

PERSPECTIVE

Effective Data Governance



Abstract

Data governance is no more just another item that is good to talk about and nice to have, for global data management organizations. This PoV looks into why data governance is now on the core agenda of next-generation organizations, and how they can implement it in the most effective manner.

Why is data governance important and challenging now?

Data has grown significantly

Over time the desire and capability of organizations, to collect and process data has increased manifold. Some of the facts that came out in various analyst surveys and research suggest that:

- **Structured data** is growing by over **40%** every year
- **Traditional content** types, including unstructured data, are growing by up to **80%** every year^[1]
- Global data will grow to 40 zettabytes (ZB) by 2020^[2]
- 85% of this data growth is expected to come from new types; with machine-generated data being projected to increase 15 times by 2020^[2]

Variety of data and increase in sandboxing culture

The next-generation analytics utilize data from all kinds of social networks and blogospheres, machine-generated data, Omniture / clickstream data, as well as customer data from credit management and loyalty management. Alongside this, organizations have now set up sandboxes, pilot environments, and adopted data discovery tools and self-service tools. Such data proliferation and the steep increase in data consumption applications demands stringent and effective data governance.

More stringent regulations and compliance

The growing regulatory and compliance requirements are making effective data governance a must have for industries like financial services and healthcare. The regulatory requirements are now more

demanding around data privacy, personal information protection, data security, data lineage, and historical data.

This makes data governance top priority for Chief Information Officers (CIOs). In fact, a survey by Gartner suggested that by 2016, **20% of CIOs in regulated industries would lose their jobs** for failing to implement the discipline of Information Governance, successfully^[3].

Data to insights to actions: Need for accurate information

Today's managers use data for decisions and actions. In our experience, many managers feel that the data they are accessing is inaccurate or incomplete, and hence, both confidence and adoption of business intelligence and analytics systems is low. Hence, data governance initiatives need to generate good confidence in the data managers see and use for decision-making.

What are organizations doing to establish effective data governance?

Role of the chief data officer (CDO)

Gartner predicts that 50% of all companies in regulated industries will have a CDO by 2017^[4]. The CDO will be responsible for enterprise data management, enterprise data governance, data consumption applications like data warehousing,

business intelligence and analytics systems, data quality, metadata management, master data management, data strategy, and data life cycle.

Collaborative Data Governance Organizations

It is important for Chief Data Officers to engage the right teams in the Data Governance Organization, to align data governance and business goals. Business teams are important to define vision, identify data to be governed,

support the implementation of DG policies, while actively participating in change management and monitoring. IT teams are responsible for the management of data and measurement of data quality, as well as collection and management of metadata and master data. They also provide tools and technologies for implementation of the related tools and technologies.

