Module: Data Governance, Compliance and Ethics

WEEKS 1-3: DATA GOVERNANCE AND MANAGEMENT (OVERVIEW)

PRESENTATION TO "TRAIN THE TRAINER" VIENNA,

MAY 31ST, 2022

john.bohan@ncirl.ie

Good Morning from Dublin

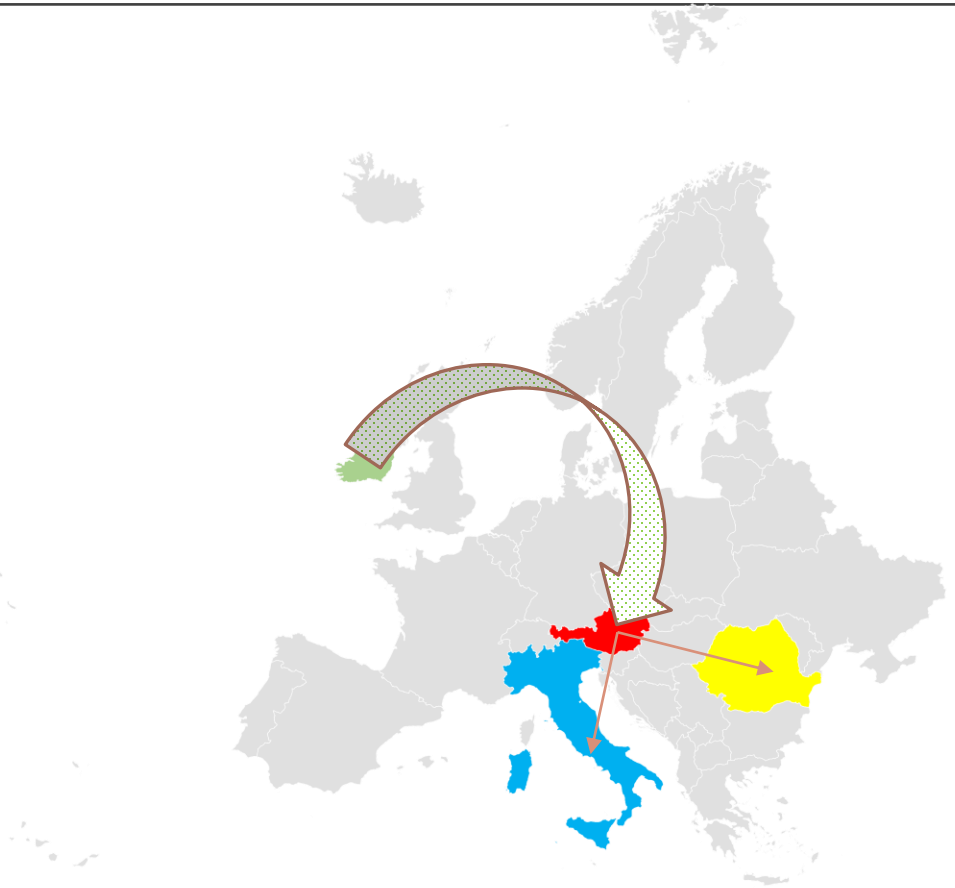
Good Morning to you all

Maidin mhaith agaibh go léir

Guten Morgen an alle

Buongiorno a tutti

Bună dimineața tuturor



About me

John Bohan (john.bohan@ncirl.ie)

Associate Faculty, NCI

- Data Governance and Ethics
- Statistics, Data Visualisation, Research project

Academic background

- Institute Public Administration, Dublin (BA Public Administration)
- Trinity College Dublin (M.Sc. European Studies, Higher Diploma in Statistics, MSc in Public Sector Analysis)
- KU Leuven, Belgium (European Master in Social Security)
- National College of Ireland (MSc in Data Analytics)

Career

- Former senior manager in Department of Social Protection, Ireland
- Member of EU Social Policy Committee and Chair of Indicators Sub-Group
- Seconded to European Commission, Irish Aid Programme (Lesotho)

Aims of presentation

Provide overview of DGM section of Data Governance and Ethics module (normally 6 contact hours)

Introduce DGM Concepts and Definitions

Provide DGM Rationale and Goals

Show strategic overview of DGM (who, how, what etc)

Illustrate DGM building blocks

- Roles and responsibilities
- Operating Model

Explain implementation method (using Ladley model)

2.1.1 Module aims and objectives

This module aims to provide learners with the knowledge and skills around the complex issues of data management and governance in an organisational context, including ethical and compliance issues that these present. Learners will explore the ethical, legal, and social implications of using data-driven technologies such as big data, analytics, internet of things, and machine learning. The students will learn how to establish processes and systems that consider best practices for data governance and adhere to ethical and regulatory requirements for data handling.

2.1.2 Minimum intended module learning outcomes

LO1 Demonstrate critical understanding of the governance and regulatory frameworks associated with the key data lifecycle stages for an effective management of data assets.

LO2 Demonstrate critical awareness and interpretation of the data privacy and data protection regulatory landscape in socio-technical environments.

LO3 Critically analyse and evaluate the main ethical, legal, and social implications of using data-driven technologies.

LO4 Investigate and appraise the interplay of fairness, accountability, and transparency in algorithmic decision-making systems and demonstrate awareness of technical solutions to enhance these concerns.

Sources and resources

DAMA International (2017). *DAMA-DMBOK: Data Management Body of Knowledge* (2nd ed.). Technics Publications. [ISBN:978-1634622349].

Ladley, J. (2019). *Data governance: How to design, deploy, and sustain an effective data governance program*. Academic Press.

Smallwood, R. F. (2019). *Information Governance: Concepts, Strategies and Best Practices*. Wiley. [ISBN: 978-1119491446].

Overview of DGM section

Topic	Lecture Topic	Lecture Detail
1	Data Governance and Data Management I	Data governance and data management; Data management principles; Data lifecycle; Data quality; Data provenance; Data integrity and security
2	Data Governance and Data Management II	Data governance frameworks; Data governance within the DAMA Wheel; Policies, principles, rules, procedures, and standards; Data governance operating models and tools
3	Data Governance and Data Management III	Roles and responsibilities; Maturity levels; Ladley's 8-phase implementation process for data governance; Data risk identification and management

What do we mean by corporate governance?

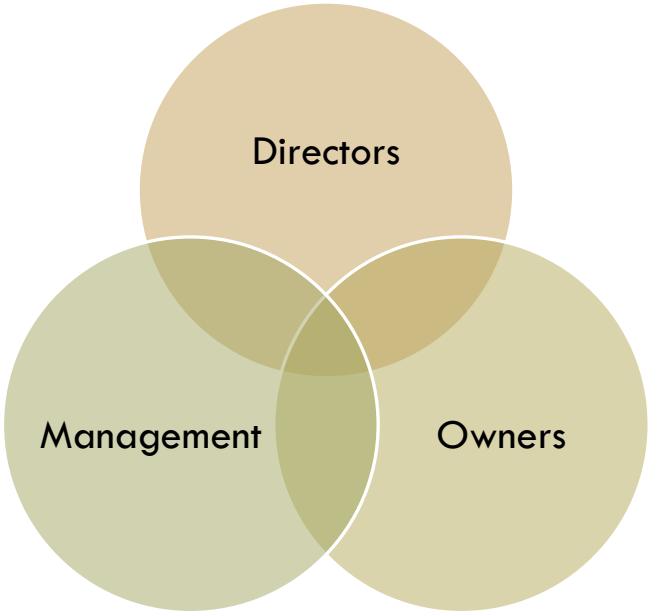
Corporate governance is the structure of rules, practices, and processes used to direct and manage a company.

A company's board of directors is the primary force influencing corporate governance.

The basic principles of corporate governance are accountability, transparency, fairness, and responsibility.

Corporate governance goes beyond day-to-day management by considering for instance ethical behavior, corporate strategy and risk management.

Bad corporate governance can cast doubt on a company's operations and its ultimate profitability.

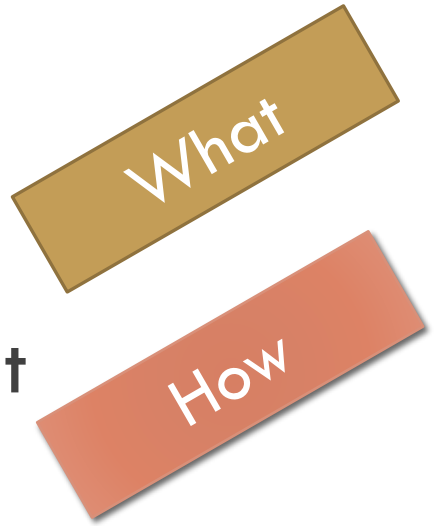


[Corporate Governance Definition: How It Works (investopedia.com)]

Data Governance: What?

Data governance (DG) is the process of

- managing the availability, usability, integrity and security of the data in enterprise systems,
- based on internal data standards and policies that also control data usage.



Source: https://www.techtarget.com/

Data never sleeps”

The amount of data in the world is estimated to double every 18-24 months

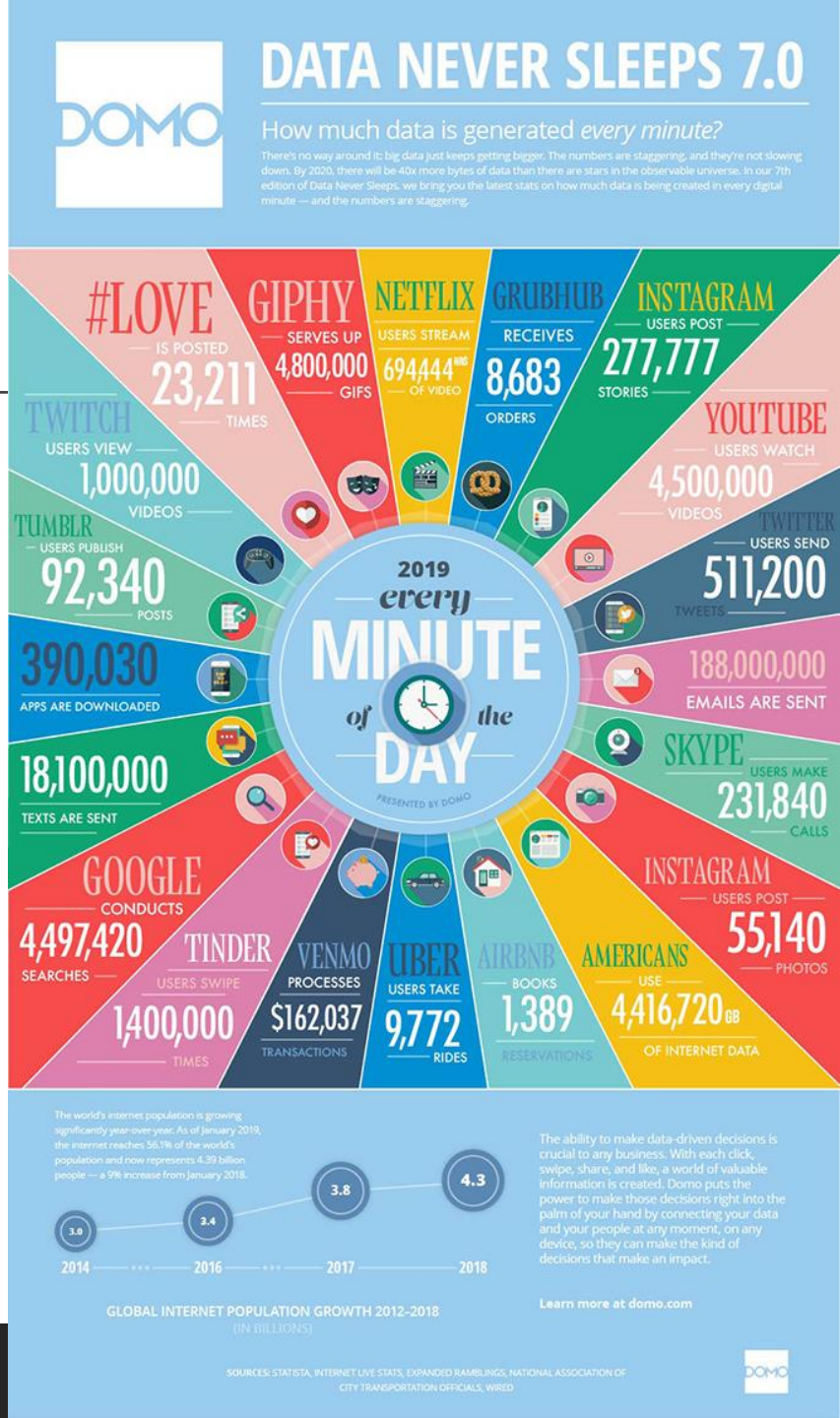
How much of it is used (or useful?)

