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 Business Intelligence Principles and Practices

Charting the Course to BI Success

The Data Warehousing Institute 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:

email: info@tdwi.org

Publication Date: February 2013

TABLE OF CONTENTS

COURSE OBJECTIVES

To learn:

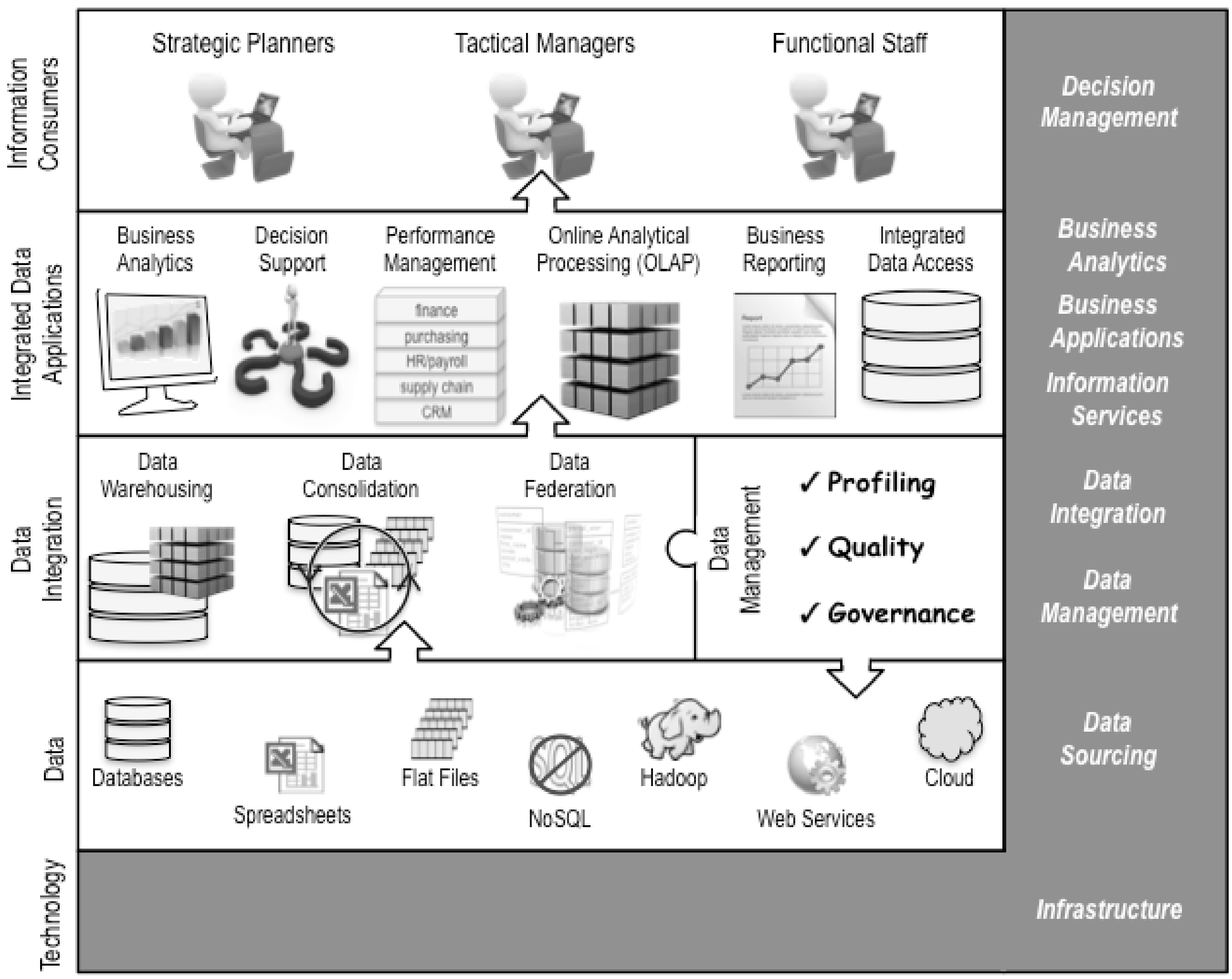
- ✓ ***Meaningful and actionable definitions of business intelligence (BI)***
- ✓ ***Effective ways to deliver BI – web, mobile, desktop, etc.***
- ✓ ***Common kinds of BI reporting – ad hoc, published, enterprise, and operational***
- ✓ ***Performance management principles including dashboards, scorecards, and KPIs***
- ✓ ***Business analyst principles including OLAP, analytic modeling, and data visualization***
- ✓ ***Advanced analytics concepts for data mining, predictive analytics, and text analytics***
- ✓ ***Data management practices including profiling, cleansing, and quality management***
- ✓ ***Data integration practices – consolidation, virtualization, data warehousing, etc.***

Module 1

Introduction to Business Intelligence

BI Components

Data and Technology



BI Components

Data and Technology

DATA

Data is the raw material from which information is created. Early data warehouses were often based on the assumption that the core business transaction systems were the entire scope of data sources. In today's BI world many other sources of data are valuable and sometimes essential. When identifying source data consider a broad range of possibilities including:

- Databases, both internal to your organization and externally available through syndication and commercial services, are key data sources. Internal databases are likely to be the foundation for much of your information services. External databases enrich the data and offer opportunities to improve data quality.
- End-user data including the many spreadsheets and occasional databases built to meet individual and departmental needs may be more current and detailed than corresponding data in corporate databases. In some instances end-user data is the only available source of data needed for analysis.
- Flat files found throughout information systems are often overlooked. These are frequently the interface files between disparate applications and may have already accomplished some steps toward integration.
- Unstructured data such as text, documents, and images may enrich the data resource particularly when data is needed for business analysis.
- A variety of data opportunities found on the Internet and embedded in e-mail communications also offer data enrichment possibilities.

Module 4: Data Integration describes more about data types and sources.

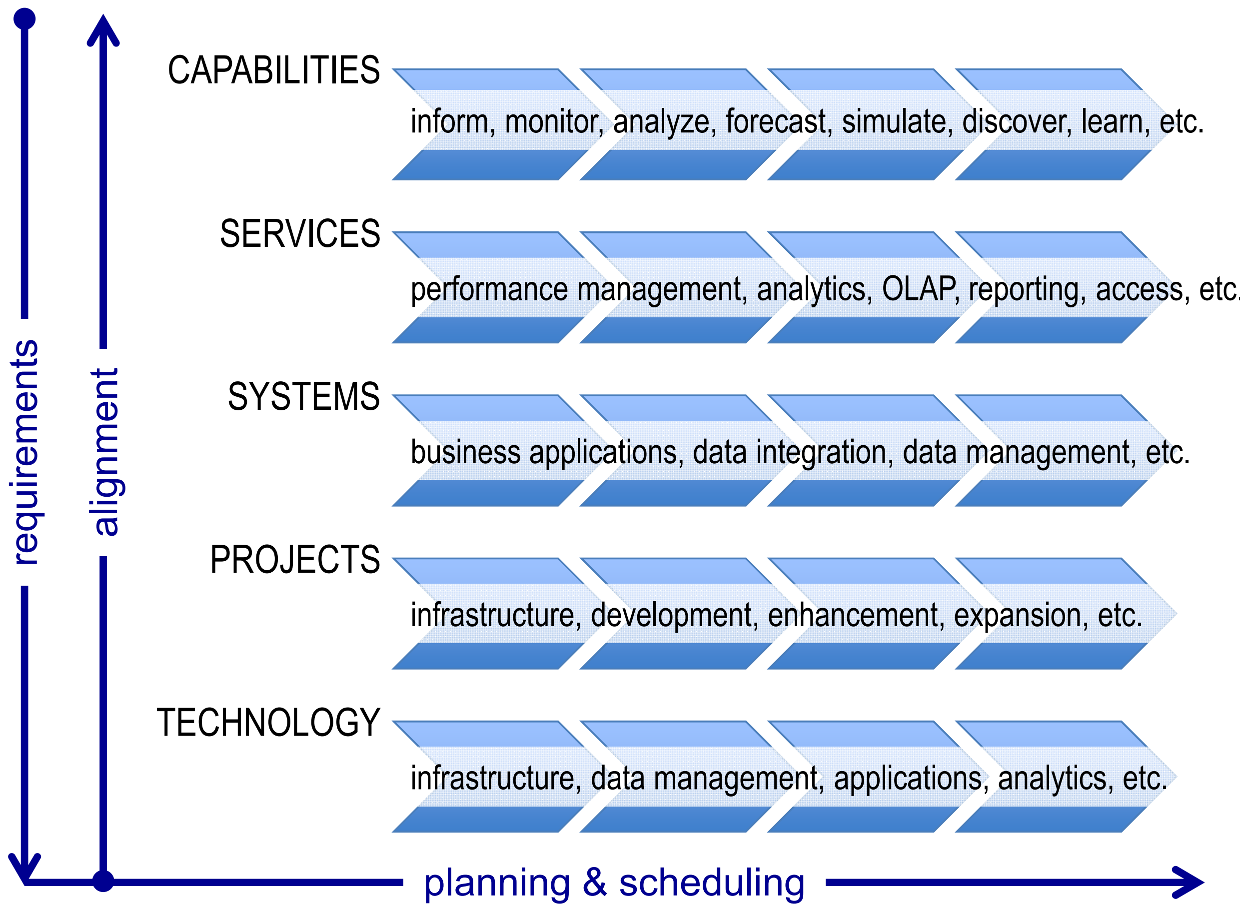
TECHNOLOGY

A variety of technology is needed to implement and operate BI systems, including support for BI infrastructure, data sourcing, data management, data integration, information services, business applications, business analytics, and decision management.

Module 6: BI Technology describes the technology types, features, and functions at each level of the BI technology stack.

