Transforming Data With Intelligence™

Previews of TDWI course books offer an opportunity to see the quality of our material and help you to select the courses that best fit your needs. The previews cannot be printed.

TDWI strives to provide course books that are content-rich and that serve as useful reference documents after a class has ended.

This preview shows selected pages that are representative of the entire course book; pages are not consecutive. The page numbers shown at the bottom of each page indicate their actual position in the course book. All table-of-contents pages are included to illustrate all of the topics covered by the course.

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY.

This page intentionally left blank.

TDWI Data Governance Fundamentals

Managing Data as an Asset

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

COURSE OBJECTIVES

To learn:

- ✓ ***Definitions and dimensions of data governance***
- ✓ ***Key considerations and challenges in building a data governance program***
- ✓ ***The practices, roles, skills, and disciplines essential to data governance***
- ✓ ***The qualities that make good data stewards and stewardship organizations***
- ✓ ***The processes of developing, executing, and sustaining data governance***
- ✓ ***Activities, issues, and options when building a data governance program***
- ✓ ***How maturity models are applied for data governance***
- ✓ ***The importance of adapting data governance for trends such as big data, cloud services, and agile development methods***

TDWI takes pride in the educational soundness and technical accuracy of all of our courses. Please send us your comments—we'd like to hear from you. Address your feedback to:

info@tdwi.org

Publication Date: May 2014

© Copyright 2010-2014 by TDWI. All rights reserved. No part of this document may be reproduced in any form, or by any means, without written permission from TDWI.

TABLE OF CONTENTS

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

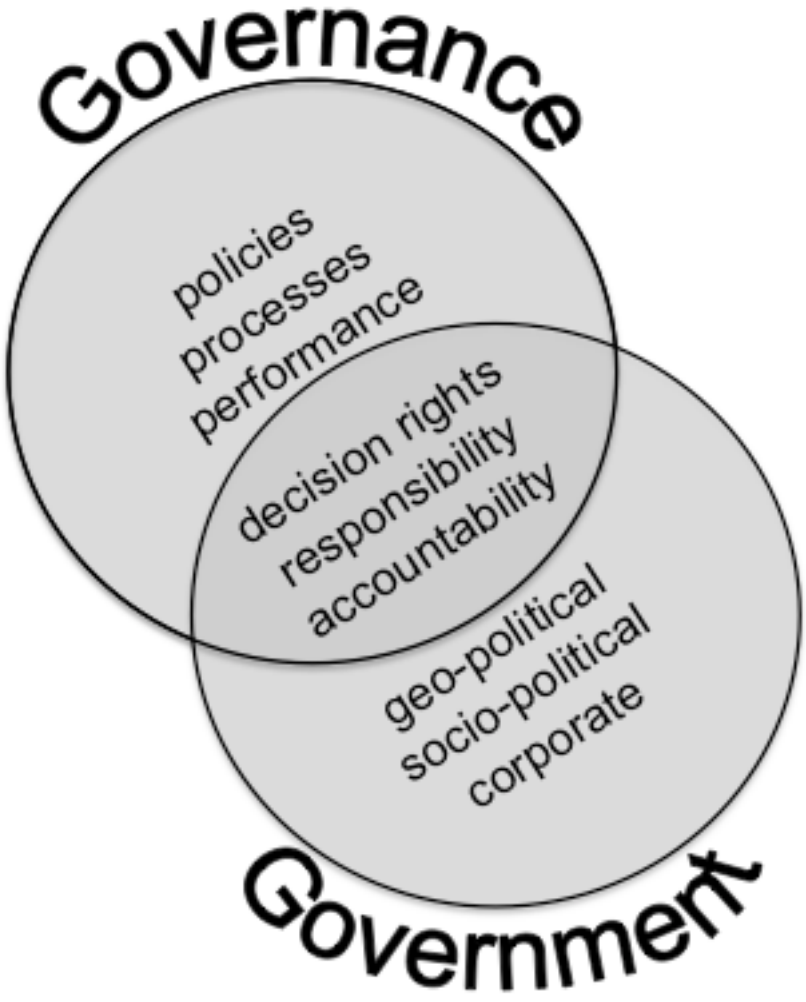
Module 1

Data Governance Concepts

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Defining Data Governance

Governance Defined



 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Defining Data Governance

Governance Defined

THE ACT OF GOVERNING

To understand data governance it is useful to first understand the term governance. Governance is the activity of governing – the collection of things that a government does. When we hear the word government, we typically think of geo-political governments such as nations and states. Yet other kinds of governmental structures also exist: socio-political structures such as tribes and families, and the corporate governments that are the power structure of business entities.

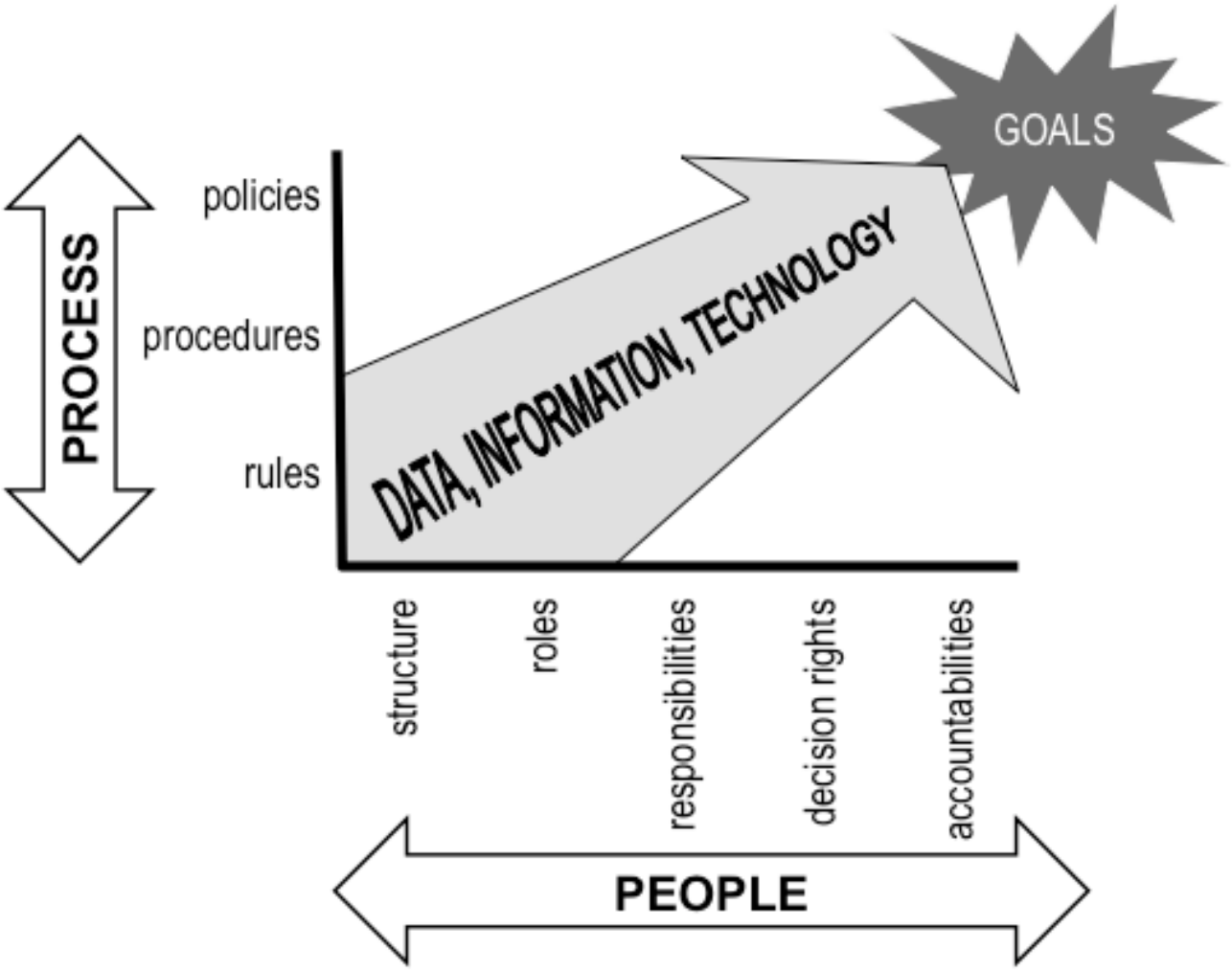
While the scope and formality of government varies widely between these bodies, they all encompass the core elements of governance:

- defining decision rights,
- designating responsibilities,
- assigning accountabilities,
- establishing policies,
- defining processes,
- managing performance.

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Dimensions of Data Governance

Aspects of Data Governance



 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Dimensions of Data Governance

Aspects of Data Governance

GOALS, PEOPLE, AND PROCESSES

Data governance is a program of managing information assets to achieve defined information management goals. Typical goals include such things as quality, security, and standardization. Processes must be defined and people must be engaged to achieve the stated goals.

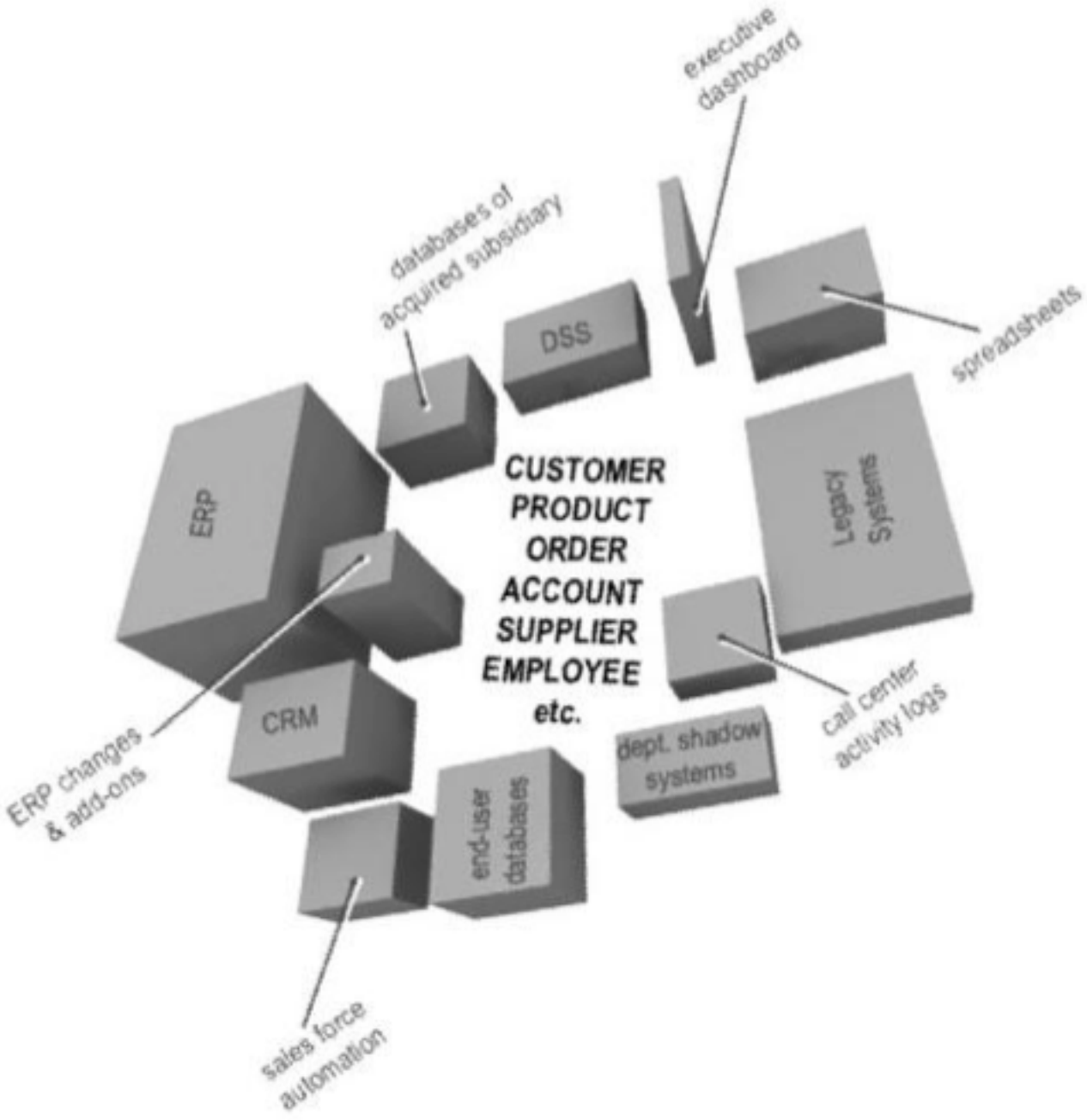
The process dimension of data governance includes policies, procedures, and rules. The people dimension of data governance includes organizational structure, roles, responsibilities, decision rights, and accountabilities

These dimensions create a management framework within which data and information are managed, and technologies are employed to achieve specific information management goals.

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Data Governance Challenges

What Data to Govern?



 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Data Governance Challenges

What Data to Govern?

SCOPE OF DATA

One of the big challenges when starting a data governance program is to determine the scope of data to be governed. Every organization has lots of data. It is present in enterprise systems and databases, data warehouses, decision support databases, departmental systems, shadow systems, end-user databases, spreadsheets, and more. Each database has different needs and considerations related to quality, security, compliance, etc. Furthermore, the data encompasses many subjects – customers, products, orders, accounts, employees, etc. – and each subject has different needs and considerations for quality, security, and compliance.

The combination of abundant and often redundant data with many data subjects makes scope of data a complex and compound set of questions. Looking at a single subject – customer, for example – the questions include:

- Do you need to govern customer data?
- What are the motivations for governing customer data – quality, security, compliance, etc?
- Where are all of the places that customer data exists, including end-user databases and spreadsheets?
- For each location where customer data exists, is data governance necessary? Is it practical?

These are not easy questions to answer, and the answers will vary from one business to another. Consider, for example, the implications of customer data in a spreadsheet on a portable USB drive. If your business is healthcare and the customer data is really patient data, the security and compliance issues are significant. If your business is media services, this may be a lower risk scenario. It may be necessary to govern spreadsheets in one instance and impractical to do so in another.

