Information as a Second Language: Enabling Data Literacy for Digital Society

Published: 21 September 2018 **ID:** G00365697

Analyst(s): Valerie Logan

Digital society expects its citizens to "speak data." Unless data and analytics leaders treat information as the new second language of business, government and communities, they will not be able to deliver the competitive advantage and agility demanded by their enterprises.

Key Challenges

- Poor data literacy is the second highest inhibitor to progress, as reported by respondents to Gartner's third annual Chief Data Officer Survey, behind culture change and just ahead of lack of talent and skills.
- An information language barrier exists across business units and IT functions, rooted in ineffective communication across a wide range of diverse stakeholders. As a result, data and analytics leaders struggle to get their message across and information assets go underutilized.
- Although academic and professional programs are beginning to address the disparity in talent and skills, in many cases they reinforce the information language barrier with narrow content focus, bias toward tools training, and lack of contextualization by role.
- While conversant in the "people, process and technology" capabilities of business models, many C-level executives and professionals do not speak data fluently as the new critical capability of the digital era. Data and analytics leaders must find new ways to bridge this gap.

Recommendations

For data and analytics leaders, in support of their programs:

- Cultivate information as a second language (ISL) across business and IT stakeholders by first establishing the base vocabulary, clarifying industry and business domain "dialects," and developing levels of proficiency.
- Drive and sustain improvements to your organization's data literacy by identifying areas where data is spoken fluently, where language gaps exist, and establish an ISL proof of concept for language development.

- Change the way you and others interact with leaders, stakeholders and peers by speaking data in context in everyday interactions with colleagues, managers and even executives and board members, and as a basis for outcomes-oriented business cases.

Table of Contents

List of Tables

List of Figures

Strategic Planning Assumptions

By 2020, 80% of organizations will initiate deliberate competency development in the field of data literacy, acknowledging their extreme deficiency.

By 2020, 50% of organizations will lack sufficient artificial intelligence (AI) and data literacy skills to achieve business value.

Introduction

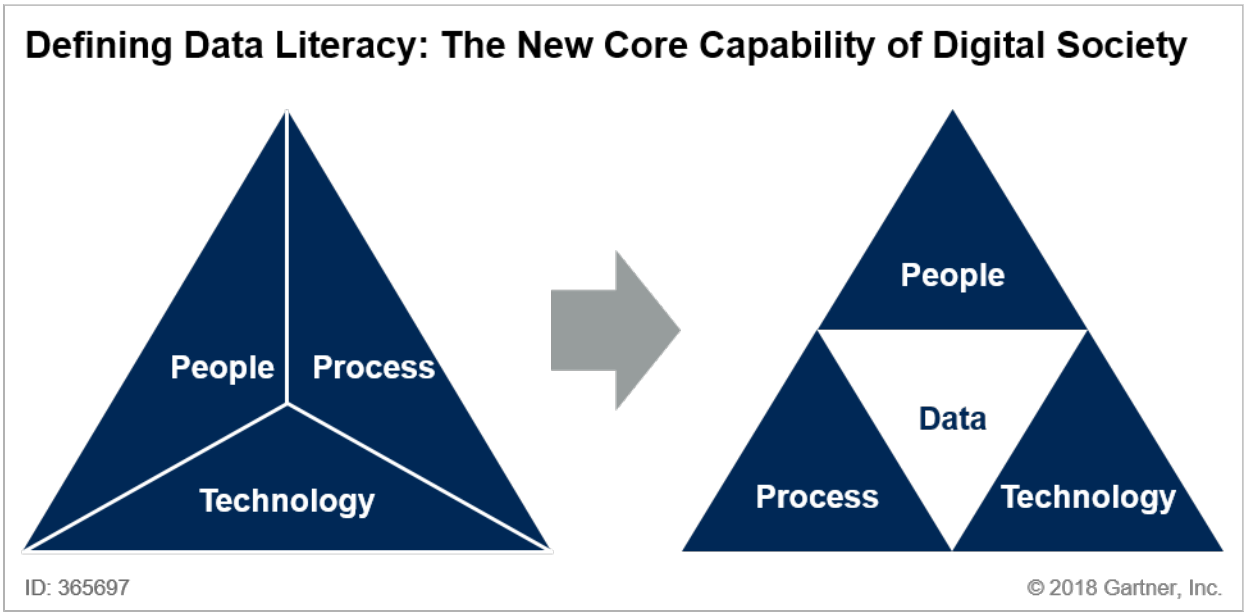
With successive waves of data and analytics initiatives spanning decades, the professional diversity among those who design these solutions ("creators") and those who use them ("consumers") has never been broader. In addition to classic business versus IT heritage, diversity now includes:

