

BDNA White Paper

Best Practices for Enterprise IT Data Quality

A Data Maturity Approach for
Chief Data Officers

B | D N A

Executive Overview

As organizations wrestle with fast-growing data volumes, they are beginning to place responsibility for managing that data in the executive suite. According to Gartner, more than 25% of global organizations will have a Chief Data Officer (CDO) by 2015.ⁱ While CDOs must control the costs and consequences of explosive data growth, they're also called on to help the organization capture, understand, and exploit big data for strategic gain. The first step in achieving this goal is to ensure that the organization is working with data that is standardized, complete, and up-to-date. This paper discusses how focusing on data quality for enterprise IT allows the CDO to improve data maturity rapidly and create early wins that support broader data quality initiatives.

Why Is Data Unreliable?

The success of IT processes and projects depends on reliable data, but much data is unreliable. In fact, Gartner estimates that more than 40% of business initiatives fail to meet their objectives due to poor data quality.ⁱⁱ Moreover, when process and project owners try to address data quality issues, they often do so in isolation, which further fragments data across the enterprise and makes it even more difficult for solutions to exchange information. In enterprise IT, three persistent issues drive this problem:

Lack of Vendor Standards

Hardware and software innovation is moving at such a rapid pace that managing an enterprise IT portfolio is an exercise in chasing a constantly moving target. Every day brings new products, versions, editions, subversions, and bundles: in January 2014 alone, 525 new vendors brought more than 4,476 products and solutions to market.ⁱⁱⁱ As these products, from security to deployment to finance, make their way into the infrastructure, organizations must categorize and uniquely identify them to manage them effectively.

Unfortunately, vendors do not provide standards for naming and documenting their products, their support for identifying them is extremely limited, and hardware and software offer only cryptic and incomplete clues for operational tools to leverage. For example, the recent Heartbleed vulnerability took advantage of the OpenSSL library — which is identified by more than 2,000 unique package and file strings across the 40 most popular operational tools.^{iv} Since few organizations have the time or skill to develop and maintain a repository for every software package and hardware device on the network, most organizations cannot entirely trust their data. This uncertainty affects all core IT processes and projects.

Market Velocity

In addition to tracking individual hardware and software solutions, organizations need to understand the shifting IT market and its impact on their IT portfolios. Clear visibility into enterprise IT data demands awareness of new companies, new technologies, and mergers and acquisitions, all of which can have a dramatic effect on rolling up total vendor spend.

Market velocity also pressures technology vendors to shrink time-to-market, regardless of the effect on data quality. Even within popular technology stacks, product groups racing to integrate applications introduce inconsistencies (for example, in feeding data from a configuration management database to a software asset management repository). Multiple representations often occur within the same

Multiple representations often occur within the same technology or even the same instance of a given technology. Multi-vendor organizations — which is to say, most of them — are left to reconcile data themselves within and across vendor offerings.

Technology Debt

The increasing cost of maintaining older systems further exacerbates enterprise IT data issues. Legacy systems that have not yet been retired contain legacy data that still has value for the enterprise. The company cannot avoid spending money to access this data — either by investing in maintaining the legacy systems, or by investing instead in translating legacy data into a form new systems can process. IT organizations, under pressure to respond to business needs while remaining within budget, often release capabilities without developing comprehensive test suites or achieving clarity about business requirements. They then iterate to add capabilities rather than refactoring applications or configurations to meet changing functional demands or timelines. As a result, data is often trapped in silos, making every integration scenario more costly and complex.

When organizations fail to invest in data issues within processes or projects, integrating this siloed data to improve efficiency creates massive process confusion. Because different systems refer to the same technology by different names, the task of resolving conflicting data slows cross-process integration and hinders business agility. This discourages further attempts at integration and reinforces the siloed approach to data just as the business is clamoring for better solutions. The more IT scrambles to meet business demands, the more data quality deteriorates.

A Faster Path to Data Maturity

More than 67% of respondents to a Forrester survey about data quality issues said the barriers to addressing the problem are moderate to significant.^v Although data maturity and data governance models address the problem of data quality at the source, a CDO is not always capable of implementing these models to address the issues of enterprise IT data quality. Lack of vendor standards and market velocity are both outside a CDO's control, while technology debt is such a large issue that it can only be addressed incrementally over a long time frame.

Given these realities, the CDO must find a way to accelerate data maturity in order to unlock the value of data for delivering greater efficiency, cost savings, and agility. One approach is through abstraction and simplification, which the IT industry has long used to solve complex issues. For example:

- **The Java programming language abstracts the complexities of cross-platform development to a reference language, and the Java Virtual Machine hides the details of the platform on which it operates.**
- **Hardware virtualization abstracts the complexity of the underlying hardware, vastly accelerating and simplifying system provisioning.**
- **Software Defined Networks abstract network topologies, shielding applications from the complexity of network design.**

Similarly, organizations can leverage the rapid growth of compute and storage infrastructure to create an abstraction model for enterprise IT data and leap from basic data integration with simple ETL tools and custom code to a cross-enterprise platform supported by a service-oriented architecture. Instead of struggling with basic point-to-point integrations, the CDO can create a clearinghouse for trusted and contextual information about enterprise IT data, accessible by any relevant stakeholder. This makes it easier to deliver the data services that support solutions for strategic insight, functional and process agility, and operational effectiveness.

