## On the Rise of the Chief Data Officers in a World of Big Data

Yang Lee<sup>1</sup>, WooYoung Chung, Stuart Madnick, Richard Wang<sup>2</sup>, Hongyun Zhang<sup>3</sup>
Information Quality and Data Science Initiative
Massachusetts Institute of Technology

"What are the roles and responsibilities of the Chief Data Officers (CDOs) in the age of the big data?" This abstract describes the emerging practice of CDOs, and how the CDO will manage enterprise data in a world of big data. In the workshop, we will have rich illustrations with examples of current practices by pioneering CDOs. As part of a stream of research on the CDO by the *MIT Information Quality and Data Science Initiative*, we have been conducting in-depth case studies in the private and public sectors in the U.S. and internationally with CDOs and Subject Matter Experts (SMEs). Already, we have conducted more than forty cases, and we are targeting more than one hundred. Importantly, many CDOs (with the actual position) have agreed to serve as our longitudinal research subjects.

We have distilled insights and developed a framework to understand the diverse roles of the CDOs in regard to big data, ranging from data service delivery to data strategy setting for a corporation's business values and performance, as shown in Figure 1. We have also identified key salient factors for additional insights for our second stage study of quantitative analysis based on a survey: 1) the posture of data in an organization and the corresponding industry, 2) the key drivers and triggering events and phenomena, 3) the inner workings of the CDO's practice, and 4) the transformative link between data practice with the overall business value and financial performance (See Figure 2).

We have identified many key issues:

- How the CDOs and their organizations respond to external and internal drivers, such as compliance and risk
- How the CDOs garner necessary resources
- How the CDOs understand and predict benefits and risks involved in their practice
- How the CDOs strategize their organization's information posture and their own roles
- How the CDOs know or measure if their practice contributes to achieving efficient and effective
  organizational performance and yielding a higher financial performance for their organizations

## **Background and Research Question**

Many companies expect that "big data", with its volume, velocity, and variety, will be a powerful strategic resource for uncovering unforeseen patterns and developing sharper insights about customers, businesses, markets, and environments<sup>4</sup>.

Indeed, hospitals are using machine learning algorithms on patient data to find unforeseen patterns and insights. Mountains of patient satisfaction survey data, a kind of unstructured big data, can now be textmined for sensitivity analysis. As a result, hospitals can now determine exactly how to improve their patient satisfaction scores, which are tied directly to the federal government's reimbursement to the hospitals for selected medical services in the US.

However, the reality is that the majority of companies has high expectations, and yet may not be ready to harness the power of big data. A recent survey of over 500 global executives reveals that most companies are still learning how to manage big data for their companies. The survey also reports that fifty companies who have a C-level executive responsible for its data management achieve higher financial performance than their peers. Why does a C-level executive's engagement make a difference? Why is hiring many well-trained data scientists at the operational business units insufficient for a corporation to harness the value from big data? These are some questions that we will present in the workshop.

<sup>&</sup>lt;sup>1</sup> Northeastern University & Visiting Professor, School of Management, Xi'an Jiaotong University

<sup>&</sup>lt;sup>2</sup> Honorary Professor, School of Management, Xi'an Jiaotong University

<sup>&</sup>lt;sup>3</sup> School of Management, Xi'an Jiaotong University

<sup>&</sup>lt;sup>4</sup> Griffin J. 2008. The role of the Chief Data Officer. *Data Management Review* 

Corporate computing history provides a valuable lesson for corporate data management. Reflecting back on the evolution of computing, companies began with the centralized computing era and evolved to the decentralized computing era when the more convenient personal computers and micro computers were introduced. On the one hand, decentralized computing opened up many long-lasting innovations and increased individual productivity. On the other hand, it initiated the creation of "islands" of disconnected IT infrastructures, the "silos", due, in part, to the fact that the business units were able to manage their computing infrastructure and applications efficiently without coordinating with other business units or users in other business units. The era of networked computing ushered in a variety of new practices and phenomena. Due to the capacity and the capability created by the connectedness, enabled by the Internet, data from various and sometimes unknown sources and quality were available at the fingertip of the massive data consumers and corporations. Corporations began establishing CIOs so that the computing infrastructure could be coordinated across the business units for implementing the enterprise vision. Establishing CIOs has been particularly noticeable since the introduction of the decentralized computing era. Today, it is common to have a CIO in a corporation to orchestrate the entire corporation's computing infrastructure.

The computing evolution created the rise of the CIOs. The data revolution, punctuated by big data is creating the rise of the CDOs. The successes and failures of the newly established CDOs' performance and their impact on organizations are yet unknown. What is known is that there is an increasing number of CDOs being established around the world in companies and government agencies<sup>5</sup>. Some CDOs are claiming early successes while also facing new challenges in their organizations.

"What should the roles of the CDO be?" is our general research question. Our operational questions include:

- What decisions do CDOs make?
- What resources do the CDOs garner and utilize?
- How do CDOs create value for business?

These questions will be answered in the context of managing big data originated internally and externally.

## Roles of a Chief Data Officer: Results from the Study

Many organizations have already established or are beginning to consider a new position called the "Chief Data Officer" (CDO). Fundamentally, the CDO is tasked with being the voice of data and representing data as a strategic business asset at the executive level. Based on the data, we can see that some organizations see the role of a CDO as a service delivery with an enterprise perspective. However, others CDOs posit that a CDO must develop the corporation's data strategy for the enterprise's growth and management. While some CDOs view their role as still technology oriented, the majority of CDOs view their role as business oriented. Regardless, the common thread is that a CDO works as a business-oriented executive who still needs to understand technology.