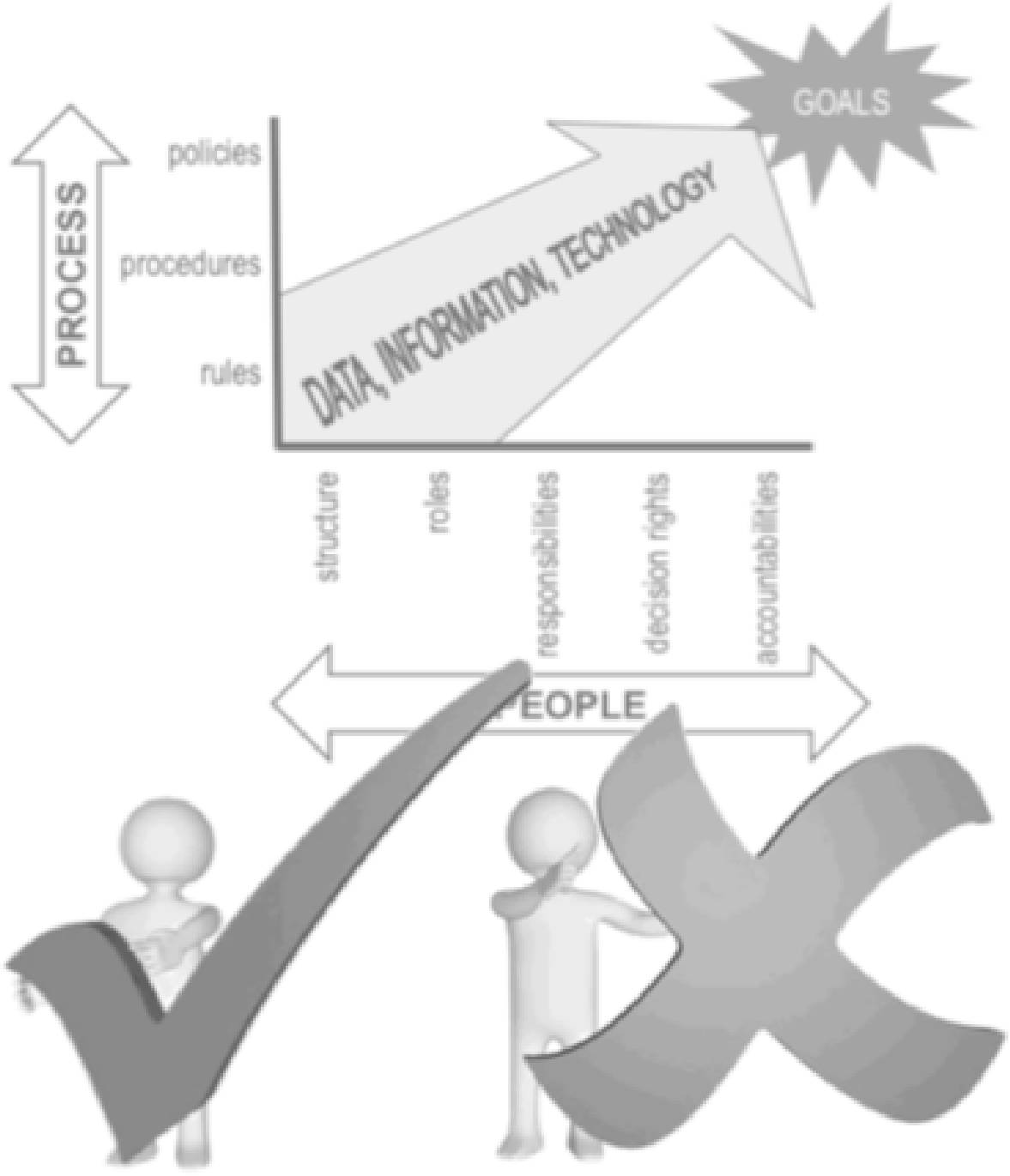
SOME GUIDELINES

Limit the scope of data to that for which there is a clear need to govern. The Data Governance Institute advises to govern “as little as will help you meet your goals.” If you’re just getting started with governance it is wise to start with small scope – one or a few subjects with a high degree of cross-functional business activity. Also consider for each subject and for each kind of database the level of business interest and participation that you can expect. What level of support and sponsorship is realistic? What level of resistance is likely?

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Fostering Participation

Participation and Resistance



 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Fostering Participation

Participation and Resistance

WHY RESISTANCE?

Data governance introduces new business processes that mean significant changes in organizations, roles, and responsibilities. As is true for any major organizational change, resistance is a certainty. Expect to experience all of the common resistance behaviors from *"the way we've always done it"* to *"what's in it for me"* as you initiate and implement your data governance program.

RESISTANCE IN BUSINESS UNITS

Business stakeholders are the consumers of data and the people most affected by data quality, security, and compliance failures. Yet they may still resist governance processes (or want them to apply to "everyone but me"). Common forms of business resistance include treating data issues as an IT or technology problem, bargaining for exceptions to policy, and demanding ROI justification for each governance policy and process.

RESISTANCE IN IT PROJECTS

System and database developers, who have much to gain from sound data management policies, are often reluctant to accept data governance. One analysis[1] indicates that 66% of development teams choose to bypass their organization's data standards groups. Common reasons to work around DG groups are *too slow* and *too difficult to work with*. Other reasons include lack of knowledge about the group and the value that it provides.

CHANGING FROM RESISTANCE TO PARTICIPATION

The first step to overcoming resistance is to acknowledge it. Don't dismiss objections, but recognize them as legitimate expressions of some people's perspectives. Then counter with proven organizational change management techniques which include:

- Clear statement of the governance vision
- Strong and visible sponsorship
- Active, visible, and respected champions
- Transparent governance processes
- Early and frequent communication that invites feedback
- Open response to feedback
- Well-defined goals with corresponding measures
- Opportunities to participate in crafting policies and processes
- Changes to job descriptions and performance measures
- Education when needed
- Process automation when appropriate
- Tools, templates, and services for easy-to-follow standards

[1] Source: *Agile/Lean Data Governance Best Practices* by Scott W. Ambler (http://www.agiledata.org/essays/dataGovernance.html). Ambler is a consultant, practitioner, and author in the field of agile software development.

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

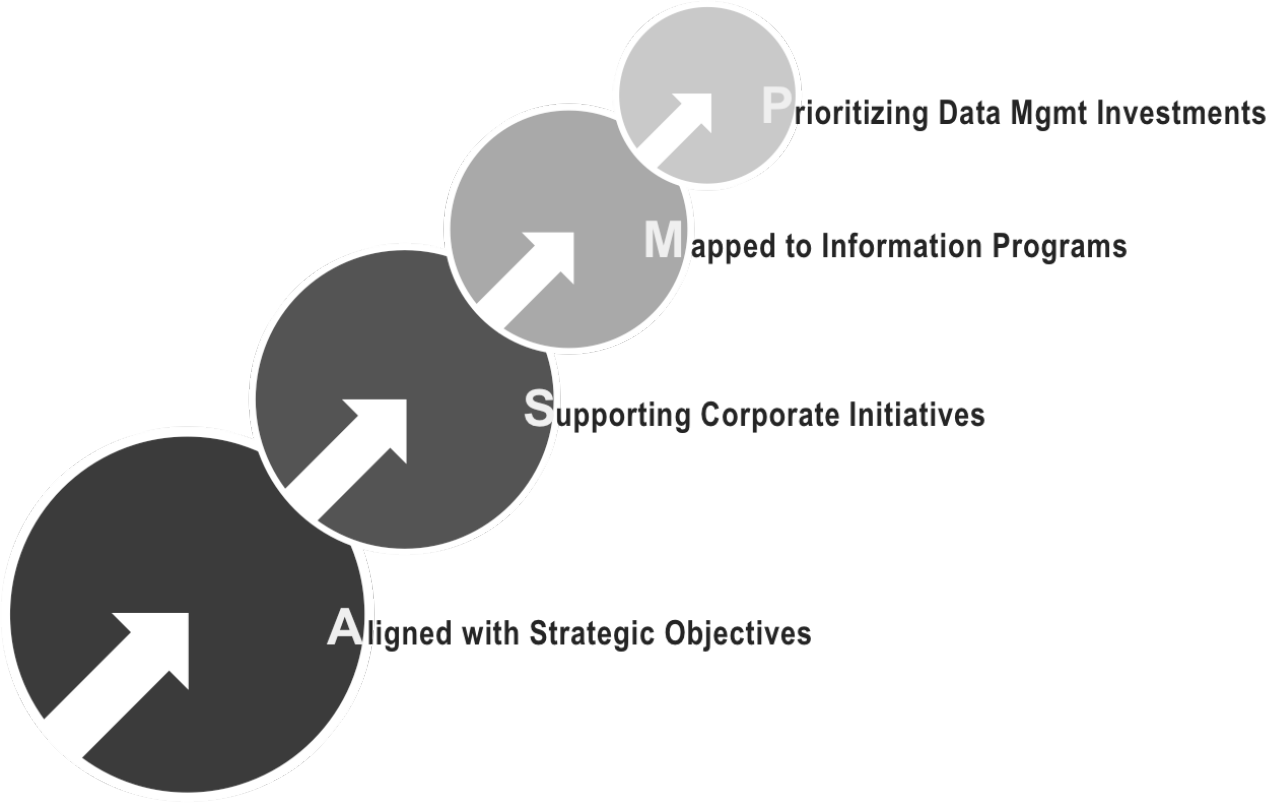
Module 2

Data Governance Organizations

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Positioning Data Governance in the Enterprise

Connecting Business Strategy with Data Management



 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Positioning Data Governance in the Enterprise

Connecting Business Strategy with Data Management

STRATEGIC ALIGNMENT

Data governance is best positioned as shared services for the enterprise supporting business goals at all levels – strategic, tactical, and operational. Although a data governance program has specific data-oriented goals (security, privacy, quality, compliance, etc.) it is important to connect governance with higher level business goals – increased revenue, reduced cost, risk avoidance, process optimization, etc. – that are the concerns and responsibilities of business stakeholders. Stakeholder goal alignment is essential to achieve buy-in and adoption. continuous realignment is necessary for long-term sustainability.

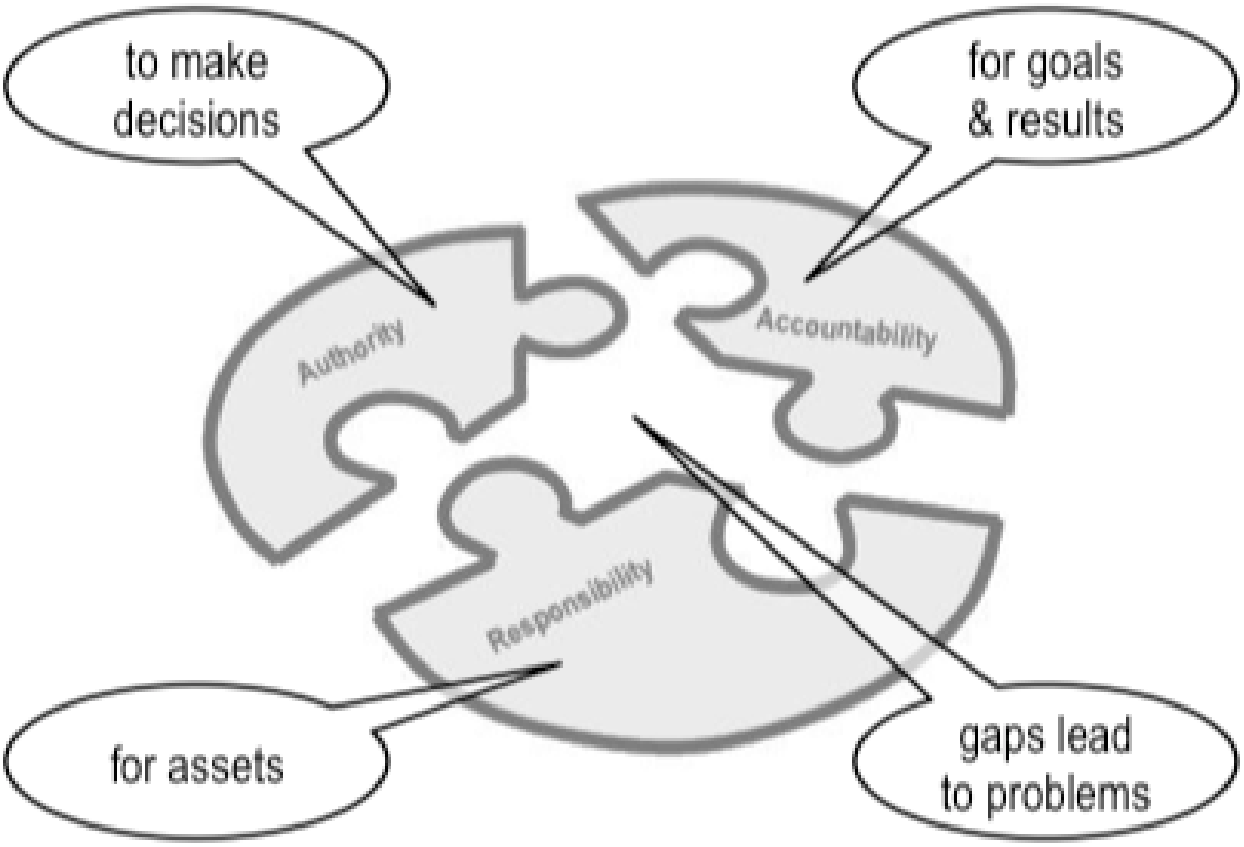
Ideally a data governance program is:

- Aligned with strategic objectives such as profit, positive customer perception, process excellence, and exceptional organizational and individual competencies.
- Extended from strategy to identify how governance connects with tactics and processes through management disciplines such as CRM, BPM, SCM, and HCM.
- Supportive of continuous process improvement initiatives such as six sigma, lean, and business process redesign.
- Supportive of cultural initiatives such as customer delight, employee engagement, agile business, social responsibility, and continuous innovation.
- Mapped to and enabling of information management programs and roadmaps such as BI, MDM, advanced analytics, big data, etc.
- Applied to systematically prioritize and oversee a portfolio of data management investments.

