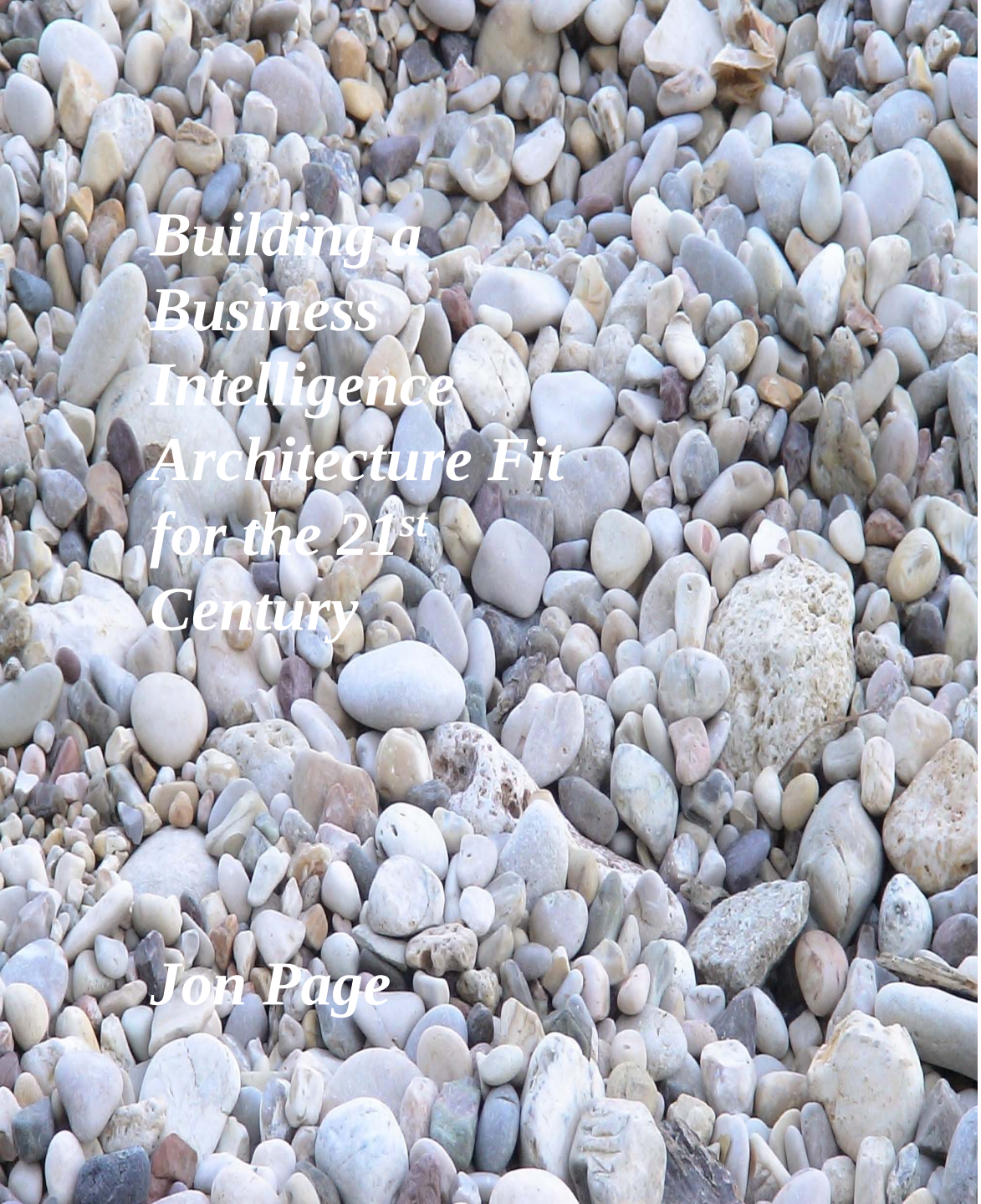
Building a Business Intelligence Architecture Fit for the 21st Century

Jon Page



This seminar is based on the Contents of this book

The premise is that:

BI is not delivering what is needed

Current BI architectures are not appropriate

Business and IT need to work Together on agreed strategies

Technology must be used to enable the reduction of copy management and the support of real time accuracy of critical data

People and organisation are Key to success in BI

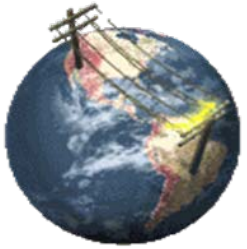
Our Objectives

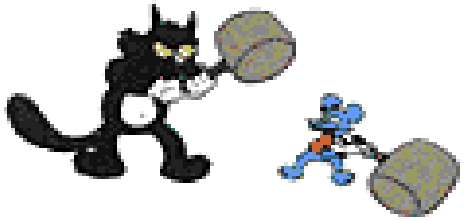
To record some history – what has happened in the past that makes the future quite challenging.
To provide real examples of BI at work – good and bad.
To illustrate the nature of data and why it has become so important in driving forward the business in the 21st century.
To outline a way to align technology with the business so that efforts and budget are spent in a way that will enable the future rather that support the past.
To propose a set of principles and ideas that can guide a company in a way to make data available to all who have the penchant to turn it into useful and valuable information.
To describe the new organisation unit that will be needed to realise the dream.

The Changing Business Climate

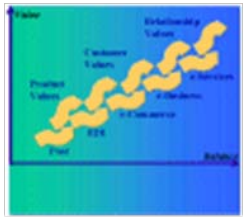
Globalisation

Deregulation





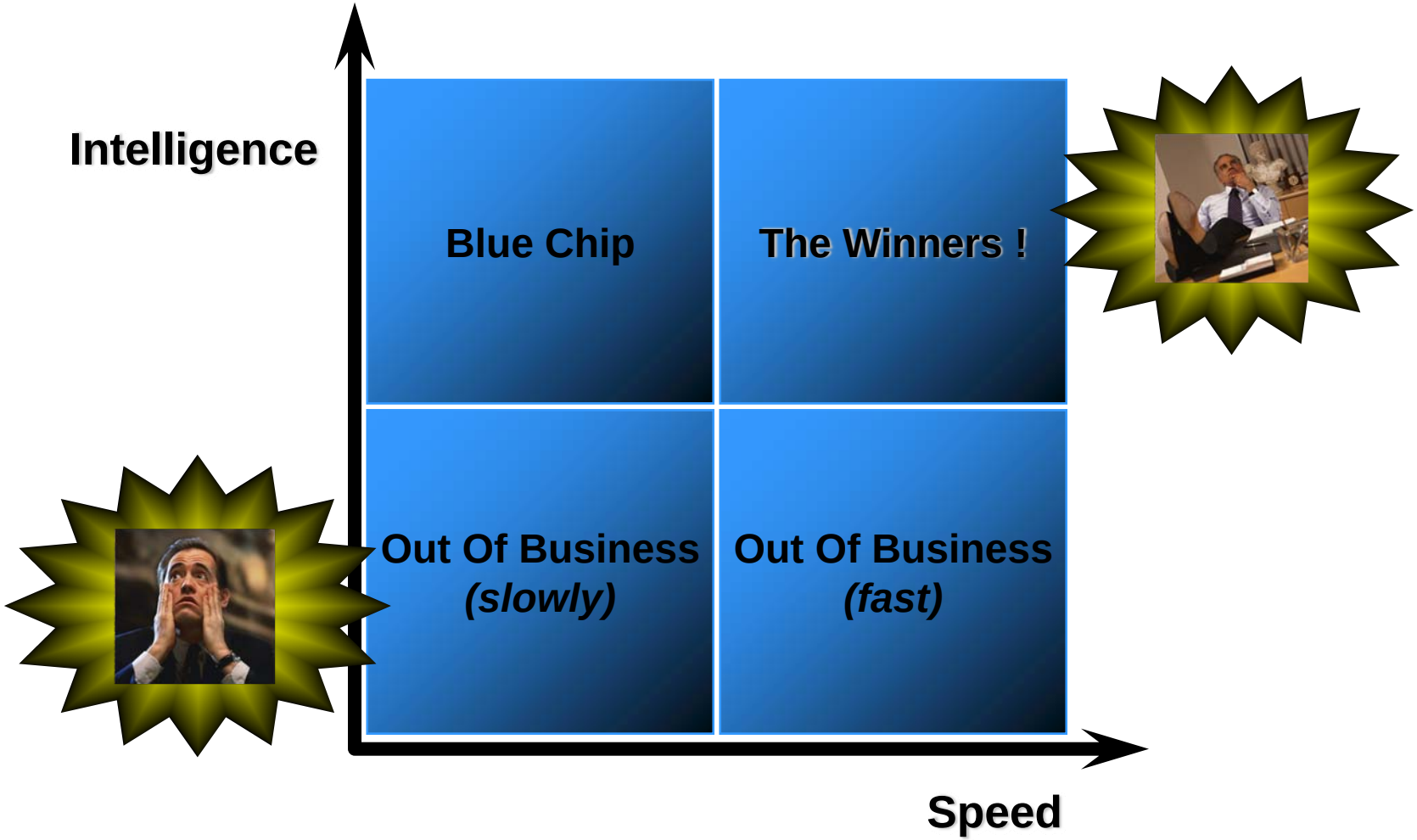
Technology



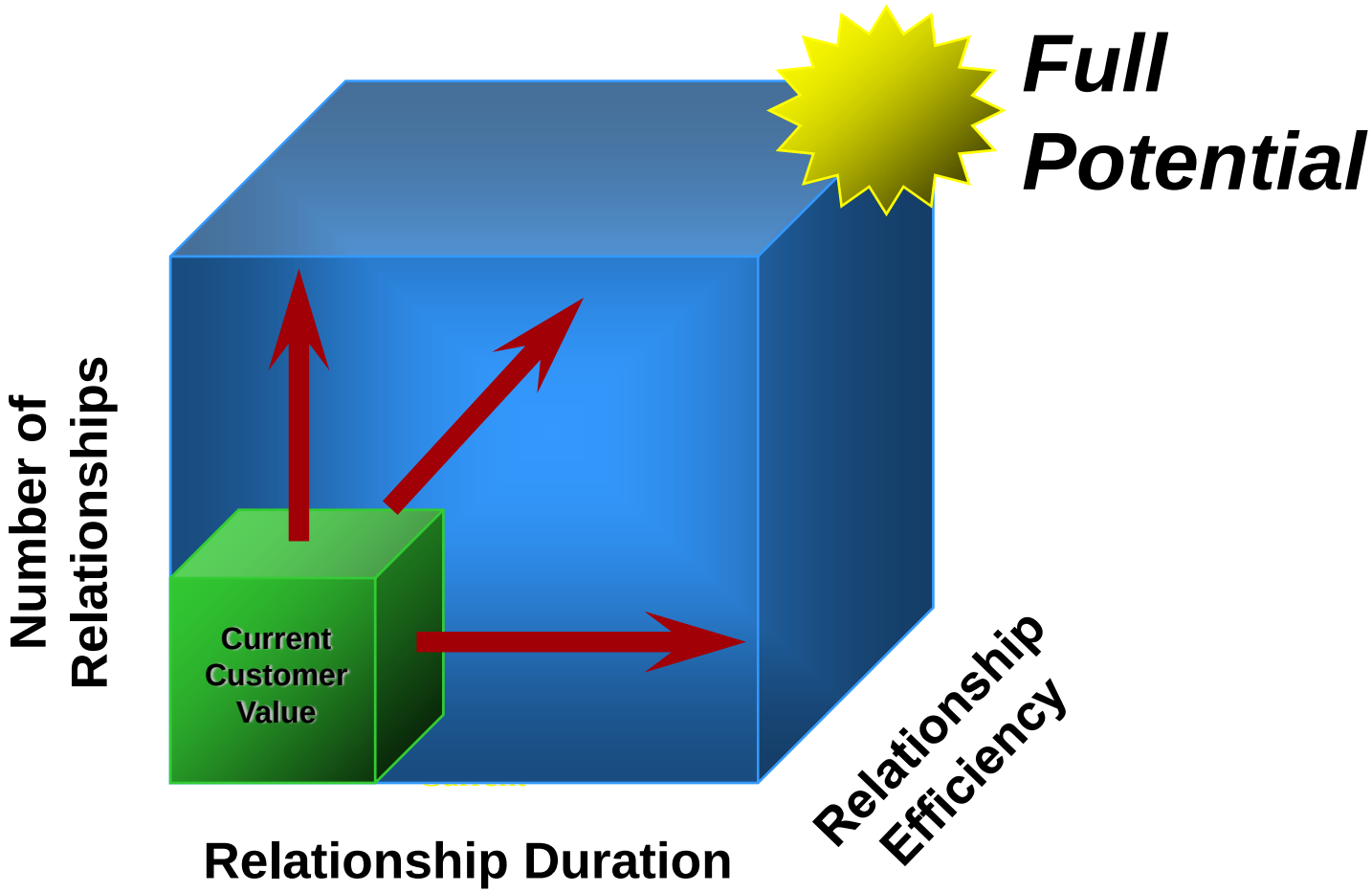
CHOICE

Competition : Very aggressive

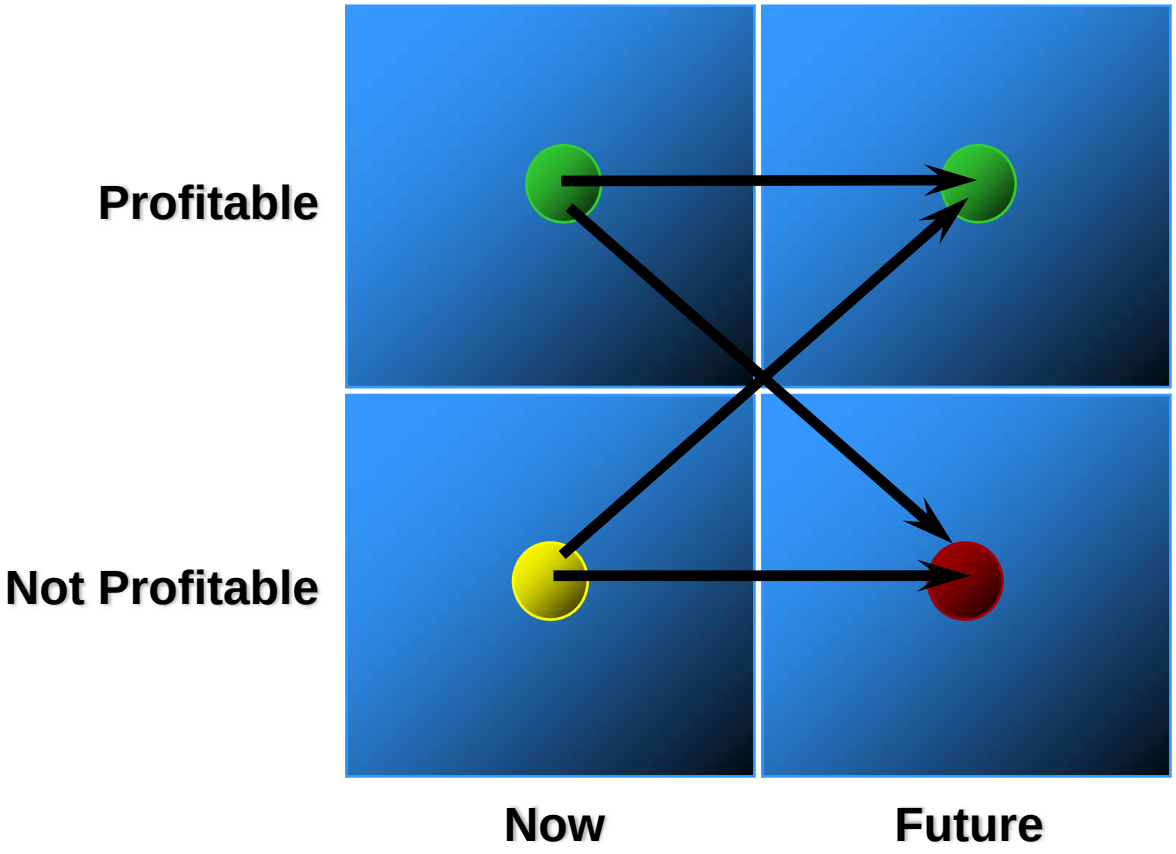
To Be Successful – Two Critical Components



Maximising Customer Value

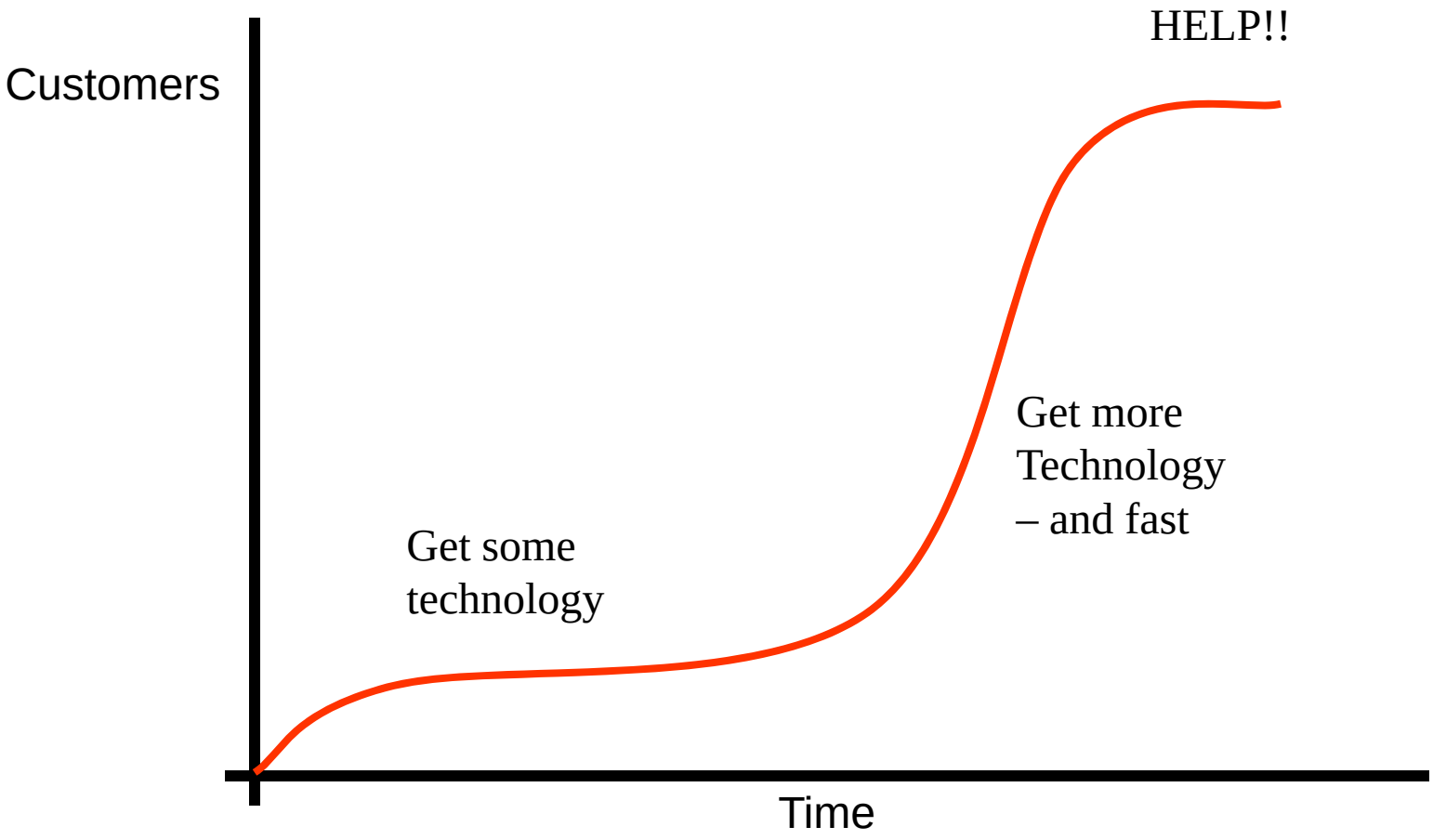


Imagine You Could Do This...