The BI Roadmap

Parallel Paths



The BI Roadmap

Parallel Paths

PLANNED EVOLUTION

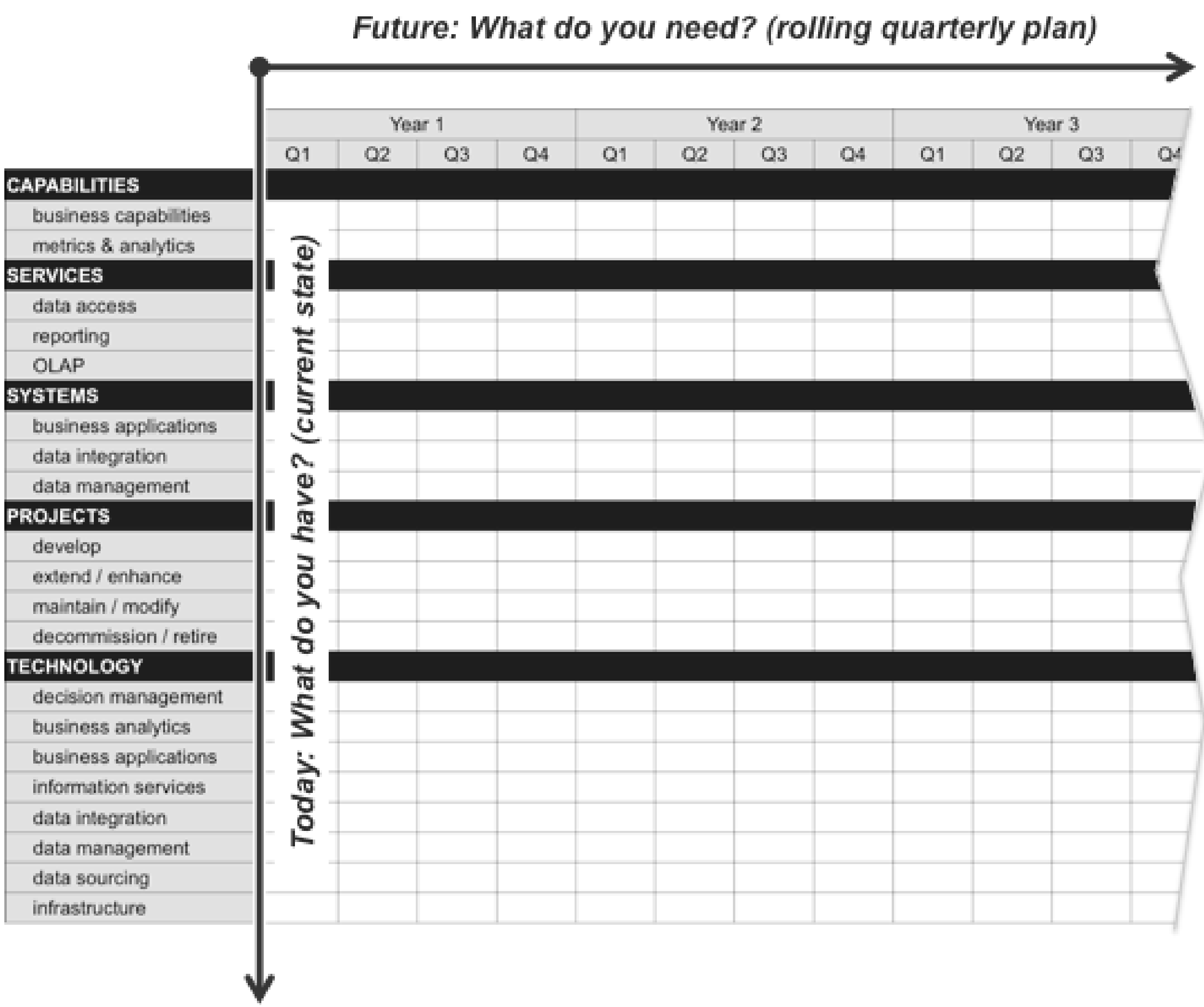
The BI roadmap is a plan for managed evolution of BI capabilities and systems. It is a timeline view of upcoming BI projects and activities needed to move from current state to future state as guided by portfolio management. The roadmap begins with planned capabilities, and is a process of prioritization, dependency management, and planning that answers questions in four areas:

- BI Services – What services are needed to fill gaps or meet future business requirements? When are they needed? In what sequence?
- BI Systems – What systems need to be developed, enhanced, or modified to deliver the needed services? When and in what sequence?
- BI Projects – What projects must be planned and executed to produce the needed systems? What project dependencies exist? When and in what sequence should projects be scheduled?
- BI Technology – What technologies are needed to support and enable planned BI projects? When and in what sequence should technology be deployed? What projects must be planned and executed to deploy the technology?

The roadmap illustrates four parallel tracks – services, systems, projects, and technology – mapped to planned business capabilities and to a corresponding timeline.

The BI Roadmap

Continuous Planning



The BI Roadmap

Continuous Planning

CURRENT STATE

The roadmap begins with a view of the current state. Planning for the future depends on first understanding what you have today. To begin the process of developing a roadmap, you may need to begin with an inventory of capabilities, services, systems, projects, and technology.

FUTURE STATE

The future view is a look ahead that is typically two to three years in duration. The plan begins by looking at needed capabilities and plotting them on a timeline. Each capability then cascades to identify:

- Information services that are needed to enable each capability
- Systems that are needed to deliver information services
- Projects to build the needed systems
- Technology that is required by the systems, services, and capabilities

ROLLING PLAN

Planning is a continuous process of responding to change in business needs and technology growth. As BI systems evolve and technical growth continues, the plan must be adapted and adjusted. A rolling quarterly plan works well for the BI roadmap. Plot the plan across quarters with greatest confidence and detail for the immediate next quarters and fewer specifics as you move further into the future. Then update the plan once each quarter by:

- Updating the current state inventory
- Adapting and adjusting to change
- Adding detail and specifics for upcoming quarters
- Extending the timeline by one quarter to sustain the 2-3 year look ahead

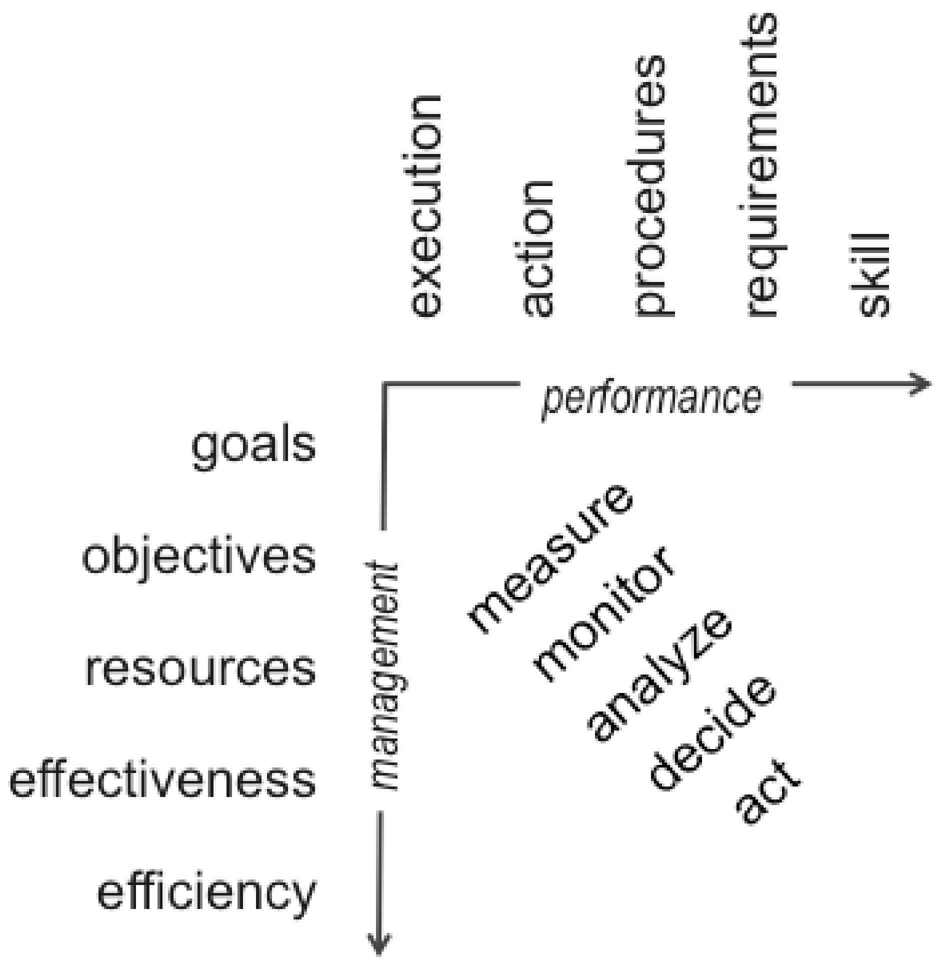
THE DATA WAREHOUSING INSTITUTE

Module 2

Business Metrics and Analytics

Performance Management

Definition and Concepts



Performance Management

Definition and Concepts

PUTTING THEM TOGETHER

Combining the key elements of performance – execution, action, procedures, requirements, and skills – with the key concepts of management – goals, objectives, resources, efficiency, and effectiveness – begins to give some shape and structure to the processes of performance management. At the intersection of performance and management we need processes to:

- Measure and monitor performance
- Analyze performance measures
- Decide what to do and take action

PERFORMANCE MANAGEMENT DEFINED

Gary Cokins defines performance management concisely as "the translation of plans into results – execution. It is the process of managing an organization's strategy"[1]

Wayne Eckerson says performance management "consists of a series of processes and applications designed to optimize the execution of business strategy… a framework that takes the long-standing task of *measuring performance* to the next level, that of *managing performance*."[2]

Frank Buytendijk says, "Performance management tries to capture an organization's business model. As it becomes clear how various business domains affect the business results, performance management provides insight into who drives results and how results are driven."[3]