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Governance and Management Practices

Responsibility, Authority and Accountability



 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Governance and Management Practices

Responsibility, Authority and Accountability

PEOPLE GOVERN

Data governance is something that is done by people. Many people in various roles interact to govern data. Roles are an important concept of governance. Every role has designated responsibility, authority, and accountability. These things – responsibility, authority, and accountability – are the cornerstones of governance.

RESPONSIBILITY

Responsibility is the obligation incurred by an individual in a specific role to perform the duties of that role. The individual is obligated to take actions and produce results that affect the organizations assets.

AUTHORITY

Authority is the power granted to an individual in a specific role to make decisions and direct others to follow those decisions.

ACCOUNTABILITY

Accountability is the individual liability created by use of authority. It is a condition of being fully answerable for results and achievement of goals.

ALIGNMENT

A governance organization (or any organization) works well when responsibility, authority, and accountability are well aligned. The authority granted to a role must be aligned with the responsibilities of that role. And the accountability of the role must correspond with the scope and level of authority. Gaps and inconsistencies between the three elements are certain to cause dysfunction in the organization.

Accountability is the logical consequence of responsibility and authority.

DELEGATION

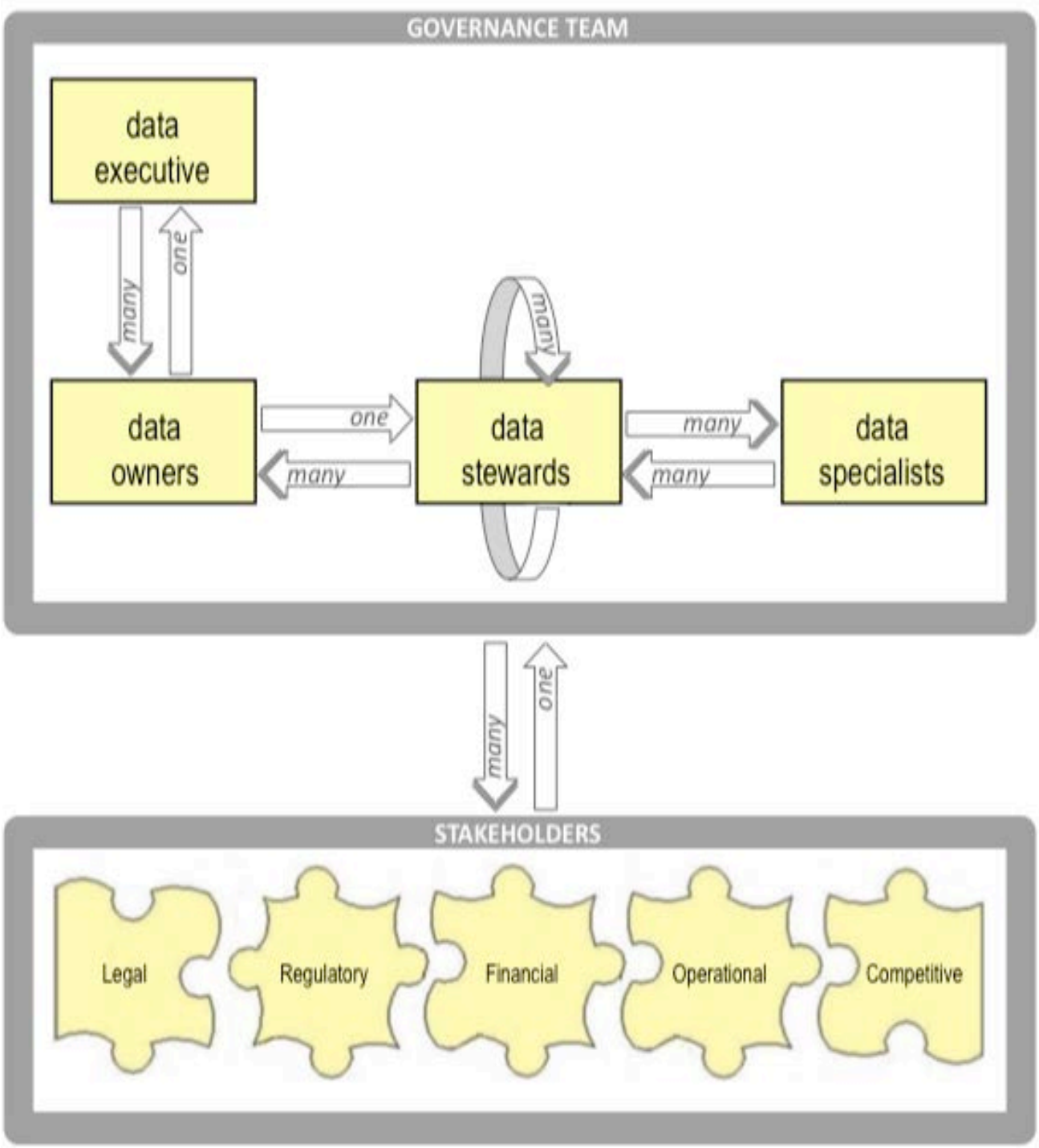
Delegation is an additional consideration to keep in mind when defining governance roles. A particular role may be assigned a particular responsibility with or without the right to delegate that responsibility. When responsibility is delegated the corresponding authority must also be delegated.

Responsibility and authority may be delegated. Accountability cannot be delegated – it remains with the delegator.

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Data Governance Roles

Governance as Teamwork



 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Data Governance Roles

Governance as Teamwork

THE GOVERNANCE TEAM

A typical data governance team includes four roles as illustrated on the facing page:

- A data executive designated as the person who provides overall leadership of a data governance program.
- Data owners responsible for access, distribution, retention, etc.
- Data stewards who facilitate consensus data definitions and foster sound data quality, data usage, and data security practices.
- Data specialists such as data architects, data modelers, database developers, and database administrators who have custodial responsibility for data.

THE DATA EXECUTIVE

The data executive role is handled differently in various organizations and data governance programs. In some instances data executive is a C-level position. The CIO or the CFO may be designated as having executive responsibility for data. Security and compliance-driven programs may designate a senior security or compliance officer. In *Managing Your Business Data*, Villar and Kushner offer an extensive discussion and make a compelling case to establish a new Chief Data Officer (CDO) position.[1]

In other instances a Data Governance Program Manager is created as a senior management position that provides a bridge for engagement of Director and CxO positions.

Whatever means makes best sense for your organization's size, needs, governance goals, and culture it is important that executive perspective is included in the data governance team.

[1] Source: *Managing Your Business Data*, pp 114-121, Villar & Kushner

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Data Governance Skills and Disciplines

Data Architecture

 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Data Governance Skills and Disciplines

Data Architecture

THE NEED FOR ARCHITECTURE

Architecture describes the way that a diverse set of components fits together and interacts to fulfill a purpose or meet needs. Enterprise data certainly fits the criteria of "a diverse set of components." Managing the data architecture is a critical skill without which any data governance team will struggle.

Data architecture is complex because enterprise data is complex. A typical enterprise has redundant data dispersed across many systems and databases. Much of the data resides in legacy and ERP systems where details of data structures may be obscure. More data exists in personal databases and spreadsheets – often unknown to the IT department. Still more data may be found in the external systems of service providers or business partners. Architecture is necessary because:

- Enterprise data is widely dispersed and redundant.
- Multiple representations of similar information are fundamentally different in format, structure, definition, quality, completeness, and security controls.
- Most of the data lacks documentation, models, and other metadata.
- *You should not govern what you don't understand.*
- The core data management processes of IT departments touch only a small percentage of the enterprise data resource.
- Much of the data that exists outside of core systems is invisible to those who govern data.
- *You cannot govern what you can't see.*

WHAT IS NEEDED?

Data architecture can take many forms and mean different things to different people. For governance purposes you need documentation that describes the data resource and illustrates how the components fit together. Minimum documentation includes:

- A subject model that describes the various data domains
- Logical models that describe business views of the data
- Subject and entity definitions
- Implementation maps showing where entity data resides physically
- A matrix map of business unit interactions (create, report, update, delete) with data entities

Beyond the minimum documentation you may also want to document data flow, data lineage, and attribute/element definitions. In a changing environment, "as is" and "as needed" models are valuable.

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

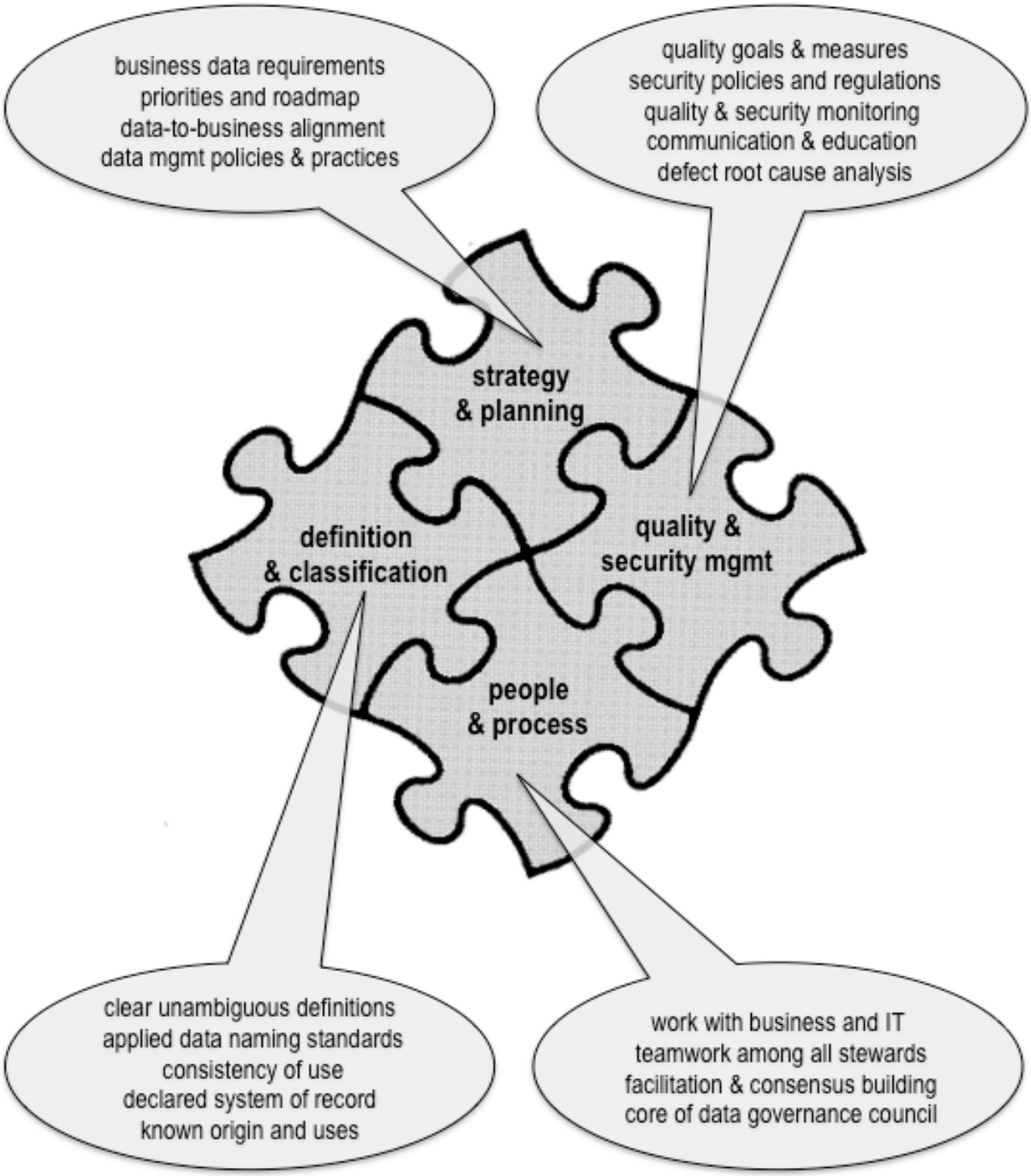
Module 3

Data Stewardship

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Stewardship Concepts

Responsibilities and Accountabilities



 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Stewardship Concepts

Responsibilities and Accountabilities

A NEW AND DIFFERENT ROLE

As we've already discussed, data governance involves three primary roles – ownership, stewardship, and custodianship. Stewardship, however, differs from the other roles in some significant ways. Data stewardship needs to be explored in greater depth because:

- Stewardship is a distinctly new and different role. Ownership recognizes and formalizes a set of responsibilities of business managers but does not redefine the job of business management. Similarly, custodianship recognizes and formalizes responsibilities of data specialists but doesn't redefine their jobs. Data stewardship is different because it is a new kind of job.

- Data stewardship is the nexus of data governance. It provides linkage among owners of different but related data subjects. And it connects business rules and requirements with data models, database design, information systems implementation, and day-to-day management and administration of data.

SCOPE OF RESPONSIBILITY

Data stewardship responsibilities fit into four major categories:

- *Strategy and Planning* identifies business requirements for data and information, helps to set priorities, and maintains a roadmap of activities to meet requirements. Continuous alignment of data with business needs, and of data management policies and practices with business goals are primary strategy and planning objectives.

- *Definition and Classification* addresses the many definitional and metadata topics previously discussed – data definitions, data naming, consistent use of data, lineage and traceability of data, etc.

