above, one of the developments that has drastically changed the landscape is that AI and Machine Learning (ML) have left the research realm, and can now be broadly deployed in a corporate analytics setup.

There are many more questions that remain unanswered. Will Artificial Intelligence (AI) finally modernize the algorithms we use today? How will AI change our fundamental thinking, and what impact will this have on business? Are there social implications, and will the customer happily interact with computer-based problem-solvers? As younger generations learn to participate in this ecosystem comfortably, how can we use their skills to our advantage?

**Are we adopting the right architecture for the long term?**

The group was unanimous, agreeing that enterprises are headed towards the data lake construct—a move which brings performance and scalability powers unseen until now. From a conceptual architecture point of view, the fundamental notion is that of a shared concept. Data must be sourced from multiple sites and brought into the data lake. Abstracted layers of data must then be instantiated to provide data to the consumer in a suitable format, for a variety of uses including management reporting, ad-hoc querying, and advanced analytics.

Will the current approach be sustainable through the increasing explosion of data volumes and types, from social and digital sources, and soon-to-come wearables and Internet of Things? Will this also be able to support the wider dispersion and adoption of AI and ML capabilities that most companies believe to be the future secret sauce? Will this architectural approach give the ultra-fast performance required by businesses to make timely and sometimes real-time decisions? Maybe there is a “rethink” required for a AI-enriched, hyper-scalable architecture, the group agreed.
What happened to big data?

The current journey to big data began around 2010. This started with the availability of data sources which became available through the social and digital interactions of consumers and businesses. This data could be used to enrich the insights into our customers and prospects by tying their external digital and other behaviour to internal enterprise interactions.

Every day we read about ground-breaking ideas—from big data for cancer tracking, satellite imagery of farms providing insights on what to plant, big data-based drug discovery processes, and the never-ending stream of ideas coming from the fintechs. Yet, there is a sense of a lack of excitement in the outputs from the financial industry through their new data-driven capabilities. How exciting have been the new products and services introduced through big data-driven insights?

One of the causes of this, the group reasoned, was that to date most of the effort by financial institutions has been towards building big data lakes that captured and stored more internal data with greater detail, more quickly. Somewhere along the way, the promise of power from external (digital and social) data integration seems to have been forgotten. Many companies have delegated the work of using and deriving insights from external data to third parties. Massive internal lakes do not contain much external digital and social data, if at all. Now that the first-generation lakes have been built, it may be time to rethink how quickly, and how deeply, should external “big data” integration be undertaken.

What is the CDO?

Those present at the roundtable discussion included CDOs and senior executives responsible for the data agenda without a formal CDO title. Each participant reported to many different leaders including risk, finance, operations and tech, IT, business units, and the administrative officer. A series of statistics illustrated the core idea that the CDO role and function is still evolving.

Evolving CDO Function

Who does the CDO report to?

<table>
<thead>
<tr>
<th>Role</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIO</td>
<td>26%</td>
</tr>
<tr>
<td>CRO</td>
<td>21%</td>
</tr>
<tr>
<td>COO</td>
<td>21%</td>
</tr>
<tr>
<td>Business</td>
<td>16%</td>
</tr>
<tr>
<td>Ops &amp; Tech</td>
<td>11%</td>
</tr>
<tr>
<td>CTO</td>
<td>5%</td>
</tr>
</tbody>
</table>

Jan '18, Accenture survey of leading FS Chief Data Officers
The statistics showed a lack of correlation between the size of the CDO organization and the size of the company (in revenues specifically). Some CDOs had less than a dozen direct reports in large companies, whereas a comparable company’s CDO had over a thousand. It was discussed that this disparity truly reflected the degree to which the company and its senior executives “believed in the power of data”—reflecting how deeply data-driven philosophy underpins their strategy.

