

ESD Working Paper Series**A Cubic Framework for the Chief Data Officer: Succeeding
in a World of Big Data****Yang Lee**

Associate Professor
Supply Chain and Information
Management Group
D'Amore-McKim School of Business
Northeastern University
Email: y.lee@neu.edu

Stuart Madnick

John Norris Maguire Professor of
Information Technology and Professor of
Engineering Systems
MIT Sloan School of Management and MIT
School of Engineering
Massachusetts Institute of Technology
Email: smadnick@mit.edu

Richard Wang

Director
MIT CDO and Information Quality Program
Massachusetts Institute of Technology
Email: rwang@mit.edu

Forea Wang

PhD Candidate
School of Medicine
Stanford University
Email: forea@stanford.edu

Hongyun Zhang

Postdoctoral fellow
School of Management
Xi'an Jiaotong University
Email: zhanghongyun@mail.xjtu.edu.cn

**A Cubic Framework for the Chief Data Officer:
Succeeding in a World of Big Data**

Yang Lee, Stuart Madnick, Richard Wang,
Forea Wang, Hongyun Zhang

Working Paper CISL# 2014-01

March 2014

Composite Information Systems Laboratory (CISL)
Sloan School of Management, Room E62-422
Massachusetts Institute of Technology
Cambridge, MA 02142

A Cubic Framework for the Chief Data Officer: Succeeding in a World of Big Data

A new breed of executive, the chief data officer (CDO), is emerging as a key leader in the organization. We provide a three-dimensional cubic framework that describes the role of the CDO. The three dimensions are: (1) Collaboration Direction (inwards vs. outwards), (2) Data Space (traditional data vs. big data) and (3) Value Impact (service vs. strategy). We illustrate the framework with examples from early adopters of the CDO role and provide recommendations to help organizations assess and strategize the establishment of their own CDOs.¹

Yang Lee
Northeastern University
(U.S.)

Stuart Madnick
Massachusetts Institute of
Technology (U.S.)

Richard Wang
Massachusetts Institute of
Technology (U.S.)

Forea Wang
Stanford University (U.S.)

Hongyun Zhang
Xi'an Jiaotong University
(China)

The Need for a Chief Data Officer

Increasingly, companies expect that big data, with its focus on volume, velocity, variety, veracity, and value,² will be a powerful strategic resource for uncovering unforeseen patterns and developing sharper insights about customers, businesses, markets and environments. For example, some hospitals are applying automated learning algorithms to patient data and insurance claims data to discover new patterns and insights. The text in mountains of patient satisfaction survey data and data from social media, a kind of unstructured big data, can now be mined to analyze patients' sentiments about a hospital. As a result, U.S. hospitals can now determine how to improve their patient satisfaction scores, which are directly tied to the federal government's reimbursements to the hospitals.

Organizations need to determine who should manage big data. Data scientist roles have emerged to capitalize on the analytical opportunities of big data, but placing these specialists

¹ Jeanne Ross, Cynthia Beath and Barbara Wixom are the accepting senior editors for this article.

² Almost everyone agrees that big data is important, but few can agree on a definition. In our research, we define big data as having five key characteristics, known as "the five Vs": volume, velocity [the speed at which data is created], variety, veracity, and value. Note that veracity and value are often omitted from big data definitions. Veracity is concerned with data quality; high volume and high velocity is of limited use if the data is of poor or uncertain quality.



in operational business units without leadership at the corporate level might be insufficient to harness the full value of big data. A survey³ of nearly 600 global executives revealed that most companies are still learning how to manage big data at the enterprise level. The survey also revealed that companies with a top executive responsible for data management have better financial performance than their peers.

To address the challenges and opportunities of big data, leading organizations have established a new breed of executive, the chief data officer (CDO). Wikipedia describes the CDO role as including "... defining strategic priorities for the company in the area of data systems, identifying new business opportunities pertaining to data, optimizing revenue generation through data, and generally representing data as a strategic business asset at the executive table."⁴ In reality, although some CDOs strive to exploit big data for business strategy, others focus solely on data preparation for external reports, overseeing compliance and establishing data governance.

Emergence of Chief Data Officers

Leading organizations have learned an important lesson—that seemingly tedious data problems are often fundamentally business problems. As such, data problems can reflect weaknesses in business strategy and operations. Traditionally, organizations have addressed data problems by assigning a small group within the IT department to clean up data. As it has become evident that data problems, particularly business problems rooted in data problems, cannot be solved by the IT group alone, organizations have appointed data managers with a variety of titles, such as data quality managers, data quality analysts and data stewards. Data-governance mechanisms, committees, councils and workgroups have also been developed to identify and solve data-related problems and resolve conflicts. Finally, enterprise architecture and data

architecture have also been employed to align data, IT, and business processes and strategies.

Despite these efforts, organizations have continued to face data issues, and their ongoing concerns have led a growing number to establish an enterprise-level, executive-rank CDO. Some might argue that traditional data-related managers and data-governance mechanisms can deliver the same results as a CDO. However, there are critical differences between the efforts of low-level data managers and those of executive-rank CDOs. The key contrast lies in organizationally sanctioned leadership and the accountability given to executive-level CDOs.