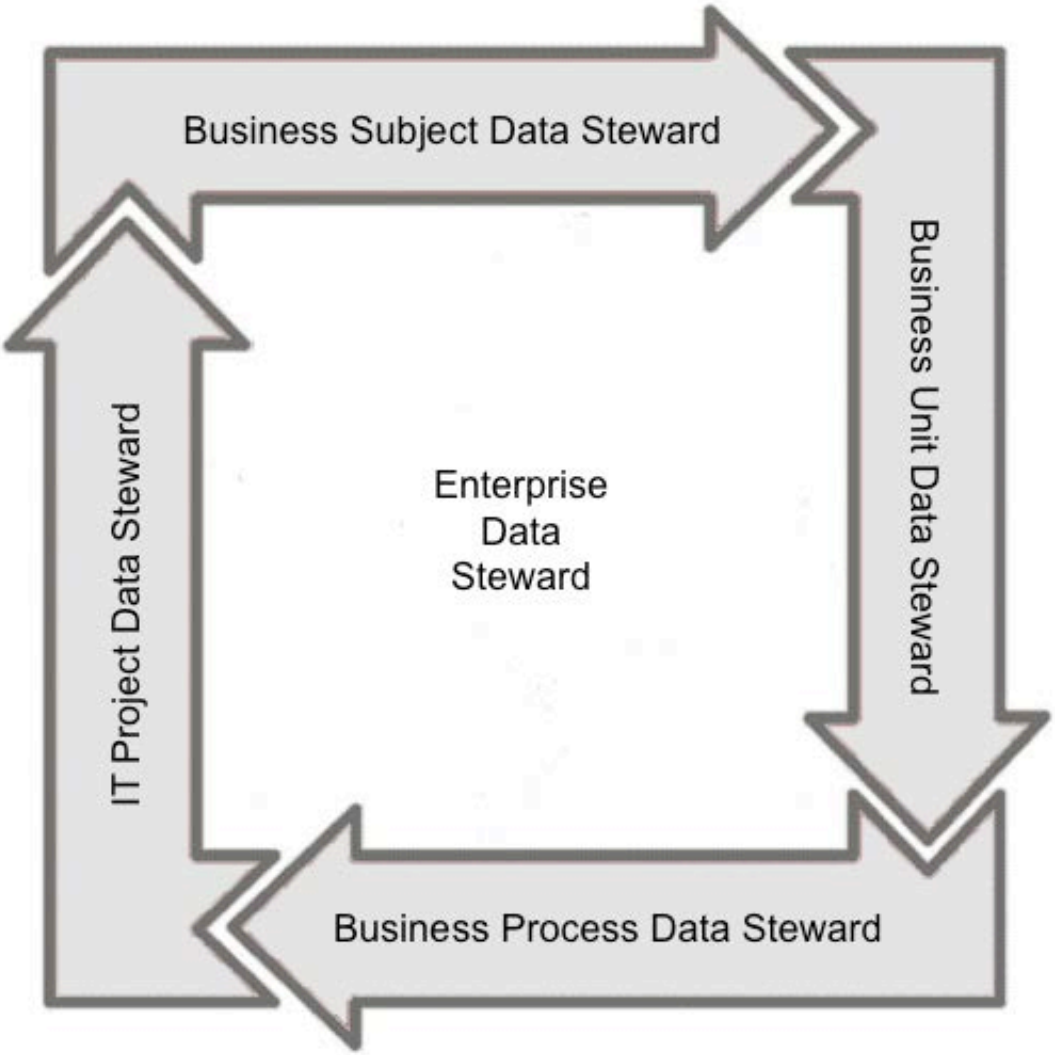
- *Quality and Security Management* is a direct connection to the goals of—and motivations for— governance. Areas of responsibility include policies, regulations, goals, measures, monitoring, communication, education, and root cause analysis.

- *People and Process* responsibilities are among the most important of data steward responsibilities. Through teamwork, facilitation, and consensus building they form the core of a governance organization.

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Stewardship Organizations

Kinds of Data Stewards



 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Stewardship Organizations

Kinds of Data Stewards

MANY DATA STEWARDS

It is not realistic to assign a single data steward. Every organization is sure to have many stewards. Maybe you'll have only a few at the start of a governance program, but expect the number to grow as data governance expands and matures.

Claudia Imhoff recognized the need for multiple stewards very early in emergence of the discipline. In 1997, Imhoff wrote, "A typical corporate Data Stewardship function should have one Data Steward assigned to each major data subject area. These subject areas consist of the critical data entities or subjects such as Customer, Order, Product, Market Segment, Employee, Organization, Inventory, etc. Usually, there are about 15-20 major subject areas in any corporation."[1]

Imhoff describes what is known as a *Business Subject Data Steward*. As the discipline has evolved several other kinds of data stewards have emerged including:

- *Business Unit Data Stewards* responsible for the data needs of a particular business department or business function. This approach may augment subject stewardship for departments with a high level of data dependency and need for data management.
- *Business Process Data Stewards* provide data oversight from a process perspective, and may supplement subject stewardship when a business process depends upon critical data flows from many sources.
- *Business Location Data Stewards* represent a location-specific data view in global and multi-national corporations.
- *IT Project Data Stewards* representing data management perspective of data-centric projects such as those to implement ERP, MDM, or data warehousing systems.

You may also find need to identify an *Enterprise Data Steward* – also known as *Lead Data Steward* or *Chief Data Steward* – as a clarifying and coordinating role among many stewards. According to Imhoff "The lead Data Steward's responsibility is to determine and control the domain of each Data Steward. These domains can become muddy and unclear, especially where subject areas intersect. Political battles can develop between the Data Stewards if their domains are not clearly established."[1]

[1] Source: *Data Stewardship: Process for Achieving Data Integrity*, The Data Administration Newsletter (www.tdan.com/view-articles/4196). Imhoff is Founder and President of Intelligent Solutions and a well-known speaker and author in business intelligence.

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Stewardship Skills and Knowledge

People Skills

 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Stewardship Skills and Knowledge

People Skills

THE HUMAN SIDE OF DATA MANAGEMENT

Stewardship can't succeed without engaging and involving people. The stakeholders of enterprise data are many and the priorities, preferences, and needs often in conflict. Successful stewards rely first on people skills – communication, facilitation, and consensus building – to achieve goals.

COMMUNICATION

A skilled data steward recognizes that communication involves both sending and receiving of information. Stewards must communicate policy and processes, and they must also seek input and be receptive to feedback. Communication is an essential foundation for the human and group management skills of facilitation, consensus building, and team building. Effective communication in all forms – verbal, written, and visual – is an essential trait for effective data stewardship.

FACILITATION

From enterprise-view definitions to policy shaping and problems solving, data stewardship drives stakeholder group work. Facilitation makes group work easy to achieve and effective in getting the right results. A facilitator provides non-directive, non-controlling leadership to get results through assistance and guidance. A skilled facilitator will:

- elicit participation and inspire creativity in the group
- distinguish process from content
- maintain objectivity and lead without bias
- separate facts, feelings, and opinions and understand the value of each
- respect individual perspectives and the wisdom of the group
- help to identify, express, and resolve conflict
- read and understand group dynamics
- allocate time to relationship building as well as to task execution
- decide by consensus and negotiation – not majority or mandate
- expect and adapt to change
- personally exhibit desired team-member behaviors

CONSENSUS BUILDING

Consensus – a critical stewardship component – is a process of problem solving by collaboration. The activities (and skills) include:

- defining the problem (facilitation and analysis)
- defining successful outcomes (using facilitation)
- engaging participation (facilitation and group dynamics)
- fact finding (analysis, systems thinking, critical thinking)
- identifying alternatives (facilitation, brainstorming)
- decision making (facilitation, negotiation, mediation)

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

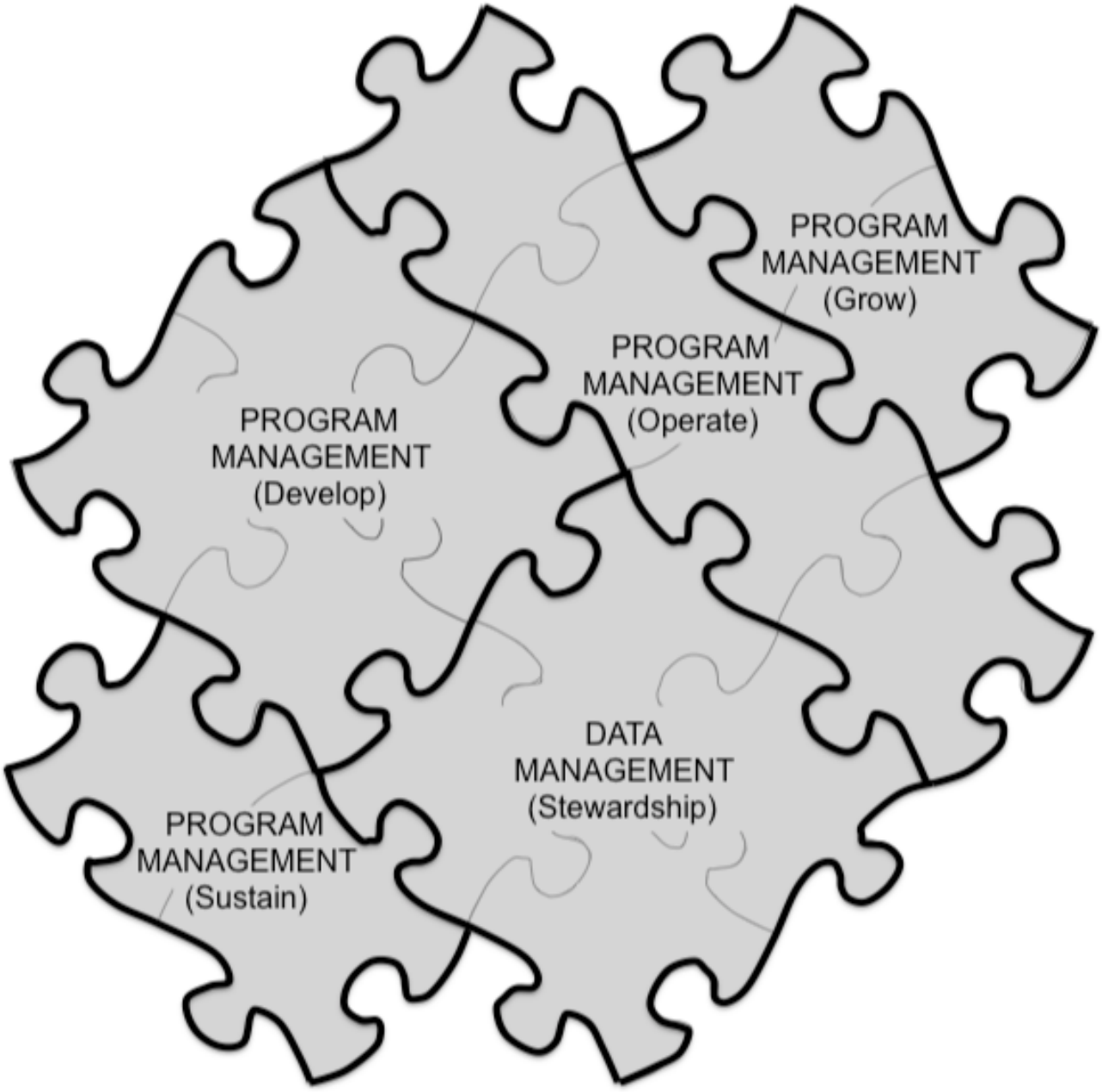
Module 4

Data Governance Processes

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Governance and Management

The Processes of Governing Data



 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Governance and Management

The Processes of Governing Data

STEWARDSHIP IN A GOVERNANCE PROGRAM

We have now discussed data stewardship extensively. We've looked at data steward responsibilities and accountabilities, the purpose of data stewardship, various kinds of data stewards, creating a data stewardship organization, and essential skills for data stewardship.

Data stewardship is important and central to data governance. But stewardship alone does not make data governance. Data governance is a program – a system of projects and services designed to manage the data resource – that coordinates the activities and efforts of data owners, stewards, and custodians.

MANAGEMENT AND A GOVERNANCE PROGRAM

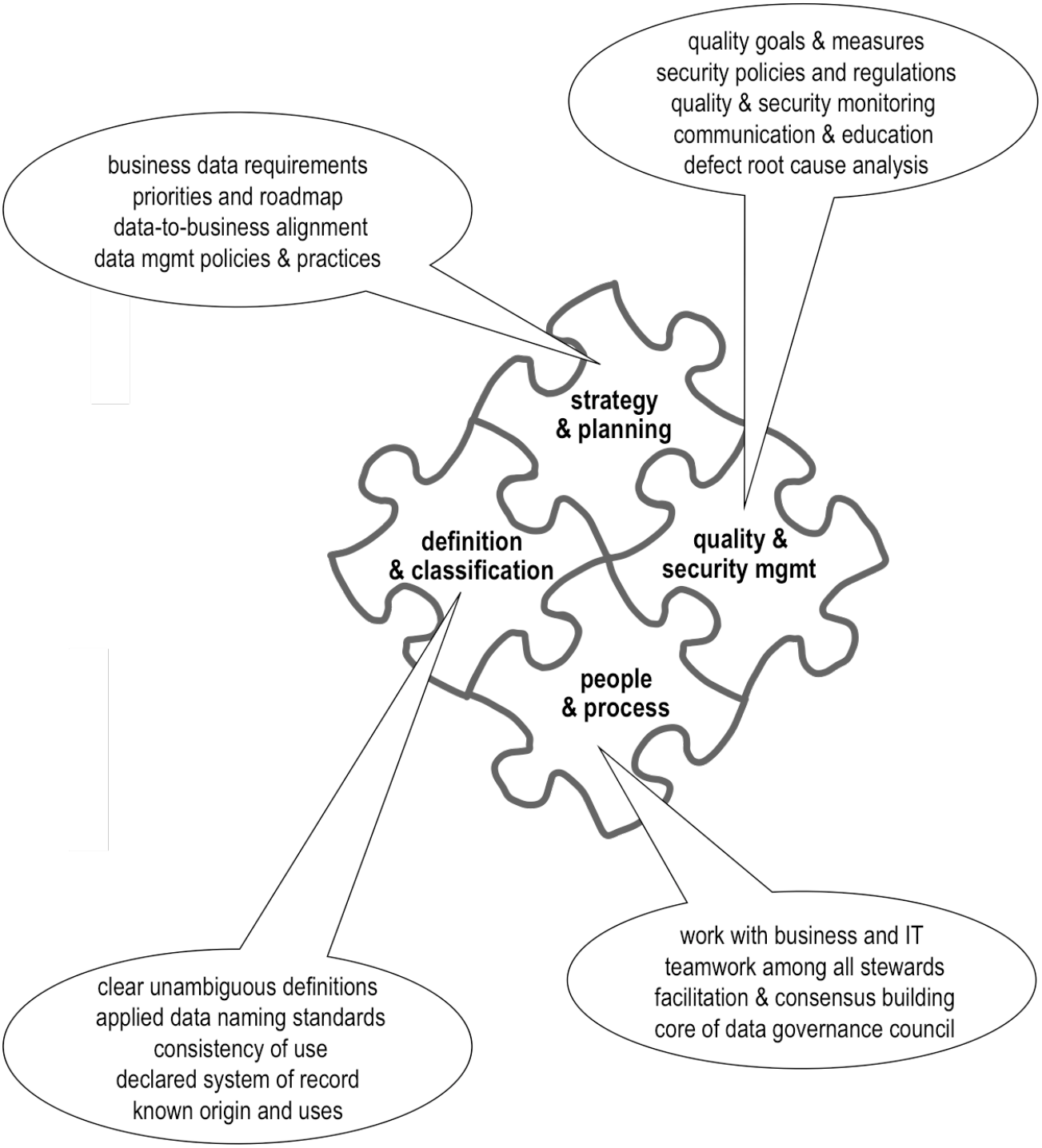
Data stewardship focuses on managing data – its quality, security, value, etc. A complete data governance program includes data management but also requires program management with distinct roles, responsibilities, and accountabilities to:

- *Develop* and establish a new governance program at its inception.
- *Operate* data governance on a day-to-day basis.
- *Sustain* the program as issues arise and scope evolves.
- *Grow* governance capabilities and data management maturity.