BI is an afterthought
BI is reporting
BI is not taken seriously
Everyone does their own ‘bi’
Lots of tools, lots of platforms
No-one trusts the data
No-one understands the data
No-one shares information
No-one owns the problem
Business and IT don’t work together
There is no BI Strategy

And on, and on, and on


HELP!!
Customers
Get more
Technology
– and fast
Get some
technology
Time

Data Warehousing Terminology

Strategic Information Systems

Data Sandpits

MIS

E-Intelligence

Data Warehouse

TNF

DSS

OLAP

Dimensional

Star Schema

ROLAP

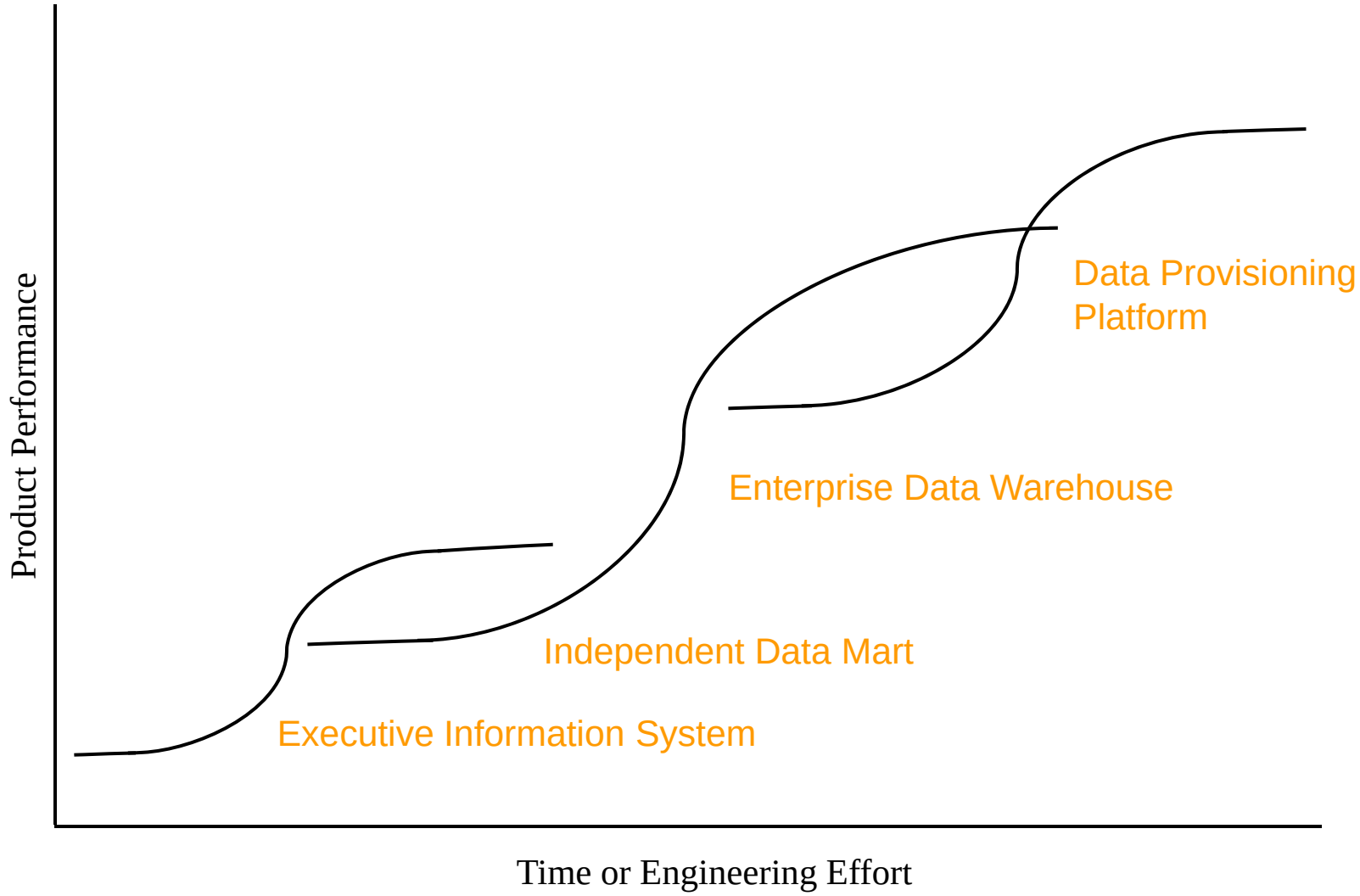
Business Intelligence

Snowflake

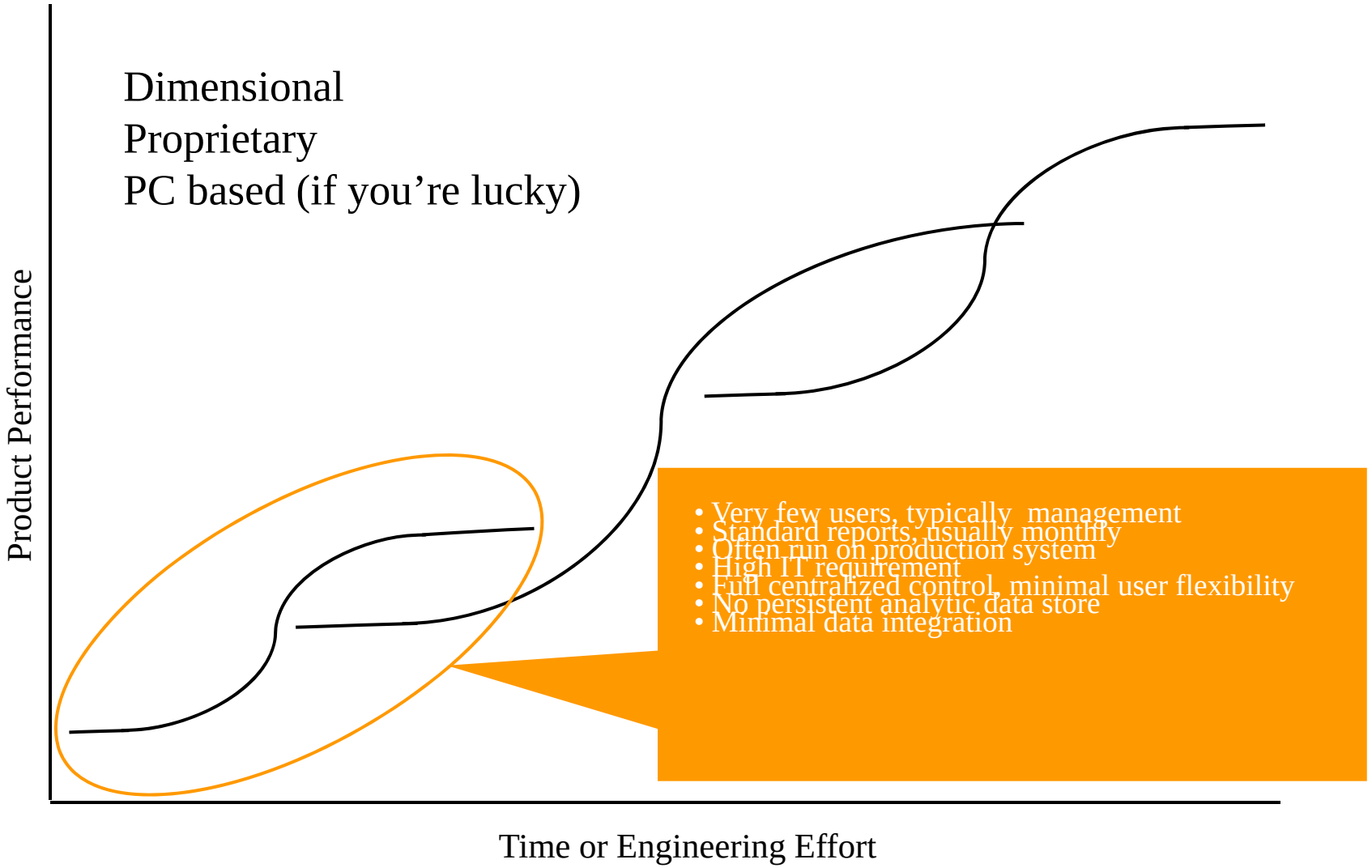
MOLAP

Data Mart

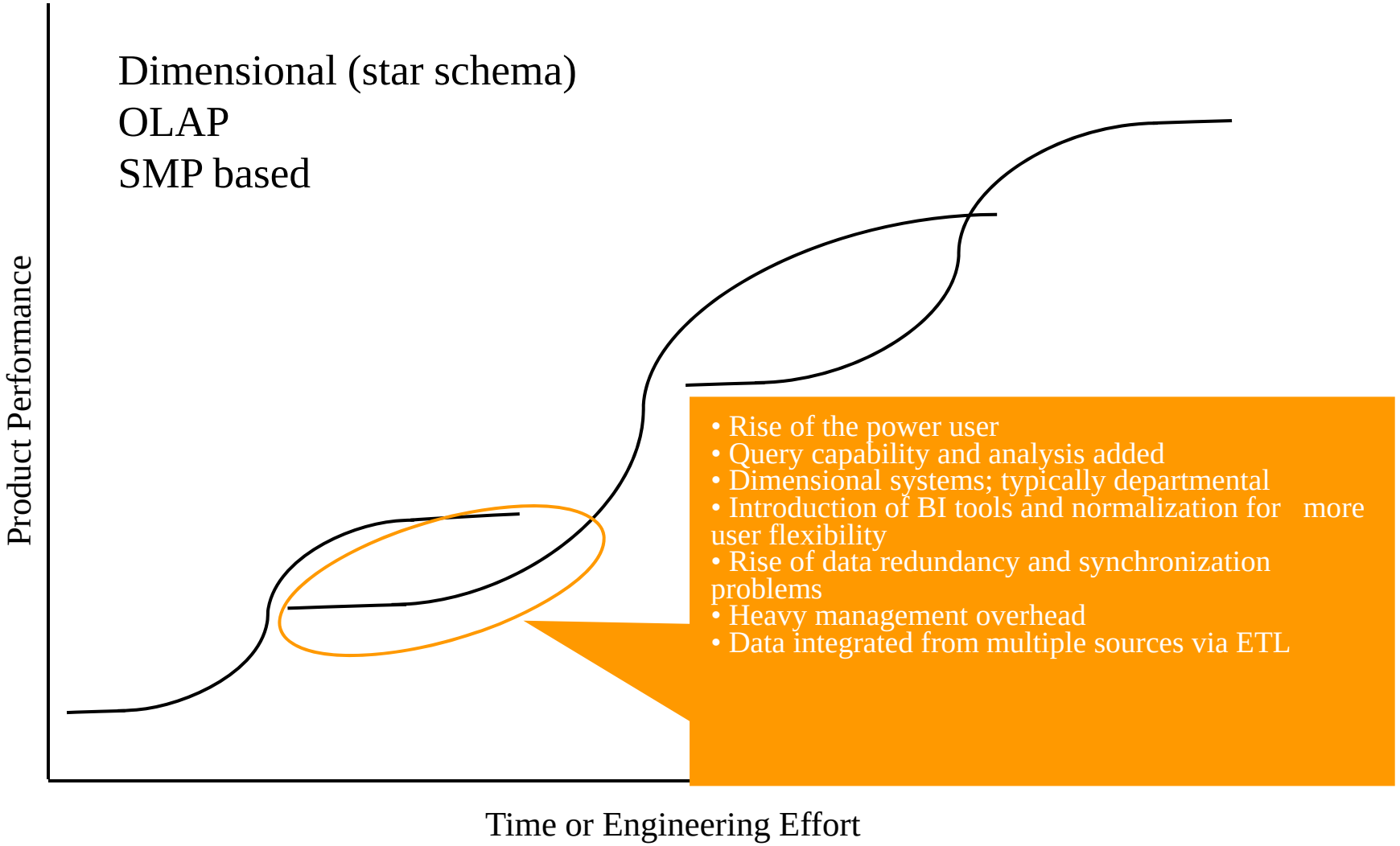
BI technology S-Curve



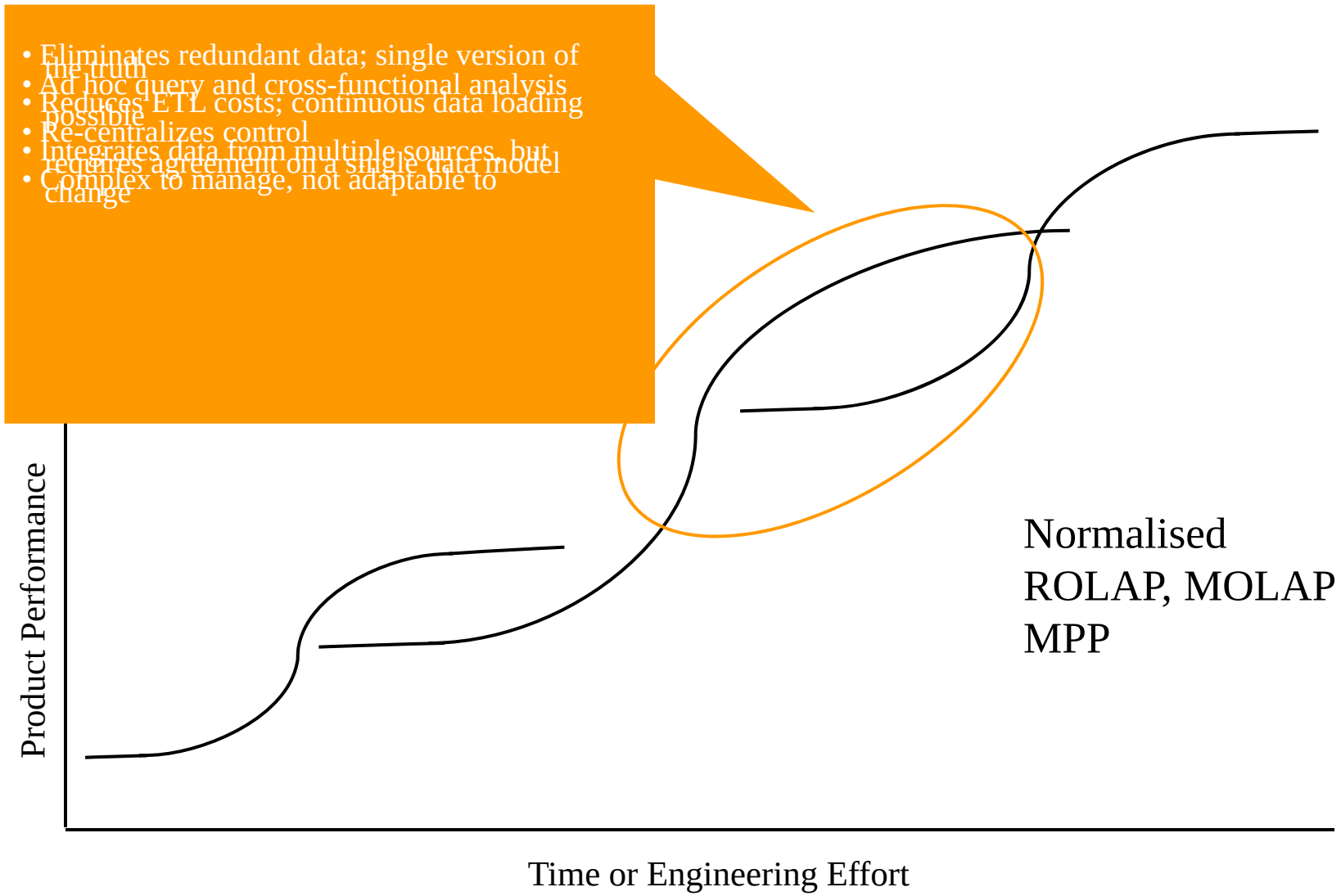
Executive Information System (EIS)