Typical Data Governance Org Structure

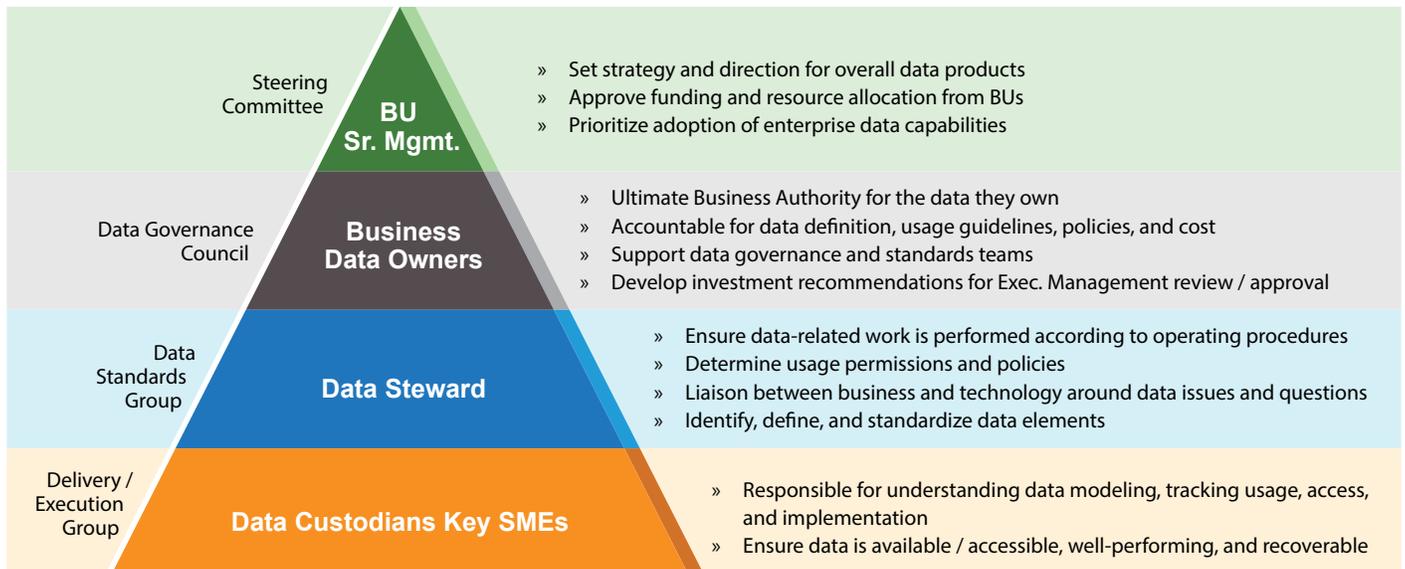


Figure 1: Role of Business Teams in Data Governance

Articulate tangible benefits from data governance initiatives

In order to keep all the stakeholders engaged and ensure continuous investments in data governance initiatives, it is important to articulate the value generated by data governance initiatives via the right metrics. A dashboard with relevant metrics can be an easy tool to assess the impact of data governance initiatives, such as improvement in data quality parameters, improvement in dollar-value impact because of improved data quality, and the ability to rationalize enterprise metrics to data definition standardization. Such tools play an important role in monitoring and controlling adherence to the data governance policies.

governance strategy includes policy and rules for data quality management, metadata management, master data management, as well as data security. The implementation of each of these individual disciplines presents the intended benefits. However, data governance strategies

and implementations that drive common goals and integrate outcomes across these disciplines enjoy benefits like improved reliability, traceability, and authenticity of data. This will also provide a common communication platform across all the stakeholders.

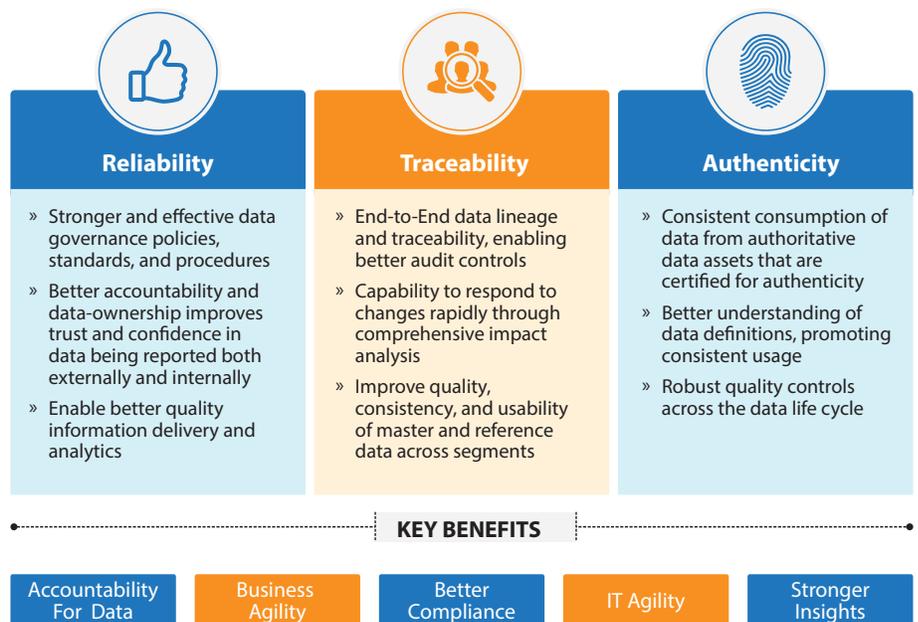


Figure 2: Benefits of Seamless Data Governance

Sum of parts is greater than the whole – seamless data governance

Organizations are now looking at data governance holistically. The Data

Big data user ratings are replacing traditional data quality (DQ) metrics

Ratings by data consumers are replacing traditional means of measuring data accuracy for non-traditional data such as sentiments and opinions. Most of the leading data platform providers such as Cloudera, Hortonworks, and Informatica (BDE) offer ways of capturing the usage of data from data lakes and user ratings about the usefulness and quality of the information.

Effective Data Governance at Infosys

Infosys holistic service offerings help next-generation organizations to implement effective data governance.