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Data Management Processes

Stewardship of Data



 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Data Management Processes

Data Stewardship

STEWARDSHIP IS MANAGEMENT OF DATA

A quick review of data stewardship responsibilities sets the stage here because stewardship encompasses the data management processes of data governance. Stewardship is the core upon which a data governance program depends.

Data stewardship responsibilities fit into four major categories:

- *Strategy and Planning* identifies business requirements for data and information, helps to set priorities, and maintains a roadmap of activities to meet requirements. Continuous alignment of data with business needs, and of data management policies and practices with business goals are primary strategy and planning objectives.

- *Definition and Classification* addresses the many definitional and metadata topics previously discussed – data definitions, data naming, consistent use of data, lineage and traceability of data, etc.

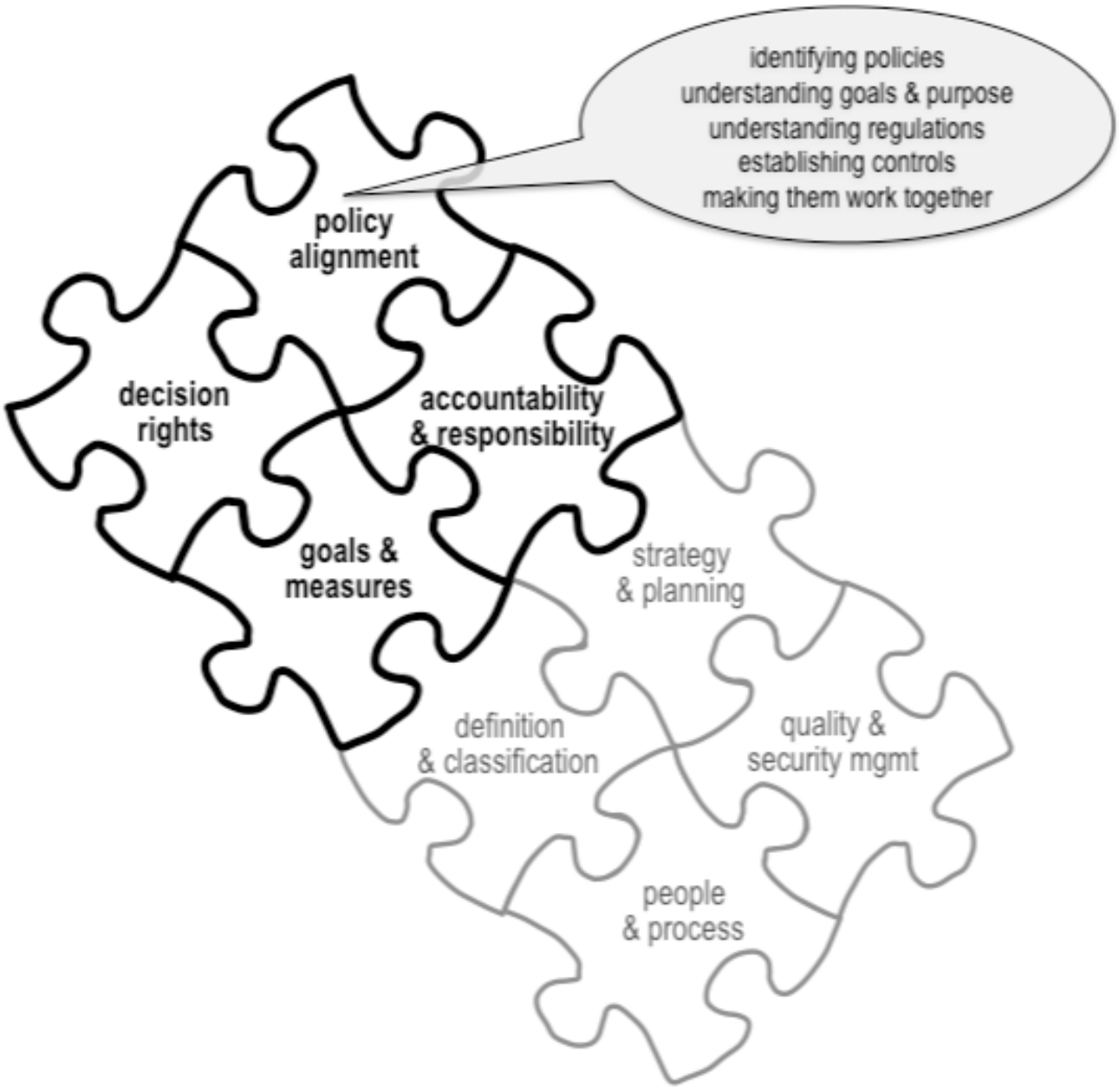
- *Quality and Security Management* is a direct connection to the goals of and motivations for governance. Areas of responsibility include policies, regulations, goals, measures, monitoring, communication, education, and root cause analysis.

- *People and Process* responsibilities are among the most important of data steward responsibilities. Through teamwork, facilitation, and consensus building they form the core of a governance organization.

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Program Development Processes

Policy Alignment



 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Program Development Processes

Policy Alignment

POLICIES THAT REALLY WORK

Data governance begins with data management policies, and developing a governance program starts with policy alignment. Understanding policies is a necessary first step before program development can address topics of decision rights, accountability and responsibility, and goals and measures.

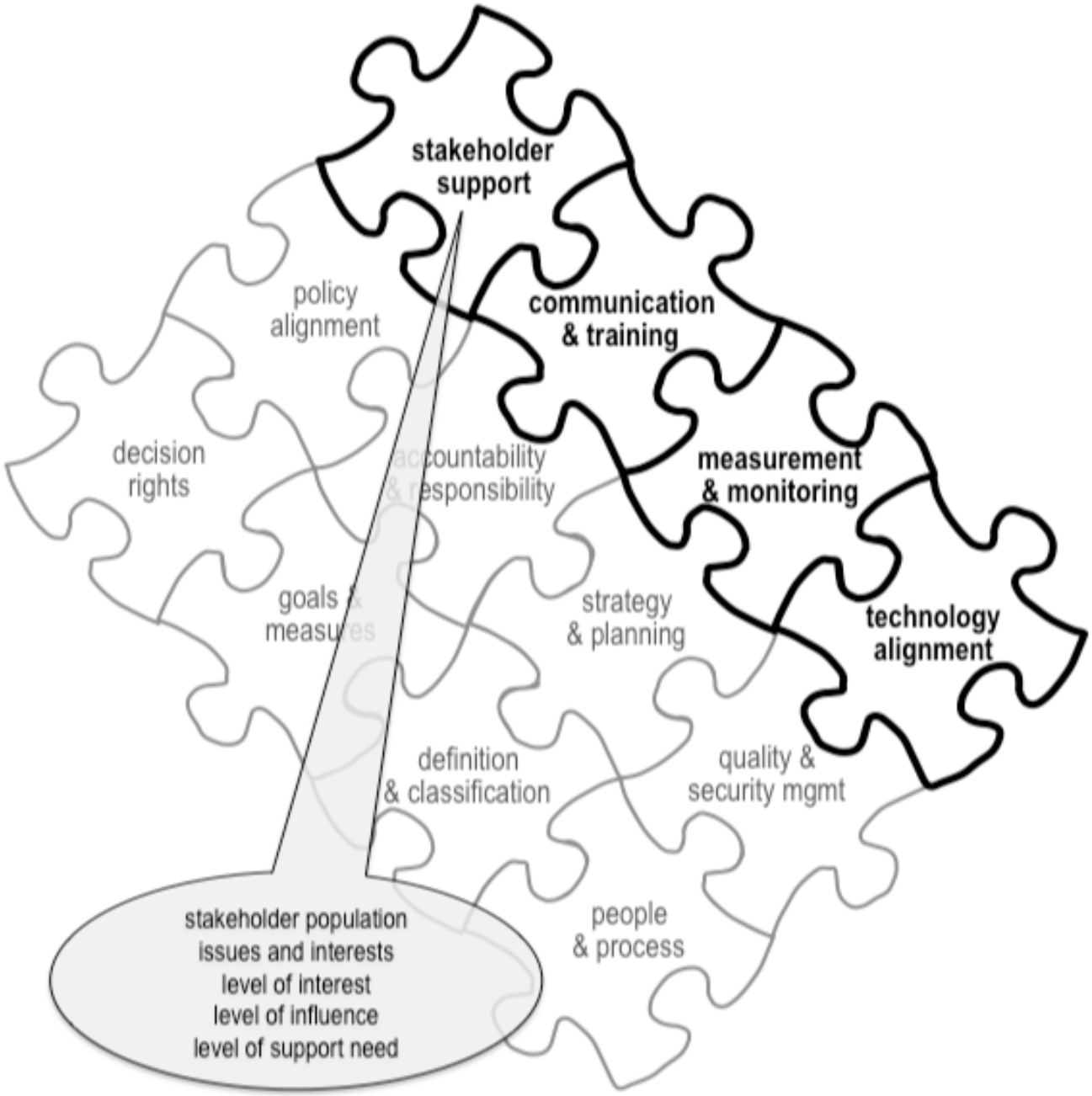
The policy alignment process seeks to:

- Identify data management policies – both those that exist and those that are needed.
- Understand the goals and purpose of each policy and relate them to the motivations for data governance – security, quality, compliance, etc.
- Understand the regulations that drive policy and be clear about those policies that are not discretionary.
- Establish the controls necessary to ensure that policies are known, that they are followed, and that compliance reviews and audits are performed.
- Ensure a workable set of policies by finding and resolving any conflicts or inconsistencies that exist between policies.

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Program Operation Processes

Stakeholder Support



 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Program Operation Processes

Stakeholder Support

CUSTOMERS OF GOVERNANCE

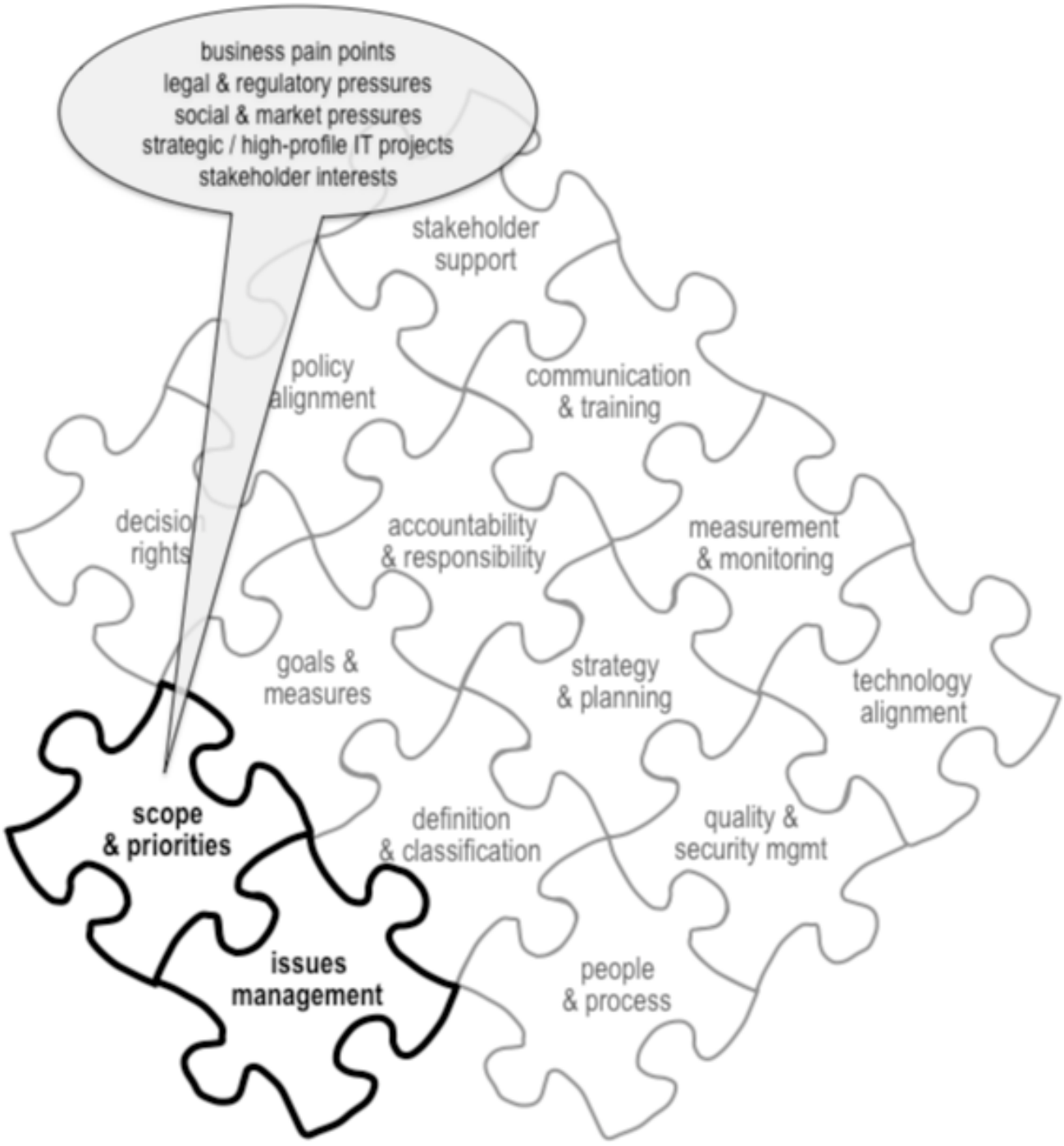
Day-to-day operation of a data governance program begins with stakeholder support – meeting the needs of those who are providers, consumers, and regulators of data. These people are the customers of a data governance organization. Stakeholder support includes processes to:

- Identify the stakeholder population – a sort of census-taking activity. The population may include any or all of legal, regulatory, financial, operational, and competitive stakeholders.
- Understand the issues and interests of each stakeholder. Regulatory stakeholders, for example, may be concerned with HIPAA, SOX, or a variety of other compliance issues. Legal stakeholders may be interested in contract compliance, risk of litigation, etc.
- Classify stakeholders by level of interest, which has strong bearing on both participation and support needs.
- Classify stakeholders by level of influence. It is politically wise to provide top-level support to those who are highly influential.
- Assess the stakeholder population to understand individual and collective support needs.