In 2019, every internet minute saw

- 4.5 million Google searches
- 55,000 Instagram posts
- 694,000 Netflix streams
- 4.5 million YouTube views

Enormous increase in corporate data

Source: https://www.domo.com/learn/data-never-sleeps-7



Data growth and the need for corporate governance

More data presents opportunities for organisations to improve…

- Business intelligence
  - More efficient and better-quality control in manufacturing and service delivery
  - More effective provision of supports within organizations such as planning, IT etc
- Customer relationship management
- Strategic and business planning

… but also to higher risks from …

- Data losses and data breaches leading to loss of reputation and financial losses including fines
- Threats of hacking, ransomware, system failure
- Existential threats to business models from competitors

Example of poor corporate and research data governance

Former Cambridge Analytica chief receives seven-year directorship ban

Alexander Nix handed penalty for 'potentially unethical' behaviour linked to scandal

▲ Cambridge Analytica's former chief executive Alexander Nix has been barred from acting as a company director for seven years. Photograph: Tolga Akmen/AFP/Getty Images

Alexander Nix, the former boss of Cambridge Analytica, has been banned from serving as a company director for seven years over "potentially unethical" behaviour linked to his position at the centre of a global scandal.

https://www.theguardian.com/uk-news/2020/sep/24/cambridge-analytica-directorship-ban-alexander-nix?CMP=Share_AndroidApp_Other

https://www.penguinrandomhouse.com/books/604375/mindfck-by-christopher-wylie/

Some examples of goals that a DG programme might seek to achieve

Improve strategic direction of organisation

- Get better value out of customer and client information

More effective operational processes

- Reduce costs by removing duplicated effort in data collection and maintenance
- All for better data analysis and better decisions

Reduce organisational friction

- Set clear data responsibilities between different business and support areas
- Protect the needs of various data stakeholders within and without organisation

Ensure organisational accountability and compliance

- Improve organisational capacity to address data issues through education and trainings
- Allow for transparency by adoption of standard, repeatable data processes

Data Governance: How?

Effective data governance operates through policies, standards and tools that seek to

- Facilitate and improve the internal and external **communication** about data
- Provide for the identification and minimization of data-related **risks**
- Establish and control internal **rules** for data use
- Ensure the implementation of industry, national or international compliance **requirements**

Data Governance vs Data Management

Data Management is the process of gathering, storing, organizing and maintaining data created and collected by an organization.

Data Governance determines the overall practice of data management. It is regarded as the "management of data management" and envisages a separation between those governing and those managing

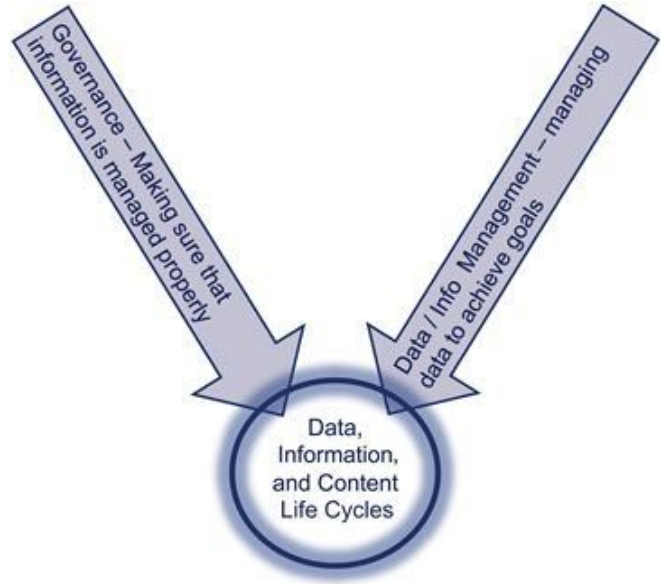
"Doing things right" vs "Doing the right thing"

Even if an organisation has no specific Data Government practices, things still get done. But are the right things being done?

The issue is whether a clear structure and roles need to defined through agreed policies, operational processes and procedure

Why the importance of documenting what these are?

- Induction and training of staff
- Accountability and audit
- External consumption (media, regulators)

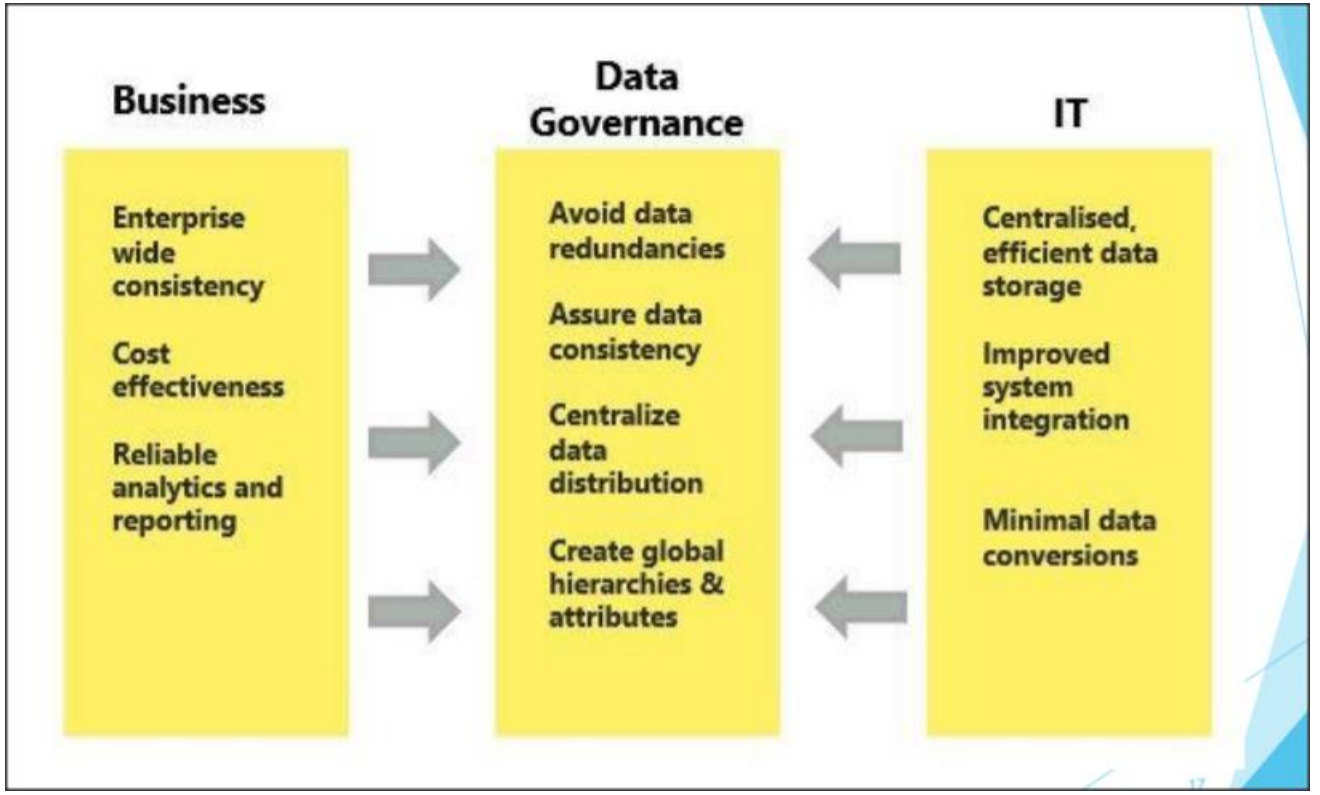


separation between those governing and those managing (Illustration from Ladley)

Data Governance as partnership

DG requires a partnership between business and IT to

- Provide an effective and consistent means of distributing data to those who need it for business purposes
- Calibrate and empower access to appropriate use of data
- Providing quality assurance about data
- Avoid retaining or gathering redundant data



Data Governance should be business led with IT input not the other way around

Data Governance should be business led with IT input not the other way around

Like other management processes, it must balance strategic and operational needs.