The goal of data governance initiative is to manage data for delivering timely, trustworthy, and relevant information to make informed business decisions. However, data governance is not only about data, but also about enabling clear ownership, business rules, operational

requirements, tools and business processes, easy decision-making process, stakeholder interaction, and business access.

Once an enterprise establishes the above process of enabling informed business decisions as the collective aim of data governance, it becomes significantly easier to define the goals and priorities for any initiative. Good data governance paves the way for enterprise policies, people, and processes to launch that initiative.

Infosys Offerings for Effective Data Governance

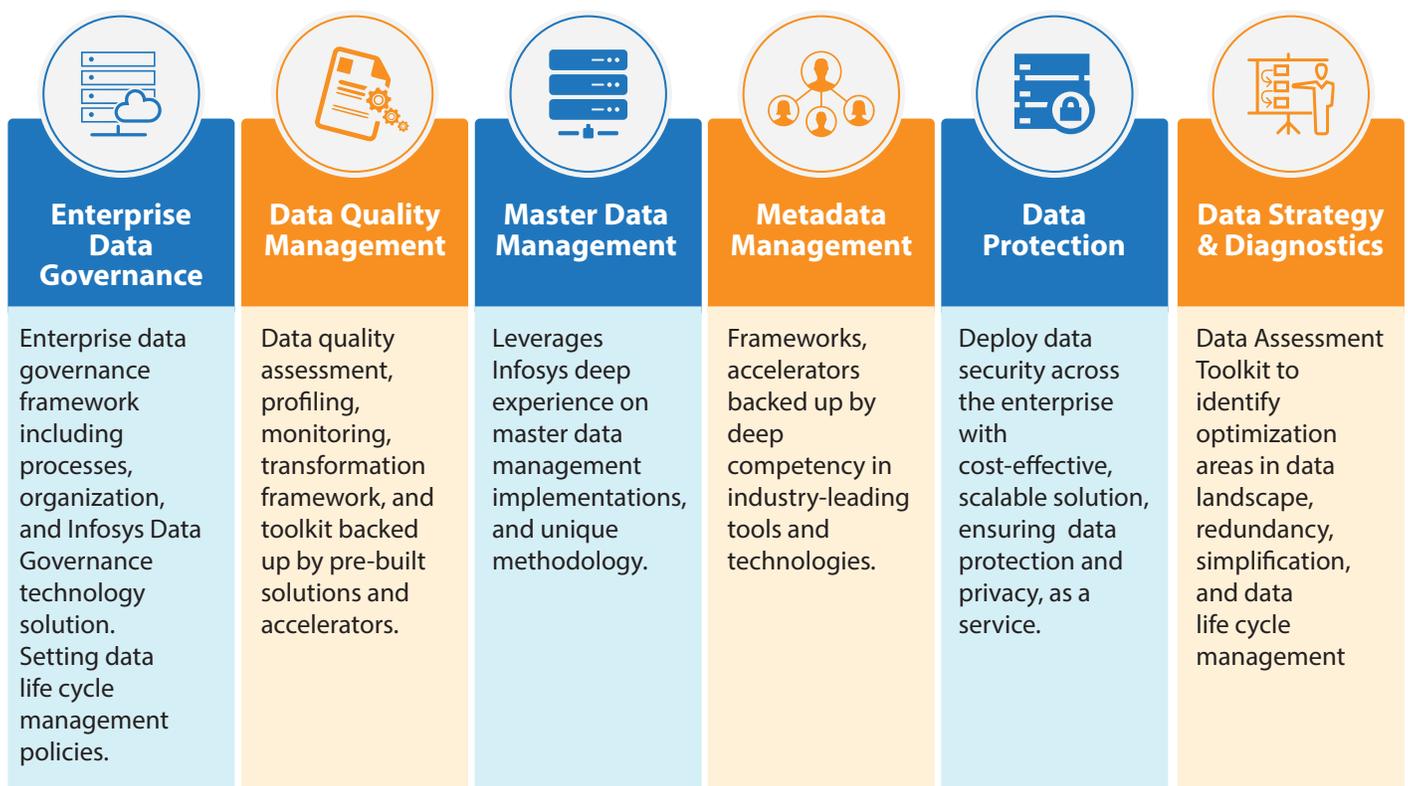


Figure 3: Effective Data Governance @ Infosys

The focus now turns to the tools and technologies to implement and ensure compliance with data governance requirements in accordance to business access.