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Program Sustaining Processes

Scope and Priorities Management



 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Program Sustaining Processes

Scope and Priorities Management

BUSINESS VOLATILITY AND GOVERNANCE

Business is volatile, so static data management policies don't work. Changing business priorities will cause change in the scope and focus of data governance. Scope and priority management matches data governance priorities and activities with business continuously changing business drivers. These drivers are the political, economic, sociological, and technological conditions that create:

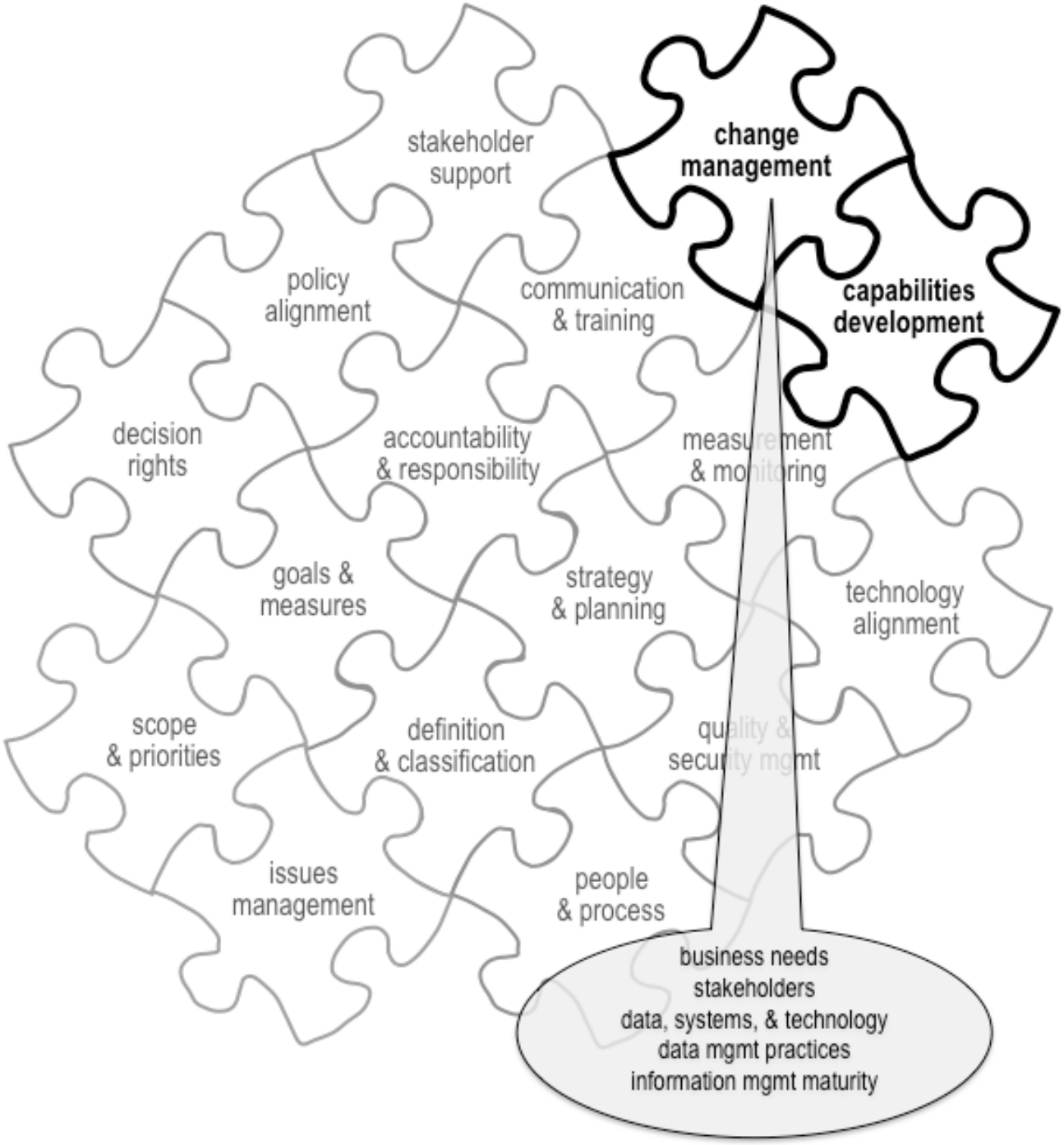
- Business pain points.
- Legal and regulatory pressures.
- Social and market pressures.
- IT strategic initiatives and high-profile IT projects.
- Rapid shifts in stakeholders and stakeholder interests.

Scope & priorities management is a process of responding to changes that are often unanticipated – a sometimes-unavoidable method of reacting to change. Shortly we'll look at proactive change management for the more predictable kinds of change.

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Program Growth Processes

Change Management



 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Program Growth Processes

Change Management

GROWTH BY ACTIVE CHANGE MANAGEMENT

Some changes are predictable and others are not. Some are within our control and others are not. The scope and priorities management process previously described is necessary to react to unanticipated change. But not all change management is reactive. We can often anticipate change, especially changing business needs, stakeholders, data, systems, and technology.

To actively manage change, keep these principles in mind:

- Planned change depends on defined strategy and architecture.
- When change is anticipated it can be managed proactively.
- When externally driven change occurs unexpectedly, it is often possible to anticipate and proactively manage internal changes.
- Change never happens without issues arising. Change management needs a good issues resolution process.
- Organizational change is highly dependent on good communication. Change management needs a good communications plan.
- Desired change (improved data management practices, improved data quality, information management maturity, etc.) doesn't happen by accident. It only occurs when change is managed.

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

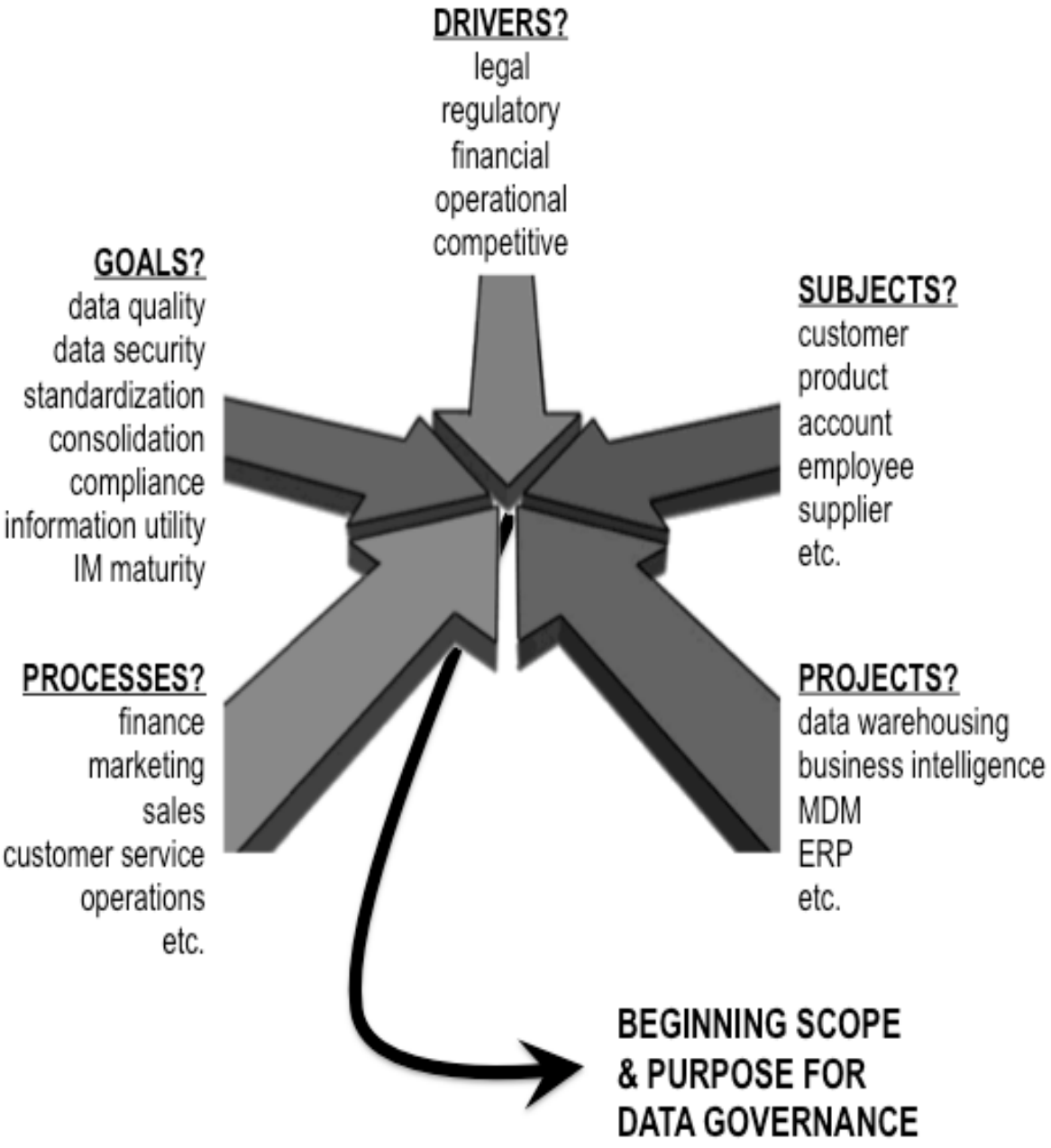
Module 5

Building a Data Governance Program

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Getting Started

Where to Begin?



 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Getting Started

Where to Begin?

START SMALL AND GROW SYSTEMATICALLY

Where to begin? This may be the most vexing question when starting a data governance program. Enterprise data governance can be incredibly complex with so many variables, dimensions, and details that it appears to be an impossible job. To avoid this sense of "boiling the ocean" you'll need to start small and grow systematically.

Define initial scope that is small enough to be achievable – limiting the number of drivers, goals, and subjects. Gain experience and create some success. Then grow the program by expanding the scope incrementally. To contain the scope you will have to make hard decisions – deciding to not immediately address things that are important and that may be urgent. Creating a data governance roadmap – a growth plan with a timeline – can help to control creeping scope.

SETTING SCOPE

Begin scope setting with the question of motivation. Is the desire for data governance motivated by business pressures (legal, regulatory, financial, etc.) or by projects where data quality is a success factor? Consider this question carefully …

- If the answer is *both pressures and projects* then you'll need to make a decision. Choose to focus on one as the initial motivator and defer the other to a later increment.
- If the answer is *business pressures* then limit the scope to one driver, or to a small number of drivers that have natural affinity. Addressing legal and regulatory pressures simultaneously, for example, makes sense. Financial and competitive are also logically connected. Legal and competitive, on the other hand, don't have the same degree of connectedness.
- If the answer is *projects,* ideally limit the scope to a single project. If that isn't practical, don't expand beyond two projects. And be sure that the two projects are logically connected – data warehousing and BI, data warehousing and MDM – or a similar combination.