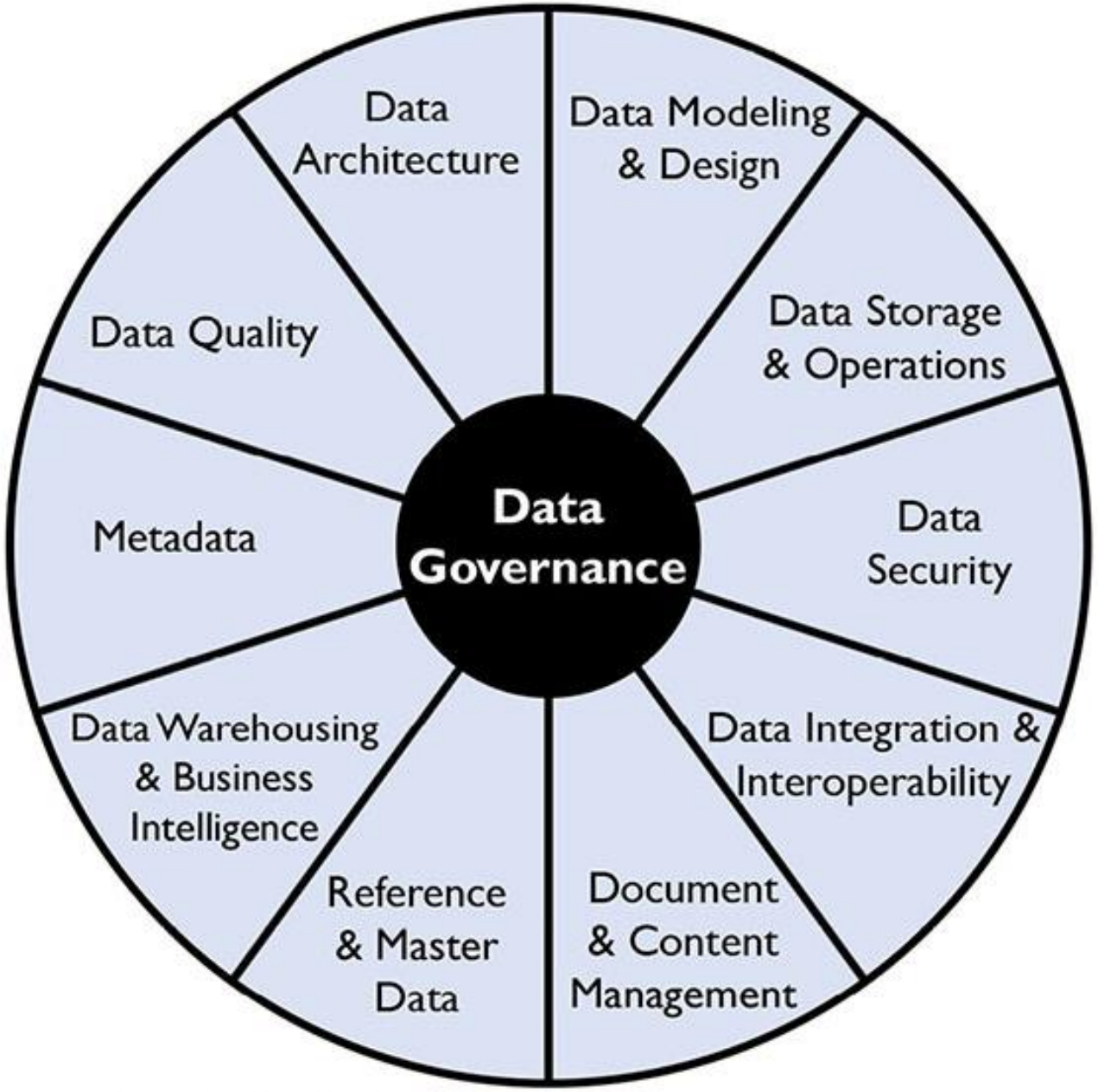
This balance can best be struck by following a set of principles that recognize salient features of data management and guide data management practice.

An organisations management should start with identifying its data principles that might include

- Effective data management requires **leadership commitment**
- Data is an **asset** that needs to be planned and accounted for like other assets
- Managing data includes **managing the risks** associated with data not just the opportunities
- Data management requires an **organisation-wide perspective** that must drive Information Technology decisions, not the other way around

Data management knowledge areas

- Data governance enables an organization to determine proper action regarding data and data processing, clear definitions, decision making rights and responsibilities and then escalation path or mediation process where issues arise.
- The Data Governance function guides all other data management functions.
- DG includes the **people, processes and technologies** required to ensure good quality data, its effective management and overall governance
- These can be grouped in a number of knowledge areas
- Should reflect the organisational circumstance; not a case of "one size fits all".



DAMA-DMBOK2 Data Management Framework

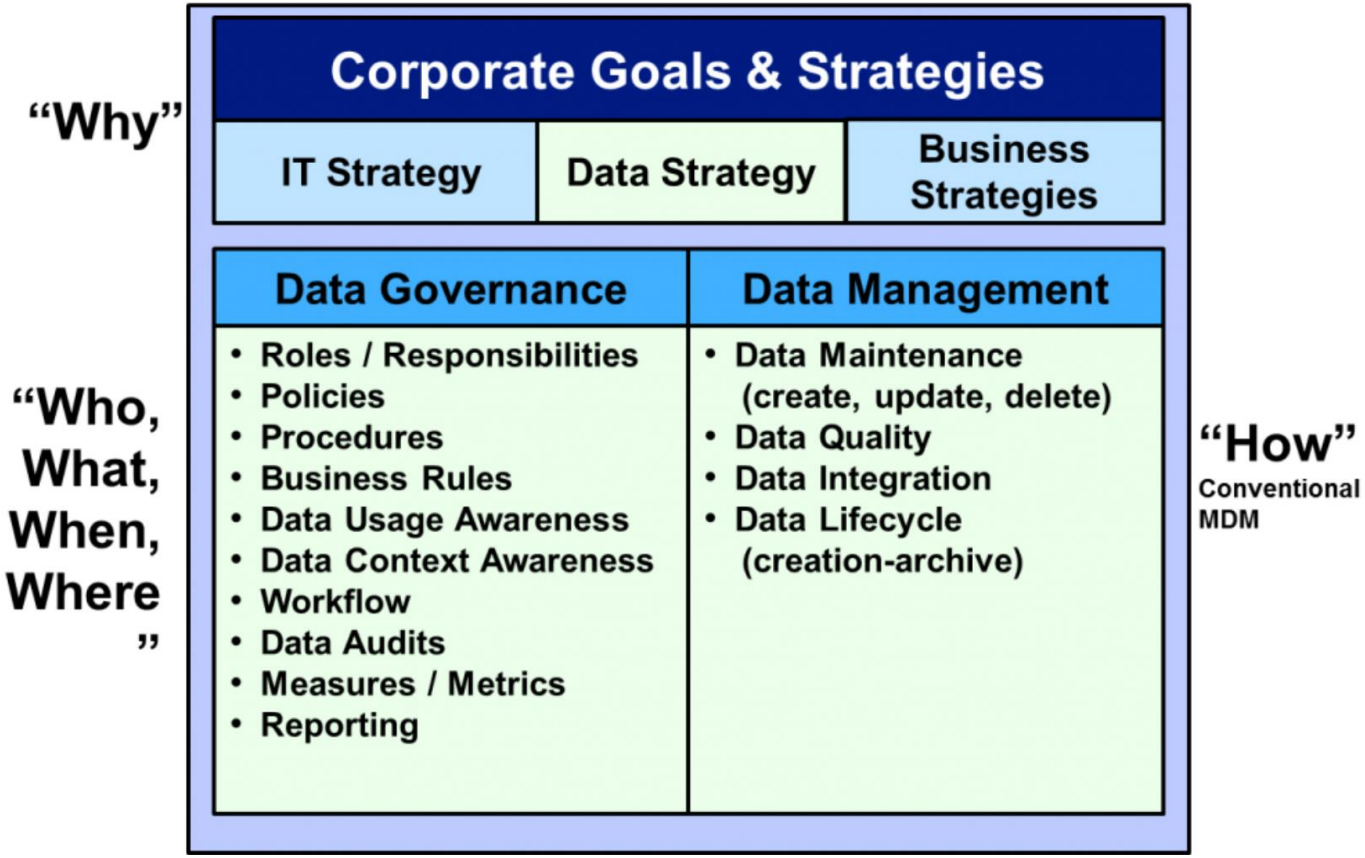
DG Building Blocks

Source: https://theodi.org/wp-content/uploads/2021/06/AdobeStock_410183096-scaled.jpeg

Data Strategy

A strategy is "a set of choices and decisions that together chart a high-level course of action to achieve high-level goals. … A strategic plan is a high-level course of action to achieve high-level goals." (DMBOK)

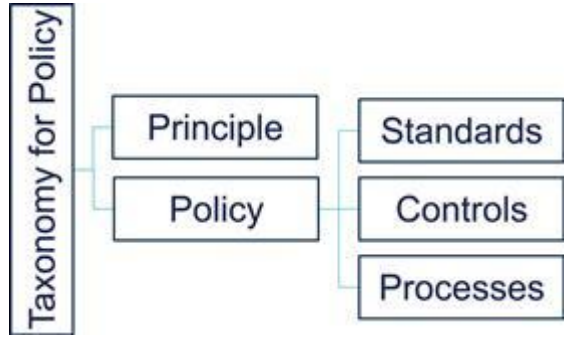
"Data strategy must come from an understanding of the data needs inherent in the business strategy: what data the organization needs, how it will get the data, how it will manage it and ensure its reliability over time, and how it will utilize it."



Data principles

The DMBOK defines (organisational) data principles as:

- A fundamental law, doctrine, premise, or assumption
- A rule or code of conduct



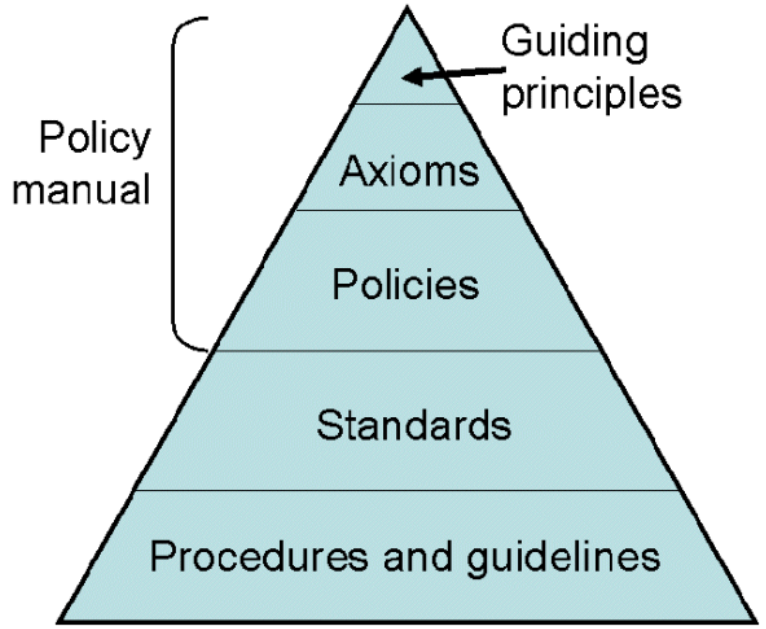
Principles are statements of philosophy (or core beliefs) around policies and behaviours around data governance to be used as guidance for procedures and decision-making efforts (Ladley)

While principles are to some extent generic (e.g. “data should be treated as an asset”), they must be adapted to the circumstances of an organisation.

They should be combined with policies to produce specific standards, controls and processes that make up DG&M operations.

What is a data management policy?

A **data management policy** addresses the operating **policy** that focuses on the **management** and **governance** of **data** assets and is a cornerstone of governing enterprise **data** assets.



Data Standards

- A standard is defined as “something set up and established by authority as a rule for the measure of quantity, weight, extent, value, or quality.”
- By adopting a standard, an organization makes a decision once and codifies it in a set of assertions (the standard). It does not need to make the same decision all over again for each project.
- DG standards should be mandatory. Enforcing standards should promote consistent results from the processes using them.
- Drafted by data management professionals but reviewed, approved and adopted by the Data Governance Council.
- Data standards must be effectively communicated, monitored, and periodically reviewed and updated.
- Examples of data standards might include:
  - Data Architecture: Enterprise data models, tool standards, and system naming conventions
  - Acceptable use policies
  - Data Quality standards (timeliness, accuracy)
  - Data security standards
  - Use of data for data analytics and presentation purposes

Text source: DMBOK (2017)

Data Procedures

Data management procedures assign responsibility for custody and security of all company data, allowing the company to control data consistently across various departments.