First, unlike data managers, a CDO can lead the effort to build organizational capability that can energize and sustain the entire organization and extended enterprise. The experience of a major U.S. healthcare institution illustrates the inherent challenges faced by data managers who lack the authority of a CDO. While attempting to re-examine the business processes that collect, store and use customer data, a data quality manager in this institution received this complaint from an executive: *"You are digging in my backyard—Get out of my backyard!"* Another data manager recalled the project as: *"A huge responsibility without authority."* As a result of these obstacles, the entire project was discontinued; the group working on the project was dismantled and some members left the company. In reality, low-level data managers are not in a position to dictate business process changes to higher rank executives, let alone external partners.

The second critical difference between a CDO and traditional data managers or data-governance mechanisms is that the CDO can be held accountable for a failure of leadership in resolving data problems. Data-governance mechanisms, such as data-quality and -governance councils, committees and workgroups, can be useful for continuous improvement of data policies or standards, conflict resolution, and for reconciling and authorizing data sources. However, because individuals have responsibilities outside of the committee or workgroup, they are usually not held accountable for governance results.

Note that the CDO does not replace the need for data managers or data governance. Rather, the CDO leads data managers and enhances the

³ A summary of current global big data practices can be found in "Big data: Harnessing a game-changing asset," *Economist Intelligence Unit*, June, 2011. It explains in detail the current practices of 586 global companies with easy-to-understand tables and figures.

⁴ The Wikipedia page on the chief data officer (http://en.wikipedia.org/wiki/Chief_data_officer) is a good starting point for scanning the industry's current trends, but it does not provide a complete picture.

effectiveness of existing governance by putting data on the organization's business agenda and in the minds of other executives and officers. Under the leadership of a CDO, business strategies reflect and exploit data, particularly big data, instead of treating data merely as a by-product of running the business.⁵

The History of the CDO

The first recognized CDO was established in 2003 at Capital One. Yahoo! and Microsoft Germany were also early adopters of the CDO role. More recently, CDOs have been established at global investment banks, consumer banks, consumer credit institutions, financial institutions, IT and data companies, healthcare organizations, U.S. federal and state governments, and U.S. military organizations. For example, the U.S. Federal Communications Commission (FCC) created in each of its Bureaus a CDO with varying rank and scope; in total, the FCC created 11 CDOs. According to GoldenSource's annual client survey, "over 60% of firms surveyed are actively working toward creating specialized data stewards, and eventually chief data officers."⁶

Many organizations recognize that they need an executive to lead data management, but not necessarily with a CDO title. These full-time CDO-equivalent executives lead enterprise-wide initiatives on data quality and analytics, data governance, data architecture and data strategy. In this article, we use the term "CDO" to refer to all executives who are carrying out enterprise-level CDO roles, even if they may not formally be titled as CDOs.

CDO Reporting Relationships

As organizations use more advanced business analytics, often there is a need to redirect the flow of information horizontally across the enterprise. Thus, many of the CDOs and executives we interviewed⁷ had the power to exert influence on company strategy. This power and authority is

often reflected in their reporting relationships, membership on senior management committees, and authority over budgets and employment.

Of the CDOs we interviewed in our study:

- 30% reported directly to CEOs
- 20% to COOs
- 18% to CFOs.

Others reported to the CIO, CTO, CMO (chief medical officer) or CRMO (chief risk management officer). Many CDOs are members of senior management committees and have the authority to establish policies and strategies. Currently, the power and authority of many CDOs is evolving from data policy toward business strategy.

The Three Dimensions of the CDO Role

To provide more structure and a better understanding of CDO roles, we identified three key dimensions, as shown in Figure 1: (1) Collaboration Direction, (2) Data Space and (3) Value Impact.

1. Collaboration Direction Dimension: Inwards vs. Outwards

The Collaboration Direction dimension captures the focus of the CDO's engagement, either inside or outside of the organization. Collaborating inwards means focusing on internal business processes associated with internal business stakeholders. In contrast, collaborating outwards means that the CDO's focus is on stakeholders in the external value chain and environment, such as customers, partners, suppliers or regulatory bodies.

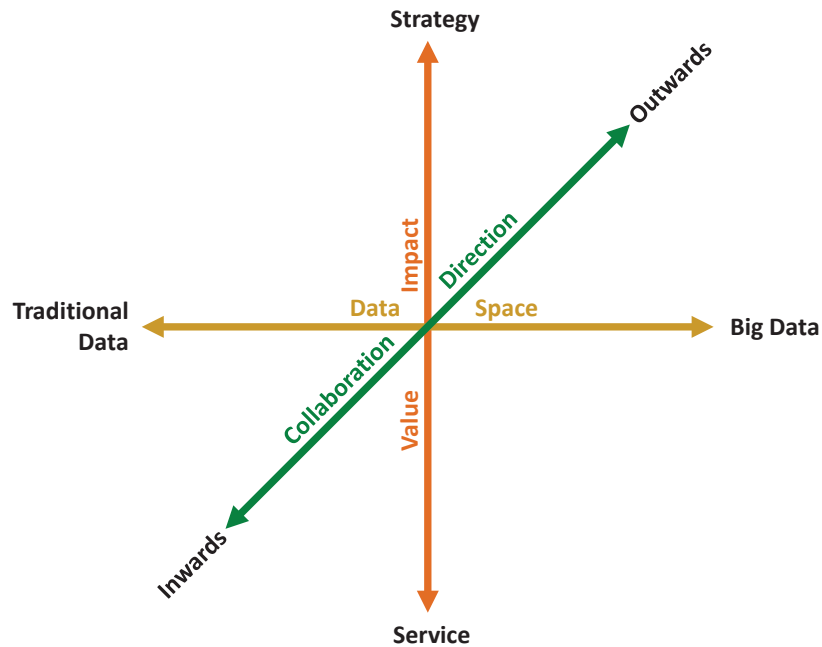
Initiatives led by internally focused CDOs typically include developing data-quality assessment methods or mechanisms; cataloguing data products, sources and standards; creating processes for managing metadata or master data; engaging in information-product mapping; and establishing data-governance structures. These initiatives seek to deliver consistent data inside the organization and to address the root causes of data-quality issues. Streamlining the internal business process associated with key data flows requires cross-functional cooperation, and can result in efficient and effective business operations. The CDO's success in these initiatives