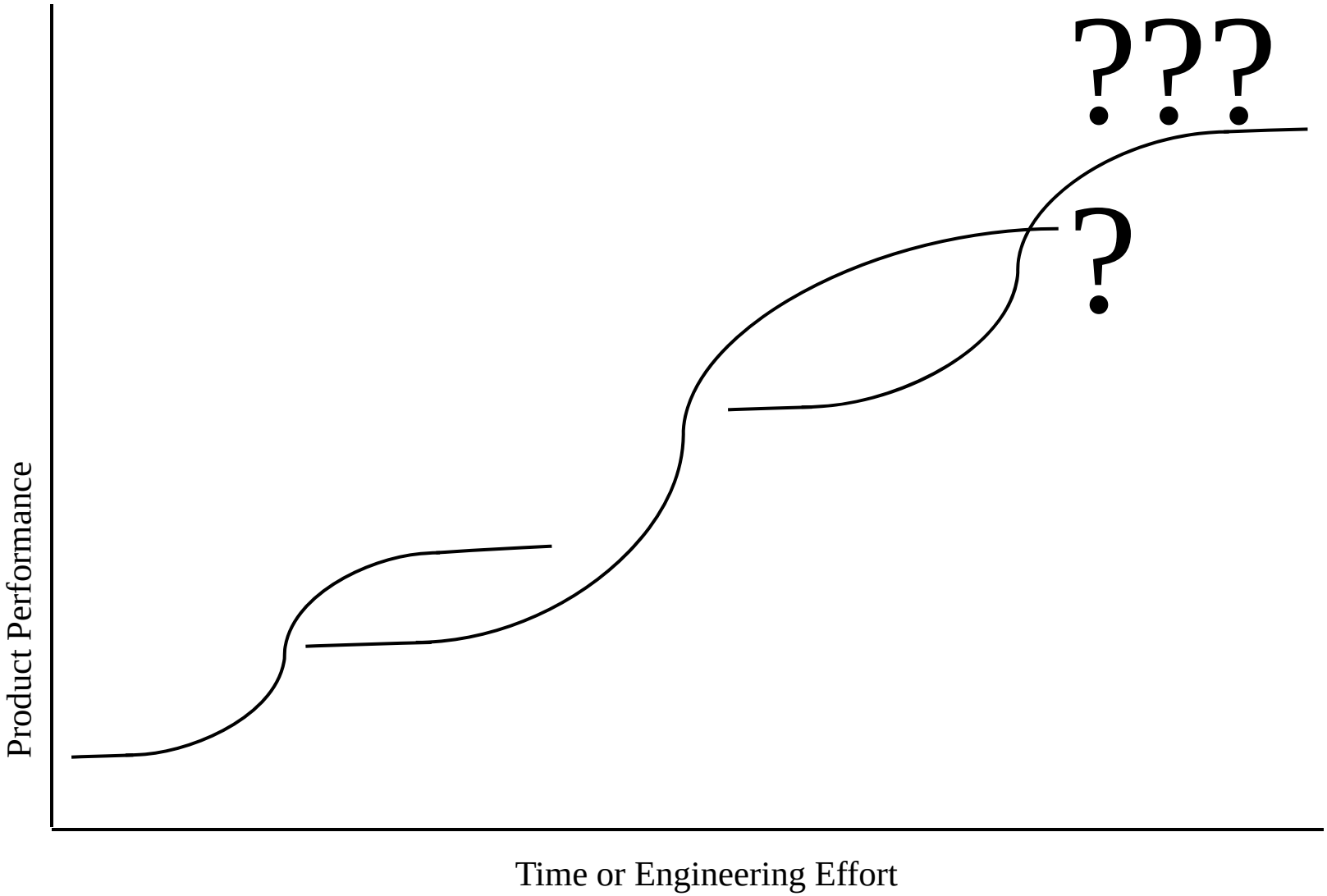
Data Warehousing: Independent DMs



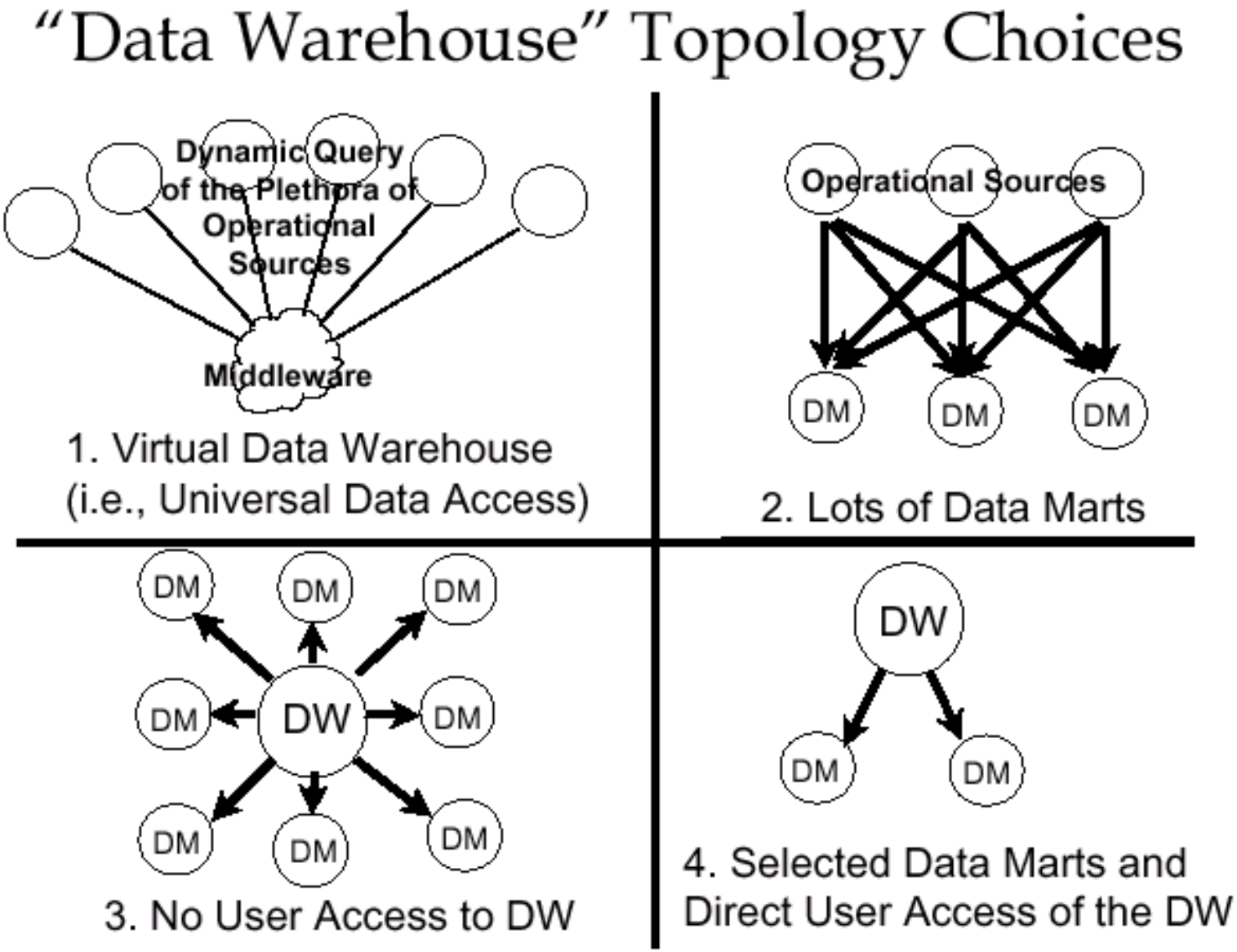
Data Warehousing: Enterprise DW



Data Provisioning? Operational BI?? Cloud???



Choosing the Right Architecture

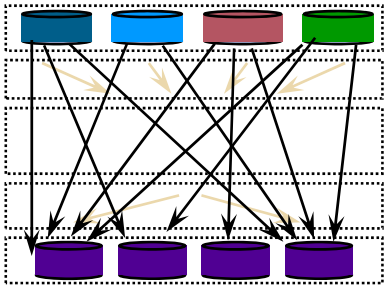


When to build data marts?

- Geographical considerations.
- Specific performance requirements.
- Specific availability requirements.
- Departmental control over reports and queries.
- Specialized applications.

- The Three P’s
- Performance, Politics and Packages

DM Topologies – Independent DMs



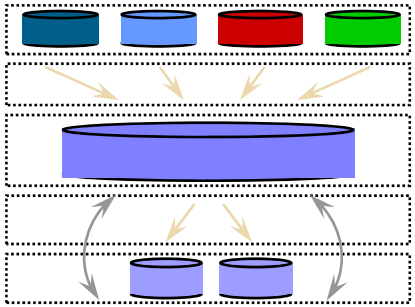
Advantages

- “Fast” implementation.
- Quick ROI.
- Departmental control.
- Not reliant on IT for data.

Disadvantages

- Multiple data models.
- No consistent corporate data model.
- Multiple interfaces to manage/maintain.
- No single version of the truth.
- Duplication of data.

DM Topologies – Dependent DMs



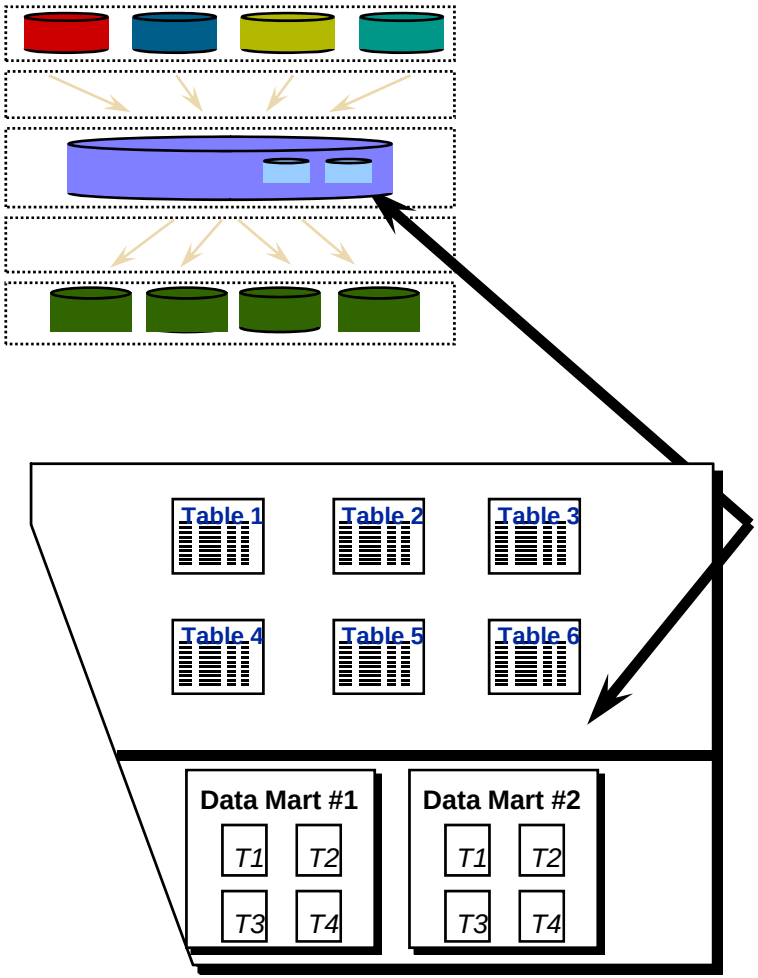
Advantages

- Single version of the truth.
- Clean/scrubbed data.
- Consistent data model.
- Robust data transformation.

Disadvantages

- Must have an existing data warehouse.
- Must fit with corporate strategy.
- Duplication of data.

DM Topologies - Dependent DMs



If you have an EDW and you want to propagate data to Data Marts, you have a choice:

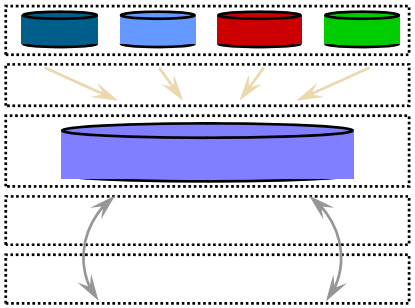
1) Propagate data to <u>distributed</u> Data Marts...

- è *Original data is stored in EDW.*
- è *Propagated data is stored in different physical sites.*

2) Propagate data to <u>integrated</u> Data Marts...

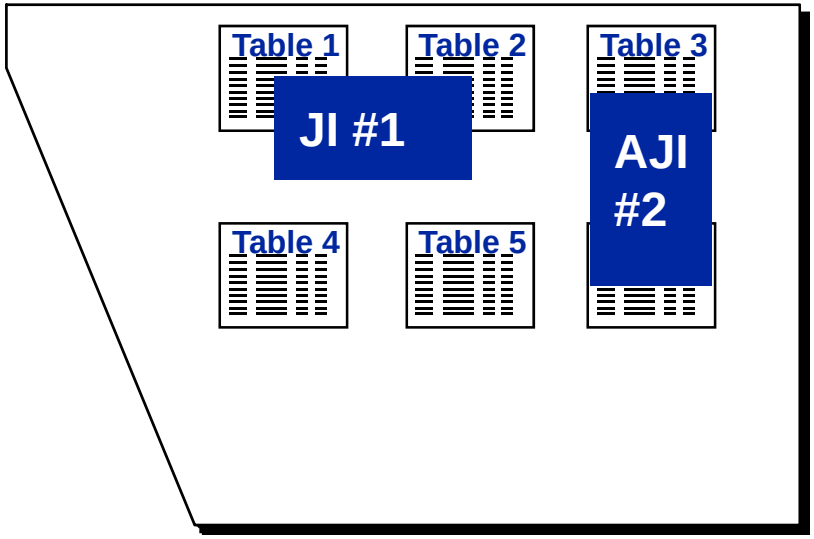
- è **Original data is stored in EDW.**
- è **Propagated data is stored in separate "databases" within the EDW environment**

DM Topologies - Integrated DMs

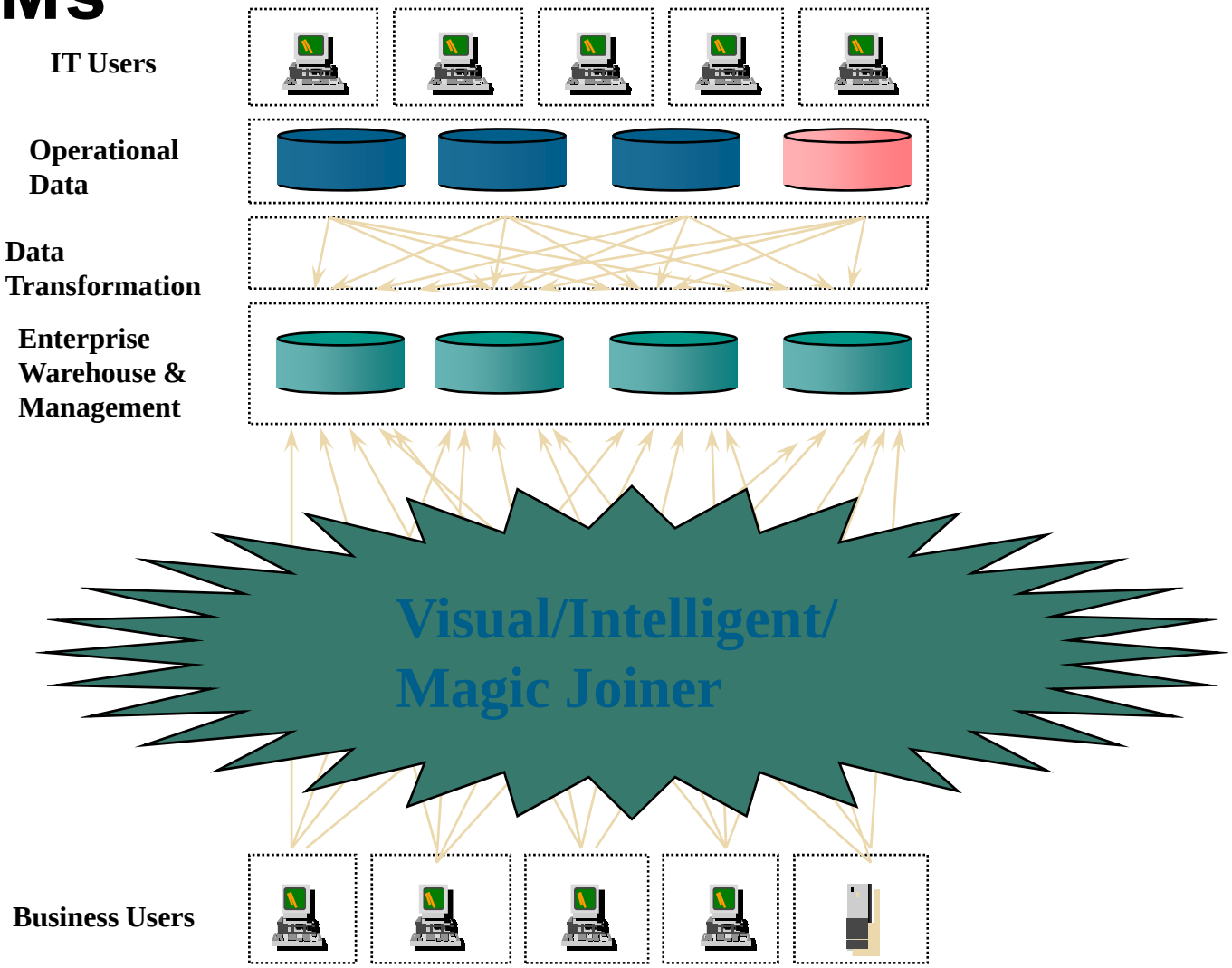


If you have an EDW and you want to achieve the Data Mart result <u>without creating more databases to maintain</u>, you can create Materialised Views within the EDW.

- **èThe data remains stored in the EDW.**
- **èMV's can provide superior performance.**
- **èApplications can use logical model of the data.**



DM Topologies - Federated DMs



Question???

What do all of these architectures have in common?

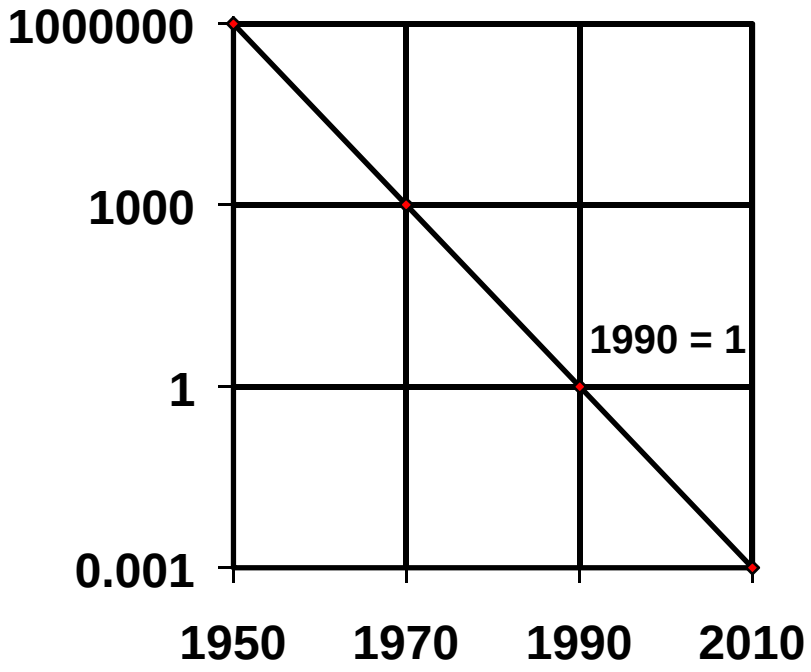
Failure is Always Blamed on Technology

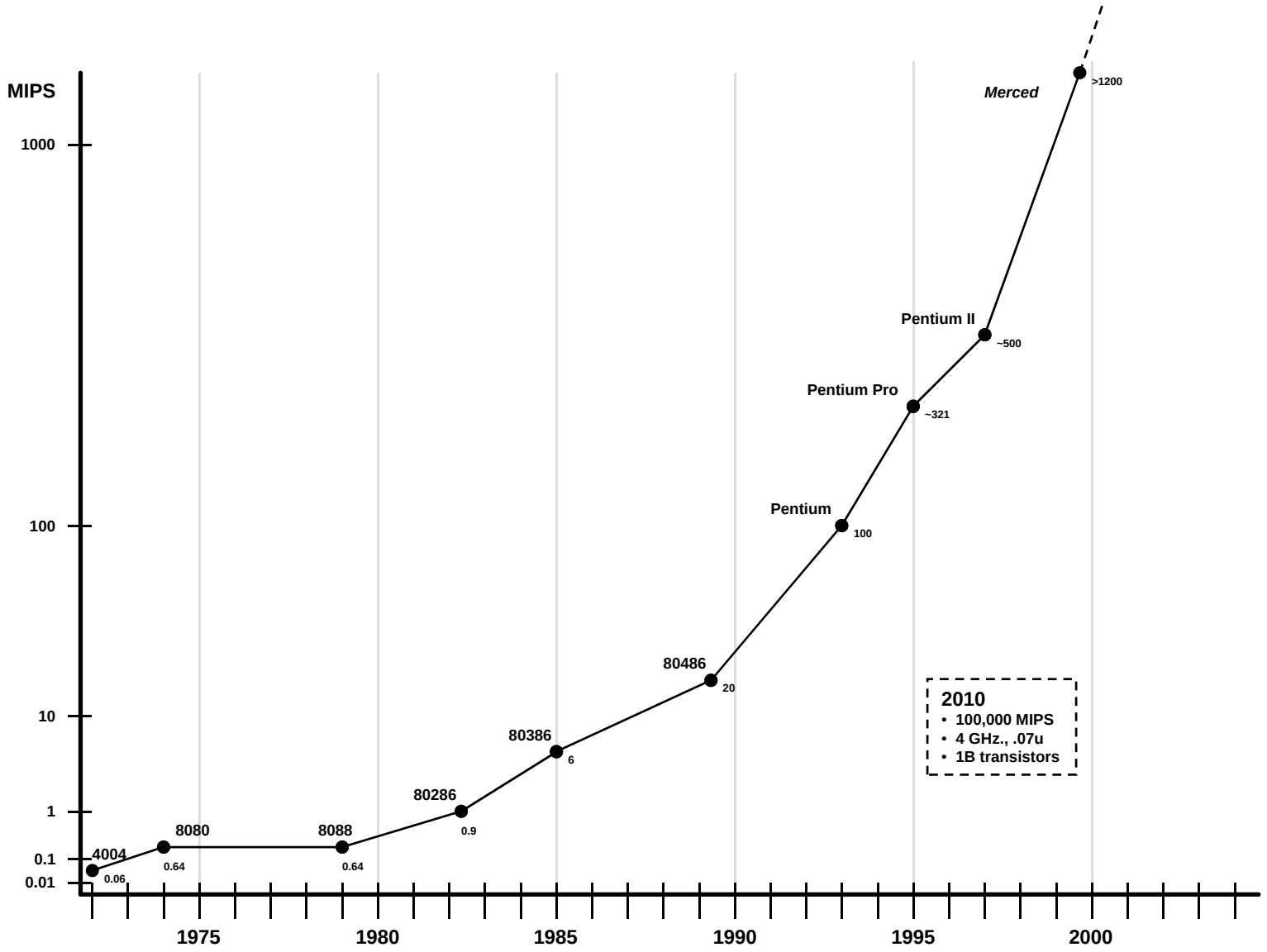
- The Computer is too slow
- The Database can’t scale
- SQL is not powerful enough
- The Database doesn’t have the functionality I need
- The data is not to be trusted (not upto date)

- Nearly always technology is blamed for faults in people/organisation
- There’s really nothing wrong with technology today…

Moore's Law Illustrated

'Every 20 years the cost of processing a single bit of information declines by a factor of 1,000'




MIPS
1000
100
10
1
0.1
0.01
1975
1980
1985
1990
1995
2000
4004
0.06
8080
0.64
8088
0.64
80286
0.9
80386
6
80486
20
Pentium
100
Pentium Pro
~321
Pentium II
~500
Merced
>1200
2010
• 100,000 MIPS
• 4 GHz., .07u
• 1B transistors

Questions?