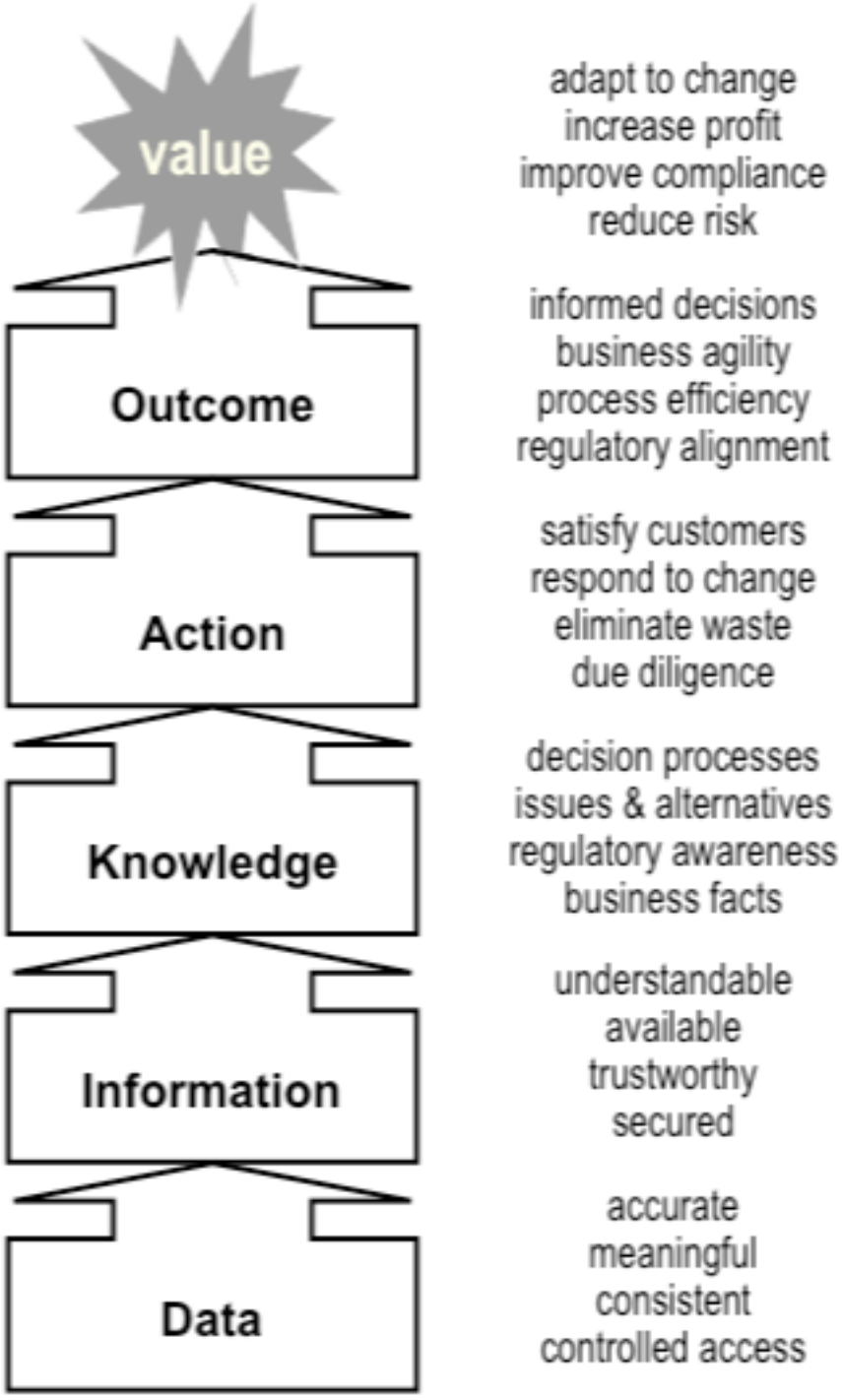
LIMITING GOALS, SUBJECTS, AND STAKEHOLDERS

Whatever your decisions about pressures and projects, it is likely that you'll want to address multiple subjects. Again, limit initial scope to one or two subjects and defer others until later. Similarly, it is probable that you'll have multiple goals. Once again, set limits and seek affinity among multiple goals. Finally, consider the number of processes affected and the corresponding size of the stakeholder population.

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Planning and Preparation

The Business Case



 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Planning and Preparation

The Business Case

GOVERNANCE RATIONALE

The data-to-value chain describes a structure that is useful to define the business case for data governance. Start at the top of the chain with business value and trace downward to arrive at business dependency on data:

- Business value is created by increasing profit, reducing cost, improving regulatory compliance, reducing risk, and adapting to change. Which of these are among your data governance goals?
- Positive business outcomes that drive business value depend upon informed decisions, business agility, process efficiency, and regulatory alignment. Which of these resonate with your business executives and managers?
- The right actions cause positive business outcomes – satisfying customers, responding to change, eliminating waste, auditing compliance, performing due diligence for legal, regulatory, and financial decisions, etc. Which of these are among the goals of your business executives and managers?
- Reliable knowledge leads to confident decisions and taking the right actions. Good decisions occur when they are fact-based decisions with awareness of rules, regulations, risks, issues and alternatives. To what extent does this create a challenge for your decision makers?
- Good information is essential to aware and fact-based decision processes? Information must be understandable, available, and trustworthy. But it must simultaneously be secure and sensitive to privacy considerations. How well does your information measure up?
- Data is the raw material from which information is derived. Good information is only possible when data is meaningful, accurate, and consistent. Information security and privacy needs are only met when access to data is controlled. How effective are your data management practices in meeting these needs?

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Building the Team

Organizational Structure

Where is the organization sponsored?
How large is the initial organization?
What growth is expected?
How formal should the organization be?
What reporting structure – hierarchy, matrix, network?

What should the organization be called:

Data Governance Office?
Data Governance Council?
Data Governance Team?

 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Building the Team

Organizational Structure

DEFINING THE ORGANIZATION

Formal data governance needs to have a defined organizational structure. The organization needs to be named – Data Governance Office, Data Governance Council, or whatever language fits with that of similar or peer-level organizations. Do you have a Project Management Office (PMO)? If yes, then DGO may be a good fit. If you have Compliance Council, Ethics Council, or similarly named organization follow the existing pattern.

Further define the organization by addressing:

- *Sponsorship*

  From what executive is the authority to govern derived?
  Where does the political will to govern reside?
  Who holds ultimate decision rights for data management issues?

- *Size*

  How many data owners to start?
  How many data stewards to start?
  How many data custodians to start?
  What is the expected growth rate and timeline?

- *Structure*

  How formal is the organization?
  Does it have a designated leader such as a Program Manager?
  Does it have assigned workspace?
  How much top-down data management?
  How much bottom-up data management?
  What is the reporting structure?
  Is an organization chart needed?

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Building the Infrastructure

Technology

 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Building the Infrastructure

Technology

SUPPORTING PEOPLE AND PROCESS

The infrastructure of a data governance program begins with technology. Some argue that governance is about people and process – that introducing technology shifts the balance from a business-driven program toward excessive IT influence.

To put technology in perspective, let's go back to the definition of data governance as "effective management of information assets."[1] Data governance is a complex business process of asset management. The business processes for management of financial assets would be difficult without use of technology. The business processes to manage human resources depend on technology. Neither of those examples is perceived as excessive IT influence.

Business processes to manage information assets and data resources also need technology. Imagine trying to track a complex network of regulations, policies, standards, processes, databases, and systems manually.

Technology doesn't govern data; people govern data. Technology helps people to govern effectively – and it must be technology that is more advanced than spreadsheets and word processors. Minimally, data governance requires these kinds of technology:

- Metadata management tools for all types of metadata – classification, description, guidance, and control.
- Data quality technology to examine and cleanse data, and to audit and measure quality.
- Data profilers to capture "real and true metadata"[2] used to understand data content and structure.
- Data security tools to protect data from intrusion and loss.
- Workflow and workgroup tools enabling process integrity, teamwork, and collaborative data management.
- Wikis and portals supporting two-way communication among data governance participants and stakeholders.

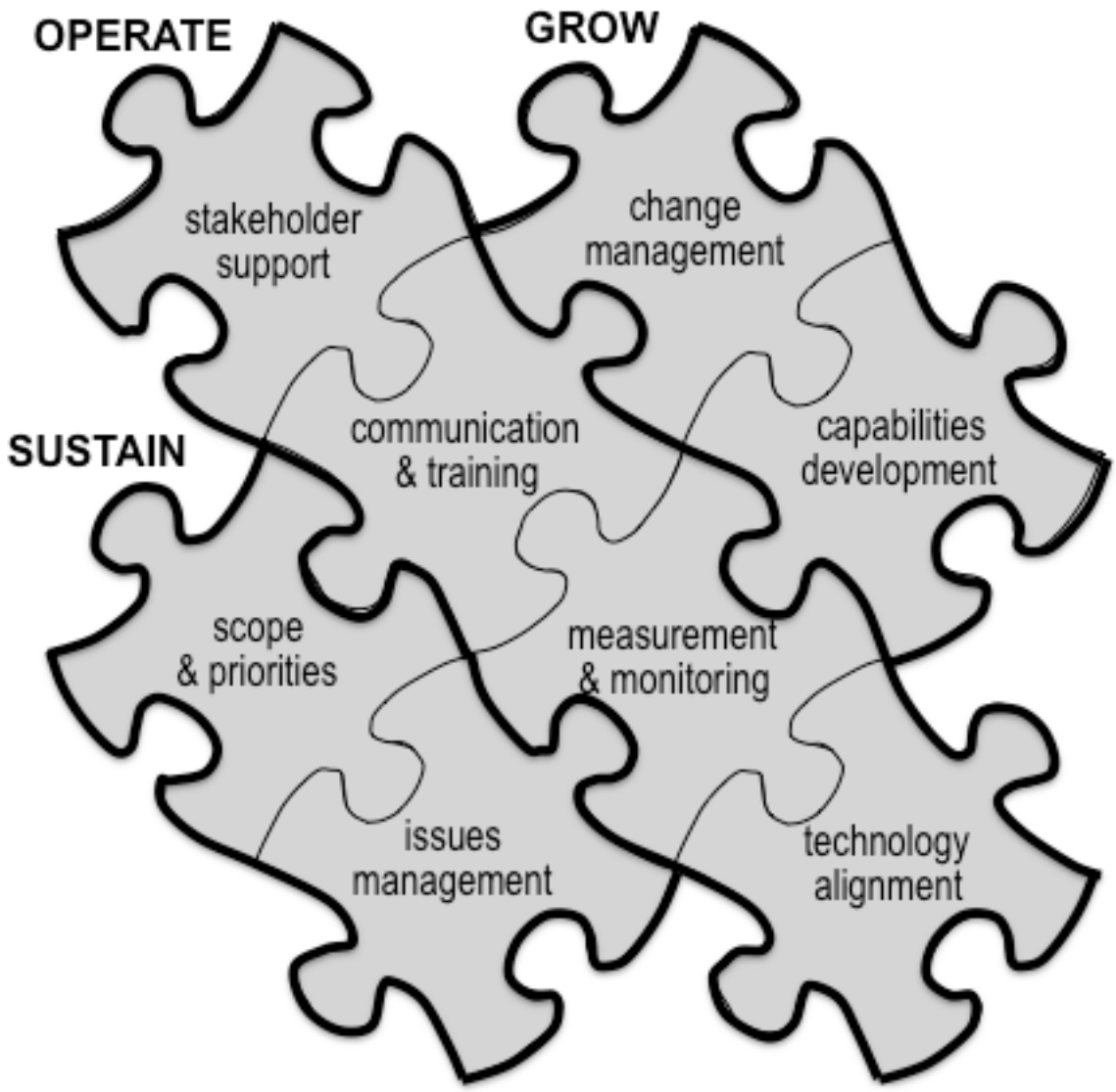
[1] Source: *Executing Data Quality Projects,* McGilvray. The definition is credited partially to John Ladley.

[2] Source: *Three Dimensional Analysis: Data Profiling Techniques,* Lindsey.

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Executing Governance

Program Execution



Policies & practices continuously aligned with business needs
Engaged and involved stakeholders
Knowledgeable and capable data owners, stewards, providers & consumers
Fast & effective resolution of data management issues
Measurement-driven achievement of data governance goals
The right technology for the right reasons

 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Executing Governance

Program Execution

RUNNING THE PROGRAM

Program execution encompasses all of the data governance processes that are not the responsibility of data stewards. Stewardship attends to the business processes of data management. Program execution includes all of the program management processes.

Effective program execution results in:

- Data management policies, standards, and practices that are continuously aligned with business needs and priorities.
- Stakeholders who are actively engaged and involved in governance processes and activities.
- Knowledgeable and capable data owners, data stewards, data providers, and data consumers.
- Fast and effective resolution of data management issues.
- Measurement driven achievement of the stated goals of governance.
- The right data management technology used in the right ways.

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

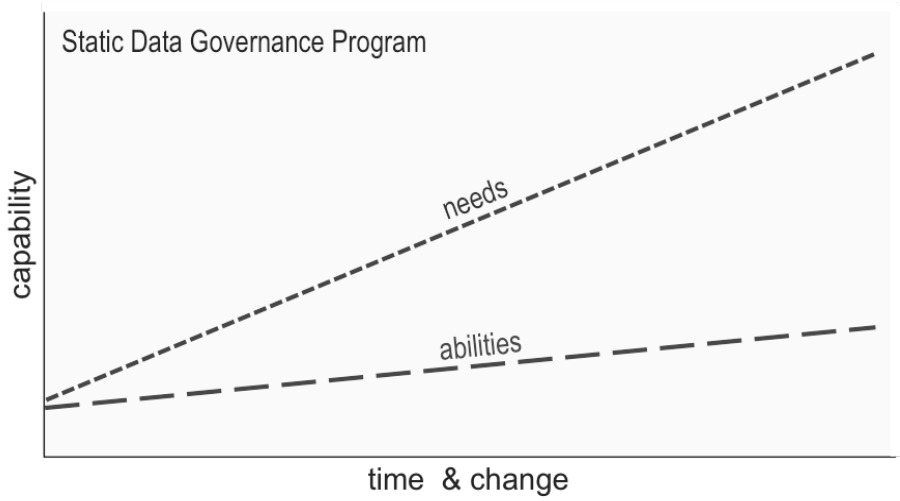
Module 6

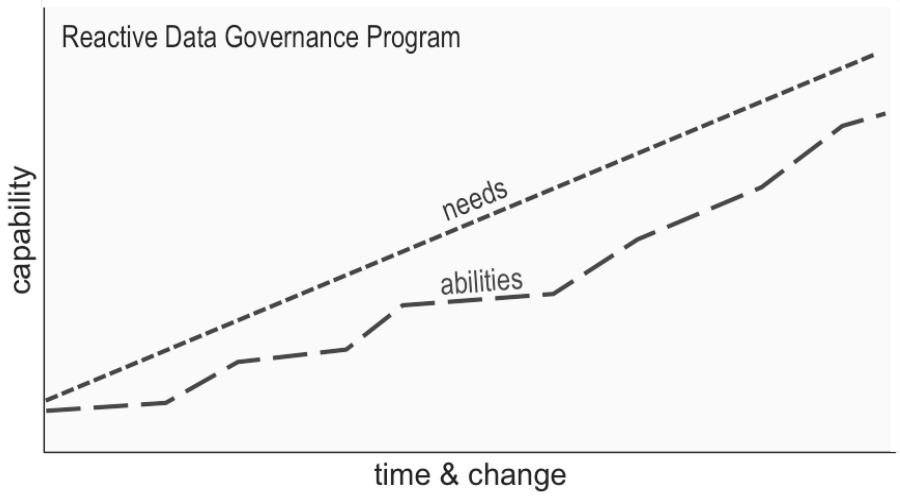
Evolving Data Governance

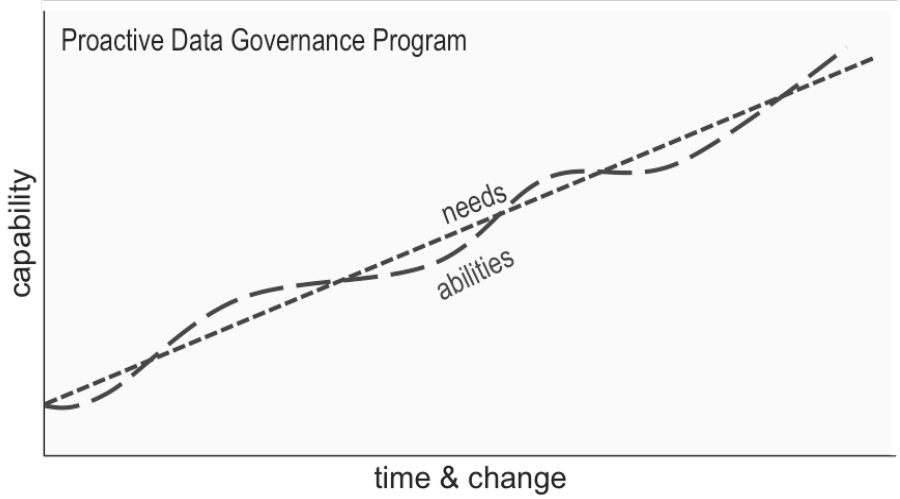
© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Modernizing and Maturing Data Governance

Responding to Change







 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Modernizing and Maturing Data Governance

Responding to Change

CONTINUOUS EVOLUTION

Data governance programs must continuously evolve because the world of data experiences frequent change. The environment in which we create, consume, and manage data experiences change from many sources -- business, data, systems, stakeholders, expectations, regulations, processes, practices, etc. A sustainable data governance program must evolve with the changes

DRIVERS OF MODERNIZATION

There will always be something new and emerging in the field of data and information management. Today's hot topics of data governance interest include cloud technologies and deployments, big data opportunities, and agile BI processes and practices.