Can take several forms depending on data type and purpose (regulations, guidelines, constraints, documentation, access, disposal, standards)

DG Operational Model

Source: https://theodi.org/wp-content/uploads/2021/06/AdobeStock_410183096-scaled.jpeg

DG operating model

So far, we've considered a range of artefacts that might be developed as part of a DG programme such as data strategy, principles, policies and standards.

But how are we going to implement a DG/DM system in practice?

We need to consider an **operating model** or framework that sets out

- Defined roles and responsibilities,
- Clear decision-making processes.

In short, we need to describe how people and functions will collaborate to deliver the aims and principles of the DG programme.

Remember, "one size doesn't fit all" and while the DG have common goals, their operating models will vary according to an organisation's functions, its geographical scope, its culture, its organisational structure, the "maturity" of its data systems, its motivation for change etc. etc.

Importance of assigning roles and responsibilities to DG

"Like any other activity within a company or government entity, there needs to be a **formal statement of roles**. The official designation of accountability and responsibility are key factors to the survival of DG. Most important to new DG programs is the concept of **accountability for data**. This is most likely a very new role. … In many organizations, the responsible parties have a formal role as designated "stewards" or "custodians."

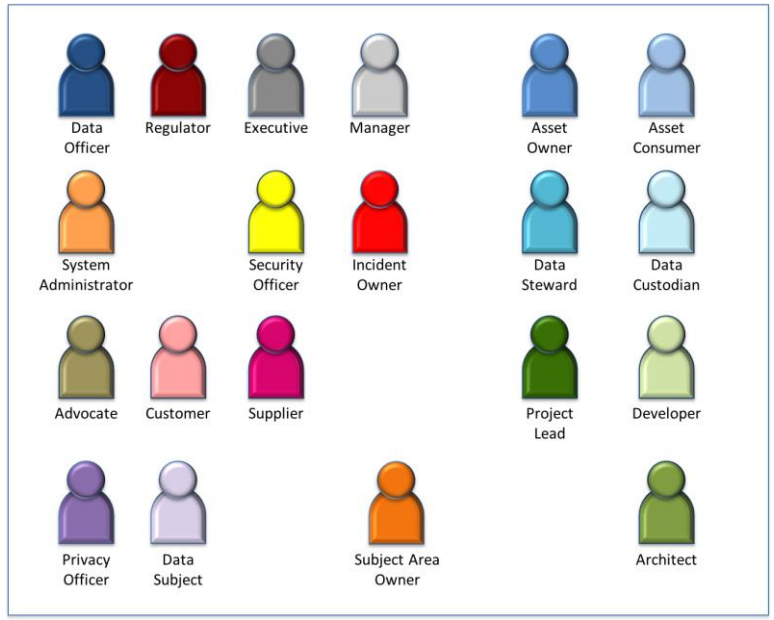
The organization around DG also requires a **hierarchy** of some sort to enable issue resolution, monitoring, and direction setting. Rarely does this hierarchy of DG become a stand-alone area (i.e., there is rarely a data governance "department"). Most of the time, the DG organization is a **virtual organization** made up of business and IT personnel." *(emphasis added)*

Chapter 3. Overview of a data governance program | Data Governance (oreilly.com)

Wide variety of Roles and responsibilities around data

An organisation's data roles will vary depending of many factors including size of organisation

- The smaller the organisation, the more likely it is that roles will multi-task.
- In a larger organisation, you might find a range of roles that are business, IT or data management specific:
  - Chief Data Officer – covered in following sides
  - Data Protection/Privacy Officer – to be covered in Privacy and Data Protection
  - Data Steward – covered in following sides
  - Data Architect - designs technical plans for a construction of data
  - Data Policy Officer - implements new regulations or processes that help improve efficiency or ensure an organization remains compliant

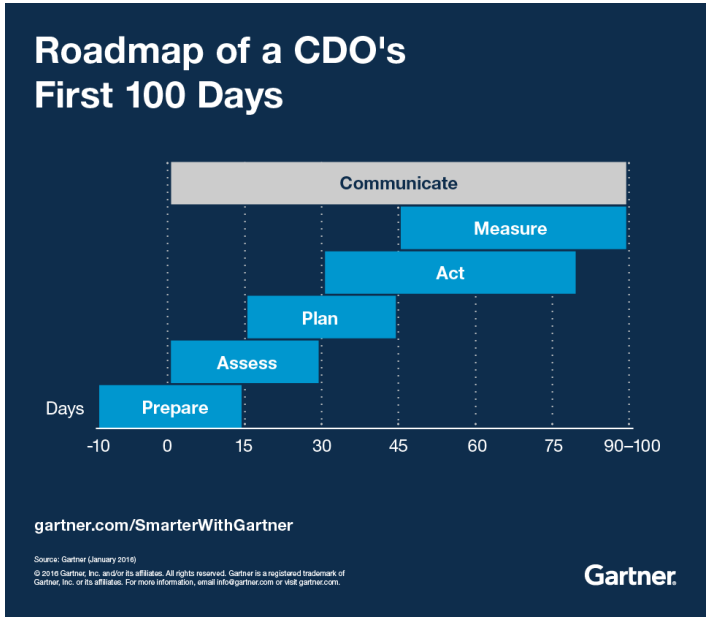


Chief Data Officer

Not all (or indeed many!) organisations have a CDO but industry analysis suggests increasing interest in the role especially in regulated sectors

Research shows a range of functions for CDO

- Establishing an organizational data strategy
- Aligning data-centric requirements with available IT and business resources
- Establishing data governance standards, policies and procedures
- Providing advice (and perhaps services) to the business for data-dependent initiatives, such as business analytics, Big Data, data quality, and data technologies
- Evangelizing the importance of good information management principles to internal and external business stakeholders
- Oversight of data usage in analytics and Business Intelligence



In effect, CDO (and their division) provides an interlocutor role between the data infrastructure functions (typically IT) and data consumers (typically business users)

Source: DATAVERSITY Launches Research Paper, "Status of the Chief Data Officer: An Update on the CDO Role in Organizations Today" - DATAVERSITY

Data Stewardship

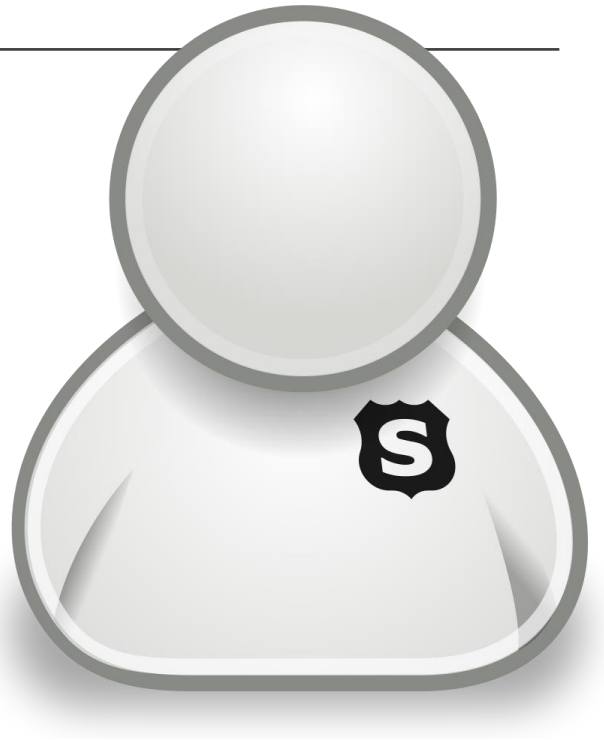
A steward is some who manages property on behalf of others.

Data Stewards therefore represent the interests of all stakeholders to ensure enterprise data is of high quality and can be used effectively.

Effective Data Stewards are accountable and responsible for data governance activities and have a portion of their time dedicated to these activities.

Data steward roles activities will change depending on organisational circumstances but may include:

- **Creating and managing core metadata** such as business terminology, data dictionary.
- **Documenting rules and standards** in order to have organisational consensus and consistent use.
- **Managing data quality issues** such as their identification, analysis and resolution
- **Executing operational data governance activities** such as participating in data management projects, data governance monitoring



Text Source: DMBOK (2017) 1.3.4

Maturity levels of Data Governance

Reactive Governance Model (driven by issues arising)

- Tactical effort to respond to a problem arising such as a data leak or disaster

Preemptive Governance Model (driven by upcoming planned event)

- Organisation facing major change (merger, legislative change)
- Approach tries to blunt issues

Proactive governance (driven by organisational capability and structure)

- Driven by an information management strategy
- Continual Focus on improving capabilities (processes, technologies) to address and resolve anticipated risks and issues
- Alignment of governance and management process formalising a culture of issue and risk awareness