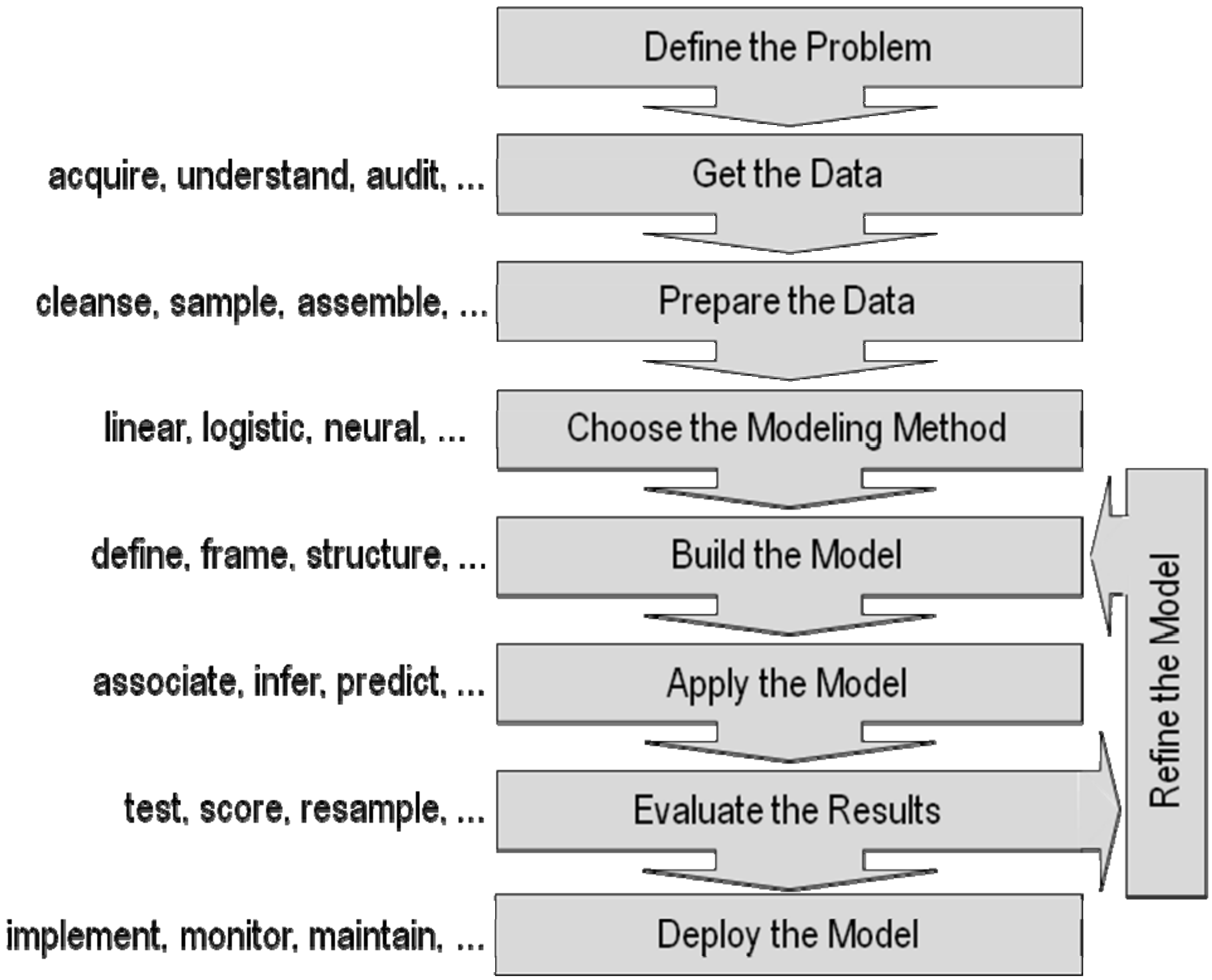
While each of these definitions offers a slightly different perspective, they all support some common themes about performance management:

- Performance management focuses on execution of strategy.
- Measurement is only a part of performance management.
- Analysis, especially cause-and-effect analysis – is a key part of performance management.
- The right actions and results are the goals of performance management.

[1] *Performance Management,* pp. 9, Cokins
[2] *Performance Dashboards,* pp. 11, Eckerson
[3] *Performance Leadership,* pp. 17, Buytendijk

Business Analytics

Analytic Modeling



Business Analytics

Analytic Modeling

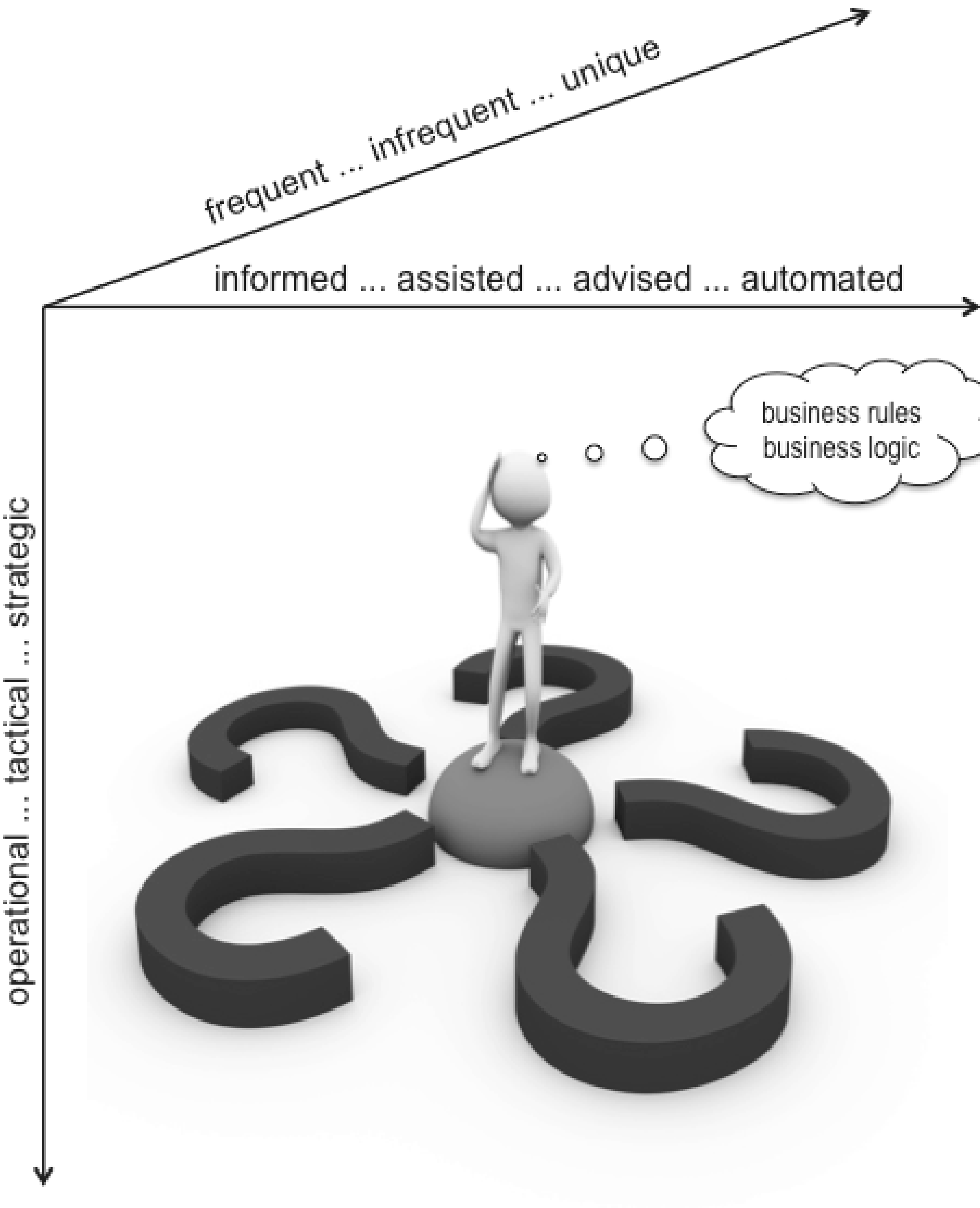
DEVELOPING AND USING ANALYTIC MODELS

The range of skills needed to develop and apply analytic models is quite extensive. The diagram on the facing page illustrates a typical modeling process and identifies some of the skill areas at each step of the process:

- Define the problem
  - Knowledge of how the business works
  - Problem framing – who, what, and where
  - Influence diagramming
- Get the data
  - Data acquisition
  - Data profiling
  - Data quality assessment
- Prepare the data
  - Data cleansing and quality improvement
  - Data sampling – bias, statistical significance, etc.
  - Data structures, migration, and manipulation
- Choose the modeling technique
  - Knowledge of statistical methods
  - Application of various methods
- Build the model
  - Define the purpose(s) of the model
  - Frame the model – variables, functions, constraints, etc.
  - Structure the model – spreadsheets, visualizations, outputs, etc.
- Apply the model
  - Association and inference
  - Prediction, forecasting, simulation
  - Hypothesis affirmation or contradiction
- Evaluate the model
  - Model testing – reality and test cases
  - Model scoring – precision, accuracy, reliability
- Refine the model
  - Functional refinement
  - Calibration and confidence
- Deploy the model
  - Production implementation for recurring execution needs
  - Change monitoring and model maintenance

Advanced Analytics

Decision Management



Advanced Analytics

Decision Management

ACTION FROM ANALYTICS

Decision management systems are among the most advanced forms of business analytics – putting analytics to work because value is created through action, not through information alone.

DECISION TYPES

Business decisions may be strategic, tactical, or operational. Some decisions are frequent and recurring, while others are infrequent and occasional or unique. Operational decisions tend to be more repeatable – frequent and recurring – than strategic and tactical decisions.

DECISION-MAKING TYPES

Decision-making processes rely on BI in a variety of ways. They may be:

- Informed – BI provides relevant information to decision makers.
- Assisted – BI provides relevant information to and simulates. outcomes of alternatives for decision makers.
- Advised – BI provides relevant information, predictive analysis, and recommendations to decision makers.
- Automated – A decision management system (a component of BI) is the decision maker

DECISION MANAGEMENT SYSTEMS

Decision management systems target recurring operational decisions that are specific to a customer or a transaction. Analysis and modeling of a group of similar decisions identifies the logic, business rules, and data points that shape the decisions. Decision services are implemented to encapsulate logic and business rules, and to access and apply data to the decision processes. Logic, rules, and data are applied to determine which of the decision-making types is applied for each individual decision.

