

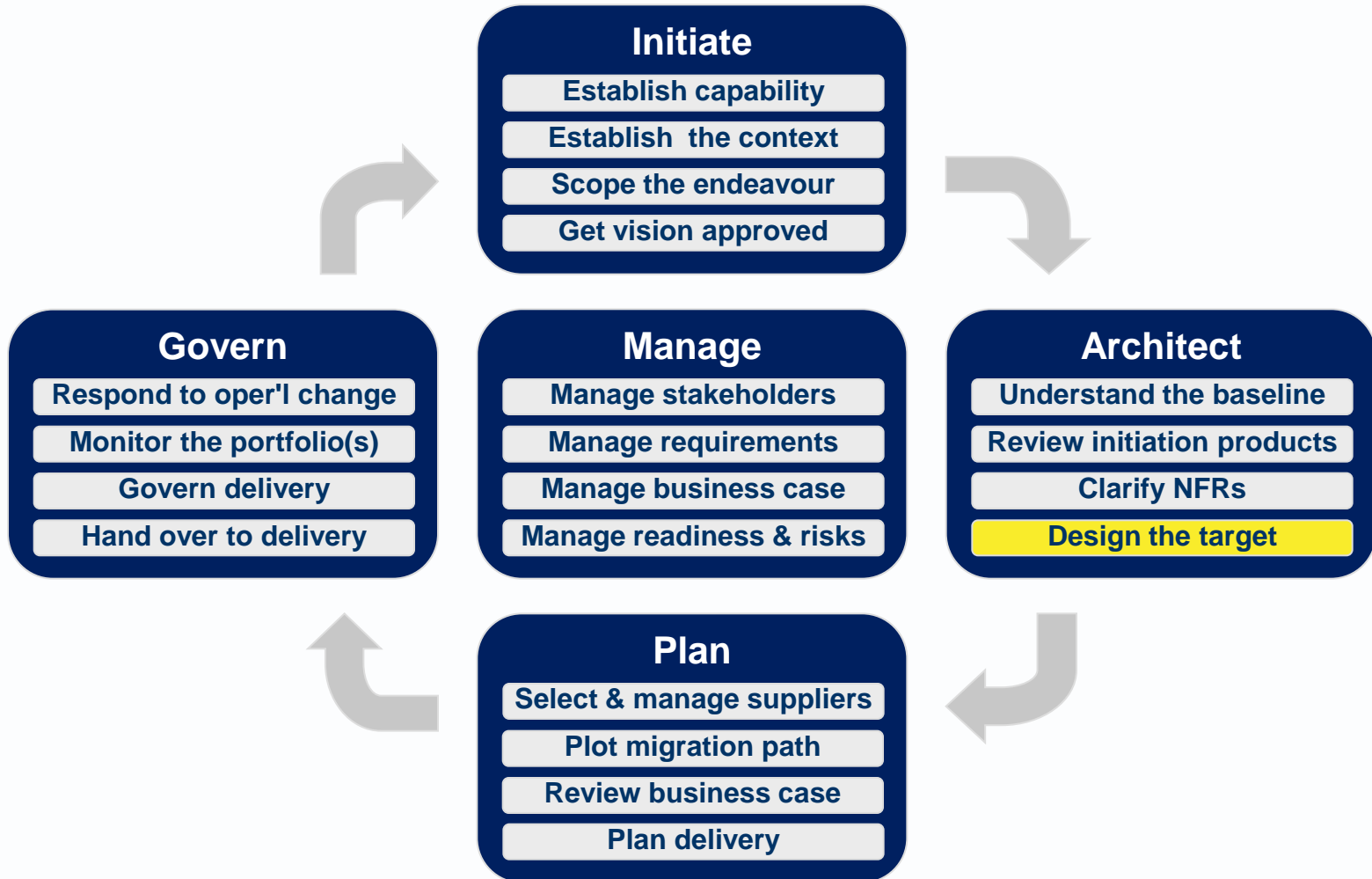
Avancier Methods (AM)

Data Architecture

Define the business context for data
Define data flows and dictionaries

It is illegal to copy, share or show this document
(or other document published at <http://avancier.co.uk>)
without the written permission of the copyright holder