The CDO’s scorecard was discussed—the success elements clearly indicating where the company thought the CDO must play their role. This didn’t always encompass the transformational capabilities that could be driven by data. The role of the CDO across the variety of measures of success, and associated decision rights, were in themselves topics worthy of deeper discussion. The participants tabled this for a separate standalone discussion at a later time.

DIFFERENT PHILOSOPHIES TO DATA

How does the size of a CDO Org compare to the overall size of the company?

Jan '18, Accenture survey of leading FS Chief Data Officers

What defines the CDO is really who you are accountable to, not so much about reporting structure.
A survey of participants indicated that all companies have three areas where they spend their time, effort, and money. These are their traditional data warehousing environments, the first generation of the lake environment already built or being built, and planning for the continued evolution of the lake to deliver revolutionary capabilities in the future. Nearly two-thirds of the total spend was towards the new lake, enhancing that lake, and extracting greater value from it.

THE JOURNEY TO THE “NEW”

Jan '18, Accenture survey of leading FS Chief Data Officers
Augmenting the “old”

According to the group, their traditional data warehousing environments continue to play a critical role in the data that powers their businesses. While a significant portion of the data and its capabilities have been transferred into the lake, substantial capabilities are still located in the “old”, where they will remain for the near future.

Three broad areas—metadata, data quality and governance—remain the focus of the CDOs on their traditional data environment. However, the search now is to explore opportunities for AI to drive improvements in these areas. With the multitude of technologies now being used in the traditional and new lake data environments, the questions arise: How can AI techniques, and graph databases, be used to track multi-system metadata? How can AI track and improve data quality in the flow of data into and out of the repositories? How can AI techniques (and even blockchain) be used to map out data lineage across the various touchpoints in the data and reporting flows?

It was agreed that with the tailwinds of regulatory demands tapering away, the impetus and budget to continue the effort and spend in these three areas is likely to wane. AI and machine learning could very well provide a more efficient and less labor-intensive way of addressing the challenges in metadata, quality and governance. Much more needs to be done—and even in the traditional data environment there remains room to explore the “new” AI-based way of operating.

AI-BASED CAPABILITIES

AI-ENRICHED METADATA,
DATA QUALITY,
GOVERNANCE, AND
MODEL MANAGEMENT
Accelerating the “new”

Forget ‘Time to Insight’, it’s really ‘Time to Revenue’.

Every participant agreed that they are either building, or have built, the “first-generation” of their new lake environment—the data lake, and the associated data layers for use by the business and data scientists. Some have built and deployed a number of AI/ML-based analytical capabilities, whilst others are just beginning. Regardless, each participant concurred that they share two major priorities: increasing the value extraction from the investment they have made, and accelerating the build and deployment of the “new.”

Four major thrusts were discussed as a means to accelerate the journey to the “new”:

1. Decommission legacy data and analytical environments
2. Migrate the analytics (models) embedded in the old data and analytical environments
3. Migrate talent from the statistics-oriented modelling mindset to the more AI and ML mindset along with the concomitant new “programming” languages (e.g., Python and R)
4. Democratize new data and AI-based analytics capabilities.

All four are required, regardless of where the company is in its journey to the “new.”

FOUR FOCUS AREAS

- DECOMMISSION LEGACY
- MIGRATE ANALYTICS @SCALE
- TRANSFORM TALENT & CULTURE
- DEMOCRATIZE USE, ESP. FOR AI
Legacy environments

Legacy data environments will continue to carry out targeted functions for the foreseeable future—alongside broad use for high-volume ad-hoc querying, there will be finance, risk and regulatory reporting for instance. For other use cases however, it makes sense to quickly migrate from old platforms, minimizing cost and user overhead from having to use multiple platforms, and reducing the potential for operational risk introduced by reporting off two separate environments. The biggest challenge is the investment case. It costs money to switch things off. As such, a well-articulated and phased decommissioning roadmap and investment case is critical.