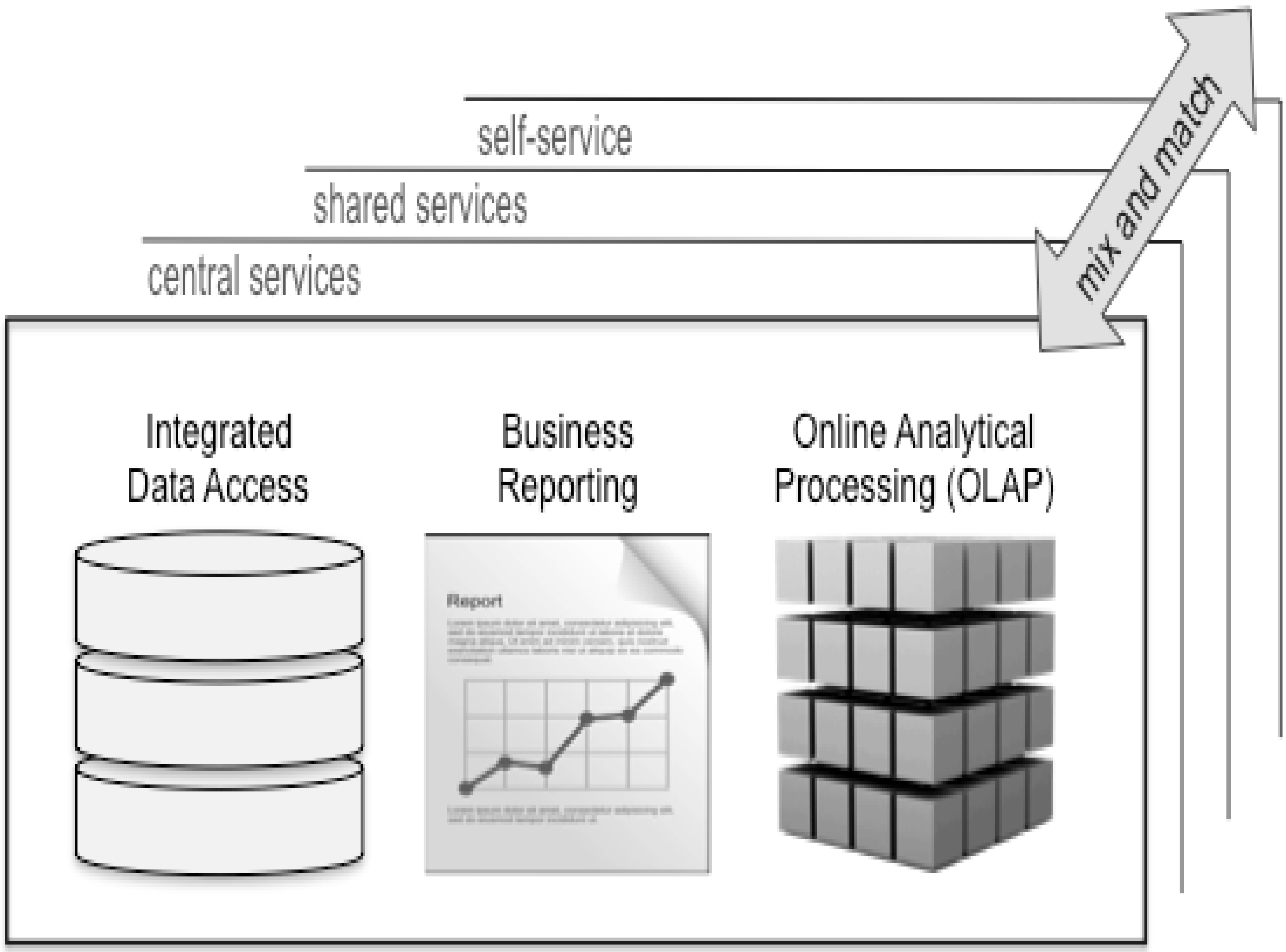
According to James Taylor, CEO of Decision Management Solutions, "Decision Management Systems are agile, analytic and adaptive. They are agile so they can be rapidly changed to cope with new regulations or business conditions. They are analytic, putting an organization's data to work improving the quality and effectiveness of decisions. They are adaptive, learning from what works and what does not work to continuously improve over time." (Decision Management Systems Platform Technologies Report Version 2, Update 4, January 4, 2013 by James Taylor, © 2012 Decision Management Solutions)

Module 3

Information Services

Information Services Layers

Three-Layer Service Model



Information Services Layers

Three-Layer Service Model

CENTRAL SERVICES

Central services are the "we build it for you" model that works well for standard reports and routinely published information. In the central services model, standards, processes, and technology are prescribed. A single centralized team is responsible for development, deployment, and management of information services. This model works well when goals are exceptional consistency, strong governance, rapid delivery, and managed costs. The central services model may be challenged to scale up to meet high demand for services.

SHARED SERVICES

Shared services is the "we build the Legos" model where a central team builds and publishes reusable data components that are accessed, configured, and assembled by distributed teams to meet their local needs. With published interfaces it is practical for local data to be appended to or integrated with central data. The shared services model defines processes, standardizes architecture, and maintains a centralized team for shared work, but much project and process work occurs in individual project teams and distributed business units. The blend of centralized and decentralized resources achieves good efficiency of resource utilization.

SELF-SERVICE

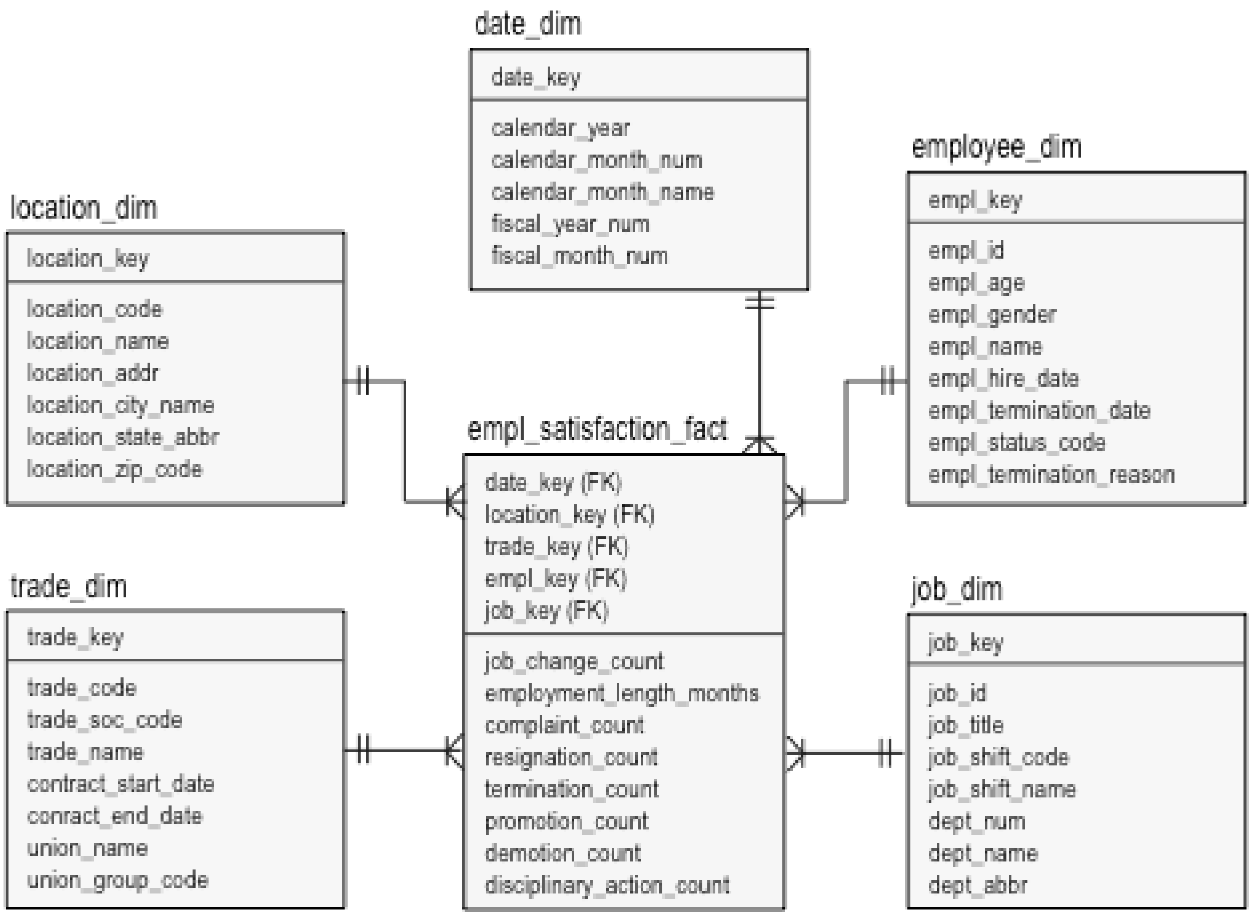
Self-service is a "we provide data access" model. It is an environment where business units meet their own information needs with access to high-quality integrated data and support of BI standards, architectures, frameworks, guidelines, examples, templates, etc. This model is suited to well-defined problem domains where data consumers have a desire for autonomy and a relatively high level of BI skills.

HYBRID SERVICE MODELS

As a practical matter, many organizations evolve to a mix-and-match hybrid of the three service models. Good guidelines and clear understanding of the criteria by which projects and service models are matched is important to ensure appropriate use of each level.

OLAP

Dimensional Data Marts and Star Schema



OLAP

Dimensional Data Marts and Star Schema

DATA MARTS

Dimensional data marts are the data foundation for OLAP. A data mart is a set of data tailored to support the analysis and reporting requirements of a business unit, business function, or work group. A dimensional data mart organizes the data as a collection of related facts (typically business measures) that are associated with analysis dimensions.