Infosys Data Governance Solution offers a robust framework and technology solutions that leverage partnerships with leading product vendors in this space.

Infosys Data Governance Framework with measurable KPIs efficiency and efficacy

This framework helps next-generation organizations to assess the data governance maturity and needs, define the strategy accordingly, and subsequently implement the same.

Data Governance Maturity Assessment

- Understand data assets of an organization
- Understand data-spread across, on premise, and Cloud
- Understand master data management, metadata management processes, data quality processes, and data security setup and processes

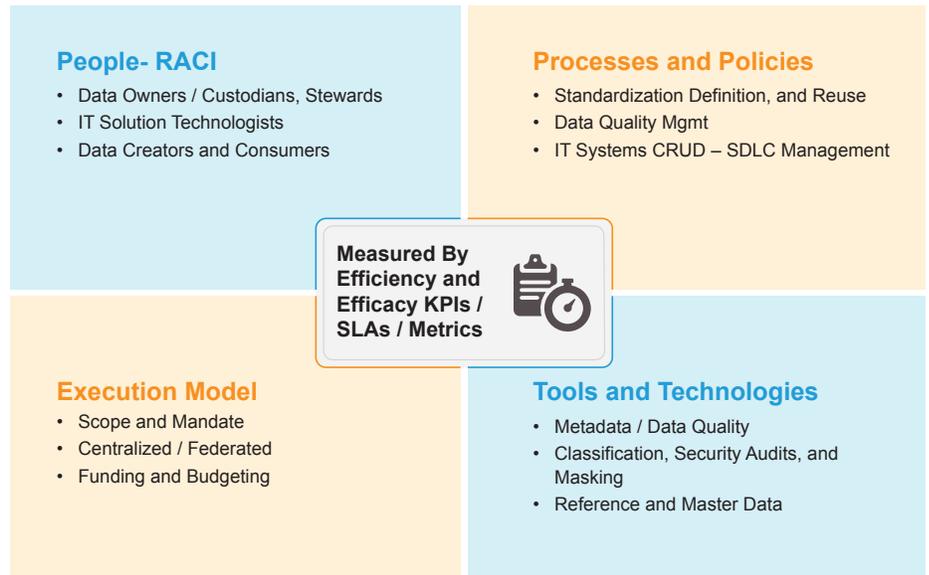


Figure 4: Infosys Data Governance Framework

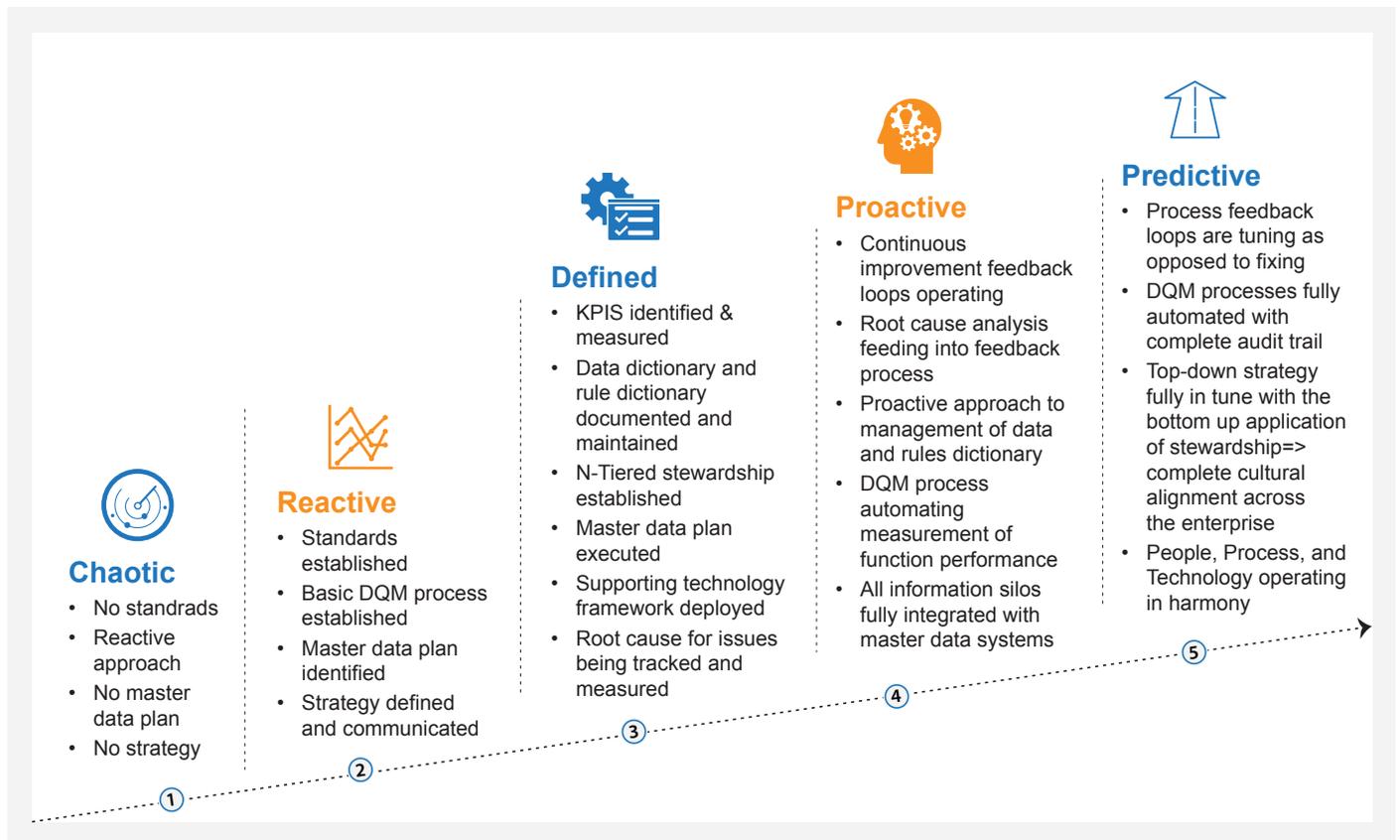


Figure 5: Data Governance Maturity Phases

Establish Goals and Objectives

- Understand the key drivers of data governance, and the pain points of data management
- Define the objectives to be achieved with data governance
- Define the business case of data governance

Data Policies, Procedures, and Standards

- The policies and standards for data management, data availability, transparency, and security
- Define the metadata standards and policies for representation of data lineage, data models, and data dictionaries
- Define the policies for capturing metadata, and Information Life cycle Management
- Define the policies for data transparency – audit, data security, authentication, and authorization

Metric Control and Framework

- Define the control metrics to evaluate the effectiveness of data governance
- Define the process of capturing the metrics and how they will be reported
- Establish the link between the metrics and the business goals of the organization
- Define the roadmap for the implementation of data governance, its specific objectives, and the timelines for achievement

Organization Structure and Rollout

- Define the data governance organization structure to maintain accountability

- Define the responsibilities of key roles such as application and data stewards
- Define the review frequency for policies, standards, governance structure, and roles
- Define mechanisms to enforce policies and procedures
- Communication plan to keep the data stewardship community informed
- Data stewardship reports to enable the effective management of data stewardship

Infosys Data Governance Tool

This is a robust data governance tool that provides single point access for data quality rules management, stewardship

and data quality metrics reporting, master data management, lineage and traceability visualization, and data issue investigations. Some of the key features of the tool are:

Single point access for measurable DG metrics for all the DG stakeholders

- Overall data health for executive sponsors
- Dollar-value impact of data quality for business leaders
- Impacted data consumers – reports, feeds, and analytics due to bad data quality
- Correlation analysis of data quality issues across the application for data stewards, and many more