⁵ Company examples and discussions on managing information as product vs. by-product can be found in Lee, Y. W., Pipino, L. L., Wang, R. Y and Funk, J. D. *Journey to Data Quality*, MIT Press, 2006.

⁶ Fernandez, T. "Chief Data Officers Becoming Crucial: GoldenSource," *Securities Technology Monitor*, June 18, 2012; available at <http://www.securitiestechologymonitor.com/news/GoldenSource-Chief-Data-Officers-30775-1.html>.

⁷ See the Appendix for a brief description of the interviews conducted.

Figure 1: Three CDO Dimensions



depends heavily on the ability to effectively lead the relevant internal stakeholders and map out the transformation journey.

In contrast, outwardly focused CDOs strive to persuade and collaborate with external partners. For example, the outwardly focused CDO of a global manufacturing company led a business-process-embedded “global unique product identification” initiative, aimed at improving collaboration with external global partners. Such CDOs may also focus on external report-submission activities, particularly if the company has experienced an external embarrassment or a sizable disaster created, for example, by poor-quality reports.

2. Data Space Dimension: Traditional Data vs. Big Data

The Data Space that a CDO focuses on can either be transactional data, typically in relational databases, or the newer and more diverse big data.

Many CDOs focus on traditional data, as it is the backbone of the organization’s operations. Without a strong foundation in traditional data, an organization’s most basic capabilities are hindered, and thus the need arises for a CDO focused on traditional data-management activities.

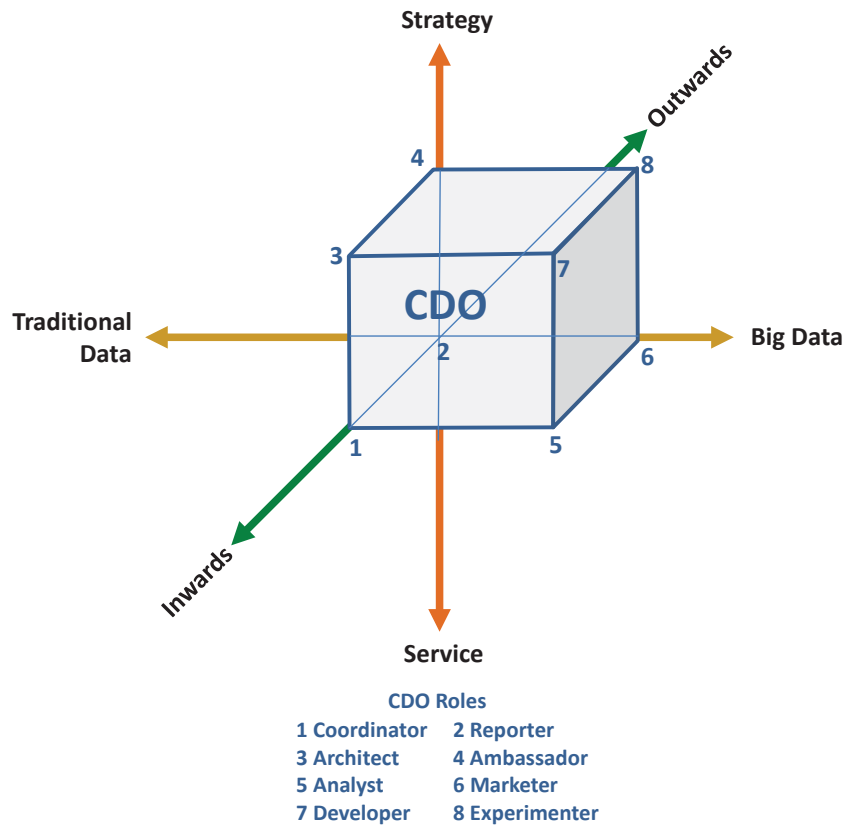
In contrast, big data is usually not connected with the organization’s transactional data or database systems, but offers innovative opportunities to further improve operations or develop new business strategies based on new insights that traditional data cannot provide. CDOs focused on big data provide the leadership to adapt to and manage the analysis of this new, diverse type of data and to gain insights from these analyses.

3. Value Impact Dimension: Service vs. Strategy

The CDO’s role can focus on improving services or on exploring new strategic opportunities for the organization. This dimension reflects the impact desired from a CDO. In many cases, the CDO role is a direct response to the on-going need for executive oversight and accountability to improve existing organizational functions. Increasingly, however, organizations require CDOs who can add strategic value by taking advantage of new tools such as data aggregators⁸ or other data products based

⁸ Madnick, S. and Siegel, M. “Seizing the Opportunity: Exploiting Web Aggregators,” *MIS Quarterly Executive* (1:1), 2002, pp. 35-46, explains web aggregators and their strategic business opportunities.

Figure 2: The Eight CDO Roles



on digital data streaming.⁹ These CDOs also explore ways to develop new market niches, or to transform the company so it can develop smarter products and services.