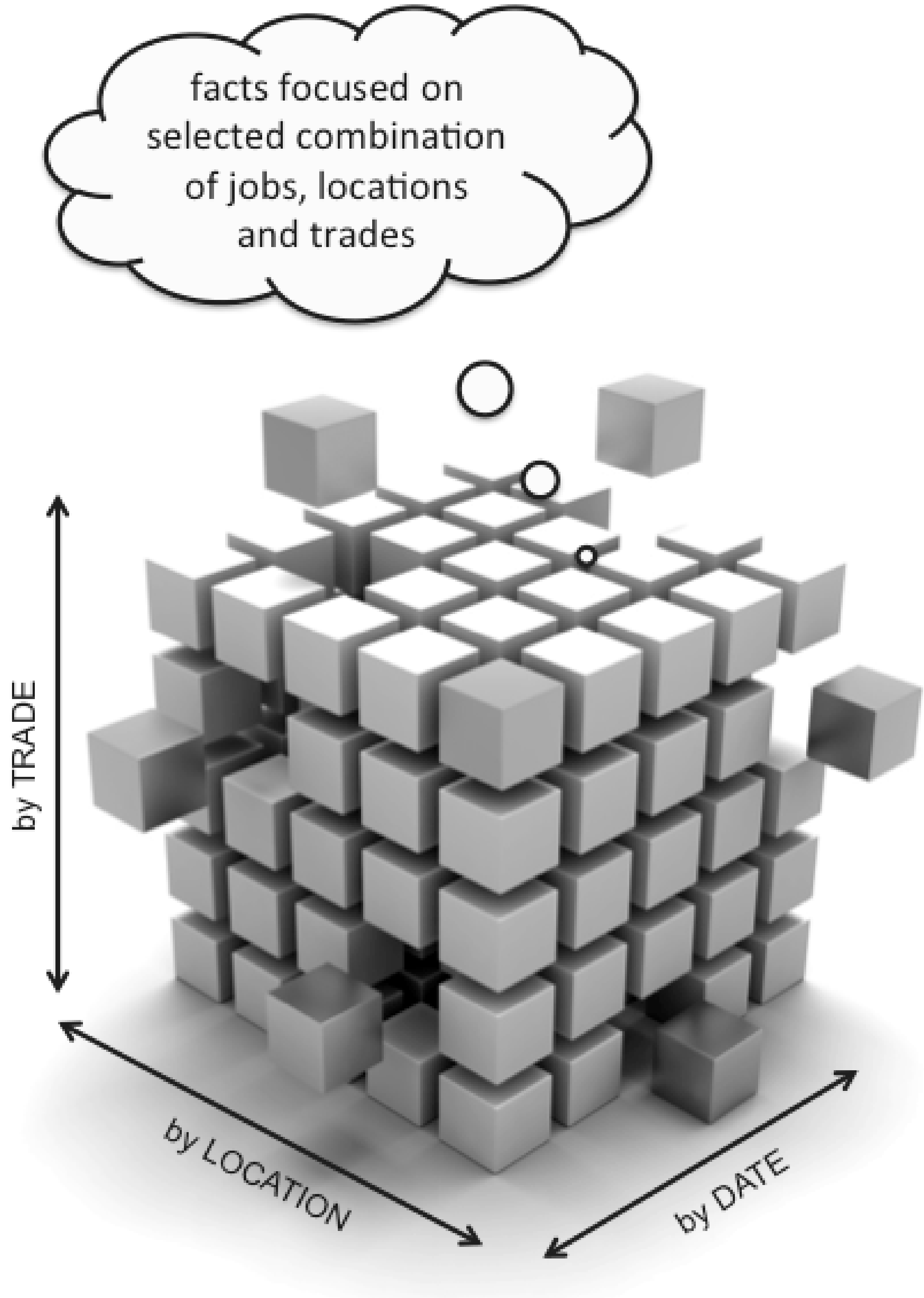
In the example on the facing page *empl_satisfaction_fact* contains facts about employee satisfaction. Those facts can be selected, grouped, and summarized by any combination of date, location, employee, trade, and job as supported by the dimensions.

STAR SCHEMA

A star schema is the design of relational tables to collect and store multi-dimensional data. Dimensional data marts are built upon star schema design. A star schema consists of a single fact table surrounded by a set of dimension tables. The example on the facing page has one fact table – *empl_satisfaction_fact* – that is associated with five dimension tables: *date_dim, location_dim, empoloyee_dim, trade_dim,* and *job_dim.* The schema can be implemented using RDBMS, but foreign-key relationships are constrained to those that associate the fact table with each of the dimension tables.

OLAP

Slice-and-Dice Analysis



OLAP

Slice-and-Dice Analysis

THE OLAP INTERFACE

Selecting which dimensions are in view and on which axis is achieved using a pivot-table-like interface. The OLAP cube (or star schema) with this interface supports several operations:

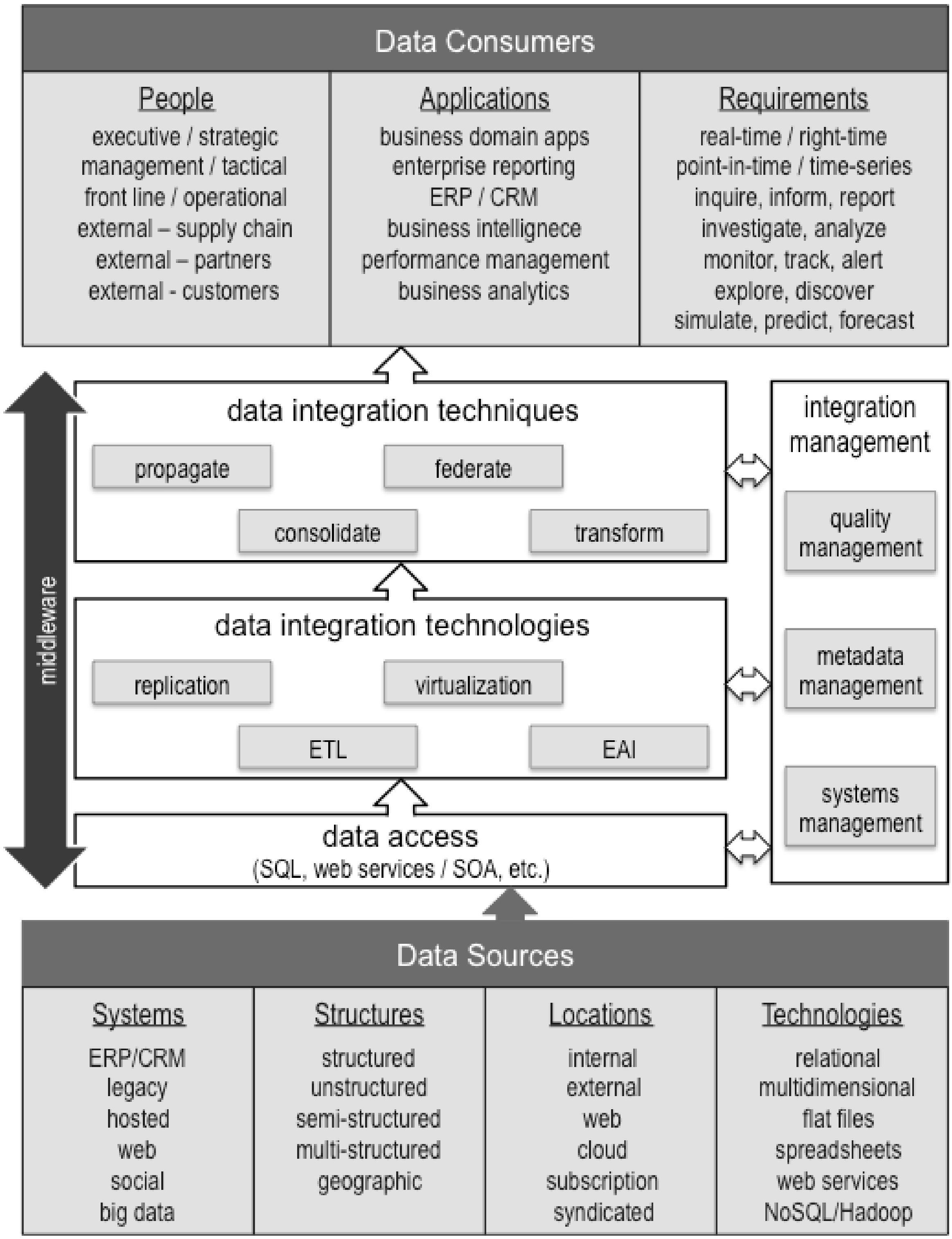
- Slice – Slicing selects a subset of the data by limiting data to a single value for one dimension. Slicing produces a two-dimensional array.
- Dice – Dicing selects a subset of the data by limiting the values that are active in two or more dimensions. Dicing produces a three-dimensional array that is in effect a smaller cube.
- Drill Down – Drilling moves from summary to greater detail, resulting in finer grain for the data that is in view. Drill is an important analysis function as most analysis begins by looking at summary data then drilling to detail where questions arise.
- Roll Up – Rolling up summarizes data to show totals at varying levels along a hierarchical dimension.
- Pivot – Pivot rotates a cube in space to change which dimensions are positioned on the x, y, and z axes of a three-dimensional view.

Module 4

Data Integration

Data Integration Architecture

Integration Techniques and Technologies



Data Integration Architecture

Integration Techniques and Technologies

TECHNIQUES

Several techniques are combined to achieve data integration goals:

- Propagation creates copies of data according to a set of rules. Source data from an operational database, for example, might be propagated to a staging area in preparation for data warehouse processing.
- Federation creates integrated views of data that is not physically integrated. Overlap and inconsistency among disparate data sources is rationalized and reconciled as views are created. Federated data sources are able to operate autonomously yet appear to be integrated. Federation might be used, for example to build a quick on-the-fly data mart without the time and cost of physical integration.
- Consolidation builds a physical data store of integrated data that is a distinct separate copy of the original data sources. Overlap and inconsistency among disparate data sources is rationalized and reconciled as data is copied and moved. Data warehouses consist primarily of consolidated data.
- Transformation is the processing that changes data during federation and consolidation. The primary role of consolidation is to rationalize and reconcile inconsistencies among data sources.

TECHNOLOGIES

Data integration technologies support the techniques described above. Replication technology creates copies of data and is used in propagation. Virtualization technology is abstraction- and view-based data integration that is used to federate data. Extract-transform-load (ETL) technology is used to consolidate data. Enterprise application integration (EAI) is a message-bus means of communicating among systems that can be used for extended capabilities in propagation, federation, and consolidation.