- Veterans versus rookies (e.g., a seasoned organizational veteran with a 30-plus-year career vs. a brand new data science PhD university graduate)
- Data versus analytics backgrounds (e.g., an enterprise data architect vs. a business intelligence developer or operations research specialist)
- Industry vertical experience (e.g., healthcare analytics vs. banking analytics specialists)
- Business domain experience (e.g., marketing analytics professionals vs. supply chain analysts)
- Scope of experience (e.g., geographical scope of local, regional or global; organizational scope of function, business unit, division, enterprise or ecosystem)

While natural and healthy, this professional diversity is creating an environment in which there is no shared common language, resulting in a fundamental communication challenge when sponsoring, proposing, leading, designing or using data and analytically based solutions.

While conversant in the "people, process and technology" capabilities of business models, most executives and business and IT professionals do not "speak data" fluently (see Figure 1).

Figure 1. Data Literacy as the New Core Capability of Digital Society



Source: Gartner (September 2018)

As data and analytics become pervasive in all aspects of businesses, communities and even our personal lives, the ability to communicate in this language — that is, being data literate — is the new organizational readiness factor.

Gartner defines data literacy as follows (see also Note 1):

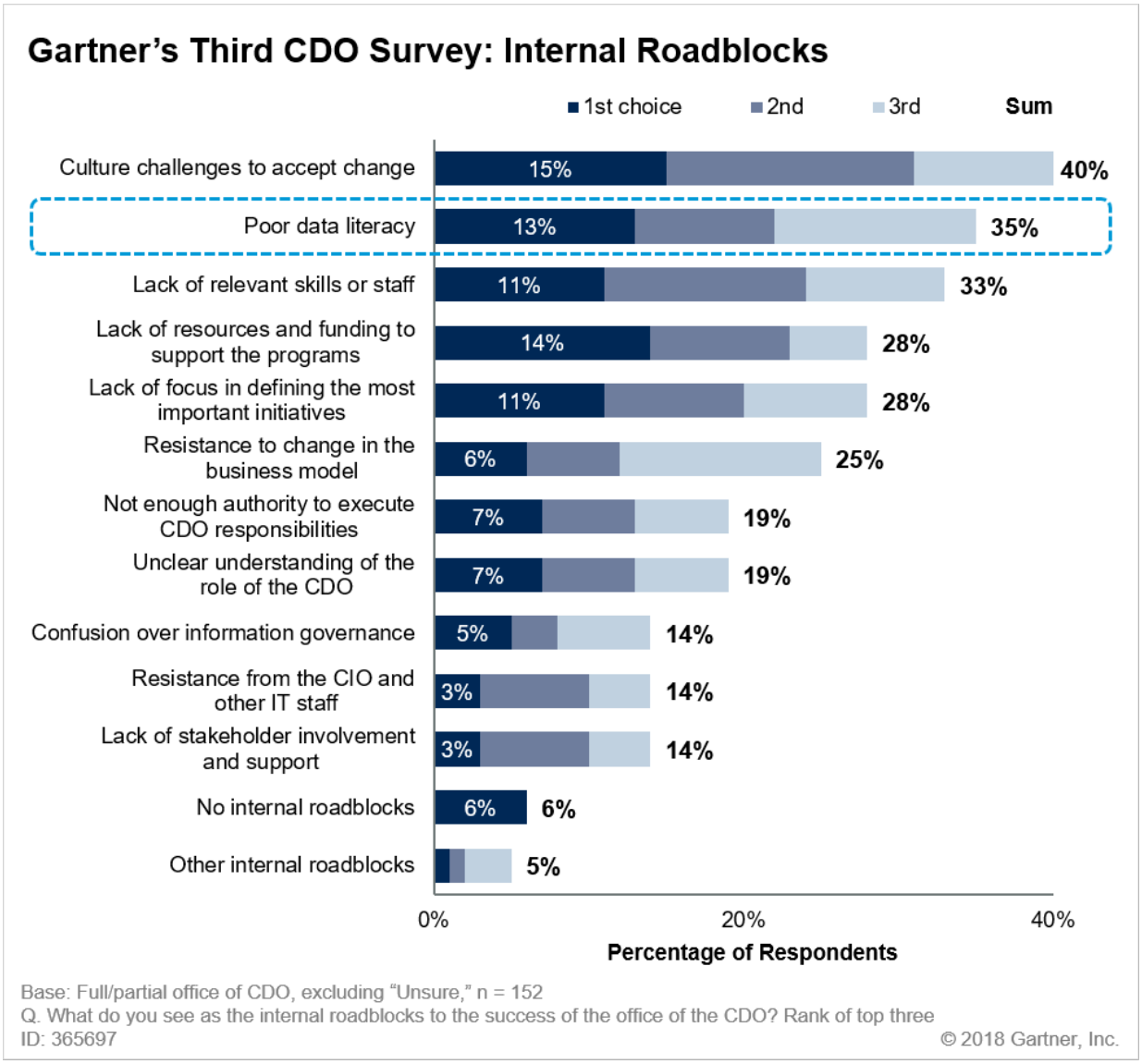
> The ability to read, write and communicate data "in context," including an understanding of data sources and constructs, analytical methods and techniques applied, and the ability to describe the use case application and resulting value.