For example, one strategy-focused CDO led an initiative to identify new information products and advance the company’s position in the financial industry. This CDO led a cross-organizational collaboration initiative to create a strategic vision for managing the new information products at the enterprise level. We have observed that CDOs who are positioned higher in the organization are better suited for taking on a strategy-focused role.

CDO Role Profiles

We have identified eight different CDO role profiles based on the three dimensions described above. These roles correspond to the eight corners of the CDO cube depicted in Figure 2.

For convenience, we have labeled the eight roles as “Coordinator,” “Reporter,” “Architect,” “Ambassador,” “Analyst,” “Marketer,” “Developer,” and “Experimenter.” “Coordinator,” for example, corresponds with the corner defined by Inwards on the Collaboration Direction dimension, Traditional Data on the Data Space dimension and Service on the Value Impact dimension.¹⁰ However, these names should not be taken too literally; they are simply a short-hand notation for each of the corners. Each of the eight roles is explained below.

It is important to note that, at any one time, a CDO may take on multiple roles. However, a CDO inevitably has one primary role. Moreover, it is common for a CDO to take on several different primary roles over time during his or her tenure as a CDO. Many CDOs that we interviewed noted that the evolution of their primary role was triggered by changes in the environment or the broader marketplace, as described below.

9 Picoli, G. and Pigni, F. “Harvesting External Data: The Potential of Digital Data Streams,” *MIS Quarterly Executive* (12:1), 2013, pp. 53-64, explains new value-creating opportunities from digital data streams. One of the five value archetypes is aggregation of digital data.

10 Note that “Coordinator” is much shorter than saying “Inwards Collaboration direction dimension, Traditional Data Space dimension and Service Value Impact dimension.”

1. Coordinator CDO: Inwards/Traditional Data/Service Focuses

The Coordinator CDO manages enterprise data resources and sets up a framework that optimizes collaboration across internal business units (inwards focus). This enables the delivery of high-quality data to data consumers in the organization for their business purposes, thereby improving business performance (service focus). The Coordinator CDO works with traditional data, such as customer information and other transactional data (traditional data focus).

For example, the CDO at a U.S. government agency identified common critical data elements across the enterprise; these elements provided the foundation for data sharing and integration at the agency level. The agency then led an initiative to identify authoritative sources for these critical data elements. This work on common data elements set the stage for other data-improvement initiatives. Part of this CDO's responsibility was to oversee the governance process for data management.

In another example, the CDO of a U.S. healthcare institution established data-governance councils and workgroups. She also led the group responsible for enterprise-wide data quality assessment and improvement initiatives.

2. Reporter CDO: Outwards/Traditional Data/Service Focuses

In heavily regulated industries, such as finance and healthcare, an emerging trend in the CDO role is a focus on enterprise data to fulfill external reporting and compliance requirements. Like the Coordinator CDO, the Reporter CDO fulfills a business obligation (service focus) through the delivery of consistent transactional data (traditional data focus). However, the Reporter CDO's ultimate goal is to deliver high-quality enterprise data services for external reporting purposes (outwards focus).

For example, the CDO-equivalent at a U.S. healthcare institution oversaw the preparation of a selected set of data for regular reporting to the state government. She worked closely with other corporate officers, such as the chief medical officer and chief financial officer, as well as with external officials, to ensure that reports were delivered in a timely manner and that they accurately and effectively represented the activities of the institution.

Similarly, Reporter CDOs are often found in financial service organizations, working with compliance or risk-management groups to fulfill external reporting requirements. Typically, these CDOs are established when the company has experienced difficulties in producing external reports, and often they play an important role in integrating the data and information silos of recently merged companies.

3. Architect CDO: Inwards/Traditional Data/Strategy Focuses

The Architect CDO's Collaboration Direction and Data Space are the same as the Coordinator CDO (inwards and traditional data focuses), but the value impact comes from using data or internal business processes to develop new opportunities for the organization (strategy focus).

As an example, the CDO of a data company was responsible for establishing an enterprise architecture that would yield value-added customer data products. Under the CDO's leadership, the company developed a blueprint that described the business processes for delivering a new data product, the time required for each process and the individual responsible for each process. This blueprint, which we call the "map,"¹¹ was used to encourage members of the organization to collaborate on a daily basis. This CDO recalled: *"We made [the map] everybody's map. Everyone knows their data role in the company."* Suggestions for improvement to data products were also attached to the "map." This CDO reported that the "map" reduced time to market for new products by 50%. In addition, the company produced better data products, and did so before competitors could, thus gaining strategic advantage in the market.

4. Ambassador CDO: Outwards/Traditional Data/Strategy Focuses

The Ambassador CDO promotes the development of inter-enterprise data policy for business strategy and external collaboration (outwards and strategy focuses) and focuses on traditional data. For example, the CDO in a financial services institution defined common datasets for risk management. He promoted a set of data standards and data-assessment measures

¹¹ At the request of the company, we have used a pseudonym for the specific artifact.

for financial data exchange among peer financial institutions.

A second example comes from an international bank in South America, which went through a strategic transformation that required significant process improvements and the establishment of data-governance mechanisms. During the transformation, the CDO, reporting to the CFO, led a close collaboration with other financial institutions to improve data security for electronic international money transfer processes and information exchange. This transformation was critical for the bank's business strategy and opened up opportunities to provide its customers with new services, which were previously not possible due to data-security weaknesses.