ACCESS

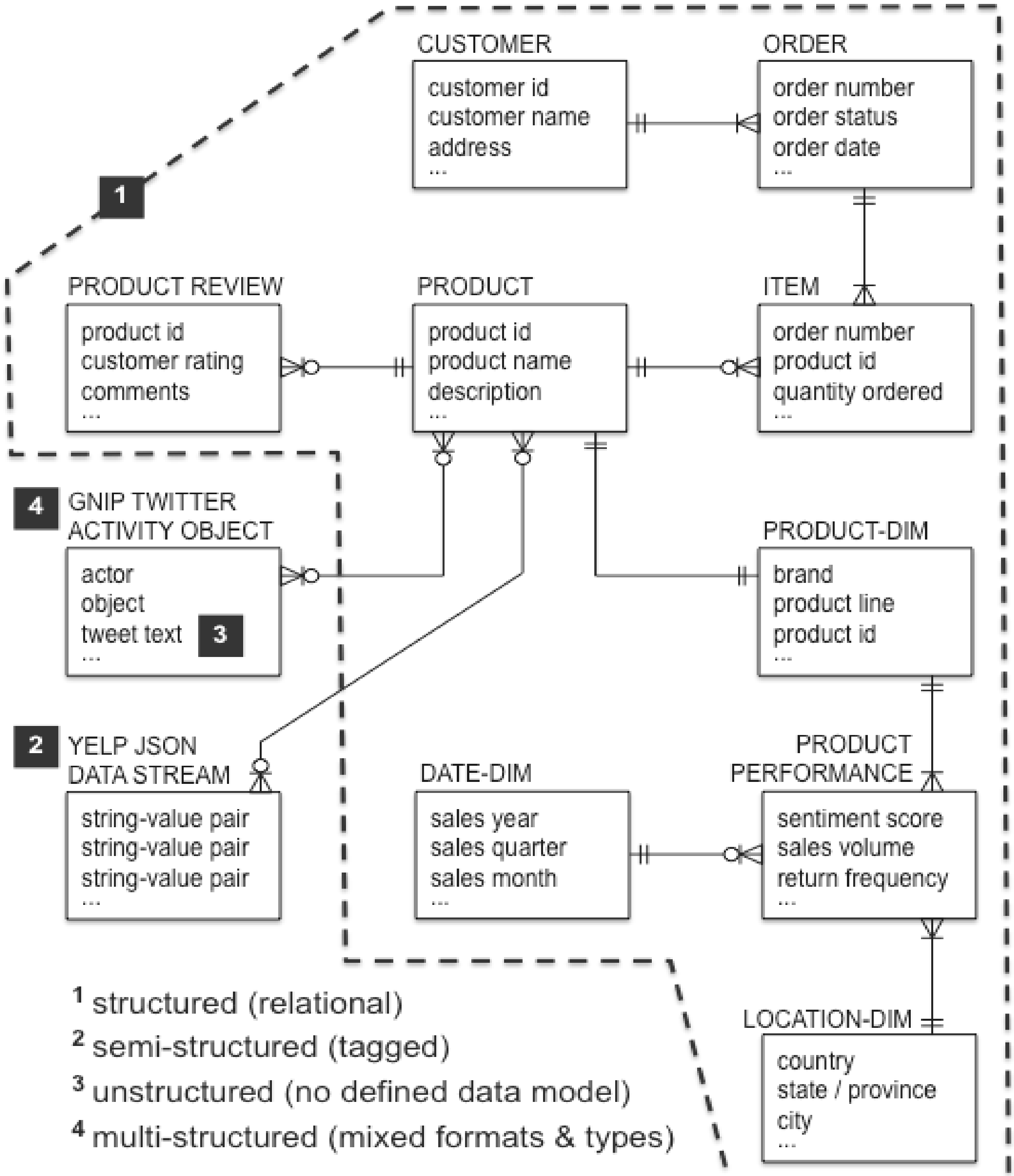
Data access includes all of the technologies, formats, and protocols that are necessary to connect with and acquire data from sources. The variety of query languages, connectors, and services depends on the kind of data sources that are accessed.

MANAGEMENT

Data integration is achieved through complex systems that require management of quality, of metadata, and of the systems themselves. Data integration systems are particularly sensitive to change – change in data sources, in target destinations of data, in business requirements, and in technology. The management components are essential to sustainable data integration systems.

Data Types and Sources

Structured and Unstructured Data



Data Types and Sources

Structured and Unstructured Data

STRUCTURED

Structured data can be described using the relational data model. Flat files, spreadsheets, relational databases, and multi-dimensional databases are examples of structured data.

UNSTRUCTURED

Unstructured data doesn't necessarily mean data without structure. Some argue that all data has structure with the exception of a string of randomly generated bits. Instead, unstructured data does not have a defined data model or doesn't fit the relational model. Text and images are common examples of unstructured data.

SEMI-STRUCTURED

Semi-structured data is a form of structured data but does not fit the relational model, is not stored in tables, and does not conform to a pre-defined schema. Semi-structured data is also known as self-describing data because the expression of structure is embedded in the data, not externalized in schema. XML is perhaps the most widely used semi-structured data format.

Important distinctions between structured (schema-based) and semi-structured data for data integrators focus on attributes of entities. In a structured database every occurrence of an entity has the same attributes with the same sequence and data type. With semi-structured data the set of attributes, the data types, and the sequence may be different between occurrences of the same entity type. With structured data, disparity occurs between tables and databases. With unstructured data, disparity is possible among entities of the same class.

MULTI-STRUCTURED

Multi-structured data contains a variety of data formats and types in a single database or data store. Embedding an XML string or a text string in a relational database table is a simple multi-structure form.

E-mail is an example of more complex multi-structured data. The data is a collection of messages that may be grouped as conversations. Each individual message contains structured data in the header with attributes for sender, recipients, date, subject, etc. Yet the body of the message is typically text that is considered to be unstructured data. Attaching an image, document, table, or spreadsheet compounds the variety of formats.

Data Warehouse Implementation

Agile Data Warehousing

Data Warehouse Implementation

Agile Data Warehousing

WHAT IS AGILE?

If you look in the dictionary you'll find that agile means "moving quickly and lightly" or "mentally quick." Synonyms include nimble and spry. Yet the term has quickly taken on particular meaning in systems and software development. Agile development has become synonymous with concepts of business collaboration and test-driven development.

THE AGILE MANIFESTO

Several of the pioneers and leading practitioners of agile development (too many to list here) collectively authored the following *Manifesto for Agile Software Development.* Following these principles will make us quick and nimble in systems development whether called "agile" or not.

"We are uncovering better ways of developing software by doing it and helping others do it. Through this work we have come to value:

- *Individuals and interactions over processes and tools*
- *Working software over comprehensive documentation*
- *Customer collaboration over contract negotiation*
- *Responding to change over following a plan*

That is, while there is value in the items on the right, we value the items on the left more." (http://agilemanifesto.org)

AGILE DATA WAREHOUSING

Agile data warehousing applies the concepts and practices of agile development to data warehouse implementation. Collaboration with business and data subject experts, combined with rapidly repeated test-driven development is used to understand source data, map source data to target data, discover and apply data transformation logic, and evolve the warehouse data model.

Module 5

Data Management

Data Profiling

Analyzing Data Profiles

Customer Table Values Frequency

Table: Customer Table
Column: email

value	frequency
lee@pacrimtrading.com	2
van@juno.com	2
abbyanne@yahoo.com	1
amwilson@skyline.net	1
bartolocci@simtech.biz	1
bobjones@comcast.net	1
brain@bnovak.com	1
cmb@gmx.com	1
dfchen@sprynet.com	1
eddieanders@gmail.com	1
fatcat@bigdeal.biz	1
greenguy@ecolabs.org	1
jrm123@gmail.com	1
kim@iconic.org	1

Customer Table Column Profile

column name	row count	distinct values count	percent unique values	null values count
cust_num	30	30	100%	0
last_nm	30	28	93%	0
first_nm	30	26	87%	0
middle_nm	30	11	37%	14
gender	30	2	7%	1
age_grp	30	7	23%	0
income_grp	30	13	43%	1
email	30	28	93%	0
mail_addr	30	29	97%	0
mail_city	30	22	73%	1
state_abbr	30	21	70%	0
zipcode	30	27	90%	0
last_tx_date	30	25	83%	0

Customer Table

cust_num	last_nm	first_nm	middle_nm	gender	age_grp	income_grp	email	mail_addr
916232	Morgan	James	Robert	M	2	3	jrm123@gmail.co	6120 Langley Avenu
[illegible]	Smith	Karen	L	F	5	[illegible]	[illegible]@earthlin	1885 Desales Street NW
[illegible]						[illegible]	[illegible]	
916244	Roberts	Thomas	W	M	5		troberts999@gmail.	6550 Pineridge Rd N
916245	Kim	Lee		M	3	14	lee@pacrimtrading.	7277 Lynn Oaks Drive
916246	Harper	Sandra		F	3	5	sharp@mindspring.	480 S. Terrence St.
916247	Chen	Daniel	Feng	M	2	3	dfchen@sprynet.co	24815 Sunburst Lane
916248	Barker	Thomas		M	3	7	tbarker@alaskanet.c	612 W. Willoughby Avenue
916249	Jacobs	David	Nathan	M	4	9	pingme@mobile.co	87 W. 24th Street
916250	Andropov	Olga		F	6	5	russianpride@senio	3120 Roxbury Drive
916251	Sato	Kyoko		F	5	6	ksato@tigertech.co	9243 Gordon Way
916252	Lee	Kim		M	3	14	lee@pacrimtrading.	7277 Lynnoaks Dr.
916253	White	Shirley		F	4	7	sabarnes@mindsp	7806 N. Altamont Street
916254	Bartolocci	Mark	M	M	4	8	bartolocci@simtech	4812 N. Plymouth Avenue
916255	O'Neil	Abigail	Anne	F	3	6	abbyanne@yahoo.	5170 N. Girard Avenue
916256	Jacobs	Kimberly		F	5	17	kim@iconic.org	8172 Chapman Ave.
916257	Van Syke	Steven		M	3	4	van@juno.com	318 W. Cary Street
916258	Vansyke	Stephen		M	3	5	van@juno.com	318 W. Cary Street
916259	Wrczynski	Tracy			2	3	polishdog@petshop	2261 Townsend Str
916260	Anderson	David		M	3	7	eddieanders@gmai	960 Oglivie Street
916261	Novak	Brian		M	2	6	brain@bnovak.com	5008 50th Avenue S.