“Even if you’re on the right track, you’ll get run over if you just sit there.”

Will Rogers

Building a Business Intelligence Architecture Fit for the 21st Century Part 2

Jon Page

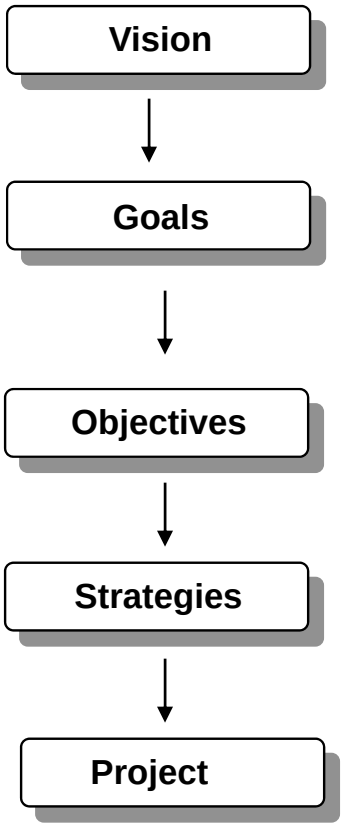
So it seems that we still have a problem? The trouble is that BI doesn’t stand alone – it feeds some very complex processes

Linking BI Strategy and Corp Strategy

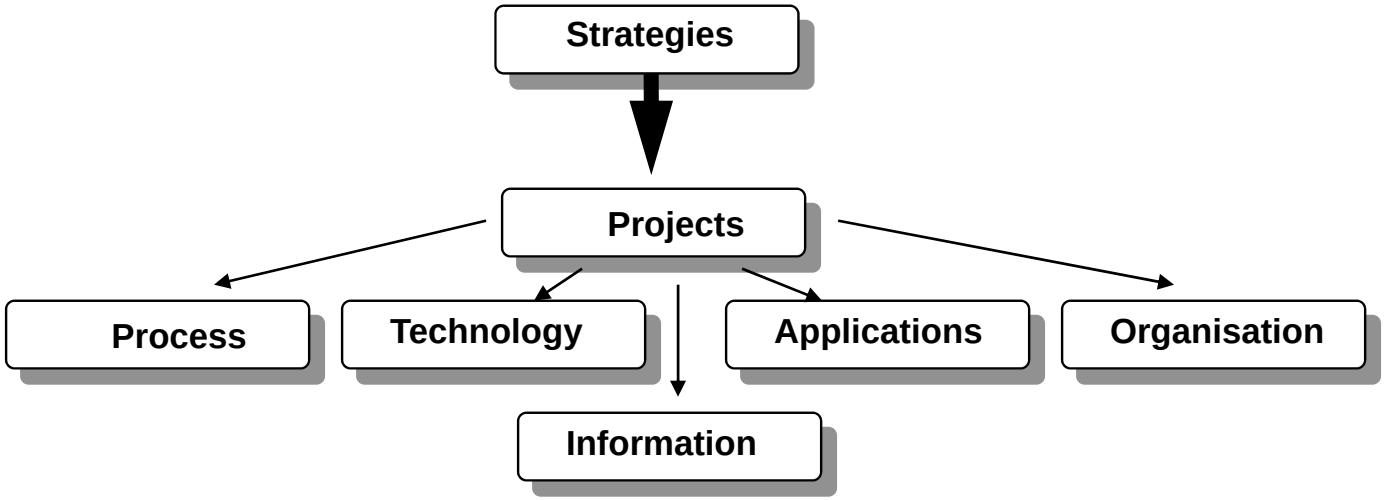
WHY???

- Raise visibility of BI
- Build required organisation
- Ensure BI solves REAL business issues
- Qualify and Prioritise BI activities
- Get funding
- Get recognition for a job well done

Corp Strategy Looks Something Like This:

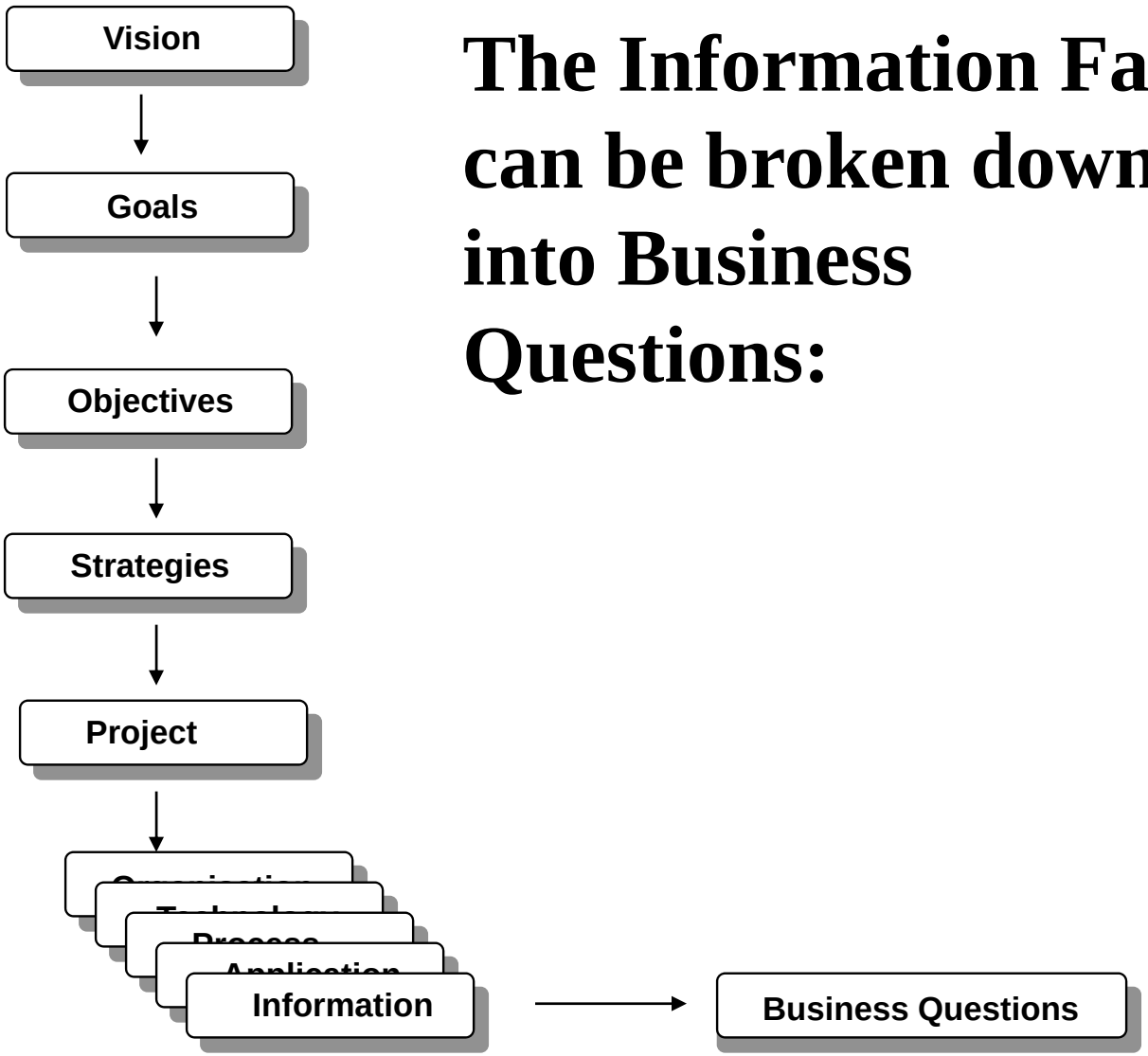


Projects can be broken down into Facets:

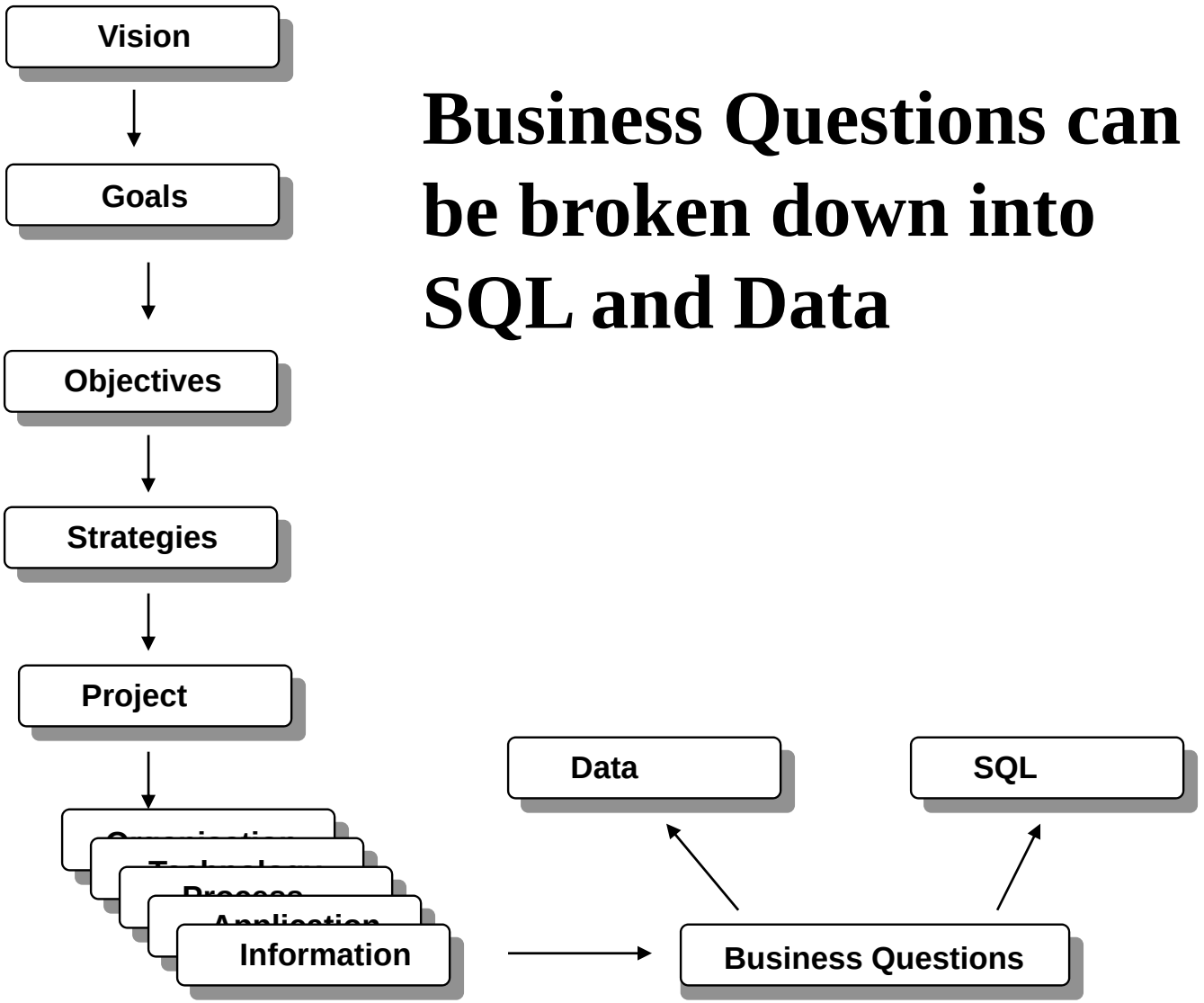


(Now we're moving into BI Strategy)

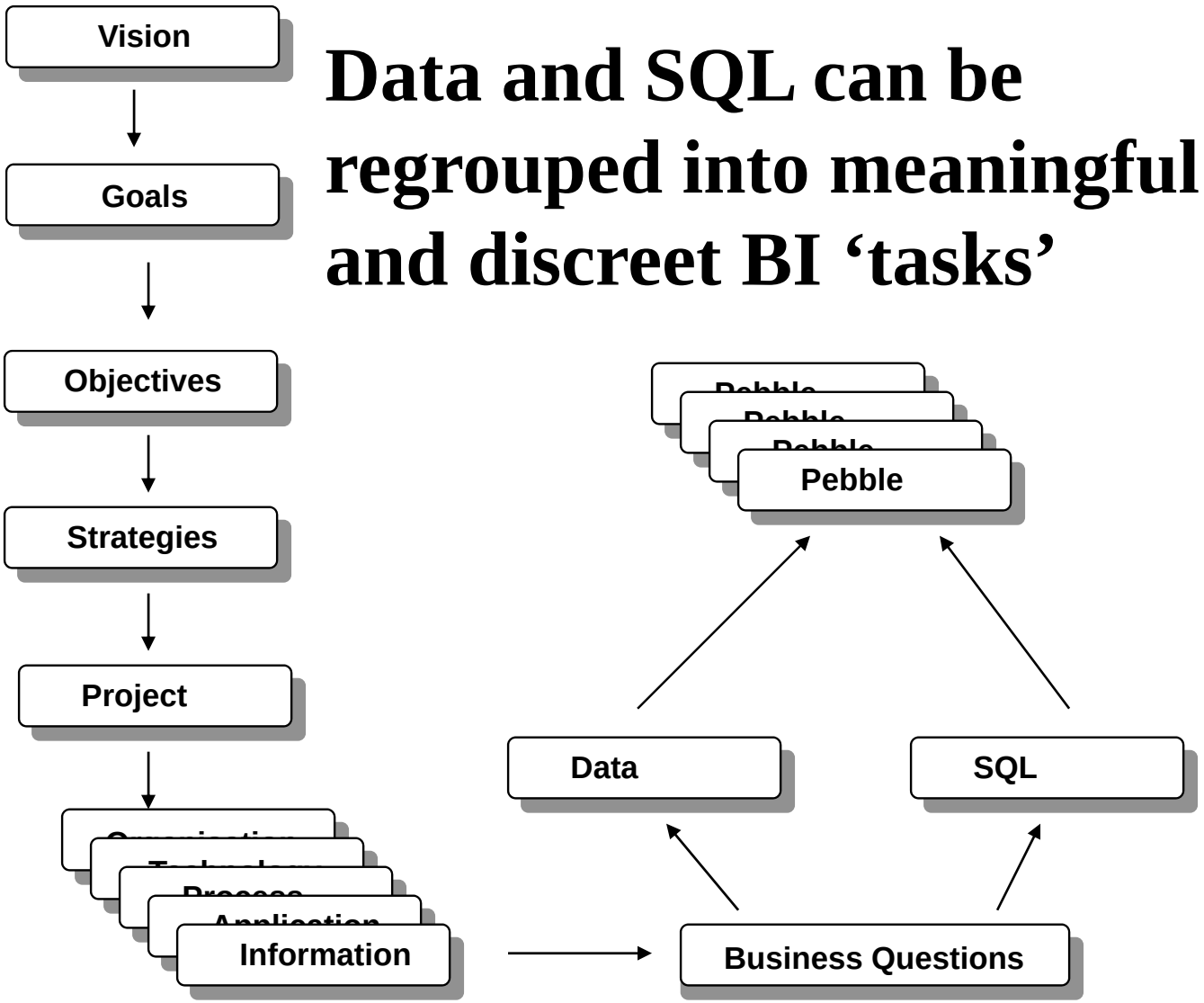
The Information Facet can be broken down into Business Questions:

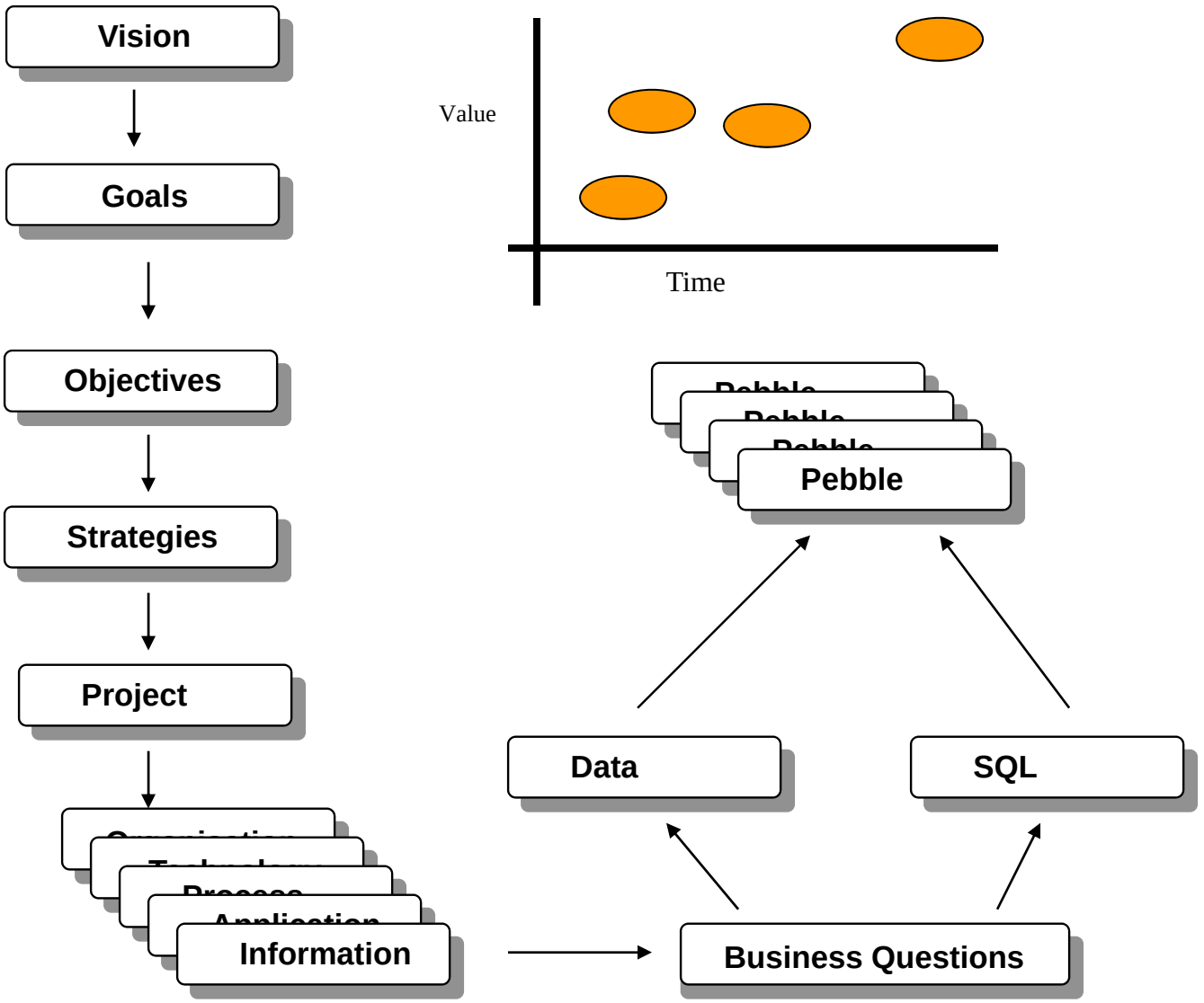


Business Questions can be broken down into SQL and Data



Data and SQL can be regrouped into meaningful and discreet BI ‘tasks’




Vision
Goals
Objectives
Strategies
Project
Information
Value
Time
Pebble
Data
SQL
Business Questions

What is a Pebble?

They are opportunities that:

can be enabled by a technology based solution
need to exploit information
can be realised in no more than six months
are well understood by business and technology
are of high priority
are measurable in terms of success whether this be ROI, TCO or some other internalised mechanism
have a descriptive and meaningful name.
solve at least some part of the information facet of a project.

Perhaps the most important characteristic of pebbles is that they support agreed and strategic projects

Real ‘Pebbles’

- Management of churn in a specific segment of customers – let’s reduce churn by 5% in our customers who make more than 5 calls per day
- Revenue recognition – lets make sure that we improve by 5 points the ratio between service usage, billing and collection across the board
- Increase acquisition rate of customer in a specific segment – let’s increase rate by 2% for corporations with 50 to 300 employees.
- Decrease fraud – let’s reduce subscriber fraud by 3 points across all customers
- Understand product affinity – lets’ determine which products are bought in combination to allow us to more effectively deploy ‘loss leaders’.