The “intelligence of the enterprise” is hidden in the analytical models that have been built (or bought and integrated) over the past several decades. These have been deployed into many operating layers—cross-sell prompts in the contact center, fraud models in the transaction approval systems, customer servicing treatment models built into the service platforms. In the “new”, many of these models can be replaced by more powerful or timely AI/ML-based models. In this context, the group discussed the need to develop a holistic migration strategy. They reviewed their experiences in building AI/ML models to replace the old analytics. They questioned whether new models should be built only for new applications, and let both old and new co-exist, and which models should be replaced in the deployed operating channels? They agreed that this is a critical area—where the rubber meets the road to realize the value from all the data and analytics work—and an explicit strategy and roadmap is required.

A distinction was made between migration and transformation, labeling migration as the moving of code and data, and transformation as learning to deploy. One of the major challenges of migration is the work that goes into transferring legacy controls across to a new, big data environment. It was suggested that the journey from warehouse to big data should not be viewed as the entirety of the transformation needed. Talent, organization structure, remits, and the decision rights of data scientists and the CDO function are also critical to the transformation journey.

Talent migration goes hand-in-hand with legacy decommissioning and analytics migration thrusts. The advanced analytical communities are some of the most qualified staff, with the deepest insights into what makes the business tick. They need to understand how a migration program will successfully bring them into the AI/ML era. How will their day-to-day work be impacted when their model codes must be recoded? The participants understood that the messaging around this migration is important—new approaches to modelling and analytics should not dismiss the prior paradigms that have made the enterprise successful to-date.
Analytics migration

The group agreed that the journey to the “new” critically includes democratizing access to data and insights for a broad set of advanced and basic users. Unlike the data warehousing era—when tools and techniques had evolved over multiple decades—the new lake constructs are still evolving, with the tools and techniques still not as easy and quick to use as in prior model. This raised the question: How can a CDO plan for a broad deployment of such capability in this evolving context, and what’s the support model? What is the transition plan to gradually adjust the support skills and programs?

The CDOs collectively agreed that a complete switch-over does not happen quickly. As there is so much already invested in the warehouse, it becomes harder to justify additional investments on the new without an associated revenue generation impact. Not everything can be obliterated and started afresh. Some use cases must also help fund the decommission of legacy platforms, the group agreed.

ANALYTICS MIGRATION – MORE THAN CODE

Multi-dimensional Program

Analytics Migration
Deployed-Models Migration
Analytics Platform Migration
Talent Transformation
User Support @Scale

Objective number one for us is to have all our data in the new platform. Then everything else flows from this.
In reference to talent migration, CDOs commented that the battle is between financial services and tech companies. There exists a need to diversify across states and regions. Once geographies expanded and recruitment happened from outside the industry, the question was not about gaining, but maintaining talent. Retention is difficult, and necessitates developing career and learning paths that coexist in the context of decommissioning and acceleration to the new.

**Talent transformation helps us move to the future.**

Collectively, the group agreed that a holistic marketing and communications approach to migration may have to be adopted. Having conversations with the CMO and introducing this journey to the new as a formal marketing campaign were good steps to take. Many CDOs agreed that their teams need to improve their marketing acumen.

**No cloud provider is going to come in and do legacy mapping for you.**
PLANNING AHEAD FOR THE NEXT

Responsible AI

Due to the large number of data sources involved and the critical decisions that could be made or influenced from its use, with and without human oversight, the group agreed with the emerging view that AI should be transparent. With increased use of AI models and wider deployment, transparency demands will rise. Thus, the idea of creating guardrails for AI was posed by the group. These included:

- It is important to equally represent everyone, including minority groups, in the data sets and learning models
- All applications must go through a series of human intervention and tests before they are released – the reinforced learning concept coming into vogue in AI
- “User beware”—while use of biased algorithms can be the responsibility of the user, model developers must also take responsibility

If you hit ‘like’ for social irresponsibility, is that Facebook’s fault?