Data Profiling

Analyzing Data Profiles

THE STORY IN THE PROFILES

Many things about data are discovered through data profiling. Findings from column profiling include:

- Distinct values analysis finding
  - Constants – only one value that is not blank and not zero
  - Empty columns – only one value that is either blank or zero
  - Indicators – number of distinct values exactly 2 (y/n, t/f, or 0/1)
  - Codes – number of distinct values in single or low double digits
- Null values analysis finding
  - Unused columns – 100% null values
  - Optional columns – percent of null values is relatively high
  - Missing data – percent of null values is relatively low
- Value distribution analysis finding
  - Consecutive numbers –
row count = maximum value – minimum value + 1
  - Outliers – exceptionally high or low values, useful to look at top-10 and bottom-10 lists
  - Skew – substantial difference between mean and median
  - Default – exceptionally high frequency of a single value
  - Ranges and clusters – apparent ranges, clusters or gaps
- Distinct patterns analysis finding
  - Overloaded columns – two or three distinct patterns
  - Non-conforming columns – many distinct patterns such as phone numbers

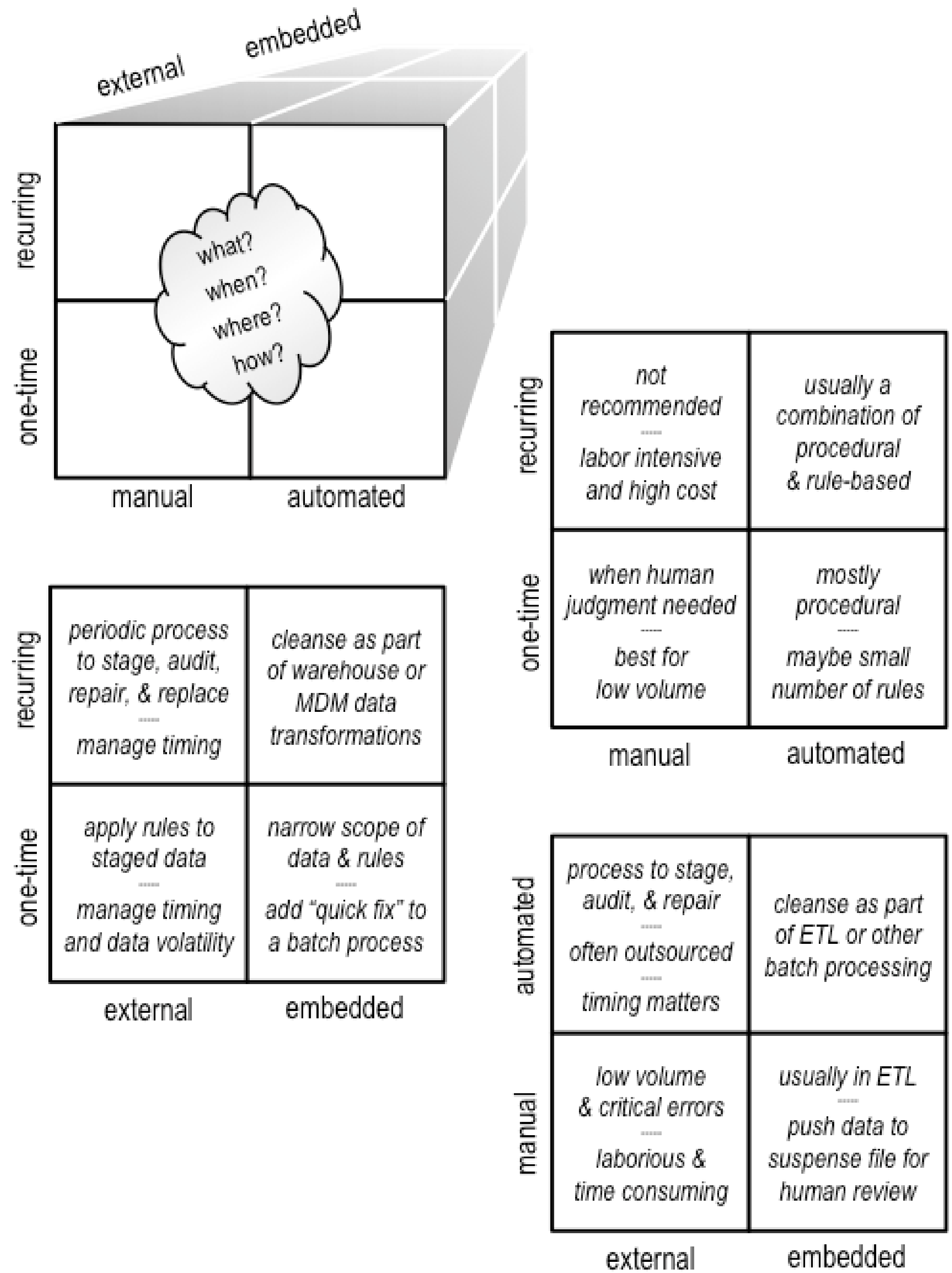
Table and cross-table profiling yield even more discoveries including:

- Definition rules
  - Data naming – synonyms for customer identifier
  - Data typing – where declared and inferred data types are different
- Content and structure rules
  - Allowed values – many examples of domain constraints
  - Uniqueness – especially as applied to primary keys
  - Optionality – where nulls indicate non-mandatory data
  - Consistency – based on column dependencies both within and across tables
  - Referential integrity – where foreign key dependencies are found
  - State dependency – where values are constrained by status, such as canceled orders that have null ship date

These are but a few examples among many possibilities.

Data Quality

Data Quality Improvement



Data Quality

Data Quality Improvement

DEFINITION AND CONCEPTS

Data cleansing is the act of detecting and correcting or removing corrupt or inaccurate records from a record set, table, or database. It is a process of finding and removing data quality defects. Cleansing may involve removing defective data from the collection, obtaining correct data from an alternate source, or adjusting defective data to comply with data quality rules.

Data cleansing may be:

- manual (performed by people) or automated (performed by computer)
- one-time (a single-instance repair) or recurring (periodic processing)
- embedded (integrated into existing processes) or external (performed as a stand-alone process).

These options combine in some interesting ways – embedded, automated, recurring for example; or external, manual, one-time. A complete data cleansing solution typically mixes-and-matches several options. High-level questions for each cleansing activity include:

- What to cleanse – which data and which defects?
- When to cleanse – at what point in business and systems schedules?
- Where to cleanse – at what point in the flow of data and processes?
- How to cleanse – using what methods and workflow?

CLEANSING VARIATIONS

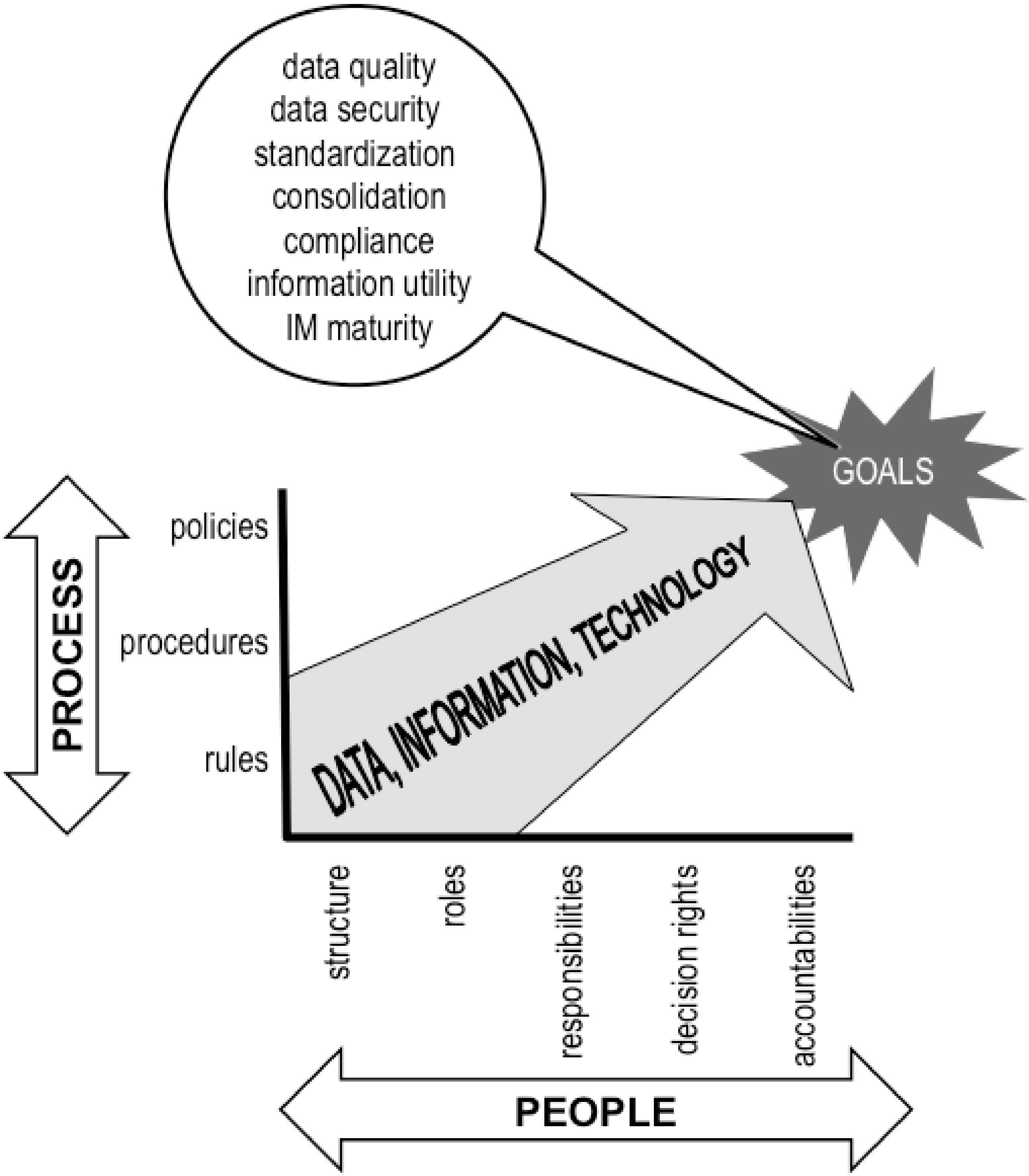
Using two-dimensional views for a closer look at combinations offers some guidelines for the use of each. One-time/manual/external, for example, yields these guidelines:

- Effective for low-volume cleansing where human judgment is needed.
- Stage the data separately from production data, and then apply cleansing rules.
- Be conscious of the time that manual processing takes and consider the matching issues that may occur when you're ready to merge staged and cleansed data back into a volatile production database.
- Manual processing is best suited to low-volume, critical errors because it is laborious and time consuming.