Technologies for Success

To successfully accelerate data maturity for enterprise IT data, CDOs must have three core capabilities: Reference Language, Data Abstraction and Content Refresh Processes.

Reference Language

The primary requirement for addressing data quality in enterprise IT is to establish a reference language, or product catalog. Organizations have historically attempted to create these on their own, with varying levels of success. Today, however, they can turn to commercial content providers who create, manage and curate the data on behalf of customers.

In choosing an appropriate reference language, CDOs should evaluate several key areas:

- **Breadth:** Does the solution cover all technologies used in the enterprise as well as potential new technologies? Does it provide sufficient historical information and enumerate older hardware and software still actively used within the enterprise as part of the technology debt?
- **Granularity:** Does the solution provide the detail needed for identified use cases? Will it provide sufficient detail for operational use cases while allowing for aggregation for strategic use cases?

- **Market Intelligence:** Does the solution provide market intelligence that is aligned to the hardware and software items to give context to the information? Does the market intelligence address both tactical and strategic use cases?
- **Availability:** Is the solution freestanding, or is it embedded as part of a separate tool or solution? Is the solution available for integration via modern APIs and other methods?

Data Abstraction

Once an organization selects a reference language, it needs an abstraction or normalization engine that insulates tactical collection of enterprise IT data from the data's strategic use. Key capabilities should include:

- **Reference Language:** Does the solution normalize enterprise IT data to the selected reference language?
- **Data Sources:** Does the solution support all enterprise IT data sources currently in use? Is there a roadmap or capability to extend it as sources are added?
- **Data Types:** Does the solution support data from financial systems (e.g. purchase orders) as well as operational systems (e.g. hardware and software deployments)?
- **Integration:** Can the solution readily integrate the normalized enterprise IT data into consuming systems? Does it have pre-built integrations for consuming systems?

Content Refresh Processes

Any organization with its own reference language or product catalog is aware of the difficulty of keeping the content up-to-date. Keeping pace with market velocity and falling behind it each carry their own risks. CDOs should assess a solution for all the following capabilities:

- **Reference Processes:** Does the solution have mature processes for content refresh? Do these processes capture daily change?
- **SLAs:** Does the solution offer SLAs for addressing gaps in content? Are those SLAs acceptable to the planned processes?
- **Customer Input:** Does the solution require direct customer involvement (e.g. access to SMEs within the organization) to address content gaps?
- **Market Intelligence:** Does the solution offer SLAs that extend to market intelligence as well as catalog items?

Adoption Strategies

Data quality issues are especially visible in IT — increasing operational costs, raising business risk, and hindering processes and projects from configuration management to vendor audits. This makes IT the ideal setting for establishing an appropriate technology foundation with which to demonstrate the value and impact of improved data maturity. By focusing on low-complexity, high-value opportunities for improving data maturity in enterprise IT, CDOs can lay the groundwork for broader adoption and greater benefit elsewhere.

Land: First Wins

Focusing on deploying reference language can deliver initial data maturity wins in many IT processes and projects. IT processes ranging from enterprise architecture to IT business management to service management rely on a definitive IT catalog. However, many reference language solutions do not provide a catalog, and those that do often provide either an outdated catalog or one that is too bare-bones for enterprise deployments. By comparison, BDNA Technopedia™, BDNA's IT reference language, is both up-to-date and comprehensive. The following use cases represent actual implementations:

Functional Area	Description	Savings / Benefits to:
Enterprise Architecture	Populated product catalog of enterprise portfolio management suite with Technopedia	Cost: 125K direct, \$150-250K annually Efficiency: Eliminated manual catalog maintenance Risk: Improved visibility to SW EOL dates Time to Value: 2 weeks
IT Service Management	Populated ServiceNow catalog with Technopedia	Cost: \$140K direct, \$220K annually Efficiency: Eliminated manual catalog maintenance Time to Value: 3 days
IT Planning	Replaced manually maintained catalog in Corporate Executive Board IT Roadmap Builder with Technopedia	Cost: \$40K direct, \$5K annually per CEB client Efficiency: Eliminated manual catalog maintenance Agility: Improved customer offering Time to Value: 3 weeks
Procurement	Established Technopedia as Master Software Title for IT procurement application	Cost: \$75K direct, \$600K/year with improved entitlement clarity Efficiency: Eliminated manual catalog maintenance Risk: Better purchasing decision and control Time to Value: 5 weeks