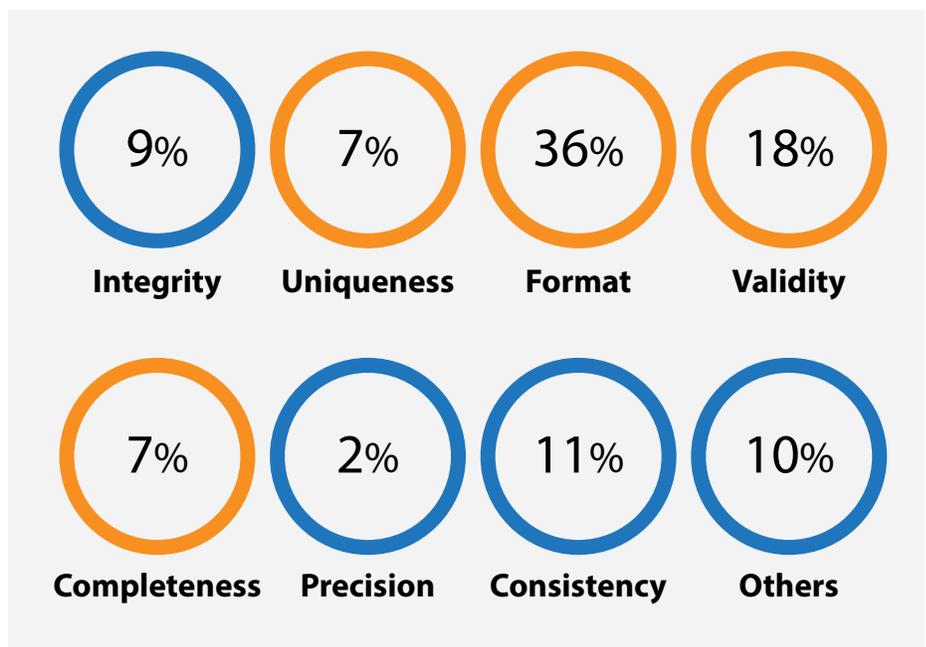


Figure 6: Infosys Data Governance Solution

Complementary to industry-leading metadata management, master data management, and data quality management tools

- Infosys data governance tool has connectors to popular third part metadata management (MM) and data quality (DQ) management tools. It integrates lineage data available from MM tools, with data quality metrics from DQ tools to enable analytics for various stakeholders.
- Covers unhandled data quality issues from support tickets.

Big data governance

Most of the leading vendors in the information management space, such as Informatica and IBM have string offerings around data profiling, cleansing, lineage, master data management, and metadata management. These vendors have already introduced, or are in the process of enhancing the capabilities of their technology stacks to address the challenges of big data.



Big data platform vendors such as Cloudera, Hortonworks, and MapR are offering components that provide capabilities around metadata management (like LOOM and HCatalog), data lineage, and data security (like Sentry, Ranger, and Knox).

The Infosys data governance solution offers a holistic approach to manage big data governance with connectors to these leading tools, and establishing traditional data quality metrics, as well as user ratings.

Leverage Strategic Partners

Infosys works with leading technology vendors in data governance, data quality, metadata management, master data management, data security, and data diagnostics like Informatica, IBM and Collibra to name a few – to bring the best of solutions and services to our customers. These partnerships help us provide truly integrated solutions and accelerate implementations.

Success Stories

Infosys helped one of the world's leading financial services company that provides asset management, portfolio management, and mutual fund services, to realize their enterprise data governance vision. We achieved this by implementing the Infosys Data Governance Solution, data quality management tools, and metadata management tools. The customer was able to visualize metadata, data quality, and the impact of data quality in a much better way; thereby improving operation efficiency and saving a lot of effort in data quality check.

Infosys helped a leading bank in South East USA to automate data quality checks for auditing data for FR FY14 compliance reporting to the Federal. The solution was then extended to other lines of business and BI applications. Infosys managed to deliver compliance reporting in time, with about 300+ rules configured, and executed using Ab Initio BRE, in a very short span of time. The customer benefited hugely from the improvement in the regulatory compliance reporting, reinforced confidence, and credibility.

For a leading national cable television operator in the US, Infosys defined and implemented a total data quality solution across data from sales and marketing, commercials, and finance lines of business. The solution monitored the health of the data landing in the enterprise data warehouse, and raised alarms in case of threshold breaches. Timely capture of key metrics variances between sources and corresponding targets, and timely automated communications across stakeholders when thresholds were met, helped in raising the confidence in data.



References

- 1 Source Gartner Information Governance: *12 Things to Do in 2012* by Gartner
- 2 *THE DIGITAL UNIVERSE IN 2020: Big Data, Bigger Digital Shadow s, and Biggest Growth in the Far East* Sponsored by EMC Corporation.
- 3 Source Gartner <http://www.gartner.com/newsroom/id/1898914>
- 4 <http://www.gartner.com/smarterwithgartner/understanding-the-chief-data-officer-role/>

For more information, contact askus@infosys.com

Infosys®

© 2015 Infosys Limited, Bangalore, India. All Rights Reserved. Infosys believes the information in this document is accurate as of its publication date; such information is subject to change without notice. Infosys acknowledges the proprietary rights of other companies to the trademarks, product names and such other intellectual property rights mentioned in this document. Except as expressly permitted, neither this documentation nor any part of it may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, printing, photocopying, recording or otherwise, without the prior permission of Infosys Limited and/ or any named intellectual property rights holders under this document.

Stay Connected    