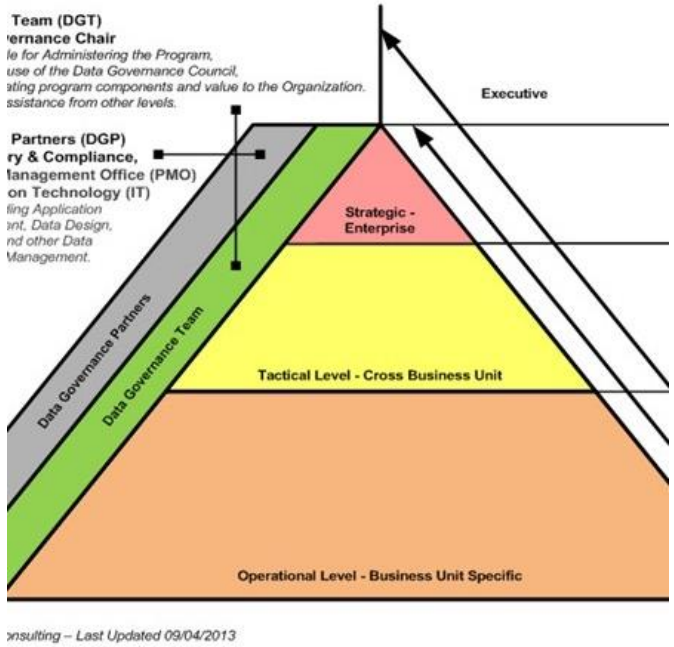
Source: Adler

Decision making processes

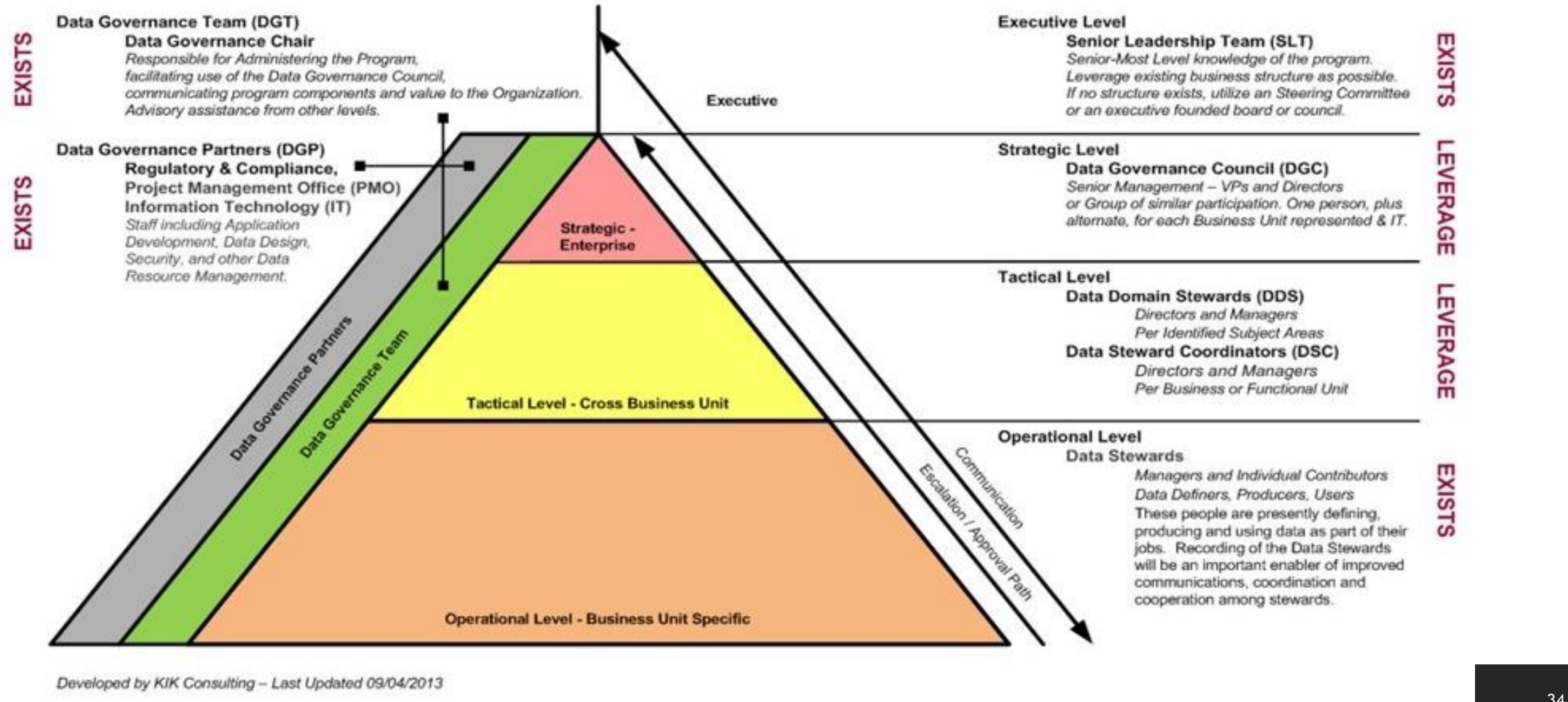
Recall that the operating model is defined by the roles and responsibilities and the processes that determine decisions

We will go into this in more detail in the lab using the operating model shown but for now it's enough to say that processes are related to

- Nature of decisions required (strategic, tactical or operational level)
- Decentralised to business areas of centralized to the Data Governance Team
- Business or technical in nature

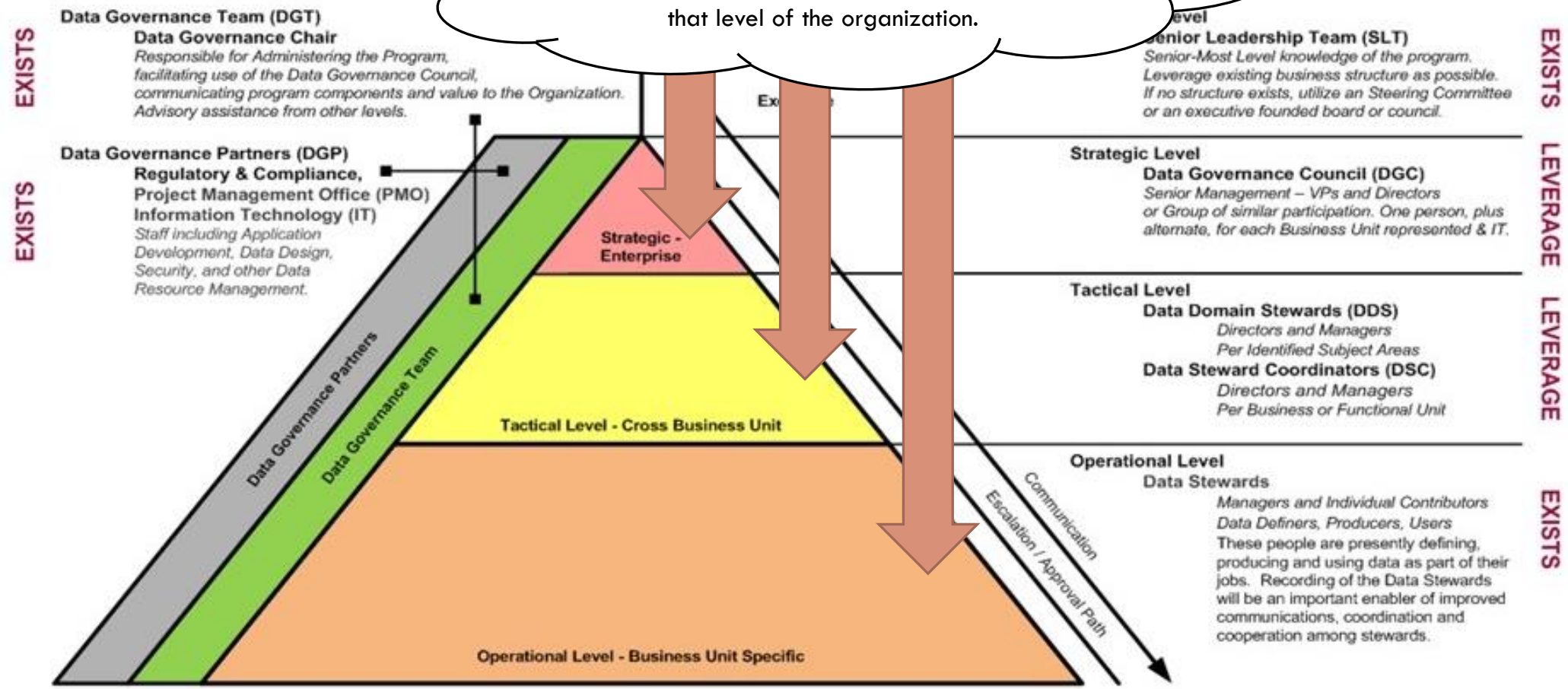


Data Governance Operating Model of Roles and Responsibilities



Data Governance Operating Model of Roles a

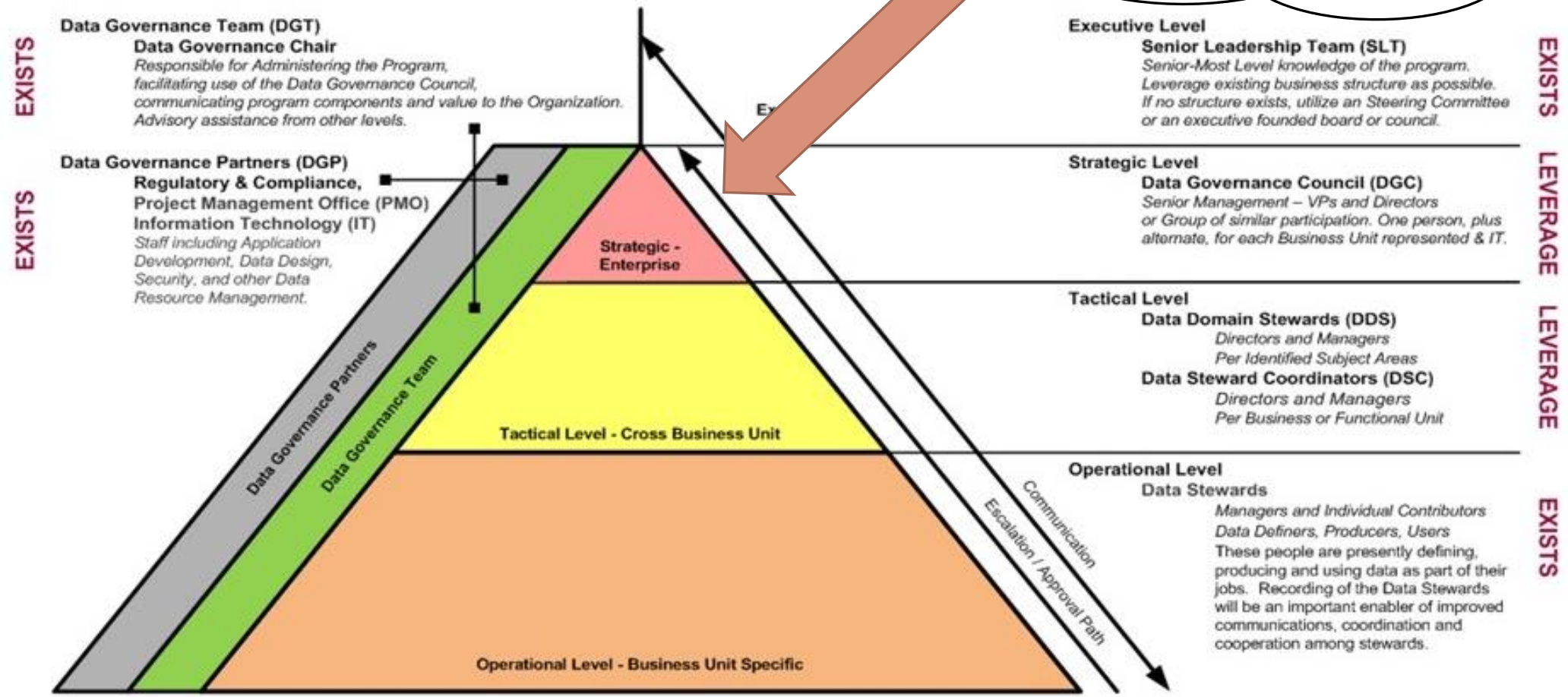
The space represents the rough percentage of instances that organizations expect decisions to be made about the data; decisions should be made at the operational level if the decisions only affect that level of the organization.