NEXT WAVE CAPABILITIES

Modelers must take responsibility. Part of philosophical aspect is to not blame the algorithm. It’s not Skynet [that does it] but it happens because of us.
Even higher performance

Companies classified as early innovators, who built their lakes over the past several years, indicated that they are beginning to face performance challenges. These heightened demands are primarily coming from the business user community. They do not see the same level of query or analytical response time that was available to them within the highly tuned and indexed data environments of the legacy data warehouse platforms.

The participants also understood that with the increasing adoption and deployment of AI and ML models, there will be amplified demand for massive quantities of data to be used in model “training”. Architecture will be required to support larger amounts of data flows. However, as the current design paradigms have revolved around optimizing for large repositories instead of flows, a pathway from today to tomorrow’s usage patterns will have to be dealt with.

EYE ON THE FUTURE

Planning for even greater scale and performance

The number one reason for going with cloud or not is gravity—the data will reside where it already resides.
Finally, the group touched upon the further explosion of data volumes they will be called upon to manage as the next wave of data-generating developments take root, in the form of drones, Internet of Things, wearables, and real-time and streaming data. A more seamless way to move data from on-premise to cloud environments—dynamically adjusting the data storage patterns in the “back” whilst insulating users in the “front”—will be needed. Deep learning models necessitating GPUs are much easier to manage on the cloud, especially at peak demand. As such, the move to cloud seems more convenient. The group agreed that while it may not be a challenge today, this scenario is only a year or two away, and it is best to develop the roadmap to the future with this eventuality explicitly baked in.

**EVEN GREATER PERFORMANCE**

<table>
<thead>
<tr>
<th>Central Processing Unit (CPU)</th>
<th>Graphics Processing Unit (GPU)</th>
<th>Field Programmable Gate Array (FGPA)</th>
<th>Application-Specific Integrated Circuit (ASIC)</th>
</tr>
</thead>
</table>

(–) **PERFORMANCE** (+)

Platforms will inevitably be commoditized in five years, which poses the question: Is it worth spending time to build specialization in Infrastructure?
Value extraction

Much has been accomplished to date—new lake, AI and ML capabilities have been built, with significant investments. The CDOs agreed that—for the scale of the spend incurred to date—there is still demand and opportunity to demonstrate greater tangible value and impact.

The search for the “killer app” remains—but until that is found, value and impact must be realized in a more sustainable and consistent manner. The richness of data in the lake, and the computational power and speed now available, has allowed for the construction of new applications that would not have been reasonably possible in the past.

Participants agreed that the fundamental power of the “new” comes from the broad deployment of AI and ML capabilities. Initially, AI and ML were applied to traditional areas such as customer service, targeting, or fraud analytics. Now, there is a desire to apply AI and ML techniques on new value adding areas with, for example, a full universe of enriched streaming and real-time data. Intelligent virtual agents that learn to serve better over time, for example, can reduce cost-to-serve in a very substantial way. They identify digital signals emanated by customers external to the company’s digital eco-system, and use those signals to provide relevant products, services and information in a timely fashion.

To accomplish this goal, broader democratization of data-access and AI and ML capabilities will be required. A structured change and communications program, along with all the elements of a support program, must be formally managed by the CDO. This is a perfect example of how the “if we build it they will come” approach will not work.
Rethinking data monetization

Data monetization has recently made a resurgence, as the hunt for topline growth returns as a priority. In general, the group believed that there is a real opportunity for CDOs to take the lead in the quest for data monetization.

If devices in factories could be tapped into for data collection, customized risk analytics could offer additional products to assist with monitoring and maintenance services. The question then arises: how can this be done at scale?

We are currently living in a “test-and-learn” phase, and there is still much to discover—exploring unfamiliar territory until a solid foundation of information and understanding has been developed.