Let’s reduce churn by 5% in our customers who make more than 5 calls per day

- Questions (SQL):
- Who makes more than 5 calls per day
- What is the churn rate in this segment?
- What do the churners have in common?
- Who in the current segment looks like these churners?
- How can I keep them?

Subscriber attributes

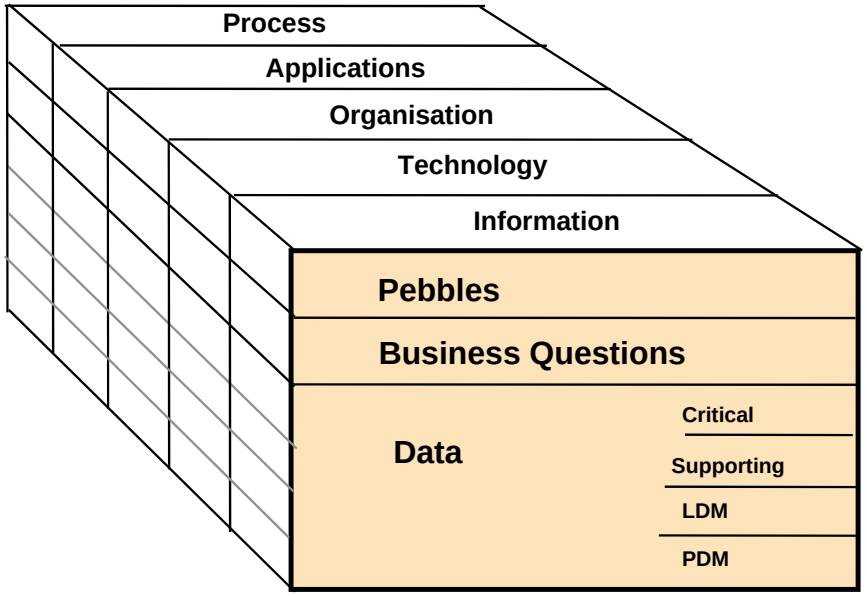
Contracts

Bills

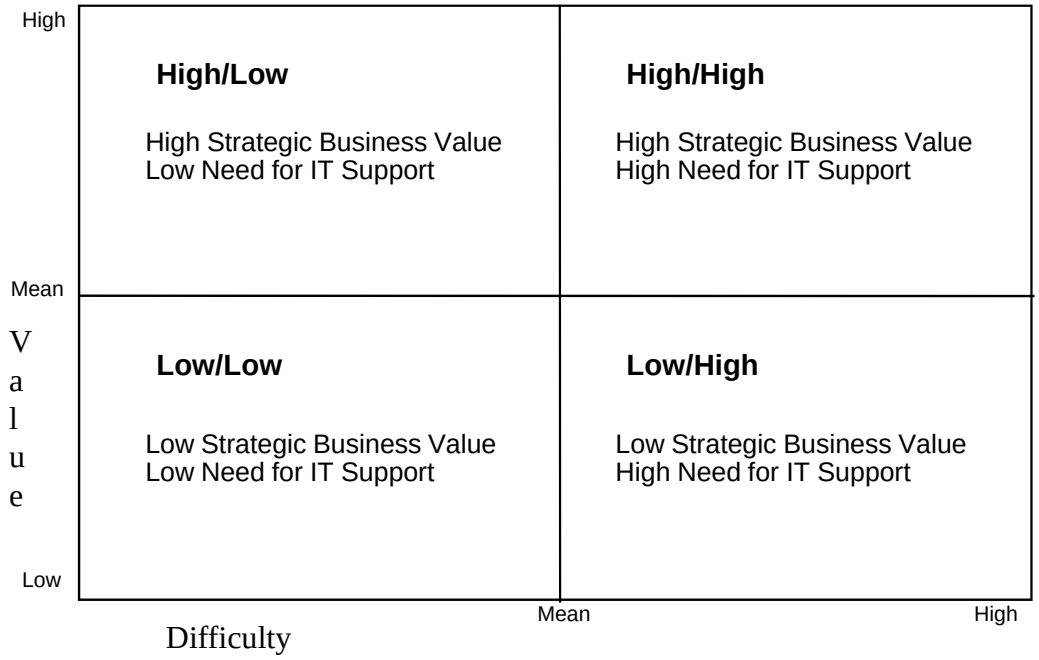
Calling Circle

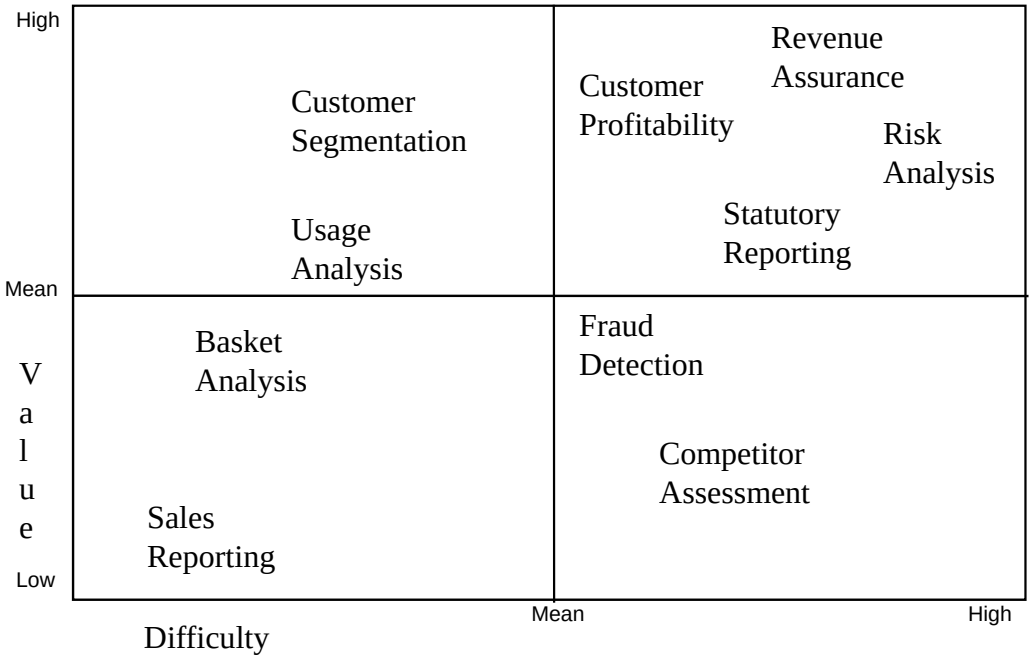
Payments

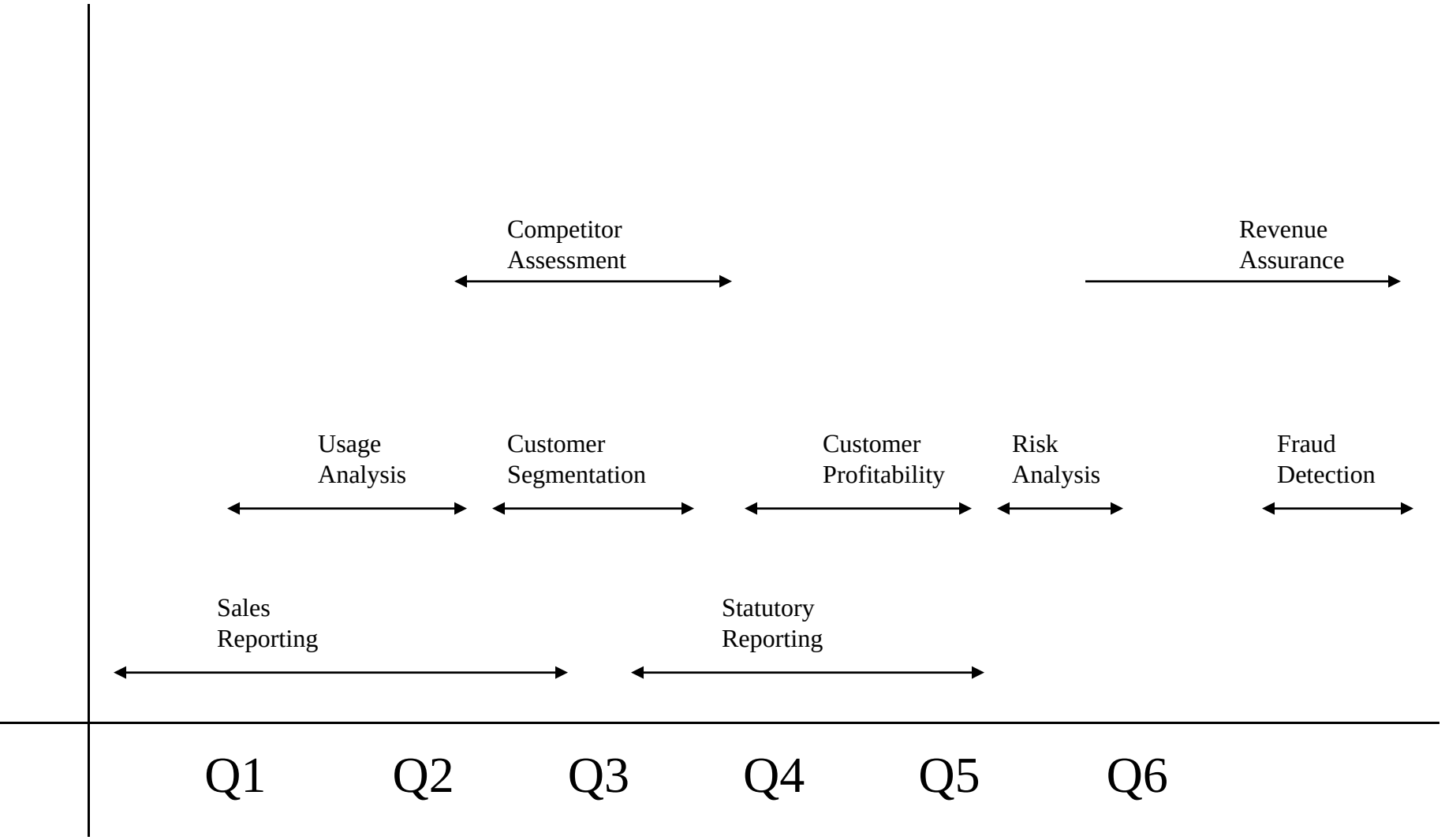
CDR’s


Process
Applications
Organisation
Technology
Information
Pebbles
Business Questions
Data
Critical
Supporting
LDM
PDM

Prioritising Pebbles




High
Mean
Low
Value
Customer Segmentation
Usage Analysis
Revenue Assurance
Customer Profitability
Risk Analysis
Statutory Reporting
Basket Analysis
Fraud Detection
Competitor Assessment
Sales Reporting
Mean
High
Difficulty


Competitor Assessment
Revenue Assurance
Usage Analysis
Customer Segmentation
Customer Profitability
Risk Analysis
Fraud Detection
Sales Reporting
Statutory Reporting
Q1
Q2
Q3
Q4
Q5
Q6

Let’s look at a worked example

Workbook 1

BI Strategy Documents and Owners

Key Information Imperatives – owner Information Master
Business Intelligence Roadmap – owner Information Mast
BI Pebble Definitions – owner Information Manager
Data Dictionary – owner Information Master
Skills Matrices – owner Information Master or Information Manager as relevant.
Security Policy – DBA
Access Policy – DB
Information Architecture – Information Architect
Data Reuse Map – Database Designer
Logical Data Model – Database Designer
Physical Data Model – Database Designer
User Requirements – Information Creator
Various technical documents - various
Various Service level Agreements – Information Manager/Master as appropriate

Building a Business Intelligence Architecture Fit for the 21st Century

Part 3

Jon Page

What Have We Learned So Far?

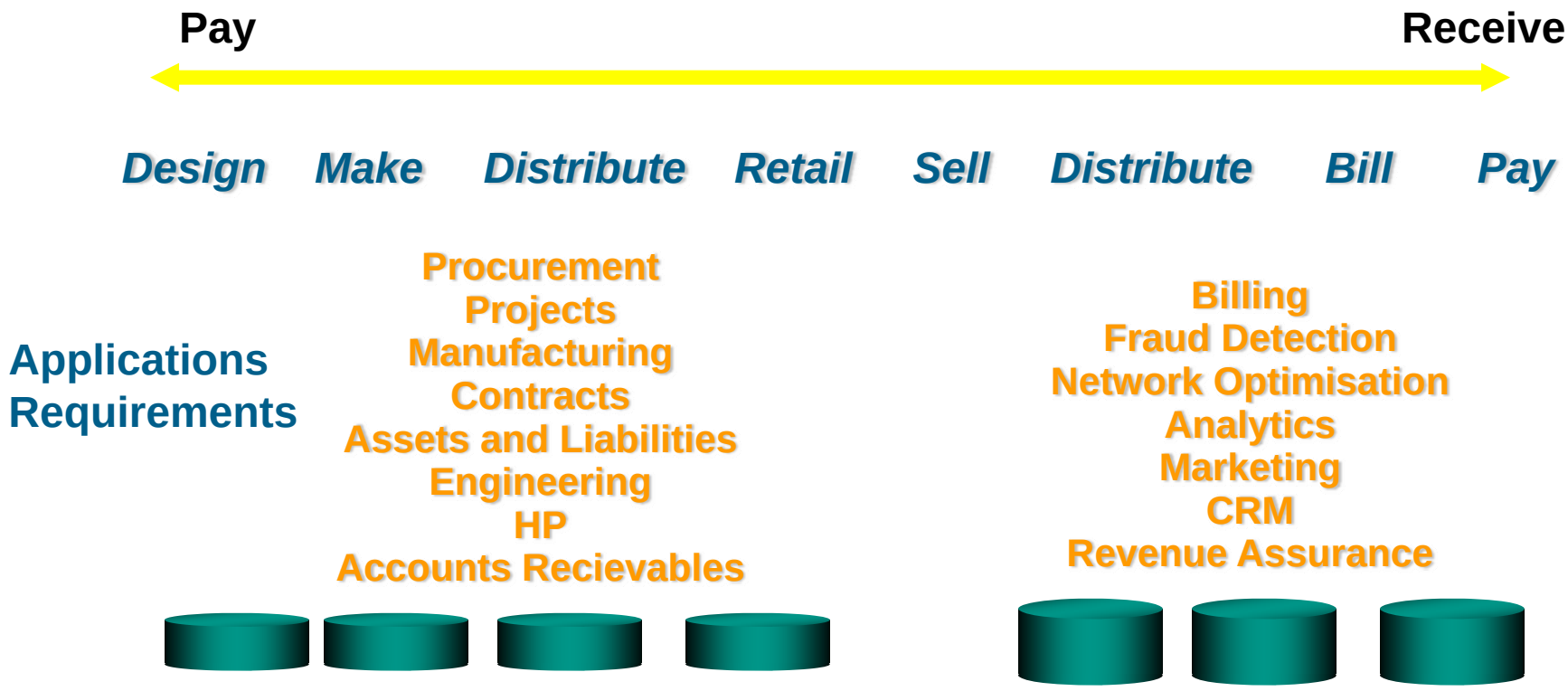
BI has not matured as fast as we maybe believe
The gap between what we need and what we have is getting bigger
The goal is to be able to predict and maybe change the future
Business and IT MUST work together
A BI Strategy linked to Corp Strategy is vital
Contemporary BI platforms – Data Marts and Data Warehouses are a good start but that's all. They're:

Full of useless inaccurate data
No one knows where the data came from or how old it is
It's copy miss-management gone crazy

We need something that:
Is trusted
Manages a single copy of master data
Can provide different 'types' of info – strategic, regulatory, ad-hoc, DSS, history, trend
Recognises that data is of many natures


Pay
Receive
Design
Make
Distribute
Retail
Sell
Distribute
Bill
Pay


Pay
Receive
Design
Make
Distribute
Retail
Sell
Distribute
Bill
Pay
Research and Design
Promote and Sell
Collect
Logistics
Logistics
Manufacture
Bill



Question 3 – What are the main differences between these two sets of data?

We need to be able to store and manage many different types of data automatically

Transactional, detailed, historical	- vast volume
Back-office, operational Data	- on-line?
Master Data	- on-line
Regulatory Data	- accurate
Summarised Data	- automatically created
KPI’s	- automatically maintained – complex algorithms
Metadata	- in one place, meaningful

What We Need to do Now:

- Build a single BI Platform that:
- Stores and Manages Critical Data (Master Data) in a real time way and as a service to applications
- Stores and makes available detailed, historical, transactional data
- Manages automatically all summary data
- Provided fresh KPI’s
- Is the basis for trusted and repeatable regulatory reporting
- Services ad-hoc complex queries and transactional queries
- Is self managed

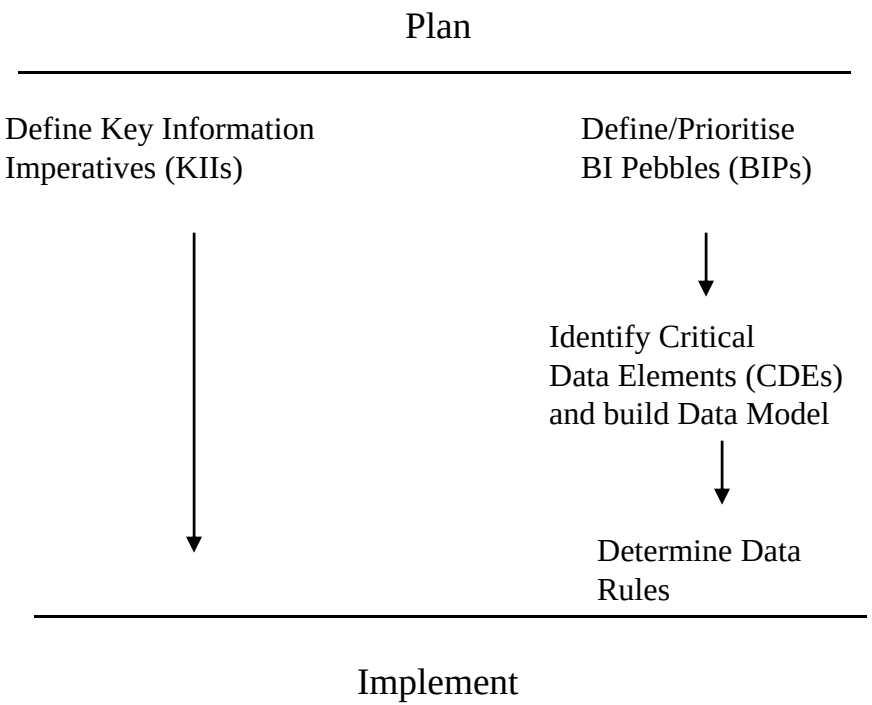
Collaborative Information Blueprint

Some fundamentals:

It depends on collaboration between Business and IT
It needs a BI strategy as described in Part 2
It builds on what we've learned in Data Warehousing
It is an incremental strategy
It is a centralised, data driven initiative
It's a set of guiding principles rather than a methodology

Depends on a centralised platform that is TIGHTLY coupled with the operational world

Steps 1 to 4 of CIB



Step One – Information Imperatives

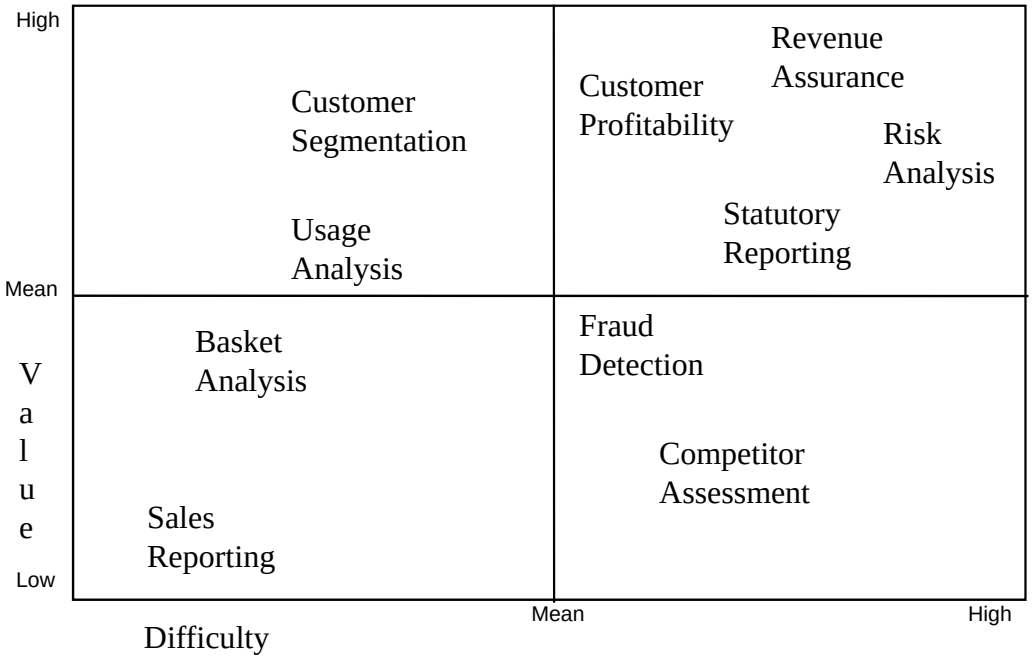
We must deploy all reporting across the web.
We will rigorously abide by all pertinent standards at all times.
We will keep a single copy of important data accessible to all.
Wherever possible we will pursue a single vendor strategy.
We will not use bleeding edge technologies to support core processes.
We will forbid the copying of data unless suitably authorised.
Non-compliant applications will only be used if a pathway to full compliance is possible within an 18 month window.
We will aggressively centralise all important data.