5. Analyst CDO: Inwards/Big Data/Service Focuses

The Analyst CDO resembles the Coordinator CDO, except that he or she focuses on improving internal business performance by exploiting big data, thus requiring different data-management and data-analysis capabilities. The need for an Analyst CDO often emerges after an organization hires data analysts or data scientists but does not assign an executive leader to provide an enterprise perspective for their efforts.

For example, a credit card company established a CDO who was responsible for overseeing internal teams evaluating and analyzing big data, such as geo-tagged data about credit card use and data from online customer surveys. This CDO collaborated with the chief risk management officer and provided direction for the data scientists. Subsequently, the company implemented enterprise-wide policies to improve risk management and fraud detection.

6. Marketer CDO: Outwards/Big Data/Service Focuses

The Marketer CDO develops relationships with external data partners and stakeholders to improve externally provided data services using big data. Marketer CDOs are often found in data product companies, where they develop working relationships with retailers, financial institutions, and transportation companies that are purchasing their companies' data.

For example, the CDO of a data product company worked closely with the company's customers, in this case healthcare institutions, to help extract insights from big data in the

form of unstructured patient feedback data. This Marketer CDO led the analysis of the data to identify ways to alleviate key weaknesses of the healthcare institutions. While few CDOs may currently fulfill this role, we observe that the Marketer CDO is an emerging trend that is important for managing supply chain partners and customers.

7. Developer CDO: Inwards/Big Data/Strategy Focuses

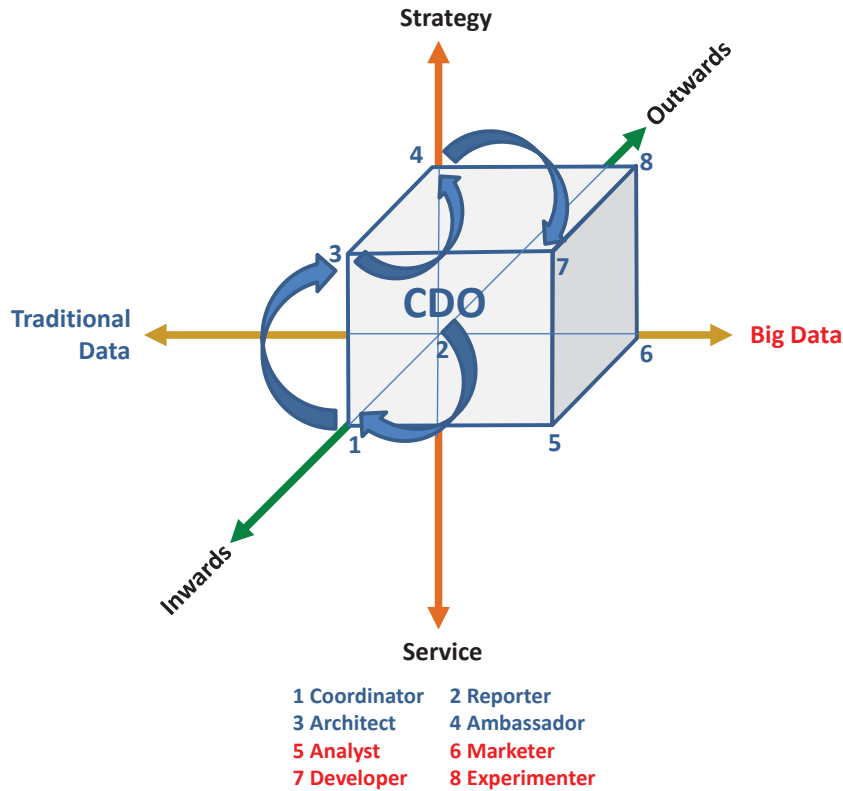
The Developer CDO interfaces and negotiates with internal divisions to develop new opportunities for the organization to exploit big data. For example, the CDO in a retail organization worked with the chief marketing officer to find opportunities for new products and services based on mining consumer behavior data from geo-tagging along with consumer feedback data taken from social media sites. Using this vast source of information, this Developer CDO developed a personalized marketing strategy for the company.

8. Experimenter CDO: Outwards/Big Data/Strategy Focuses

The Experimenter CDO engages with external collaborators, such as suppliers and industry peers, to explore new, unidentified markets and products based on insights from big data. Through strong collaborative relationships across industries, this type of CDO maintains access to various sources of data and uses them for creating new markets and identifying innovative strategies for organizational growth.

For example, the CDO of a financial institution experimented with developing marketable information products for the broader financial industry and its prospective clients. In preparation, this Experimenter CDO suggested creating new information products by transforming, integrating and reusing data from multiple sources of consumer-generated data. More importantly, he presented this new product concept to the organization's clients to gain their feedback. This Experimenter CDO subsequently developed information products based on various data sources and marketed them to client organizations. He argued: *"We should be a revenue center, not a cost center."* By taking advantage of insights from the organization's diverse datasets and guided by his knowledge of shared industry needs, he expanded the organization's

Figure 3: Example CDO Role Evolution



capability to conceive and experiment with new information products, thus adding strategic value.

Example of the Evolution of the CDO Role

Not all businesses have the same needs and priorities, and thus the role of the CDO differs from company to company. Moreover, the role of the CDO can change as the needs of the organization change.