Developed by KIK Consulting – Last Updated 09/04/2013

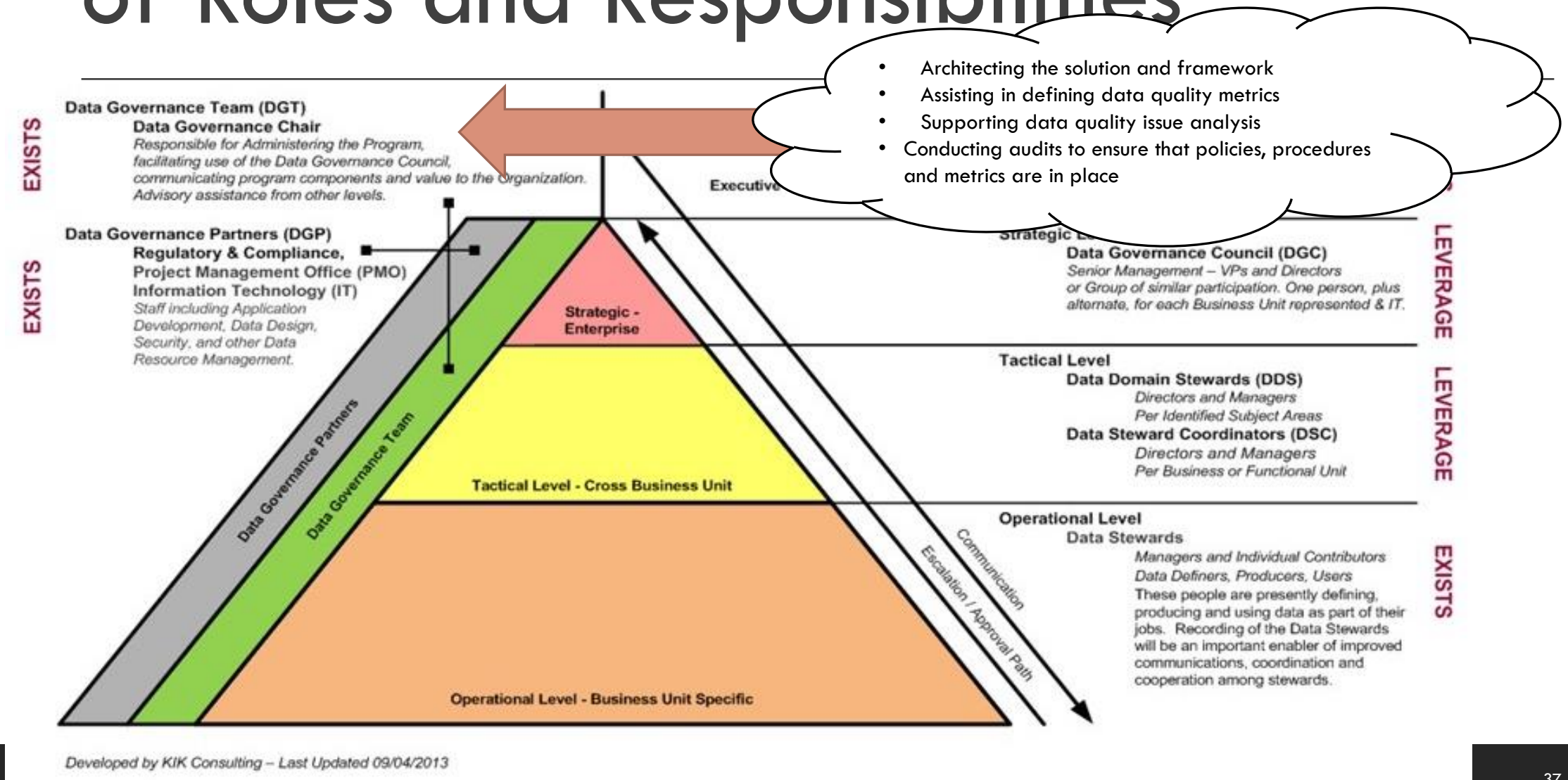
Data Governance Operating Model of Roles and Responsibilities

The escalation path does not extend into the executive level because data issues are not typically escalated to the senior-most management of an organization.



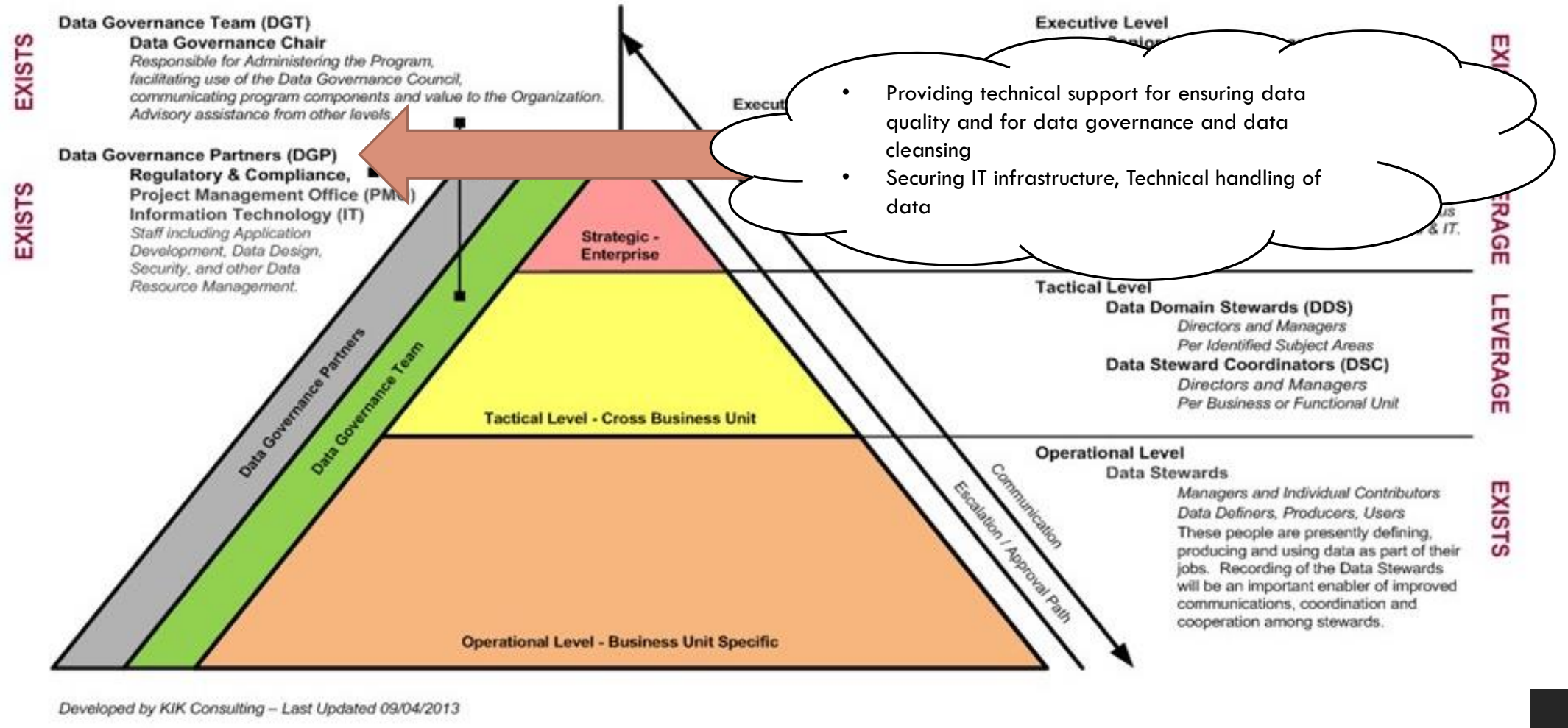
Developed by KIK Consulting – Last Updated 09/04/2013

Data Governance Operating Model of Roles and Responsibilities



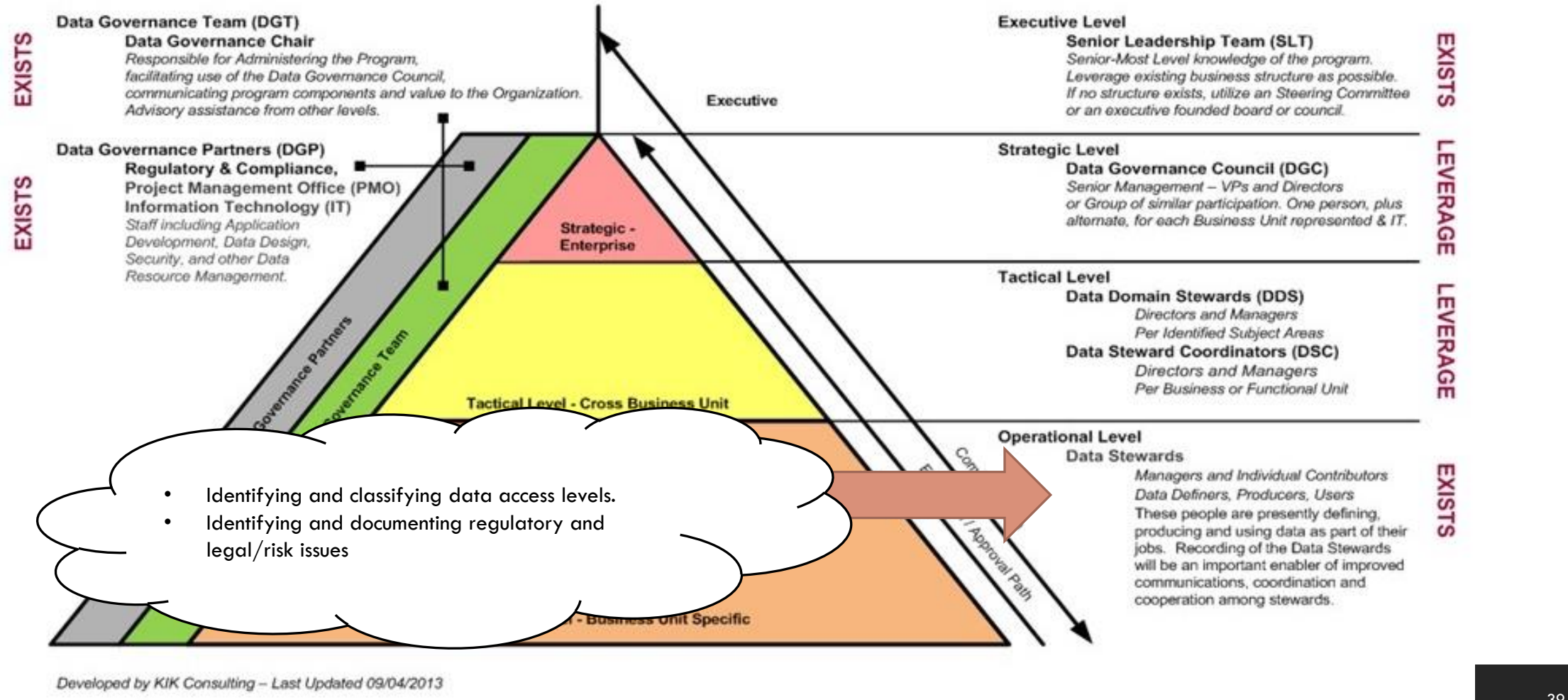
- Architecting the solution and framework
- Assisting in defining data quality metrics
- Supporting data quality issue analysis
- Conducting audits to ensure that policies, procedures and metrics are in place