We give examples of some interesting quotes from our survey in Table 1. These data are consistent with many of the themes that have emerged from literature. For example, collaborating and communicating with other managers, executives can find one role and responsibility as common for other C-suite roles.

## What are the functions of the CDO office?

Enterprise architecture defines a firm's needs for standardized tasks, job roles, systems, infrastructure, and data in core business processes<sup>6</sup>. The function of the CDO office should be consistent with and support the overall enterprise architecture. The CDO must be empowered to organize, engage and lead a team which should be a multi-functional data management team known as the CDO office, while elsewhere the new structure is labeled as the Strategic Information Management Program or Strategic Information Initiative. Functions of the CDO office include data governance, corporate data infrastructure, enterprise analytics and

<sup>&</sup>lt;sup>5</sup> Wang, R.Y., Lee, Y.W., Pickering, W.E., and Bugajski, J.M. (2013, Forthcoming). *Chief Data Officer: Enterprise Data Solution for Business Challenges*. Cambridge, MA: MIT Press

<sup>&</sup>lt;sup>6</sup> Ross J. W., Weill P., Robertson D.2006. *Enterprise architecture as strategy: creating a foundation for business execution*. Harvard Business Review Press.

enterprise data assets development, and data strategy for the organization.

**Table 1: Example CDO Research Ouotations** 

	Table 1: Example CDO Research Quotations
Issues	Example Quotes
	"A CDO is really almost a crisis manager and an innovation manager [more] than
What are the	anything else."
roles and	"The most important role of the CDO is to ensure that data stewards understand their
responsibilities?	responsibilities."
	"teach the meaning of data governance to the entire organization."
	"Acquire a practical understanding of the business's data problems."
	"Develop a process to produce visualization data product data management function."
	"Maximum effectiveness at whatever levels the leader demands that information."
	"Manage global definitions. Maintain a glossary of information products Audit data or
	information stores."
	• "we know which technology to get for analytics that works for more than one business
	functions"
	"delivers what business wants and sees if they like it"
	"ought to develop data strategies for business values and missions"
What are the	"A separate organization that was in between Operations and technology"
functions of the	"Data governance"
CDO office?	"Focal point for developing data strategies for enterprise"
	"Function of transformation"
	"Corporate data infrastructure"
	"Enterprise analytics, data governance, data architecture and enterprise data assets. Of
	those 4 things, the last 2 have more of a technology flavor, and the first 2 have more of a
	business flavor"
Who should the	"CDO is ultimately responsible for CEO"
CDO report to?	"Chief Technology Officer"
	<ul> <li>"I think it should probably be a joint reporting relationship to a CEO".</li> </ul>
	• "They report to the CEO. That's the only way to do it. The leader of the enterprise has to
	be the one that demands timely and accurate information."
	"The senior executives reporting to the Chief Risk Officer"
	• "COO", "CFO"
What kind of	"Database systems, people from IT and Operations, and business areas labor
resource the	requirements"
CDO should	"I don't think they need a lot resource if they have the authority vested in them by the
have access to?	CEO"
	"Think it would be the access to the changing rules and regulations"
	"Network of communities in the business units and the board members"
	• "Senior VPs"
2000	"I think the external communities would be vital resource."
Does CDO complement the role of CIO?	<ul> <li>"I think the CDO needs to be next to the CIO, maybe on top. It's hard to say. I think they need to have equal seats at the table"</li> </ul>
	"Data has becoming so complex and big, CIO cannot handle it anymore and require additional role."
	<ul> <li>"I've had some experience with that inmy precious company, trying to put the CDO under</li> </ul>
	CIO. What you get is a lot of architects and don't get a lot of improvement, and nothing
3371 4 -1.111 · 4	gets done."
What skill set	• "Technical issue is 20%, business issue 80% "
should the CDO	"Build relationships with their peers, and if it's a decentralized organization, build relationships with law staff in those decentralized leastings."
possess?	relationships with key staff in those decentralized locations."
CDO's role of bridging the	"CDO worked with CFO and the COO, and the heads of all the business units:  security is a converted to the cooperation of the groups that were responsible.
gap between	securitization team, the risk management team, and all of the groups that were responsible
Business and IT	for generating business and revenue."  "We had communications with the stakeholders."
Dusiness and 11	"We had communications with the stakeholders."  "We need to talk about sultival harriers, which were arganizational comparings manufactured by the stakeholders."
	• "We used to talk about cultural barriers, which were organizational, sometimes people driven sometimes the professional development of the people they just didn't have the
	driven, sometimes the professional development of the people they just didn't have the tools to understand. They were very seldom technical or budgetary barriers to get the right
	information to the right people."
	Information to the right people.



Figure 1: Roles of the Chief Data Officer

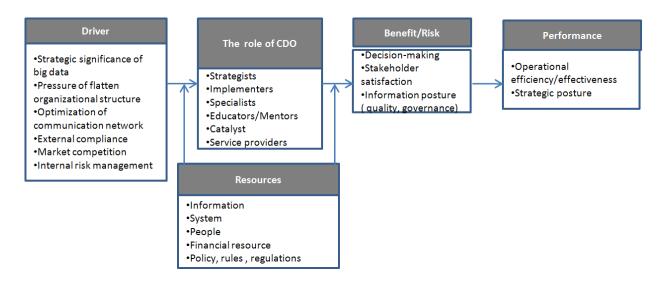


Figure 2: A Framework for the CDO Construct