Data literacy is a two-way communication dynamic — writing/speaking and reading/understanding. Whether translating to the board how data and analytics manifest in company use cases, explaining how to creatively blend internal and external datasets, or describing advanced analytics techniques including AI, speaking data is the new language of the digital economy (see "Fostering Data Literacy and Information as a Second Language: A Gartner Trend Insight Report.")

Analysis

In Gartner's third annual Chief Data Officer Survey, top data and analytics leaders, who are responsible for harnessing the value of data and insight for the organization, voiced a critical challenge of poor data literacy, alongside related challenges of shifting the business culture and lack of talent (see Figure 2).

Figure 2. Gartner's Third CDO Survey: Internal Roadblocks to Success



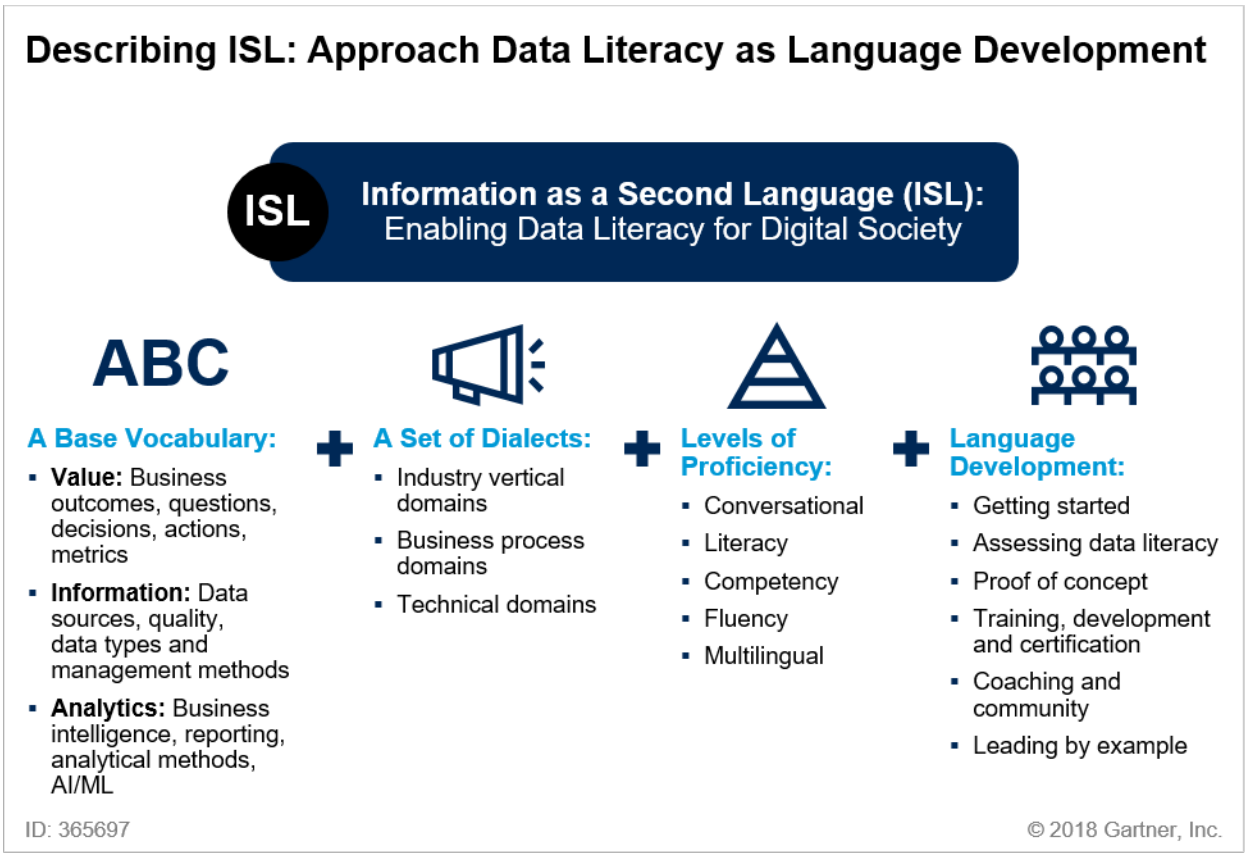
CDO = chief data officer
For more findings from the survey, see "Survey Analysis: Third Gartner CDO Survey — How Chief Data Officers Are Driving Business Impact."