Data Governance Operating Model of Roles and Responsibilities

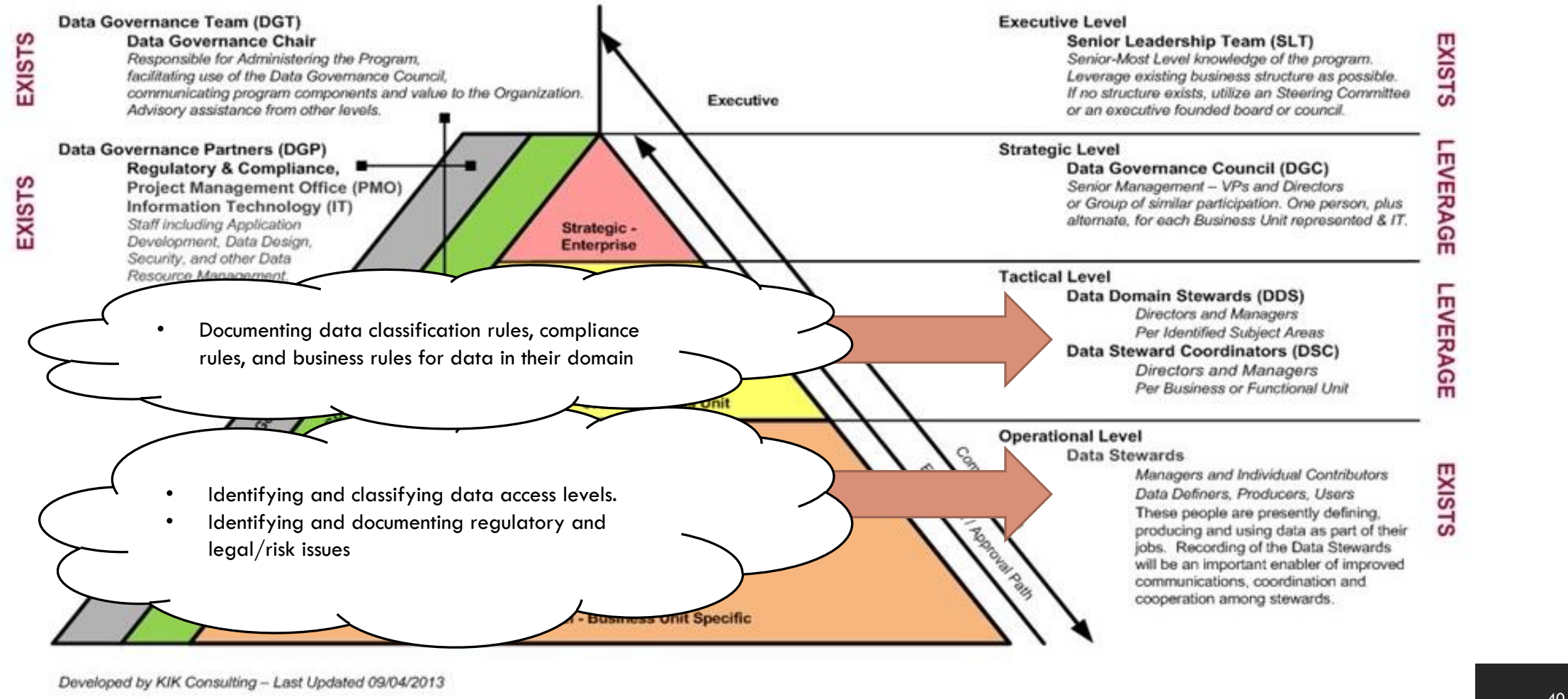


- Providing technical support for ensuring data quality and for data governance and data cleansing
- Securing IT infrastructure, Technical handling of data

Data Governance Operating Model of Roles and Responsibilities



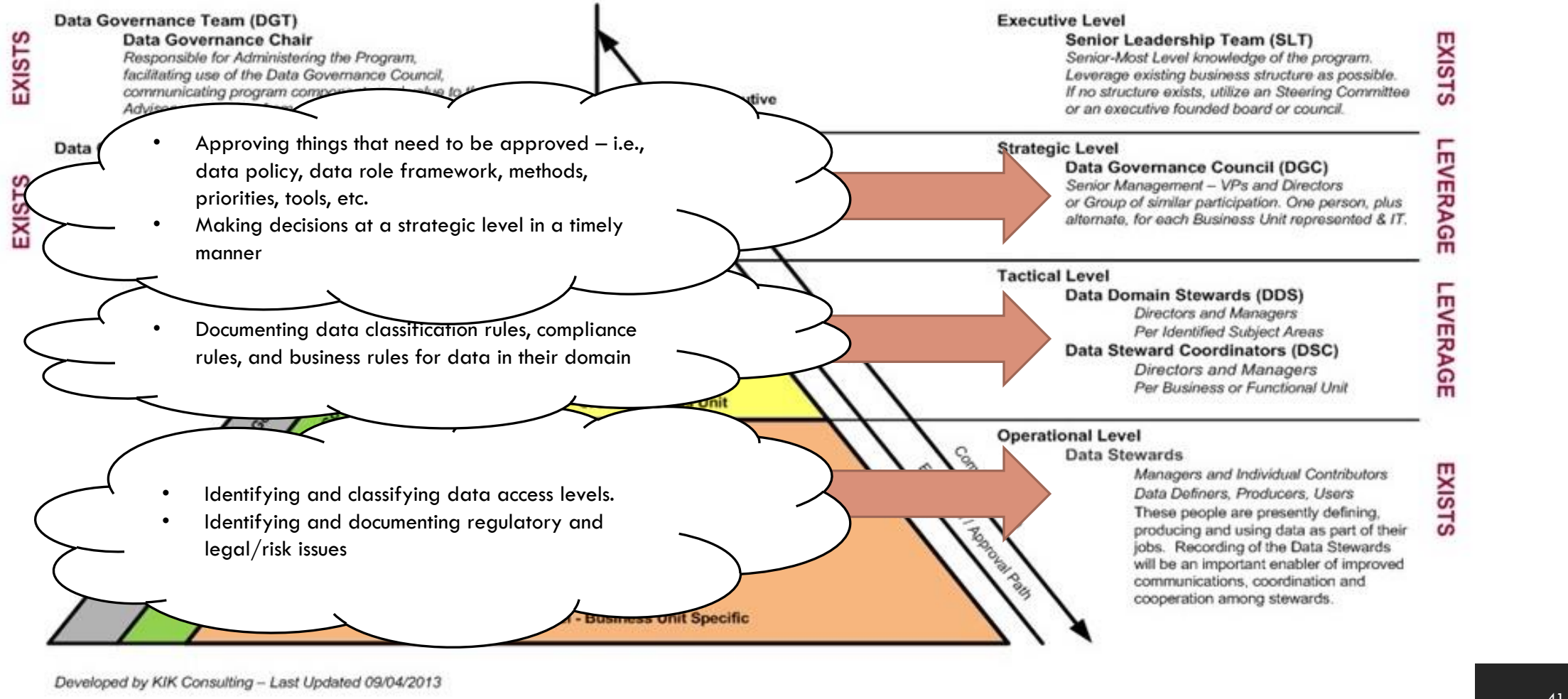
Data Governance Operating Model of Roles and Responsibilities



- Documenting data classification rules, compliance rules, and business rules for data in their domain

- Identifying and classifying data access levels.
- Identifying and documenting regulatory and legal/risk issues

Data Governance Operating Model of Roles and Responsibilities

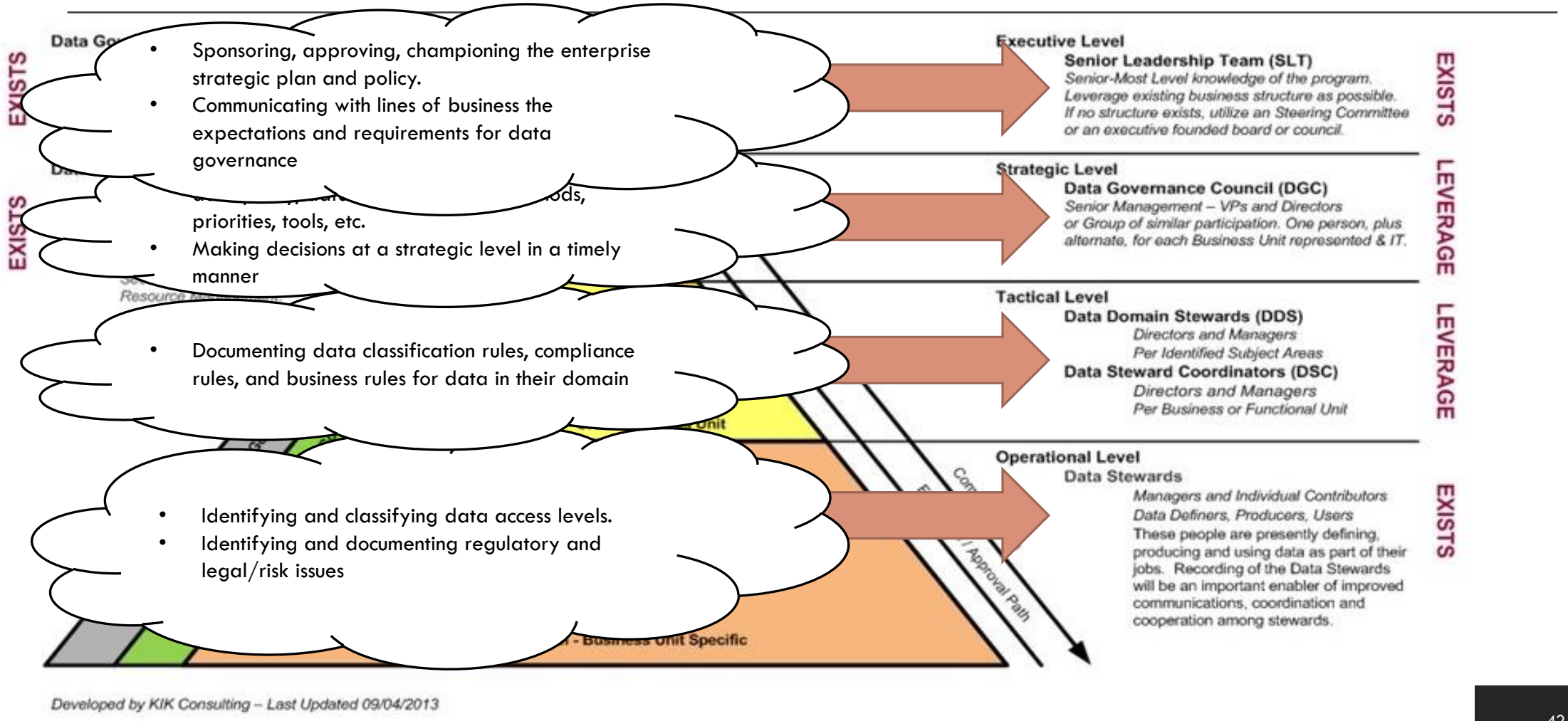


- Approving things that need to be approved – i.e., data policy, data role framework, methods, priorities, tools, etc.
- Making decisions at a strategic level in a timely manner

- Documenting data classification rules, compliance rules, and business rules for data in their domain

- Identifying and classifying data access levels.
- Identifying and documenting regulatory and legal/risk issues

Data Governance Operating Model of Roles and Responsibilities



- Sponsoring, approving, championing the enterprise strategic plan and policy.
- Communicating with lines of business the expectations and requirements for data governance

- priorities, tools, etc.
- Making decisions at a strategic level in a timely manner

- Documenting data classification rules, compliance rules, and business rules for data in their domain

- Identifying and classifying data access levels.
- Identifying and documenting regulatory and legal/risk issues

DG Implementation model

Source: https://theodi.org/wp-content/uploads/2021/06/AdobeStock_410183096-scaled.jpeg