Figure 3 depicts how the role of the CDO at a U.S. hospital evolved over a period of 10 years. In this case, the CDO started with a focus on providing good service to external recipients of traditional data. Gradually her role took on a more strategic focus, both internally and externally, and presently she is concerned with exploiting big data. Over the 10 years we studied this institution, the CDO’s role evolved from Reporter (Role 2), to Coordinator (Role 1), to Architect (Role 3), to Ambassador (Role 4) and now to Developer (role 7). Below, we briefly discuss this CDO’s role over time and explain:

1. What triggered or prompted the CDO to transition to a new role
2. Why that role was chosen
3. What was accomplished by carrying out the new role.

1. Reporter CDO Role

Initially, the CDO fulfilled the Reporter role. As such, she oversaw the provision of data to state regulators, especially for reimbursements, since these were essential to the business. This was a challenge because the data, generated internally from the hospital’s operations, often was not suitable for external reporting purposes. There were multiple sources of the same or similar data, producing inconsistent results. Several data sources were not trusted by internal data consumers, and thus some groups in the organization were reluctant to release that data for external purposes without further review. Every time there was a need for external reporting, the CDO had to go through all of the

data, cleaning it up and preparing it for external submission.

2. Coordinator CDO Role

After being fined for submitting poor-quality data to the state government, the hospital realized that, to report good-quality data externally, it needed to turn its attention to internal data quality. Given a mandate from the CEO to improve the quality of organizational data, the CDO transitioned from the Reporter role to fulfilling the Coordinator role. She established an enterprise-wide data-quality improvement framework, coordinating across functional business units to systematically address *“cleaning up and preparing the data for submission.”* In addition, she developed procedures to assess data-quality techniques periodically and established enterprise-wide dashboards for identifying and resolving data problems. Internal data consumers subsequently felt they could trust their data sources, and the external reporting process was also streamlined.

3. Architect CDO Role

Having successfully improved organizational data both for internal and external services, the CDO realized that there should be a sustainable structure and capability for data practices. This realization prompted her to fill the gap in sustainability by strengthening the alignment of data practices with business processes, thus changing her focus from service to strategy and assuming the Architect CDO role. In this role, she established governance for data quality, as well as standards committees and working groups. She also established and maintained an enterprise level data quality problem-solving process and aligned business roles with data roles for all members of the organization. She implemented a policy of assigning a specific data role to each member of the organization, such as a data collector, data custodian or a data consumer, in addition to a business role, thus strengthening business-data alignment. To reinforce the importance of data roles, each member’s contribution to the quality of enterprise data was factored into their annual bonus.

4. Ambassador CDO Role

Increased pressure from insurance companies for comparable measurements required the CDO to improve collaboration between institutions. The CDO thus evolved to the Ambassador role in which she engaged in industry benchmarking and established shared data practices through a consortium and various forums. She participated in setting the industry’s data roadmap, organizing and training other data practitioners and collaborating with other institutions to promote data quality across other hospitals. Through these efforts, the CDO transformed standards-setting for business processes and for various healthcare industry indices.

5. Developer CDO Role

The hospital’s performance from the use of its internal data eventually reached a plateau. As a consequence, the CDO took on the Developer role, where she explored the use of big data generated by patients to improve hospital performance. In particular, she focused on developing various methods for analyzing unstructured patient feedback data to identify specific factors associated with poor performance. These analyses included data-mining techniques such as sentiment analysis. In combination with analyses of standard numerical assessments, such as the Hospital Consumer Assessment of Healthcare Providers and Systems report, the methods that the CDO developed led to actionable recommendations for doctors, nurses and other units within the hospital. In further pursuing such opportunities, the CDO is currently developing new measurements to provide more tailored feedback to the clinical teams for improving patient care and safety.

Guidelines for Using the Cubic Framework¹²

Our cubic framework can be used to identify the focuses an organization’s CDO should have and hence the CDO role profile, a key to successful data practice. Below we provide a pragmatic three-step guide, based on the framework. In summary, the three steps are:

¹² The authors benefited greatly from the advice, discussion and input from the *MIS Quarterly Executive* workshop on December 15, 2012, in Orlando, Florida.

- *Assess* the current status of your organization's data-related business practices (based on the three dimensions of the CDO cube)
- *Determine* the CDO role profile needed for your organization (based on the eight roles described), and whether an executive-level CDO is required to fulfill these needs
- *Strategize* the likely path for the CDO based on a projection of organizational future needs.

Step 1: Assess the Current Status of Your Organization

Assessing the current status of your organization's data-related practices will help to highlight the weaknesses you should focus on. The CDO cube provides a framework for identifying an organization's current needs with respect to the Collaboration Direction (inwards vs. outwards), Data Space (traditional data vs. big data) and Value Impact (service vs. strategy) dimensions.

In Table 1, we provide 12 assessment statements based on the cubic framework. Each statement is assessed on a seven-point scale (ranging from 1 [strongly disagree] to 7 [strongly agree]). Statements 1-4 relate to the Collaboration Direction dimension; statements 5-8 address the Data Space dimension; and statements 9-12 investigate the Value Impact dimension. To illustrate the assessment process, we have also included sample responses in the two rightmost columns.

Note that most organizations have needs that apply to every corner of the CDO cube; the responses to these assessment statements will help prioritize which roles (i.e., corners of the cube) are most critical. Responding to the statements is also an excellent opportunity to engage many members of the organization at all levels from a variety of business units. The varied perspectives will inform discussions about what CDO role is needed and will ensure the CDO has organization-wide endorsement.