**DATA MONETIZATION**

**Examples**

- IoT
- Automotive
- Drone & Satellite Imagery

**Opportunities for FS?**

- What opportunities exist for financial services companies from these developments?
- CDOs are best placed to identify the range of new data available, and thence it’s potential use for new revenue and businesses

One CDO illustrated how a collaborative arrangement could offer a good example of testing and learning. They built a house full of connected devices in their VC wing, inviting people to interact and experiment within the living lab. This gives them opportunities to test and learn in a live environment.

Another CDO shared their “link and label” approach to achieving the proper credit for data-driven capabilities. This method helps to ensure that bottom-line impact is highly visible and credited to the data organization, so they can be more credible when seeking investments.
However, few companies are exploring data monetization to gain revenue from new businesses, products and services. Several examples from other industries were brainstormed to kick-start some ideas, including the “Internet of Things baseball,” which generates all necessary data needed—speed, angle, pitch and rotation—to understand the potential of the pitcher.

**Data is key to driving business success, and businesses must invest in data to survive.**

Experimentation in drone and satellite imagery is aiding businesses in underwriting and processing claims. However, the ability to take advantage of this technology is being hindered by the hit rate issue. How do you tag images, and match your internal processes, to increase hit rates from 40%, to 70% or 80%? AI and ML techniques for image recognition will require the architecture to support vast quantities of training data that must flow into the models—allowing for continuous learning and improved hit rates over time.

**Blockchain and Quantum**

A quick overview of the blockchain concept, and its surrounding technology, brought about a discussion on its potential application within the data realm. The group questioned whether blockchain was still an “art of the possible” concept in their data space, or if anyone was currently harnessing blockchain for practical purposes.

The group discussed research surrounding quantum computers and early R&D experiments, covering the kinds of problems that are relevant for this class of computation. The question was how this type of computing can be utilized and integrated within the data and analytics architecture within a practical time frame. Some of the attendees indicated that to address certain very large-scale problems, “skunk works” were being initiated to explore quantum and deep learning capabilities.
CONCLUSION

As the CDO Roundtable Session came to a close, the participants agreed that they are directly accountable for guiding their organization and accelerating to the “new.” The CDO must deliver on all three fronts—strategy, execution and value realization. The position is challenging and still evolving—many obstacles remain—but for those willing to push the envelope and proactively demonstrate the power of data, the possibilities are endless.

Many companies have built the first-generation big data lakes, but must now learn to accelerate the value extraction from these lakes, and drive wider democratization of the data and its use. Legacy platform decommissioning and migration of analytics to AI and ML are critically important considerations in the journey to the “new.” Though major legacy platforms will continue to exist for the foreseeable future, decommissioning these and migrating to new platforms will liberate funds, resources and management attention for additional initiatives. AI and ML will inevitably impact architecture in substantial ways including data management, data quality, culture and talent, usage patterns, development methods, and monetization opportunities.

Ultimately, to successfully accelerate organizations towards the “new”, the CDO must rely on their ability to step outside their comfort zones and seek broader remits and decision rights. The next step is to embrace the opportunities that are right in front of everyone.
AUTHORS

Ramesh Nair
Roundtable Chair
Managing Director, Financial Services Leader, Applied Intelligence
ramesh.a.nair@accenture.com

David Crawford
Manager, Financial Services
Applied Intelligence
david.crawford@accenture.com

Huendy Espinal
Manager, Financial Services
Applied Intelligence
huendy.espinal@accenture.com

Tad Slaff
Manager, Financial Services
Applied Intelligence
tad.slaff@accenture.com

ABOUT ACCENTURE

Accenture is a leading global professional services company, providing a broad range of services and solutions in strategy, consulting, digital, technology and operations. Combining unmatched experience and specialized skills across more than 40 industries and all business functions – underpinned by the world’s largest delivery network – Accenture works at the intersection of business and technology to help clients improve their performance and create sustainable value for their stakeholders. With approximately 442,000 people serving clients in more than 120 countries, Accenture drives innovation to improve the way the world works and lives. Visit us at www.accenture.com.