Scope of Typical DG programme

Strategy	Defining, communicating, and driving execution of Data Strategy and Data Governance Strategy
Policy	Setting and enforcing policies related to data and Metadata management, access, usage, security, and quality
Standards and quality	Setting and enforcing Data Quality and Data Architecture standards
Oversight	Providing hands-on observation, audit, and correction in key areas of quality, policy, and data management (often referred to as stewardship)
Compliance	Ensuring the organization can meet data-related regulatory compliance requirements
Issue management	Identifying, defining, escalating, and resolving issues related to data security, data access, data quality, regulatory compliance, data ownership, policy, standards, terminology, or data governance procedures
Data management projects	Sponsoring efforts to improve data management practices
Data asset valuation	Setting standards and processes to consistently define the business value of data assets

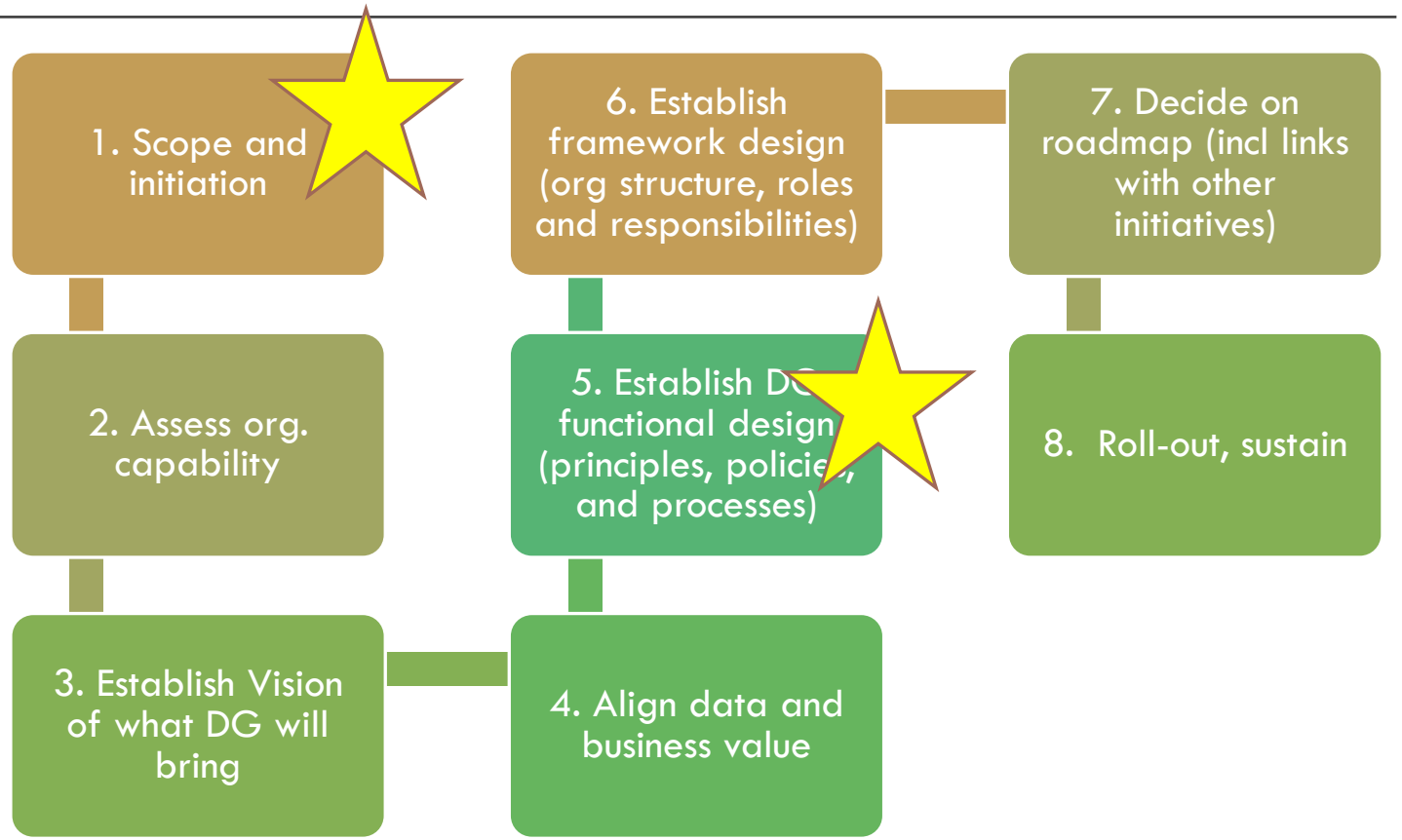
DG implementation 8-phase process

Your organisation thinks that it has a data problem (e.g. poor data quality, risk of data breaches)

Yet, in DG, we've said that "It's not one size fits all"

So how do you "do Data Governance" for your organisation?

Where do you start and how do you know when its done?



DG implementation 8-phase process (Derived from Ladley (2012) Chapter 5)

1. Scope and initiation

Understand options/Decide on Span and Depth

Is there a full understanding in organisation of nature and scope of DG

What business areas are to be covered in the programme?

- DG is ultimately an organisation-wide process but may need to take into account a realistic assessment of effort required, timing, market conditions, government regulations

How deep into the organisation does the programme need to go?

Set- up temporary, organisational steering group important to ensure "not seen just as IT project", right balance of skills and politics is available, at senior enough level to ensure taken seriously (led by executive DG sponsor)

Is there a process to approve scope, acknowledge and address constraints?

Ladley (2012) Chapter 6

5. Establish DG functional design (principles, policies, and processes)

Addresses the what and how questions of DG – moving from concept to operational ("logical model design")

Established principles that provide the foundation of effective policies and therefore processes, protocols and standards.

Identify functions for plan, design, management and operation (e.g., plan assessment info maturity, design data standards, manage data architecture, operate data controls)

Seek buy-in from organisation/DG business leadership

Ladley (2012) Chapter 10

Some examples of GAIP (Generally Accepted Information Principles)

- Data and content of all types are assets with all the characteristics of any other asset.
- They should be managed, secured and accounted for as other materials or financial assets.

Data content as assets

- There is value in all data and content, based on its contribution to an organization's business/operational objectives,
- And to its intrinsic marketability, and/or its contribution to the organization's goodwill (balance sheet) valuation.

Data assets as having Real Value

- There is risk associated with data and content.
- This risk must be formally recognized, either as a liability or through incurring costs to manage and reduce the inherent risk.

Data as presenting risk

- The relevance, meaning, accuracy and lifecycle of data and content can affect the financial status of an organization.

Asset value as determined by Data Quality

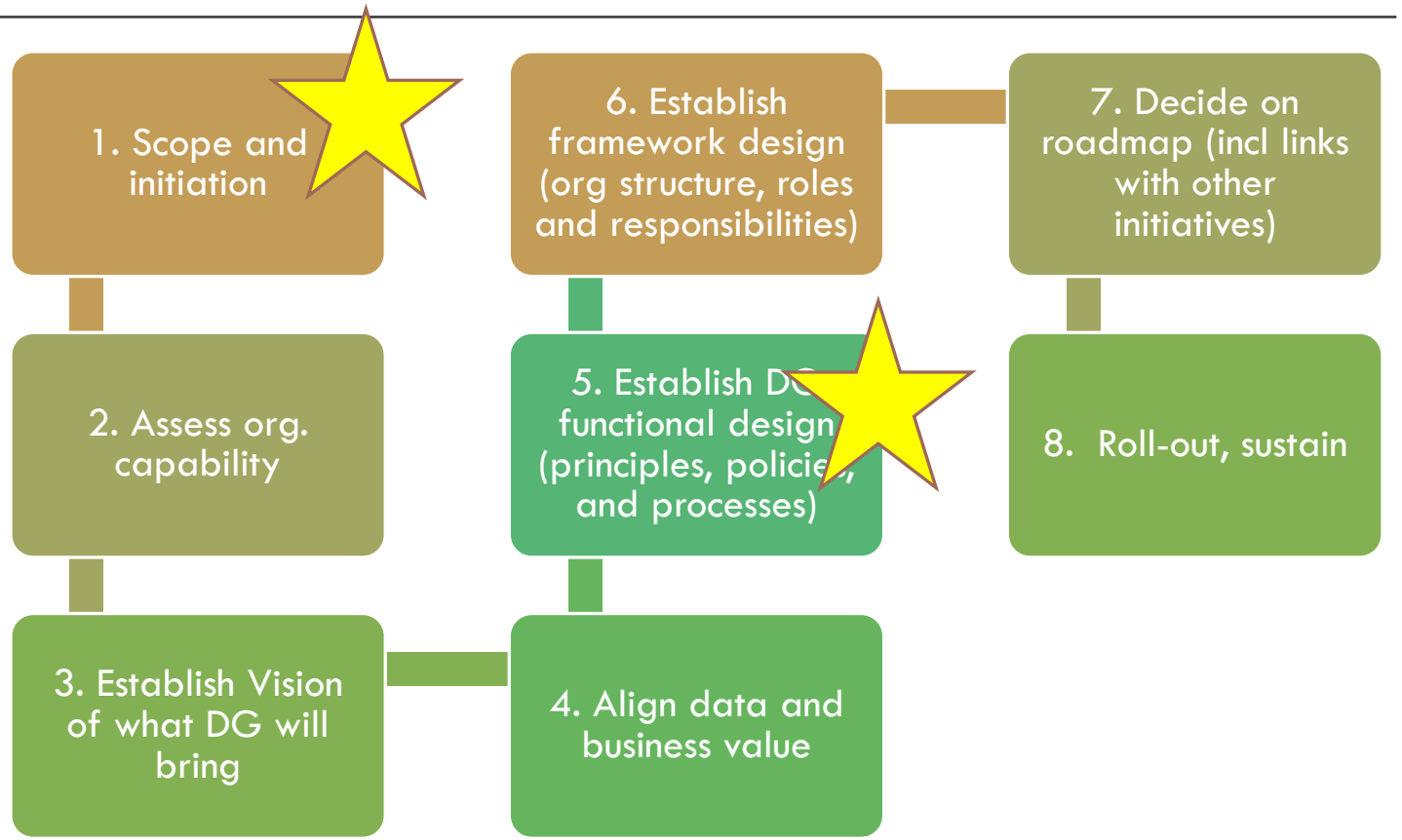
DG implementation 8-phase process

Your organisation thinks that it has a data problem (e.g. poor data quality, risk of data breaches)

Yet, in DG, we've said that "It's not one size fits all"

So how do you "do Data Governance" for your organisation?

Where do you start and how do you know when its done?



DG implementation 8-phase process (Derived from Ladley (2012) Chapter 5)

Conclusion

Source: https://theodi.org/wp-content/uploads/2021/06/AdobeStock_410183096-scaled.jpeg

Aims of presentation

Provide overview of DGM section of Data Governance and Ethics module (normally 6 contact hours)

Introduce DGM Concepts and Definitions

Provide DGM Rationale and Goals

Show strategic overview of DGM (who, how, what etc)

Illustrate DGM building blocks

- Roles and responsibilities
- Operating Model

Explain implementation method (using Ladley model)

Any Comments or Questions?