Step Two – Determine and Prioritise Pebbles
What is a Pebble?

They are opportunities that:

can be enabled by a technology based solution
need to exploit information
can be realised in no more than six months
are well understood by business and technology
are of high priority
are measurable in terms of success whether this be ROI, TCO or some other internalised mechanism
have a descriptive and meaningful name.
solve at least some part of the information facet of a project.

Perhaps the most important characteristic of pebbles is that they support agreed and strategic projects


High
Mean
Low
Value
Customer Segmentation
Usage Analysis
Revenue Assurance
Customer Profitability
Risk Analysis
Statutory Reporting
Basket Analysis
Fraud Detection
Competitor Assessment
Sales Reporting
Mean
High
Difficulty

Real ‘Pebbles’

- Management of churn in a specific segment of customers – let’s reduce churn by 5% in our customers who make more than 5 calls per day
- Revenue recognition – lets make sure that we improve by 5 points the ratio between service usage, billing and collection across the board
- Increase acquisition rate of customer in a specific segment – let’s increase rate by 2% for corporations with 50 to 300 employees.
- Decrease fraud – let’s reduce subscriber fraud by 3 points across all customers
- Understand product affinity – lets’ determine which products are bought in combination to allow us to more effectively deploy ‘loss leaders’.

Step Three) Identify CDE and Model

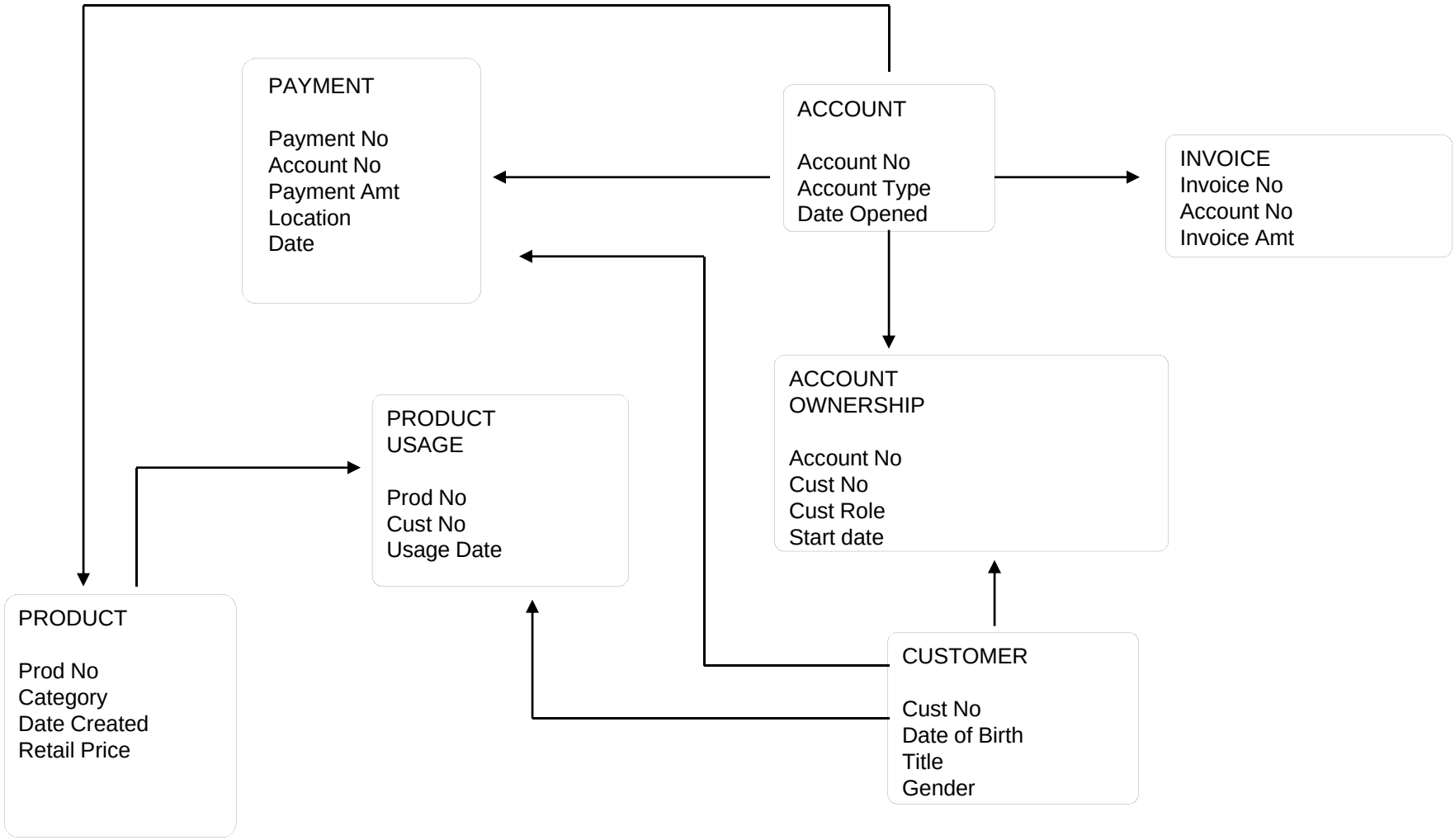
	Customer	Invoice	Bill	Office	Payment	Supplier	Product	Credit Score	Employer
BIP1	y	y	y		Y				
BIP2	y					y	Y		
BIP3	y	y	y	y	y		Y		
BIP4		y			Y	y	y		
BIP5	y							y	Y
BIP6	y	y			Y				
Total	5	4	2	1	4	2	3	1	1

Now We Have A Clear Justifiable View on what is OUR:

MASTER DATA – (working definition for this seminar!!!)

The Data That Supports Most of our Critical BI Requirements (Pebbles)

This is KEY – because I propose that we store this data (at least), in just one place and one place ONLY!


PAYMENT
Payment No
Account No
Payment Amt
Location
Date
ACCOUNT
Account No
Account Type
Date Opened
INVOICE
Invoice No
Account No
Invoice Amt
ACCOUNT
OWNERSHIP
Account No
Cust No
Cust Role
Start date
PRODUCT
USAGE
Prod No
Cust No
Usage Date
PRODUCT
Prod No
Category
Date Created
Retail Price
CUSTOMER
Cust No
Date of Birth
Title
Gender

Step 4 - Define Rules Around CDEs

- **Sample Business rules may include:**

- Account Open Date must be 12 years later than Birth Date of Account Owner
- Customer Name cannot be ‘null’
- A company must be allocated a SIC code from the published list
- An account is ‘open’ if any type of contact has been associated with this account in the past 20 days
- Address must always have a post code

- **Sample Technical rules may include:**

- Account Number must be 6 digits
- ‘Male’ will always be represented as ‘M’
- Value must be capitalised
- Value must be from following list……
- All table names should be singular

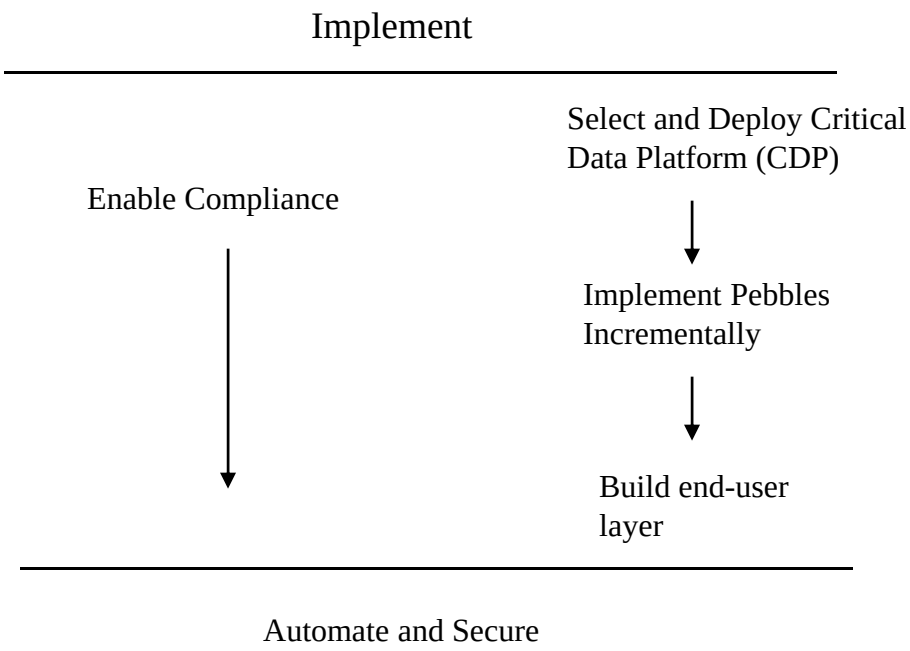
- **Process rules may include:**

- If there are more than ten of these marked ‘O’ then….
- This attribute is the result of function X on data item Y
- This cannot be deleted whilst…..

After Phase One we have:

- A prioritised map for BI implementation
- A set of IT Imperatives to guide architecture, infrastructure, purchasing etc
- A clear definition of Master Data, including rules etc
- A core Logical Data Model

Steps 5 to 8 of CIB



Step 6 - Technology must be selected to create the Critical Data Platform (CDP), and this will include at a minimum the:

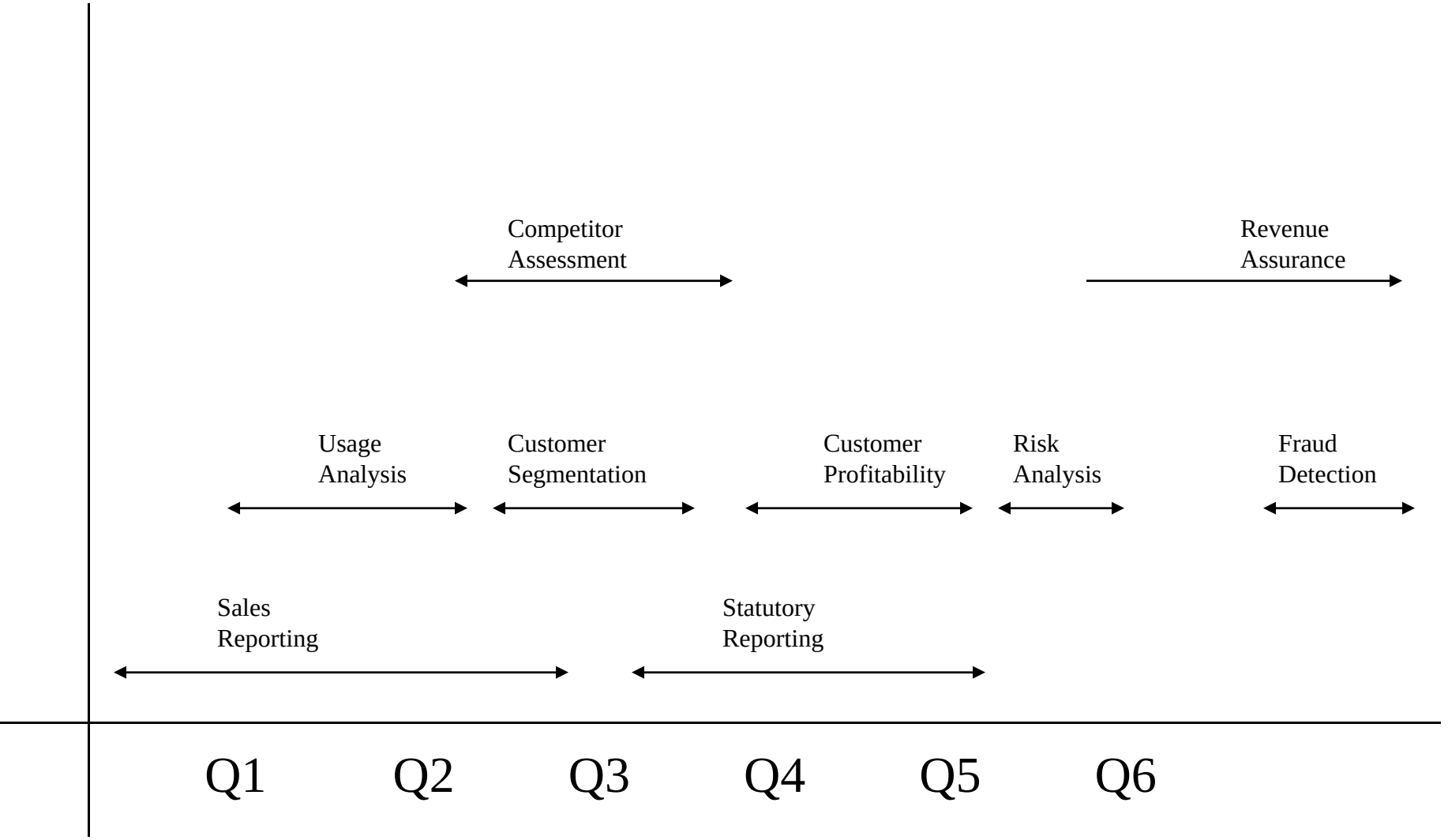
- Computing platform (hardware)
- Operating System plus management tools
- RDBMS
- Communications capability
- Business Intelligence tools (reporting, ad-hoc query, mining, OLAP)
- Integration tools
- Metadata Management tools/capability
- Backup and Recovery capabilities
- Security software

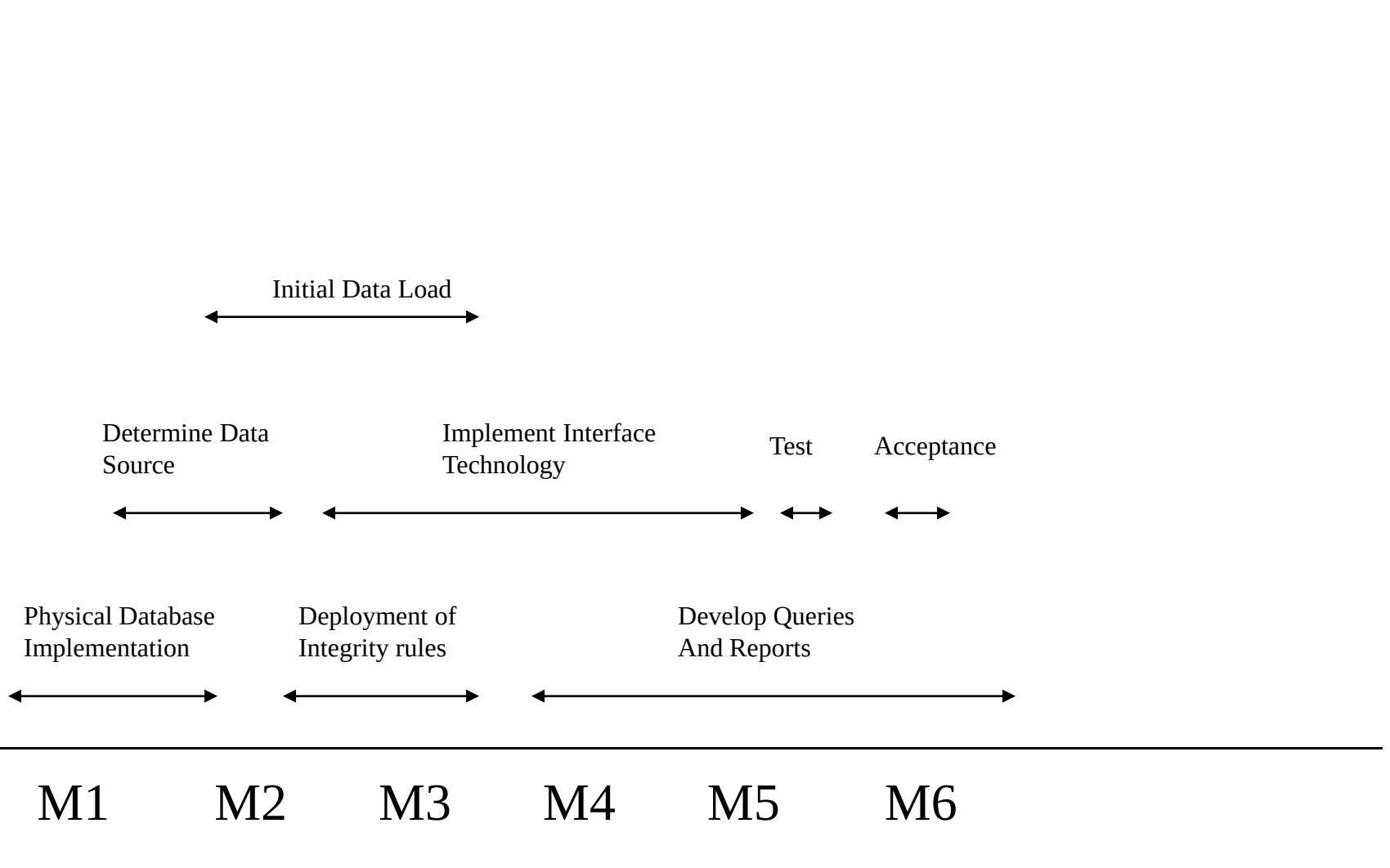
Question 5:

WHO CAN PROVIDE THE PLATFORM?

One ‘computer’
One Data Base
CDE’s on-line to operational systems
All data feeds automatic, in and out

Step 7 – Implement Pebbles Incrementaly




Initial Data Load
Determine Data Source
Implement Interface Technology
Test
Acceptance
Physical Database Implementation
Deployment of Integrity rules
Develop Queries And Reports
M1
M2
M3
M4
M5
M6

Step 8) Enable Compliance

- Get the right answers
- Get these answers quickly and reliably
- Demonstrate how numbers are calculated (audit)

- The Implement phase brings into a single database:
- Key shared data for enabling Corporate strategy (often called ‘Master Data’ but actually much more.
- Operational data needed by multiple operational applications
- Derived data for complex algorithms and scoring
- Summary data often used to form the basis of Compliance calculations as Key Performance Indicators
- Detailed, historical, behavioural data for deep mining and predictive analysis.

So in our preferred architecture, the tables we have modelled to represent our Critical Data Elements will be implemented in a single database and as a set of relational tables.