Data visualization has yet to become a governance focus, but you can be sure that it will at some time in the near future. And the future holds many other sources of change:

- What about mobile?
- What about virtualization?
- Desktop analytics?
- Social media?
- And what is emerging that is unique to your industry?

SUSTAINING MATURITY

With change comes entropy. Anything left unattended deteriorates including data governance programs. Achieving a desired level of maturity can be challenging. Sustaining that level may be even more challenging. It is easy to lose ground as time passes, emphasis on data governance fades, the pain of poorly managed data is relieved, people become complacent, and data management challenges and priorities shift.

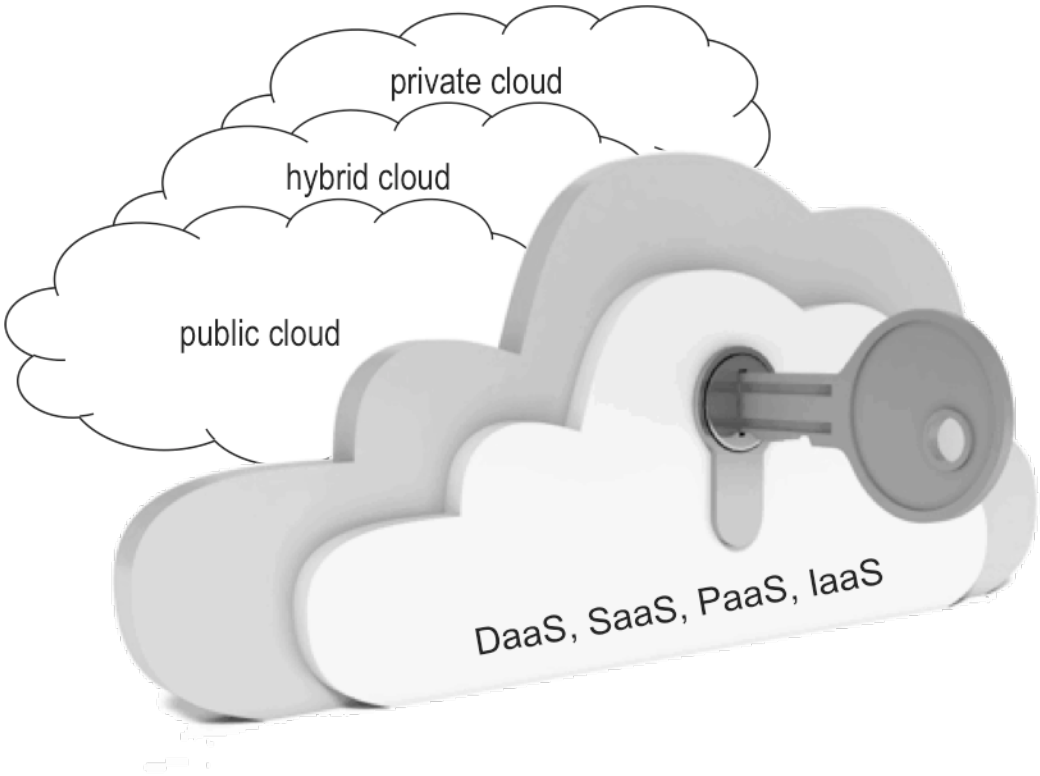
A DATA GOVERNANCE ROADMAP

Making the change from reactive to proactive implies a data governance roadmap that is aligned with your BI roadmap and with those for other enterprise data and information initiatives. Look to the future and think about which things will drive the need for change in data governance – cloud, agile, big data, mobile, visualization, etc. – then define the evolutionary plan to respond to those changes. Evolve to respond to change and also to advance data governance maturity. Consider one or more maturity models to guide the process of maturing a data governance program. Blend proactive change with steps to maturity to build an effective data governance roadmap.

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Cloud Data Governance

Changing the Data Governance Landscape



 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Cloud Data Governance

Changing the Data Governance Landscape

CLOUD SERVICES

Cloud computing is the widely used term for virtualization of computing services. Cloud technology fits into several categories:

Software as a Service (SaaS)	A web-based model for deployment of applications, SaaS allows on-demand use of applications without the need to license and install for every computer where the software is used.
Data as a Service (DaaS)	DaaS hosts data in the cloud. Data marketplace services make data available on a pay-per-use basis. The developer data hub provides web services for developers to upload data, then build web services to work with data. Store and synchronize provides services such as those of Dropbox and similar products.
Platform as a Service (PaaS)	PaaS is an approach to virtualizing the hardware, operating systems, applications frameworks, and technology stacks upon which applications are built and deployed.
Infrastructure as a Service (IaaS)	IaaS is an architecture of virtualized hardware and operating systems. On the surface it sounds a lot like another name for PaaS, but there are some subtle differences. PaaS provides developer environment as well as operations environment but is limited to web applications. IaaS delivers only the operations environment but supports a broader range of applications.

PUBLIC, PRIVATE, AND HYBRID CLOUD

All variations of cloud services are characterized by multi-tenant architecture where many customers use a single instance of servers and software. Multi-tenancy is characteristic of public cloud services. Governance concerns such as security may be mitigated with private and hybrid cloud options.

Private Cloud	Private cloud describes a virtualized environment that is exclusively used by a single tenant. The private cloud may be implemented internally behind corporate firewall, or it may be an isolated, single tenant environment hosted by a cloud service provider.
Hybrid Cloud	Hybrid cloud uses a combination of public and private to gain efficiencies without compromising security and privacy.

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Big Data Governance

Big Data Sources

- ✓ Web and Social Media
- ✓ Machine to Machine (M2M)
- ✓ Big Transaction Data
- ✓ Biometrics
- ✓ Human Generated Data
- ✓ Publicly Available Data
- ✓ Legacy Documents

 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Big Data Governance

Big Data Sources

BIG DATA DEFINED

"Big data" is a term that has become popular to describe rapid growth in the volume, variety, and velocity of data that is now available in business – unstructured data, semi-structured data, social media data, location data, radio frequency data, and more. These types of data tend to yield data sets that are too large, complex, or unwieldy to work with traditional data management and analytics technologies.

SOURCES OF BIG DATA

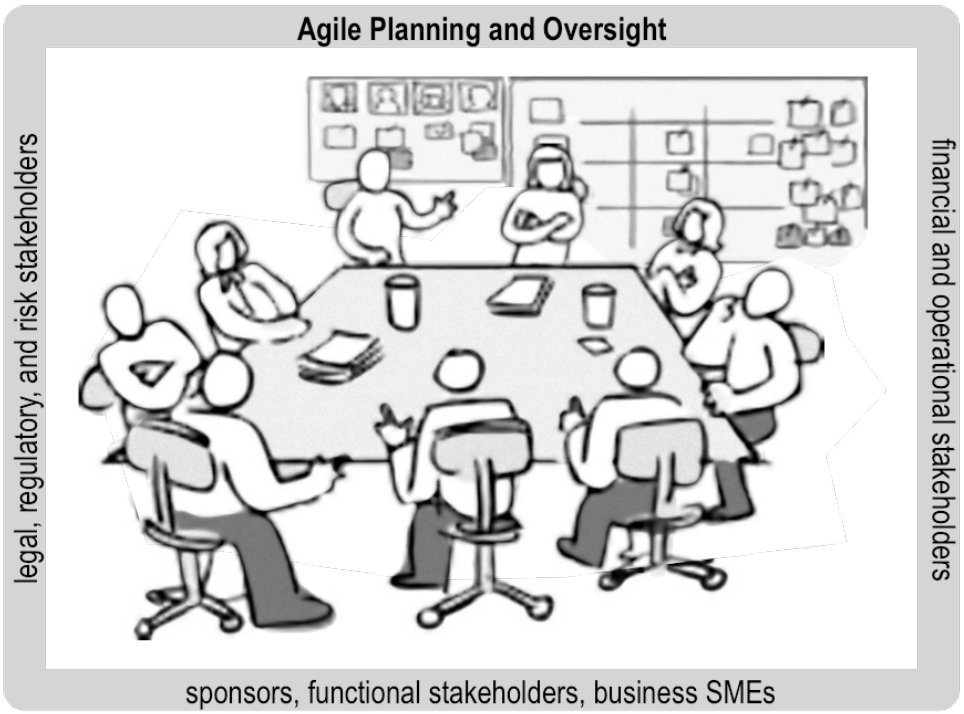
Big data can be acquired from a variety of sources in many different forms that include structured, semi-structured, multi-structured, and unstructured data. Text, images, audio, geospatial data, and tagged data are common forms of big data. Among the common big data sources are:

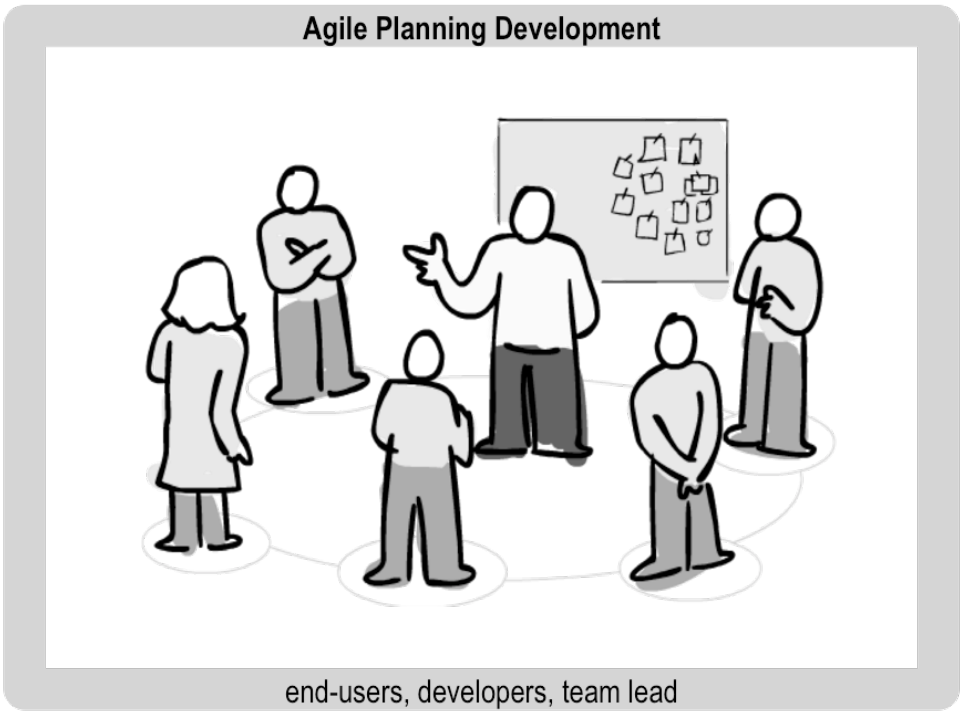
- Web and social media data
- Machine-to-machine and sensor data
- Big transaction data
- Biometrics
- Human generated data
- Publicly available data
- Legacy documents

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Agile Data Governance

Agile Teams





 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

Agile Data Governance

Agile Teams

TEAM AND PROCESS ALIGNMENT

As we just saw, the agile process has planning activities (strategy and release) and development activities (iteration and continuous building). It makes sense to organize agile teams to align with the agile process with a planning and oversight team and a separate development team.

PLANNING AND OVERSIGHT

The planning and oversight team is made up of sponsors, stakeholders, and business subject matter experts. This team has responsibility for vision, goals, funding, and project chartering. Their work is focused on strategy and release activities and they participate in iteration activities as a bridge from planning to development.

DEVELOPMENT

The development team is a collaboration of end-users and developers working under the guidance of a team lead. This team is responsible to discover requirements, design solutions, and build systems. Their work is focused on continuous development, daily planning, and iteration as a bridge to planning and oversight activities.

GOVERNANCE

In this team structure, governance begins with the planning and development team. These stakeholders – functional, legal, regulatory, risk, financial, and operational – are positioned to understand needs of governance and to express those needs as constraints, expectations, and requirements for the development team. The development team is responsible to include governance needs in the development process, and communicate with stakeholders to fully understand governance goals and constraints.

© TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY

 © TDWI. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. DO NOT COPY