Source: Gartner (September 2018)

In their role, chief data officers (CDOs) are tasked with enabling a community of creators and consumers of data-driven solutions. In doing so, they encounter a wide range of data and analytics skills levels and language abilities. The information language barrier can exist locally or systemically, regardless of program scope or organizational maturity. Addressing it requires a mindset shift, and deliberate acknowledgment and intervention to course-correct. Similar to the days of Six Sigma or business process re-engineering, some will naturally understand the foundational role that data

plays in business transformation, but most will not. Skilled leadership and deliberate change management discipline are required to course-correct, beginning with the acknowledgment of information as the new second language of the ongoing digital revolution (see Figure 3).

Figure 3. ISL: Enabling Data Literacy for Digital Society



AI = artificial intelligence; ML = machine learning

Source: Gartner (September 2018)

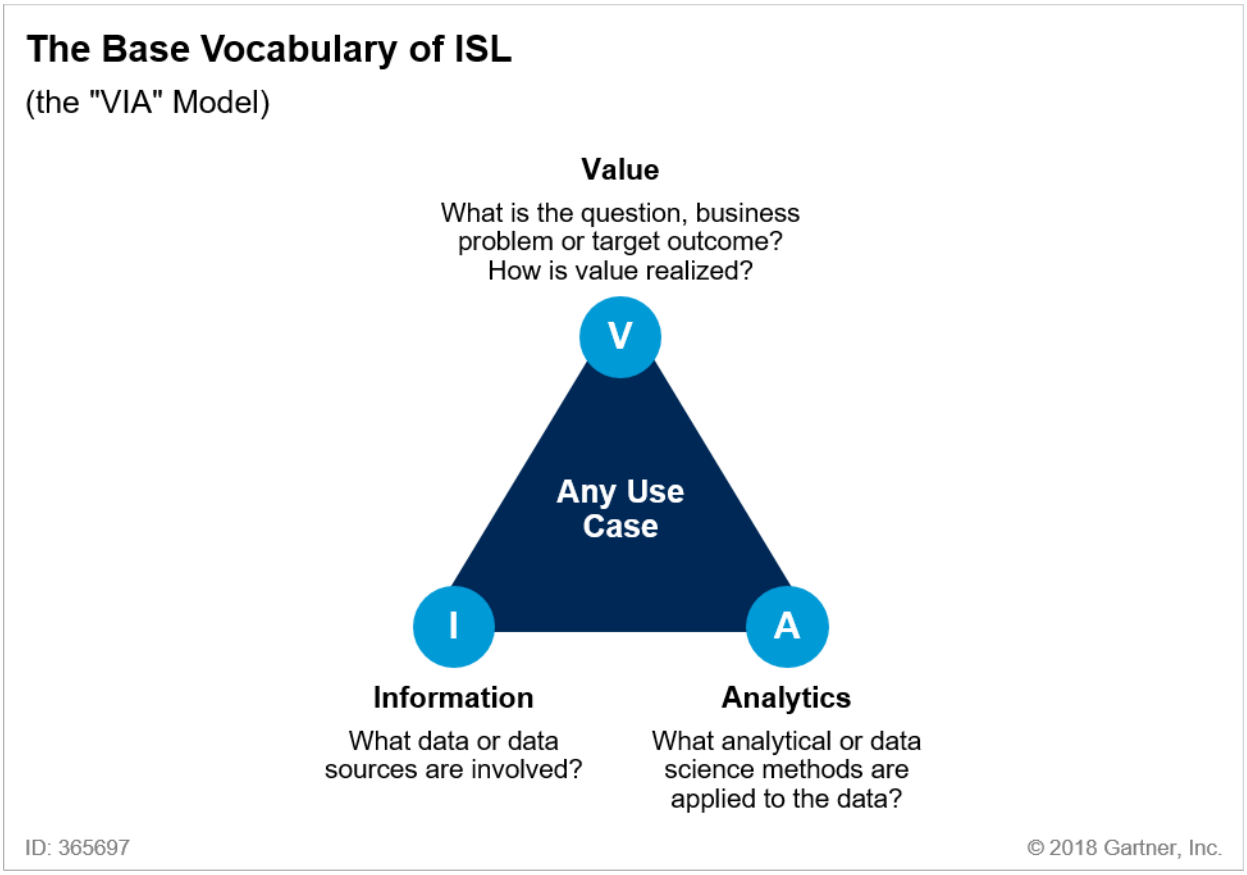
Cultivate ISL Across Business and IT Stakeholders