Collectively the guidelines indicate that one-time/manual/external works for a small number of critical errors needing human judgment to correct. There is risk if production data changes while a copy is separately staged for cleansing.

Data Governance

Data Governance Concepts



Data Governance

Data Governance Concepts

GOVERNANCE DIMENSIONS

Data governance is a program of managing information assets to achieve defined information management goals. Governance establishes the processes that are needed and designates the responsibilities of people to achieve the 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.

GOVERNANCE GOALS

Goals are the driving force of data governance – the reasons to govern data and the foundations upon which governance processes are built. Common goals include such things as:

- data quality
- data security
- data standardization
- data consolidation
- regulatory compliance
- information utility
- information management maturity

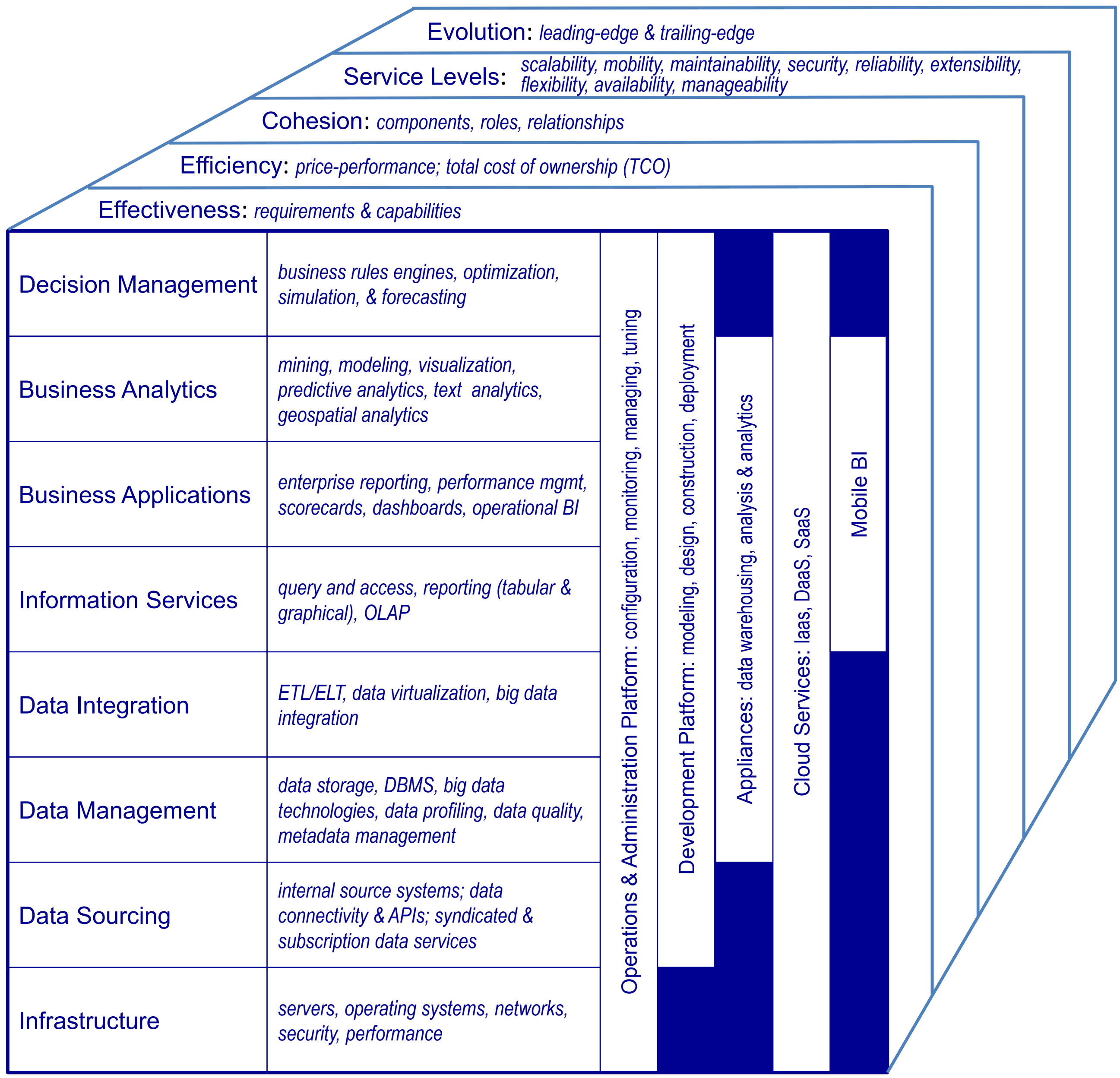
As with any program, data governance goals are not static. They change over time as the business evolves and the governance program matures.

Module 6

BI Technology

Technology Architecture

The Right Technology – Present and Future



Technology Architecture

The Right Technology – Present and Future

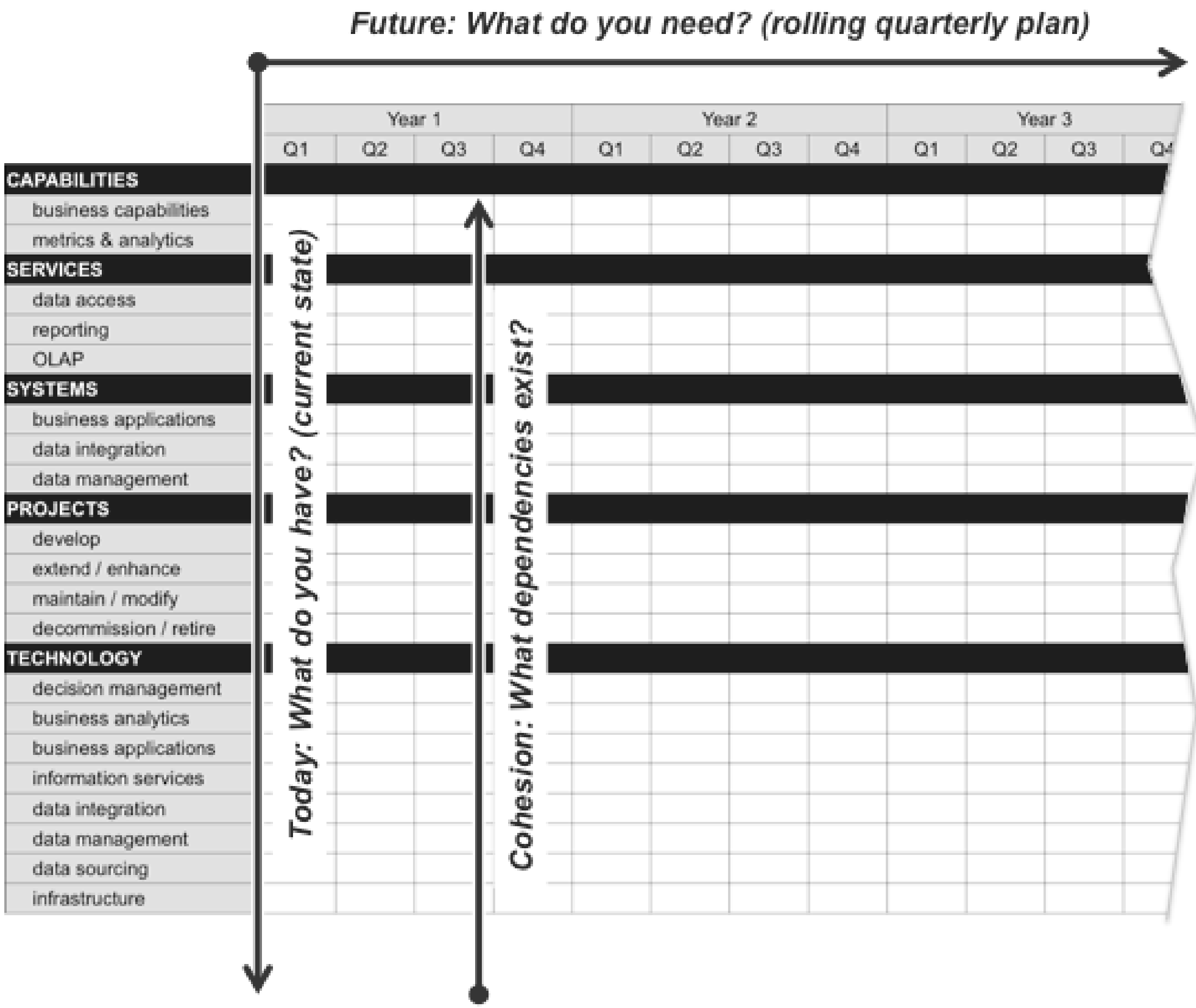
SUSTAINABLE TECHNOLOGY

Technology architecture is needed to ensure that the technology stack will adapt to change and can be sustained throughout the lifespan of enterprise BI systems. Architectural considerations and responsibilities include:

- Effectiveness of technology – the ability to meet BI requirements and enable needed business capabilities.
- Efficiency of technology including price/performance ratio and total cost of ownership (TCO).
- Cohesion of technology that is well integrated or compatible with the right components, performing in the right roles, and interconnected in the right ways.
- Service levels committed and delivered for scalability, mobility, maintainability, security, reliability, extensibility, flexibility, availability, manageability, and more. Wherever a business expectation for quality of service in BI exists, a corresponding service level responsibility exists.
- Evolution of the technology stack with attention both to extending the leading edge of technology and picking up the trailing edge. When we fail to pull the trailing edge forward, the range of technologies expands and the degree of integration and cohesion erodes. Deteriorating cohesion of technology ultimately affects service levels, user satisfaction, and business value.

Technology in the BI Roadmap

Current and Future States



Technology in the BI Roadmap

Current and Future States

MAPPING THE TECHNOLOGY

The facing page illustrates the parts of the BI roadmap that relate to BI technology. Review the current state inventory and identify future state requirements for the entire technology stack. Plot technology onto the roadmap as needed to support planned capabilities, services, systems, and projects.

The BI roadmap now represents a logical progression from need for business capabilities, through layers of services, systems, and projects to the technology that is needed. It is a plan for continuous and incremental evolution of BI systems and growth of BI maturity.