Table 1 can be used both quantitatively and qualitatively. A simple quantitative analysis involves assigning a score (on a seven-point scale) for each response. Comparing the sum of the first two scores and the last two scores

for each dimension will reveal a bias in each dimensional space. In our example, statements 1 and 2 (emphasizing inwards) each have scores of three, and statements 3 and 4 (emphasizing outwards) have scores of six and seven. The sum of statements 1 and 2 (6) is less than the sum of statements 3 and 4 (13), suggesting that collaborating inwards is less critical than collaborating outwards. This same process can be repeated for statements 5-8 to determine whether the focus should be on traditional data or big data, and for statements 9-12 to determine whether the focus should be on service or strategy. Taken together, these comparisons give an indication of which CDO role is the most critical.

A qualitative analysis considers the "why" in the "Assessment Discussion Notes" column for each of the statements. This helps to determine the criticality of each dimensional direction. The examples shown in the rightmost columns of Table 1 are very terse; more comprehensive notes could be used for further elaboration.

Step 2: Determine Whether a CDO is Needed

Based on the assessment of its current status, an organization can move on to Step 2, which determines the CDO role profile needed and whether an executive-level CDO is required to fulfill those needs. Note that considerable discussion may be required before an organization can decide which roles are most important; the scores from Step 1 should not be taken as an immediate solution. Rather, the responses to the assessment statements should be used as a tool to initiate conversations among members of the organization on data practice and the implications for the organization's overall performance.

Establishing a new CDO role requires serious consideration because it implies a change in resource allocation and reporting relationships. Before establishing a CDO position, an organization should therefore review the effectiveness of other data-practice mechanisms, such as governance committees, workgroups and mechanisms for resolving data and business process conflicts. On the other hand, data-practice initiatives alone, without assigned accountability, often do not yield effective results.

Table 1: Example Assessment of CDO Role Based on the Cubic Framework

Collaboration Direction Dimension: Inwards vs. Outwards <i>High score for Nos. 1 and 2 implies inwards direction.</i> <i>High score for Nos. 3 and 4 implies outwards direction.</i>	Assessment Score (1-7) 1 Strongly disagree 4 Neutral 7 Strongly agree	Assessment Discussion Notes Why? Explain reason for assessment score
1. It is critical that our organization improves the effectiveness of data use for internal business operations.	3	<i>We do this well, thus not critical at this point.</i>
2. Our company has the opportunity to significantly improve internal operations.	3	<i>Maintain what we do well.</i>
3. It is critical that our organization collaborates with other value chain enterprises, such as suppliers, customers, distributors or competitors.	6	<i>We need to know our suppliers and customers much better.</i>
4. Our organization’s success is critically interlocked with other companies, market changes, external situations or environments.	7	<i>Our procurement can be vastly improved with better understanding of our suppliers.</i>
Data Space Dimension: Traditional Data vs. Big Data <i>High score for Nos. 5 and 6 implies traditional data; high score for Nos. 7 and 8 implies big data.</i>		
5. Our organization’s transactional data should be more effectively used to address the enterprise’s needs.	6	<i>We need to know more about aggregated amounts of materials for different suppliers.</i>
6. It is critical for our organization to use transactional data in an integrated fashion across different business areas.	7	<i>To negotiate with our suppliers, we must get all divisions to use our existing information in a consistent way.</i>
7. Our company needs to identify opportunities for using big data and data analytics.	5	<i>We may not be there yet to go for this direction.</i>
8. It is critical for our organization to understand external sources of data, such as social media, for engaging customers.	6	<i>Our customers might be ready for new sources in the future, and we need to explore and exploit social media.</i>
Value Impact Dimension: Service vs. Strategy <i>High score for Nos. 9 and 10 implies service focus; high score for Nos. 11 and 12 implies strategy focus.</i>		
9. Our organization’s data efforts should be focusing on maintaining the current needs of the business units.	4	<i>We are doing okay in serving the business units.</i>
10. It is critical for our organization’s operations that we improve the efficiency of the data service.	5	<i>We can still improve, but we do well on serving data for the internal business units.</i>
11. Our organization’s data efforts should be largely initiated by the need for changes in the way we do business.	6	<i>We can use the data for changing the way we do procurement planning with our global suppliers.</i>
12. Our organization must achieve its strategic business goals with better data.	7	<i>We must figure out who our best business customers are and set different strategies for different customers.</i>

Additionally, in some cases, organizations may already have leaders who can take on the role, or parts of the role, of a CDO. For example, the CFO may be able to take on the responsibilities that the assessment carried out in Step 1 would assign to a Reporter or Coordinator CDO, in which case a focus on traditional data and service may not be as critical as the assessment might suggest. We have also seen a case where a chief marketing officer has taken on the responsibilities of a Developer or Experimenter CDO role. In this organization, there was effective collaboration among senior executives, and in such cases, establishing a separate CDO role may not be necessary. More often, however, data-related collaboration among executives can be short-lived and ad hoc, and there is a need for the sustainable leadership made possible by a CDO.

Step 3: Strategize the CDO Evolution Path

Strategizing for future needs can be broken down into two processes. First, the organization should create a projected timeline for addressing the needs identified in Steps 1 and 2. For example, as illustrated in the rightmost columns of Table 1, an organization might determine that the primary need is for an Ambassador CDO role (outwards, traditional and strategy focuses). In this situation, the organization may create an 18-month plan to closely align data practices with business processes.