The foundation of any language is a base vocabulary. According to Wikipedia (Note 2), we begin with simple, generally accepted uses of the terms “data” and “information”:

- **Data** is a set of values of qualitative or quantitative variables
- **Information** is that which informs, and can be encoded into various forms for transmission and interpretation

In the case of information as a language (Note 3), the base vocabulary is composed of three elements: **V**alue, **I**nformation and **A**nalytics — the VIA model (see Figure 4).

Figure 4. The Base Vocabulary of ISL — The VIA Model



ISL = information as a second language

Source: Gartner (September 2018)

1. **Value:**
   - Business outcomes, questions, decisions, actions, metrics.
   - Integrated information and analytics applied in terms of a decision in the context of a business moment.
   - Integrated strategies, governance, use cases, business process integration and contextualized applications, approaches, change management and organizational constructs.
2. **Information:**
   - Data sources, quality, data types and management methods.
   - Related strategies, governance, technologies and organizational constructs.
   - Spans the information life cycle, from creation to archive/deletion.

3. **Analytics:**
    - Business intelligence, reporting, analytical methods, artificial intelligence, machine learning.
    - Related strategies, methods, governance, technologies and organizational constructs.

In addition to a base vocabulary, with any language, dialects form specific to a particular social group or region. In the case of ISL, natural dialects emerge aligned with:

- **Industry vertical domain dialects** (e.g., speaking healthcare data)
- **Business process horizontal domain dialects** (e.g., speaking marketing data)
- **Technical domain dialects** (e.g., conceptual, logical and physical architectures; specific programming/coding languages; mathematical and analytical techniques; data management vs. content management disciplines)

(See “Harnessing the Pervasive Nature of Domain Data and Analytics.”)

In developing data literacy, five commonly accepted levels of proficiency apply, summarized in Table 1.

Table 1. Data Literacy Levels of Proficiency

Level	Definition	Example
Conversational	Basic understanding of the concepts of data, analytics and use cases; one who "gets it" but cannot explain it to others.	A professional who has a basic understanding of an analytics value proposition and the ingredients involved.
Literacy	Ability to speak, write and engage in data and analytics programs and use cases.	A professional who can explain all aspects of an analytics use case, including the industry problem, business process moment/decision affected, data sources leveraged, and analytical methods applied.
Competency	Competent in designing, developing and applying data and analytics programs.	An experienced data and analytics program manager who has designed and delivered analytical projects from concept through outcome.
Fluency	Fluent in all three elements of the information language (VIA) across most business domains within an industry vertical.	In utilities, a smart meter registers kilowatt (kW) usage. Over time, it creates kWh averages and peak demand. This is interpreted by a billing department associate far differently than a distribution planning manager, given the nature of their roles, and the ways in which they use the data and analytics in context. A fluent data speaker in the utilities industry would be able to explain all of these use cases.
Multilingual	Fluency across all three elements of the information language (VIA) across multiple business domains, industries and ecosystems.	An experienced data and analytics strategy consultant who has designed and delivered analytical solutions across multiple industries and business domains, and can explain them to non-native speakers.
VIA = value, information, analytics		

Source: Source: Gartner (September 2018)

Drive and Sustain Improvements to Your Organization's Data Literacy

Identify areas in your business where data is spoken fluently, where language gaps exist, and establish an ISL plan and initial proof of concept (POC) for language development.