Functional Area	Description	Savings / Benefits to:
Configuration Management	Normalized data from 4 discovery sources, removing all data variability, then load into ServiceNow CMDB	Cost: \$200K/year (manual processes) Efficiency: Eliminated manual catalog maintenance Risk: Better purchasing decision and control Time to Value: 5 weeks
Governance, Risk and Compliance	Normalized data from 3 security tools, enriched with Common Platform Enumeration (CPE) data, and populated into	Cost: \$240K/year (manual processes) Efficiency: Integration between security and operations Risk: Identified over 27K previously unseen vulnerabilities
Software Asset Management	Normalized data from 2 discovery sources before loading SAM solution	Cost: \$125k/year (manual processes); \$210K savings on first negotiation Efficiency: Reduced audit response time by 70% Risk: Identifying and remediating EOL SW prior to audits Time to Value: 8 weeks
Vendor Management	Normalized purchase order data from primary procurement system to better understand vendor spend	Cost: \$375K/year (manual processes) Efficiency: Aligned financial and operational reporting to same language Risk: Aggregate spend across vendor acquisitions automatically Time to Value: 3 weeks

Expand: Process Integration

Once CDOs deliver fast, visible success with data maturity in the Land phase, they can expand into related processes to demonstrate more lasting and impactful improvements — for example, by leveraging the Catalog populated in ServiceNow to address data quality in configuration management. They can also expand into tangential processes to demonstrate broader applicability. For example, they can show the cross-functional value of a reference language and associated normalization capabilities by leveraging the same Catalog used by IT planners in Enterprise Architecture to support Software Asset Management. The following examples show some of the other possibilities for rapidly expanding data maturity: *(above)*

Standardize: Enterprise Adoption

The end goal for data quality efforts is to establish an enterprise data platform that isolates the tactical collection of data from its strategic use. The more functional processes and projects use the same data abstraction, the more value the organization can derive, and in less time — aggregating, analyzing, and acting on clean data to achieve greater strategic insight and answer questions the organization might not previously have been able to ask. Establishing a common language across processes and projects also reduces organizational friction between planning, procurement, and operations while lowering the costs of integration. At the same time, the ability to support market intelligence with relevant, contextual data accelerates process and project ROI.

Although all organizations share this common goal, the way organizations move from the Land and Expand phases to enterprise-wide adoption of data quality will vary based on their specific needs. Two very different BDNA customers illustrate how varying approaches to Enterprise Adoption can lead to the same destination:

- 1) A Fortune 100 telecom provider began by replacing its manually maintained catalog of IT data with a Reference Language to support enterprise architecture. Once the software asset managers became familiar with the Reference Language, they began to normalize their software data to conform to it. This freed the configuration management process owners, who were previously responsible for supplying software asset data, to normalize hardware data to support automation and orchestration initiatives. Once these core processes used the same Reference Language, all other IT processes followed suit, allowing the company to leverage the standardized data for increased benefits enterprise-wide.
- 2) A retailer with 8,000 employees began by normalizing and enriching data to complete a tactical Windows migration project on a tight deadline. Simultaneously, it normalized IT data from additional sources for a time-sensitive software audit. The success of those projects persuaded the retailer's IT service management team to adopt the Reference Language as its configuration item (CI) catalog. In just eight weeks, the team normalized all CIs for the company's ITSM platform, establishing a Reference Language and normalization across core enterprise processes.

Individual project or process improvements deliver significant ROI: on a project or process basis, more than half of BDNA customers achieve full ROI within 30 days. This rises to 75 percent within six months, and more than 90 percent within a year. Yet enterprise adoption delivers an even greater impact by eliminating friction between processes: more than two-thirds of BDNA customers who pursue enterprise adoption achieve ROI of 200 percent or more in the first year alone. This is a compelling argument for aggressive expansion.

Conclusion

CDOs under pressure to extract value from the vast trove of data under their purview can jumpstart that process by addressing data quality for enterprise IT. This provides a foundation for an enterprise-wide approach to data quality based on rapid improvements to data maturity.

BDNA supports CDOs' efforts by solving the primary challenges associated with enterprise IT data — the data about hardware and software that is at the core of enterprise IT processes. By ensuring that enterprise IT data is consistent, accurate, and valuable, CDOs deliver the foundational information that tactical and strategic IT processes and projects need to improve efficiency, meet governance demands, lower costs, and increase business agility.

About BDNA

To learn more about BDNA solutions visit www.bdna.com

339 North Bernardo Avenue, Suite 206
Mountain View, CA 94043, USA

T: +1 (650) 625-9530
E: sales@bdna.com

- i. Gartner Newsroom
<http://www.gartner.com/newsroom/id/2659215>
January 30, 2014
- ii. Gartner Research
"Measuring the Business Value of Data Quality,"
Report G00218962
- iii. Data from BDNA Technopedia
February 1, 2014
- iv. Data from BDNA Technopedia
May 13, 2014
- v. Forrester Research
"Master Data Management/Data Quality Survey,"
August 2013