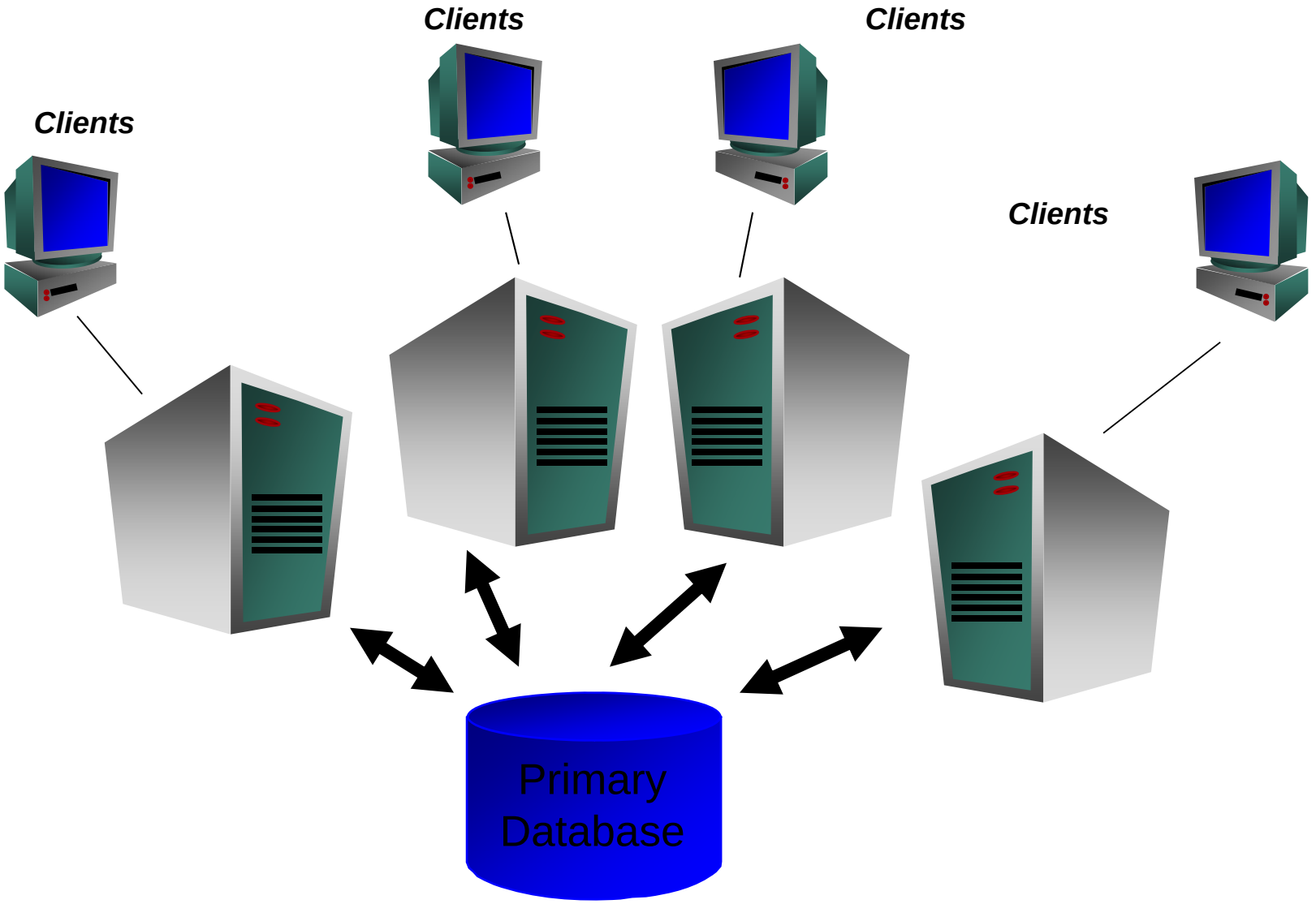
- As these tables are deemed critical to the enterprise it is imperative that:

- We have only one set of these tables in the whole enterprise.
- These tables are shared by all applications that need this data whether these applications are internal or external, operational or analytical.
- This data is protected from miss-use by a defined layer of metadata.
- The data in these tables is always up-to-date.
- The data in these tables records history.

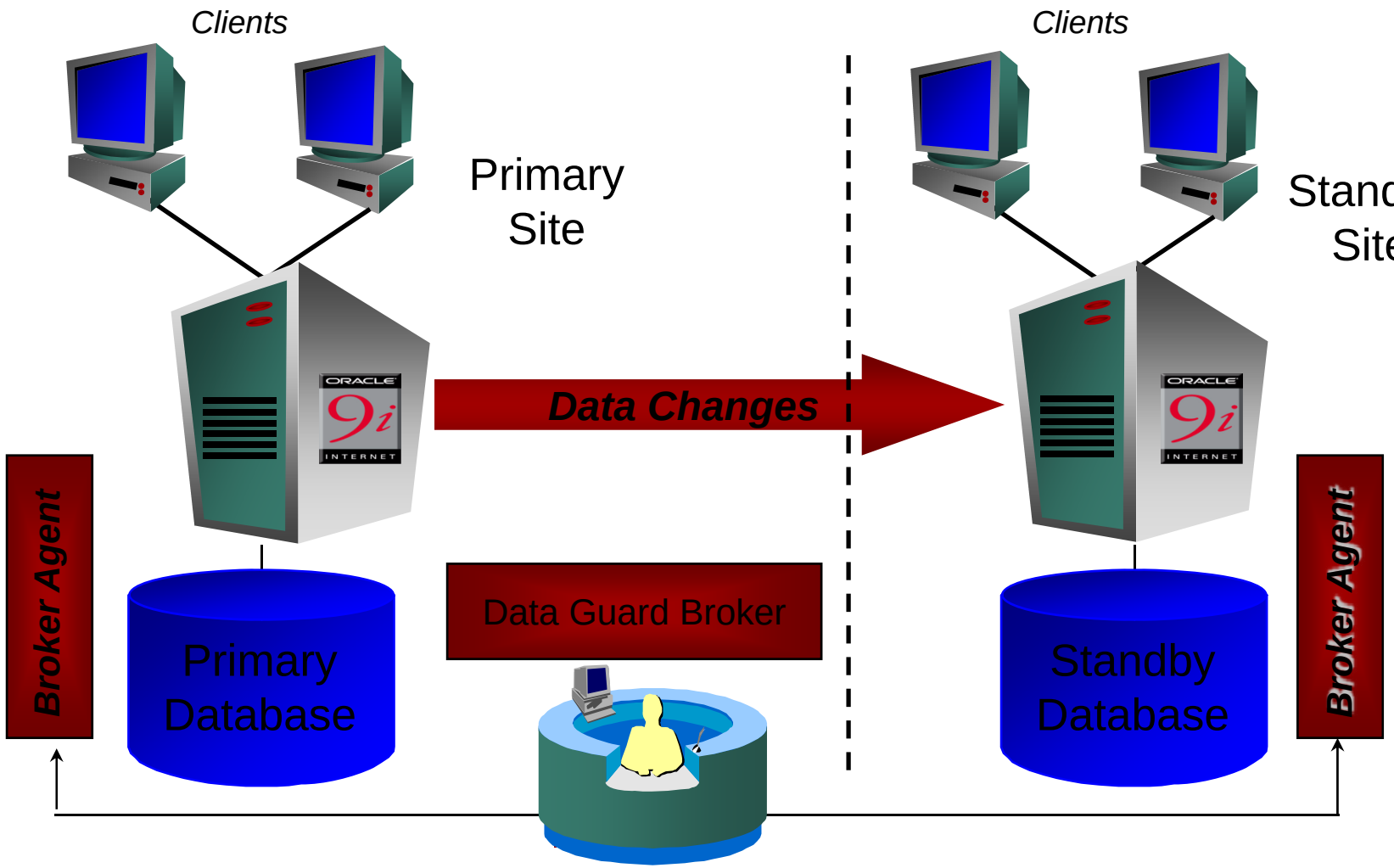
Step 9 and 10 of CIB

Automate and Secure

Ensure 100% Availability	Automate and Secure



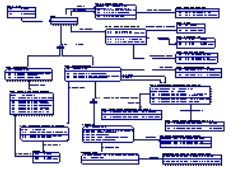
A Cluster for Reliability?



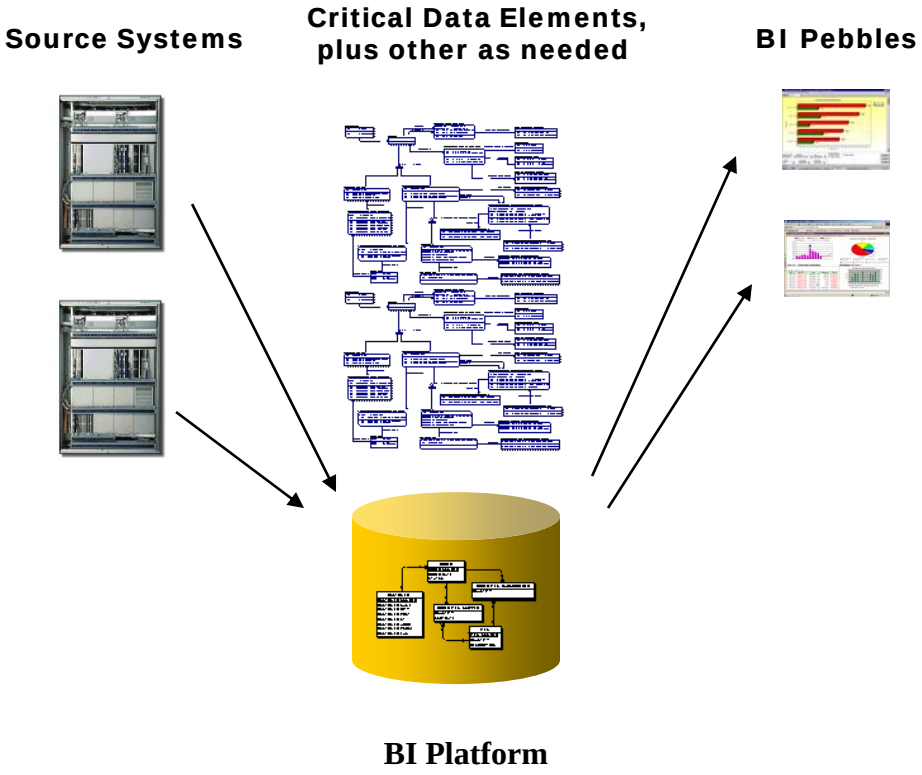
Remote Backup?

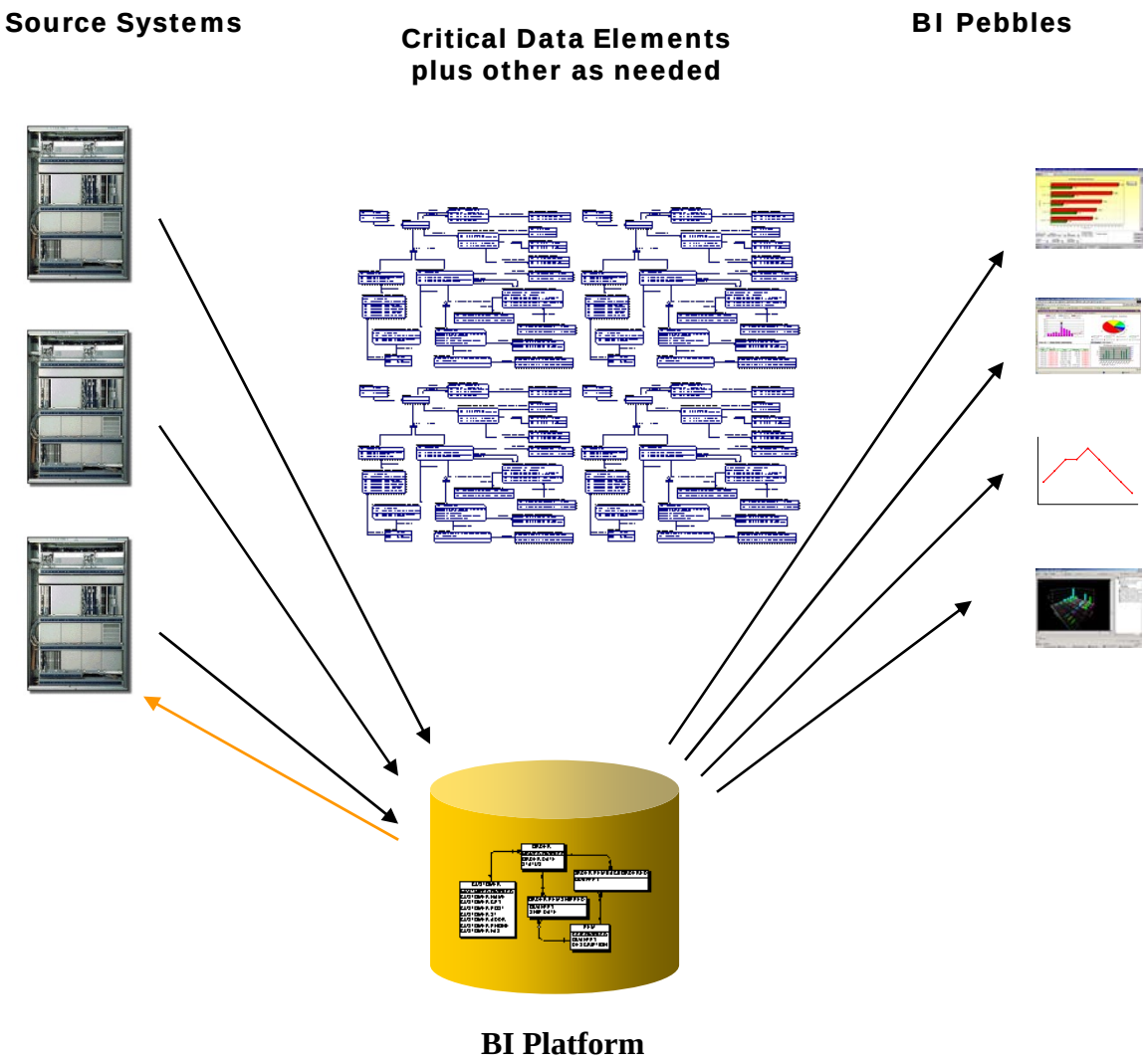
Source Systems

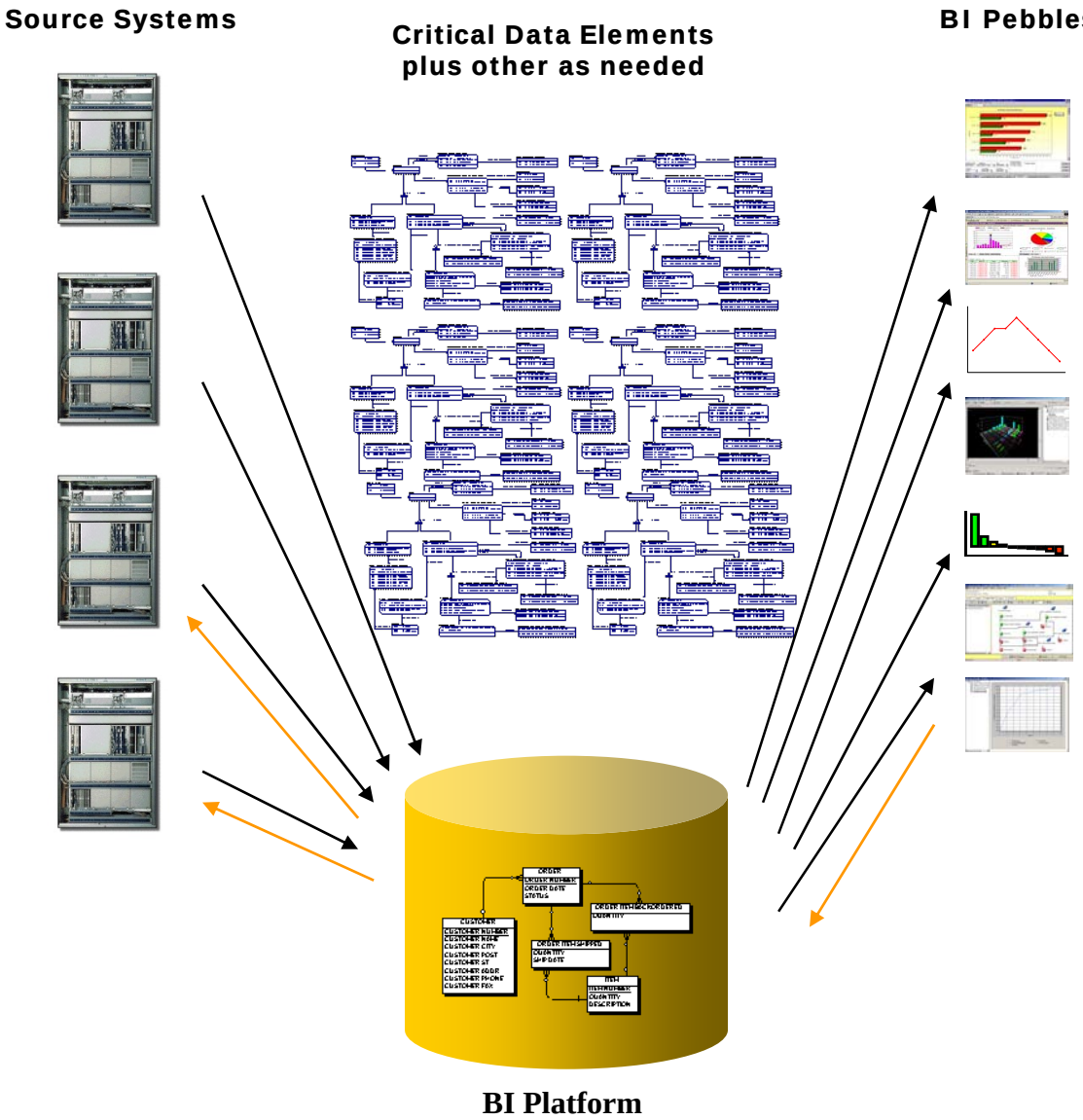
Critical Data Elements
Logical Model

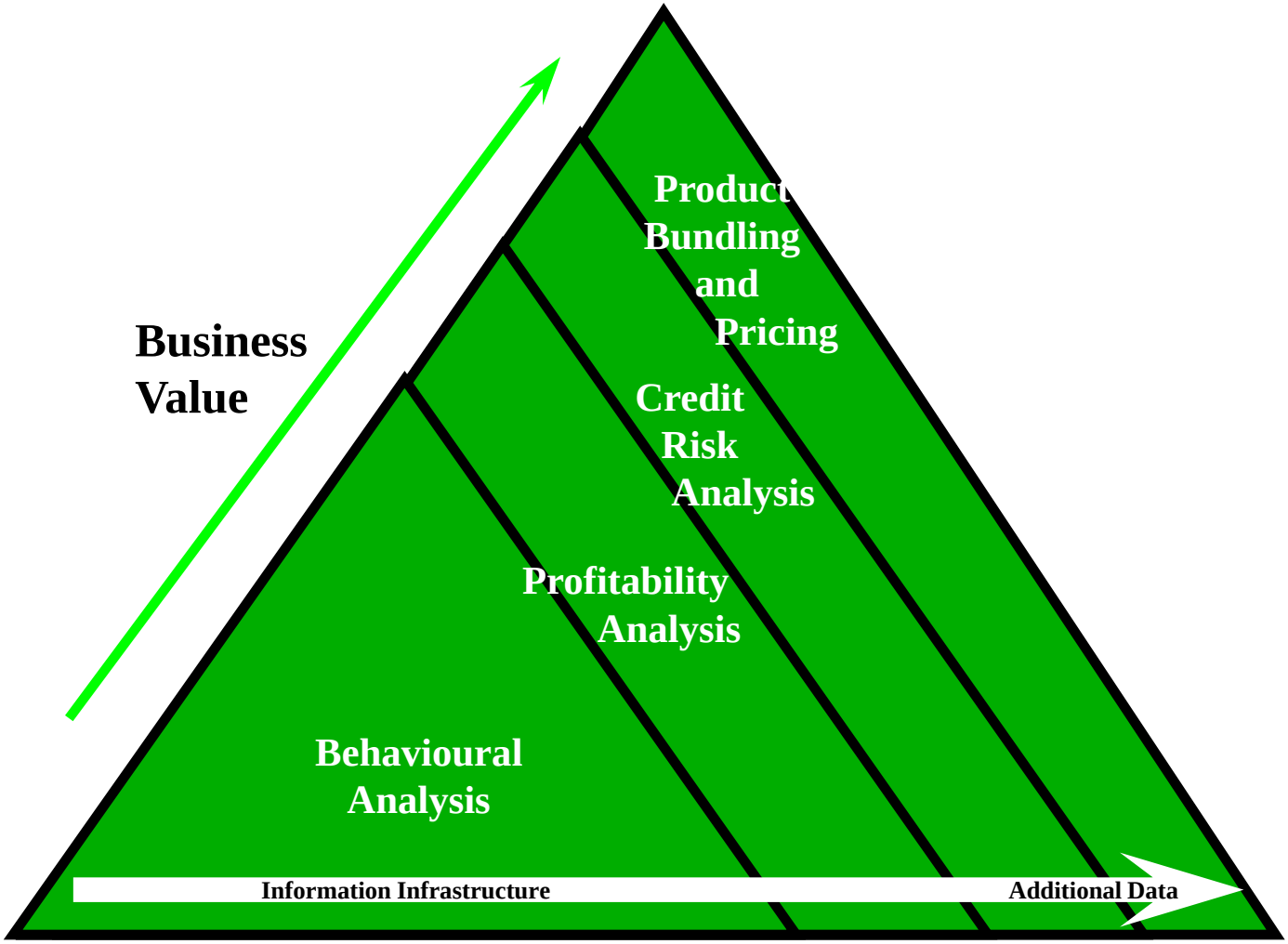


BI Pebble


Source Systems
Critical Data Elements, plus other as needed
BI Pebbles
BI Platform


Source Systems
Critical Data Elements
plus other as needed
BI Pebbles
BI Platform


Source Systems
Critical Data Elements
plus other as needed
BI Pebbles
BI Platform


Business Value
Product Bundling and Pricing
Credit Risk Analysis
Profitability Analysis
Behavioural Analysis
Information Infrastructure
Additional Data

Summary

- Define and prioritise requirements – already done in our BI Strategy
- Define first deliverable – already done in our BI Strategy (see BI Pebble Road Map)
- Build Logical Data Model for first Pebble - already done
- Build Physical Model to support initial requirement
- Deploy rules as triggers, procedures, macros etc.
- Define preferred provider system for each
- Define interface requirements of each item of data with provider system
- Perform initial data load to database
- Deploy the required data feed technologies
- Test system

How do we pay more for less

- Tailoring pre-canned applications
- Building our own logical data model
- Not using the technology that we have already paid for
- Buying data we already have
- Issuing RFIs and RFPs
- Evaluating identical technologies
- **Copying data**

Summary

- Build a comprehensive (but flexible) BI strategy linked to the Corp Strategy
- Develop IT Imperatives
- Put in place the right organisation
- Cherish Critical Data
- Build a closer but more demanding relationship with your vendors
- Use technology
- Promote yourselves – BI is the most important facet of most internal business

Building a Business Intelligence Architecture Fit for the 21st Century

Part 4

Jon Page

BI Projects fail for many reasons, we continually see:

Insufficient dedication of resources (technical and most importantly, human).