Do You Speak Data?

To gauge your organization's degree of data literacy, start by identifying where data is spoken fluently:

- Identify **fluent and native speakers** who speak data naturally and effortlessly. Start with business analysts, subject matter experts, data stewards and architects within your existing business intelligence competency centers, analytics centers of excellence and end-user organizations. Native speakers often originate from consulting backgrounds or information services companies. Fluent speakers should be adept at describing contextualized use cases

and outcomes, the analytical techniques applied to them, and the underlying data sources, entities and key attributes involved.

- Identify **skilled translators** — those who have, historically, served as the mediators across data and analytics programs, and authors of related business cases. Classic translators are often enterprise data or information architects, data scientists, information stewards and related program managers.

What We Have Is a Failure to Communicate

Identifying where language barriers are inhibiting progress is equally important to identifying where data is spoken fluently:

- Identify areas where **communication barriers** are inhibiting the effectiveness of data and analytics initiatives, paying particular attention to business-IT gaps, data/analytics gaps and veteran/rookie gaps.
- Actively listen for **business outcomes not clearly articulated** in terms of explicit action. What business moments are being enabled with enhanced data and analytics capabilities? What operational decisions are being improved? What processes are being improved with algorithms? (See "Data and Analytics Strategies Need More Concrete Metrics of Success," "Toolkit: Use Business Moments to Identify Hidden Value Opportunities for Your Enterprise" and "Toolkit: Analytics Business Opportunities From Almost 200 Use Cases.")
- Identify **key stakeholders** who require specialized translations to better sponsor, understand and support initiatives, with specific attention to C-level executives who must speak data and analytics at a conversational level to lead by example. To assess data literacy levels, ask key stakeholders to articulate the value of data as a strategic asset in terms of business outcomes, including enhanced business moments, monetization and risk mitigation.
- Identify and maintain a **list of words and phrases** used throughout the organization that are either too vague (business speak like "better decision making" or "the right info for the right people at the right time") or too precise (tech speak like "modernizing our data infrastructure for exogenous data sources"). These lists can then be used to engage the data and analytics team in crafting ways to better articulate them. Examples might include business speak like "better decision making" or "the right info for the right people at the right time" (vague), or tech speak like "modernizing our data infrastructure for exogenous data sources" (confusingly precise). Leverage key data and analytics program artifacts like a data catalog, business glossary and data dictionary as foundational resources and references. Another technique is to create/collect a set of "use this, not that" or "this is in, while this is out" pairings, as a type of game to amplify language that is relevant, and what should not be used. This type of collective gaming exercise could result in an internal guidance document for internal distribution, and reinforcing common phrases.

Set Up an ISL Proof of Concept

As you surface areas where language gaps are inhibiting progress, conduct an initial ISL proof of concept (POC) as a basis for language training and development:

- **Select an area** of the business to establish an ISL POC as a demonstration of the need and opportunity for enhanced communication and a shared language. Pick a "friendly" area where clear gaps have surfaced, and in which you have willing, diverse participants. Have each member self-assess their "data fluency" level.
- **Conduct an ISL workshop** centered on an existing analytical use case, and ask members to articulate the use case from their own point of view. Capture the similarities where language is shared, and where differences exist. After the workshop, have each member reassess their "data fluency" level. Capture lessons learned.
- **Raise awareness and understanding of the data literacy gap by broadcasting the story,** calling attention to the problem, opportunity and outcomes of the POC.

ISL Training and Development Programs

- Data and analytics leaders should **plan for and deliver deliberate competency and language development programs.** These include intentional cross-training of priority roles (business analysts, data scientists, data engineers and sponsors) with direct input from academic institutions, research organizations and consulting organizations.
- **Skilled translators and multilingual instructors will begin to emerge,** similar to early pilots and adopters of second language development programs (see Note 4), or Six-Sigma-type black belt training programs, as a sign of the deliberate and wide adoption of the new language of digital society.
- **System integrators, consulting providers and technology providers must develop related programs** for their professionals to raise their fluency in speaking data.

Change the Way You and Others Interact With Leaders, Stakeholders and Peers by Speaking Data

"Be the change you want to see in the world" — Mahatma Gandhi.