Second, based on quantitative and qualitative measures, the organization can determine how crucial other CDO roles in the cubic framework are relative to the primary role. Alternatively, the organization may determine that there are no other highly critical needs that must be addressed at this time. In either case, based on the projected timeline, the organization can either determine that the planned CDO will need to transition from one role to another, or it can decide to reassess organizational needs by repeating Steps 1 and 2 in the future.

In the example in Table 1, the scores for statements 5-8 suggest a small bias toward traditional data rather than big data (13 vs. 11). However, further analysis might suggest that big data demands are almost as critical as the traditional data needs that the future Ambassador CDO will be addressing. The organization could therefore plan for the CDO to evolve from the Ambassador role to

Experimenter role (outwards, big data and strategy focuses) at the end of the 18 months to address external needs.

An implicit, yet key strength of the three-step process is that it is a collective endeavor that engages all business units and functions. Enterprise support and approval for the establishment of a CDO lays the groundwork for the CDO to be an effective leader.

Concluding Comments

As organizations' strategies for achieving success increasingly depend on data, they must position themselves to harness the value of data. To this end, a growing number of businesses and government institutions are establishing CDO positions to exploit the critical value that data can provide. The three dimensions of the CDO cube framework presented in this article provide a guide for organizations as they analyze the need for a CDO and will enable them to determine the most appropriate profile for their CDOs now and in the future.

Appendix: Research Methodology

The study was conducted using three research methods: (1) initial informal case studies with multiple organizations; (2) detailed iterative interviews; and (3) structured surveys.

First, we used longitudinal informal case studies with 12 different organizations spanning various industries, including healthcare, finance, government, insurance, manufacturing, retail and IT. As part of our ongoing research on data practices, between 2003 and 2013, we conducted face-to-face interviews and on-site observations of these 12 organizations. The data we collected provided background on emerging CDO practices in the context of these organization, as well as their industries and the broader environment.

Second, during 2010-2013, we focused specifically on the CDOs of these 12 organizations. This entailed iterative interviews and semi-structured surveys, both on- and off-site, as well as continued onsite observation. For a comprehensive understanding of the CDO's work in the context of the organization, we also interviewed other executives and managers working directly with the CDOs on

data quality, governance, data architecture and data strategies. The interviews were semi-structured and open-ended, typically lasting one and a half hours. In total, we interviewed 65 individuals—12 CDOs, 25 other executives and 28 managers.

Third, we developed structured surveys to collect concrete and more detailed statistics on organizational practices relating to CDOs, such as reporting relationships. Between 2010 and 2013, we surveyed 95 CDOs and data practitioners and collected a wealth of data from which we could tease apart different patterns and rules of CDO practice.

Together, these three methods provided a detailed and comprehensive picture of the contemporary data practices of the chief data officer. The longitudinal research provided critical context for the study; the focused interviews with CDOs provided the activity-level details needed for devising the cubic CDO framework; and the surveys provided the statistical power to identify key trends of the CDO role.

About the Authors

Yang W. Lee

Yang Lee (y.lee@neu.edu) is Associate Professor at Northeastern University, D'Amore-McKim School of Business. She was founding Co-Editor-in-Chief of the *ACM Journal of Data and Information Quality* and co-founder of the *International Conference on Information Quality*. Her research focuses on information quality, problem solving and institutional learning, and strategic use of information. Her publications have been widely cited, translated into various languages and applied globally in the public and private sectors. She has received numerous recognitions, including the 2012 DAMA International Achievement Award and the 2005 Certificate of Appreciation from the Director of Central Intelligence, U.S. She received her Ph.D. from MIT.

Stuart E. Madnick

Stuart Madnick (smadnick@mit.edu) is the John Norris Maguire Professor of Information Technology in the MIT Sloan School of Management and Professor of Engineering Systems in the MIT School of Engineering. He received his M.B.A. and Ph.D. in Computer

Science from MIT, has been an MIT faculty member since 1972 and has headed the IT group for more than 20 years. He is Co-Director of the PROductivity From Information Technology and Total Data Quality Management research programs. He is the author/co-author of over 300 books, articles or reports. His research interests include integrating information systems, data quality and strategic use of IT.

Richard Y. Wang

Richard Wang (rwang@mit.edu) is Director of the MIT Information Quality and the MIT Chief Data Officer Research programs. He was a professor at MIT Sloan School of Management for almost a decade. He has also served as the Deputy Chief Data Officer and Chief Data Quality Officer of the U.S. Army. He is the recipient of numerous awards, including the DAMA International Achievement Award, the German Society of Information Quality Award and the International Association of Information and Data Quality Award. Wang received his Ph.D. from the MIT Sloan School of Management.

Forea L. Wang

Forea Wang (forea@stanford.edu) is a Ph.D. candidate at Stanford University School of Medicine in the Neurosciences Program. She is a recipient of the Stanford Graduate Fellowship, the National Science Foundation Graduate Research Fellowship and the MIT Information Quality Program's Decade of Outstanding Contribution Award. She received her B.S. in biological engineering from MIT.

Hongyun Zhang

Hongyun Zhang (zhanghongyun@mail.xjtu.edu.cn) is a post-doctoral fellow at the Center of Data Science and Information Quality, School of Management, Xi'an Jiaotong University, China. In addition, she is also a visiting research scholar at supply chain and information management group, D'Amore-McKim School of Business, Northeastern University. From 2008 to 2009, she was a visiting scholar in Manchester Institute of Innovation Research, Manchester Business School, University of Manchester, U.K. Her research interests include the chief data officer role, data quality and entrepreneurial orientation.