People with the wrong skills.

Multiple, unprioritised roles and responsibilities.

Lack of decision making power.

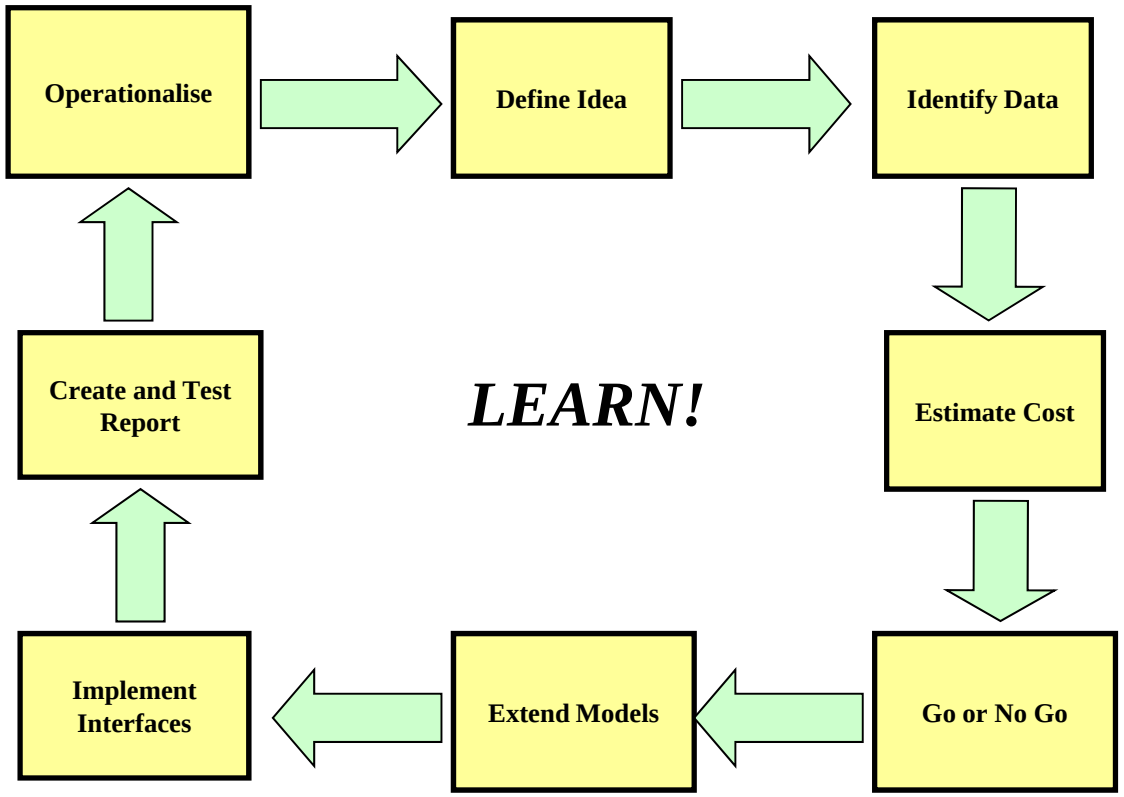
Rudimentary future planning.

Conflicting commitments.

Too many on-going projects

A Real DW Audit findings:

Lack of business involvement/ownership
No three year roadmap for expected deliverables
No formal Data Architecture or Logical Data Model
Poorly defined job descriptions/roles – confusion of responsibilities
Non-optimal deployment of Hardware
Using out of date software
No consistent application of Parallel Database features
Many important Oracle facilities not being used effectively
Lack of documentation
Non-scalable environment
Critical data was missing
No deep usage of data in terms of analytics/data mining etc


Operationalise
Define Idea
Identify Data
Estimate Cost
Go or No Go
Extend Models
Implement Interfaces
Create and Test Report
LEARN!

The Information Board

- Creation of Corporate Strategy
- Ensuring alignment of BI Strategy
- Providing support for deployment of CDP
- Providing guidance and support to the BI ‘team’
- Solving critical issues
- Arbitrating on internal conflicts
- Monitoring return on Investment and/or Total Cost of Ownership goals
- Appointment of the Information Master

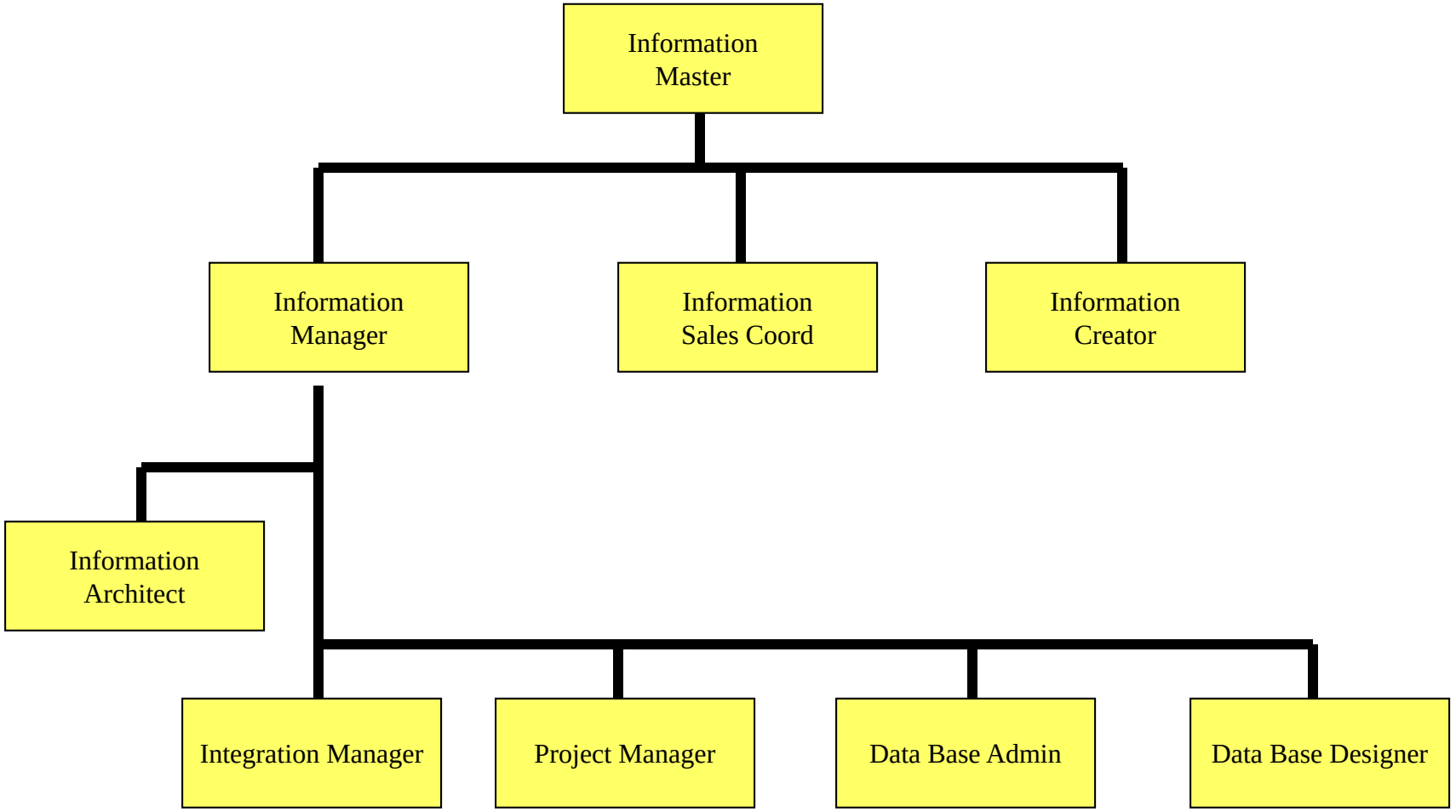
The Information Master:

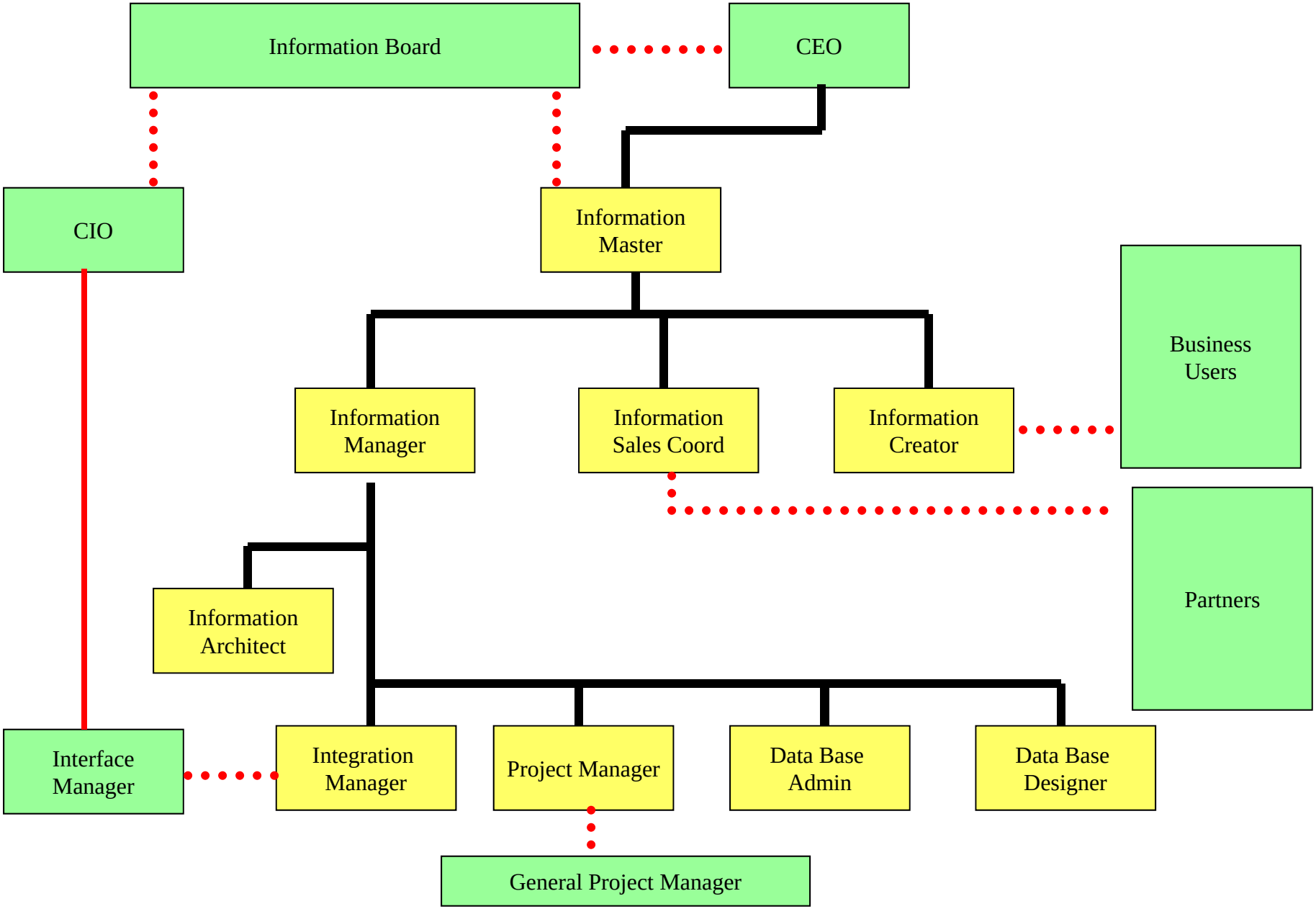
is ultimately responsible for the creation and maintenance of the following documentation:

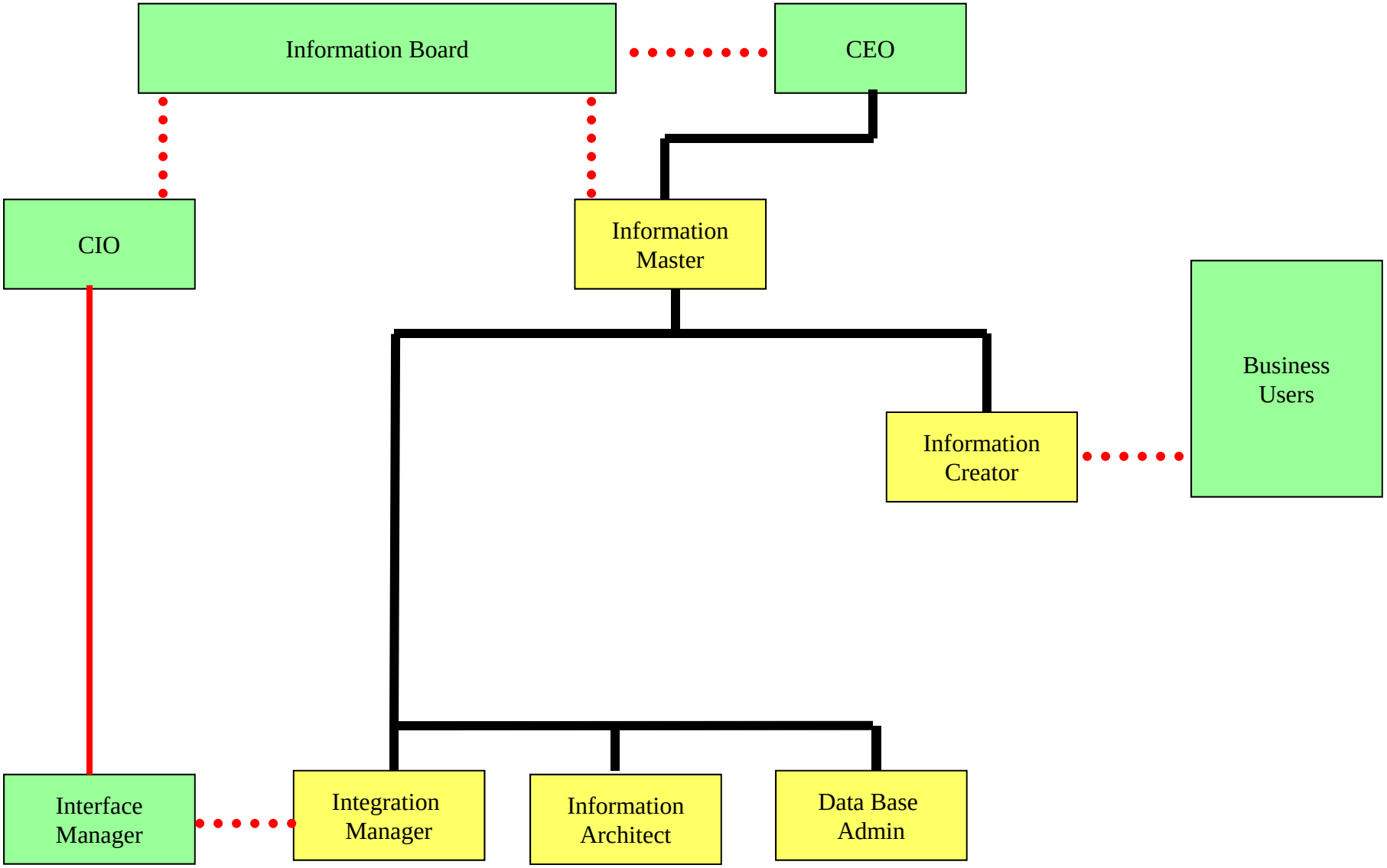
Key Information Imperatives
BI Roadmap
Data Dictionary
The complete Business Intelligence Strategy
Skills Matrices for the BI team

BI Strategy Documents and Owners

Key Information Imperatives – owner Information Master
Business Intelligence Roadmap – owner Information Mast
BI Pebble Definitions – owner Information Manager
Data Dictionary – owner Information Master
Skills Matrices – owner Information Master or Information Manager as relevant.
Security Policy – DBA
Access Policy – DB
Information Architecture – Information Architect
Data Reuse Map – Database Designer
Logical Data Model – Database Designer
Physical Data Model – Database Designer
User Requirements – Information Creator
Various technical documents - various
Various Service level Agreements – Information Manager/Master as appropriate


Information Master
Information Manager
Information Sales Coord
Information Creator
Information Architect
Integration Manager
Project Manager
Data Base Admin
Data Base Designer


Information Board
CEO
CIO
Information Master
Business Users
Information Manager
Information Sales Coord
Information Creator
Partners
Information Architect
Interface Manager
Integration Manager
Project Manager
Data Base Admin
Data Base Designer
General Project Manager


Information Board
CEO
CIO
Information Master
Business Users
Information Creator
Interface Manager
Integration Manager
Information Architect
Data Base Admin

How do we pay more for less

- Tailoring pre-canned applications
- Building our own logical data model
- Not using the technology that we have already paid for
- Buying data we already have
- Issuing RFIs and RFPs
- Evaluating identical technologies
- **Copying data**

Summary

- Build a comprehensive (but flexible) BI strategy linked to the Corp Strategy
- Develop IT Imperatives
- Put in place the right organisation
- Cherish Critical Data
- Build a closer but more demanding relationship with your vendors
- Use technology
- Promote yourselves – BI is the most important facet of most internal business