Model the language you want to encourage and nurture across your organization. Speak data "in context" in everyday interactions, with board members and team members, to set the tone for the new mode of communication. With each discussion, call attention to the three key aspects of the language: the business outcome (first and foremost), the data elements involved, and the analytical techniques that support the business outcome. Talk business — not just data or analytics.

In addition to conducting a POC, initiate cross-team training for information language development. Consider fun ways to advance awareness and learning (e.g., techno-speak Bingo). Challenge your current data and analytics team members to teach their "dialect" of the language to other members of the team to ensure everyone receives a comprehensive view of the use case, analytics and data. After this cross-training, test each member of the team by insisting they teach/translate to someone else outside of the group.

In addition to information language development, data and analytics leaders should use visualization and a maturing set of communication techniques — storytelling, customer journey maps, glossaries and infographics, for example — to convey the business impact of applying data and analytics to the business moments that matter most.

Additional research contribution and review was provided by Mark Beyer, Erick Brethenoux, Frank Buytendijk, Matt Cain, Lydia Clougherty-Jones, Melissa Davis, Alan Dayley, Alan D. Duncan, Cindi Howson, Carlie Idoine, Saul Judah, Doug Laney, Deb Logan, Michael Moran, Thomas Oestrich, Jamie Popkin, Mike Rollings, Rita Sallam, Kurt Schlegel, Svetlana Sicular, Andrew White.

Acronym Key and Glossary Terms

ISL	information as a second language (alternate acronym: I2L)

Gartner Recommended Reading

Some documents may not be available as part of your current Gartner subscription.

“Fostering Data Literacy and Information as a Second Language: A Gartner Trend Insight Report”

“Leadership Vision for 2019: Data and Analytics Leader”

“Survey Analysis: Third Gartner CDO Survey — How Chief Data Officers Are Driving Business Impact”

“Break Through the Four Barriers Blocking Your Full Data and Analytics Potential — Keynote Insights”

“Harnessing the Pervasive Nature of Domain Data and Analytics”

“100 Data and Analytics Predictions Through 2022”

“How to Establish a Data-Driven Culture in the Digital Workplace”

“Toolkit: Use Business Moments to Identify Hidden Value Opportunities for Your Enterprise”

“Toolkit: Analytics Business Opportunities From Almost 200 Use Cases”

“Applied Infonomics: Why and How to Measure the Value of Your Information Assets”

Evidence

This research is based on hundreds of client inquiries and interactions with many analysts across all aspects of data literacy, and related data and analytics topics, across industries and business domains.

The data for this report comes from Gartner’s third annual Chief Data Officer Survey, conducted during July, August and September 2017 by phone and online. The survey included 287 CDOs,

chief analytics officers and other high-level data and analytics leaders from across the world. The purpose of the survey was to test a set of five hypotheses about the CDO role and the office of the CDO, in order to understand how this rapidly growing business function is maturing and the resulting business impact.

Note 1 About Literacy (Data Literacy)

Data literacy is the ability to read, write and communicate data in context, including an understanding of data sources and constructs, analytical methods and techniques applied, and the ability to describe the use case, the application and resulting value.

Information literacy is the ability to know when there is a need for information, to be able to identify, locate, evaluate, and effectively use that information for the issue or problem at hand. (Source: The United States National Forum on Information Literacy.)

Note 2 About Data (Data, Information, Knowledge, Wisdom)

Excerpts from Wikipedia include the following simple, generally accepted definitions:

- **Data —** A set of values of qualitative or quantitative variables.
- **Information —** That which informs, and can be encoded into various forms for transmission and interpretation.
- **Knowledge —** A familiarity, awareness or understanding of someone or something, such as facts, information, descriptions or skills, which is acquired through experience or education by perceiving, discovering or learning.
- **Wisdom —** The ability to think and act using knowledge, experience, understanding, common sense and insight.

Note 3 About Language (Information as a Language)

Language is a system of communication. The study of language is linguistics.

Note 4 About Second Languages

A person's second language is a language that is not the native language of the speaker, but that is used in the locale of that person. In contrast, a foreign language is a language that is learned in an area where that language is not generally spoken. (Source: Wikipedia.)

GARTNER HEADQUARTERS

Corporate Headquarters
56 Top Gallant Road
Stamford, CT 06902-7700
USA
+1 203 964 0096

Regional Headquarters
AUSTRALIA
BRAZIL
JAPAN
UNITED KINGDOM

For a complete list of worldwide locations,
visit http://www.gartner.com/technology/about.jsp