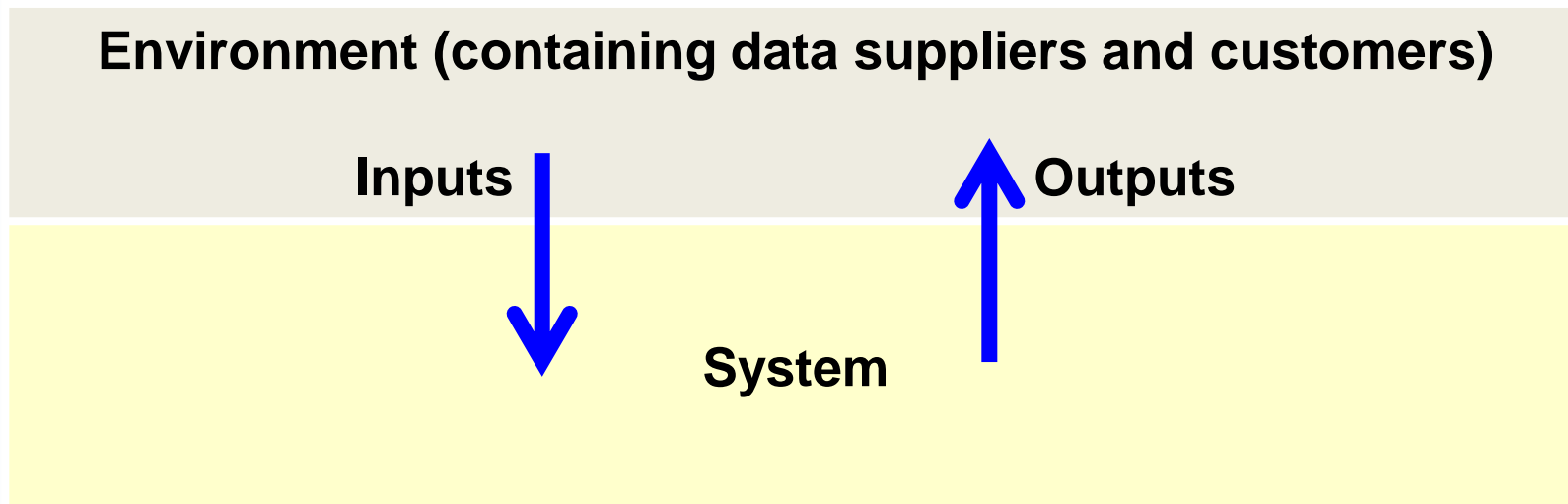
Design the target (AM level 2)



Design target data architecture (AM level 3)

- 1. Define the business context for data creation and use**
2. Define data flows (I/O messages, displays, forms and reports)
3. Define data dictionary or canonical data model
4. Define data store(s): relational and document stores
5. Address data quality issues

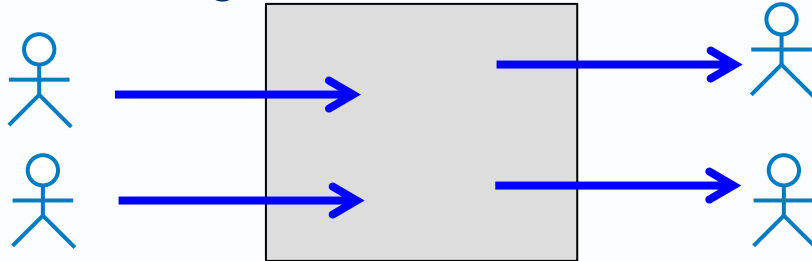
- ▶ Architects define a system from out to in
- ▶ Starting with what crosses the system's input/output boundary.



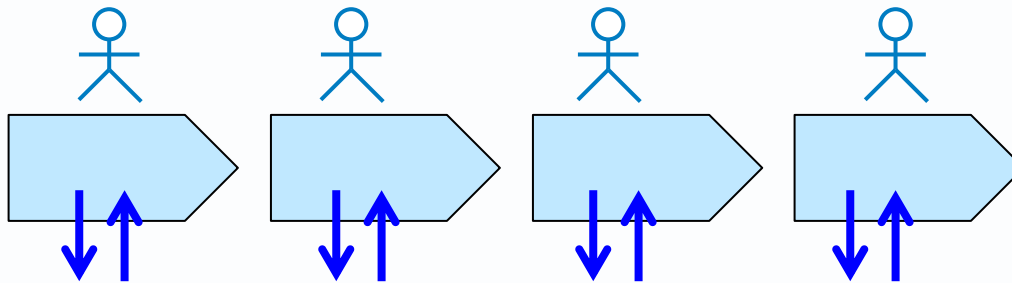
- ▶ The system can be divided into layers and/or subsystems that are definable in the same way.

Define the business context for data creation and use

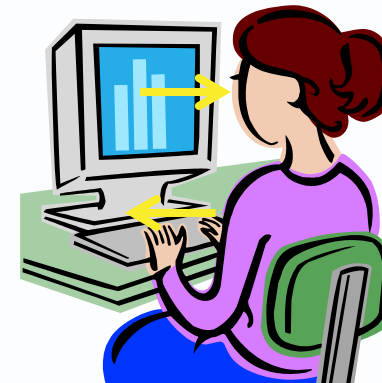
▶ Context diagram



▶ Value stream / scenario diagrams (showing OPOPOT activities)

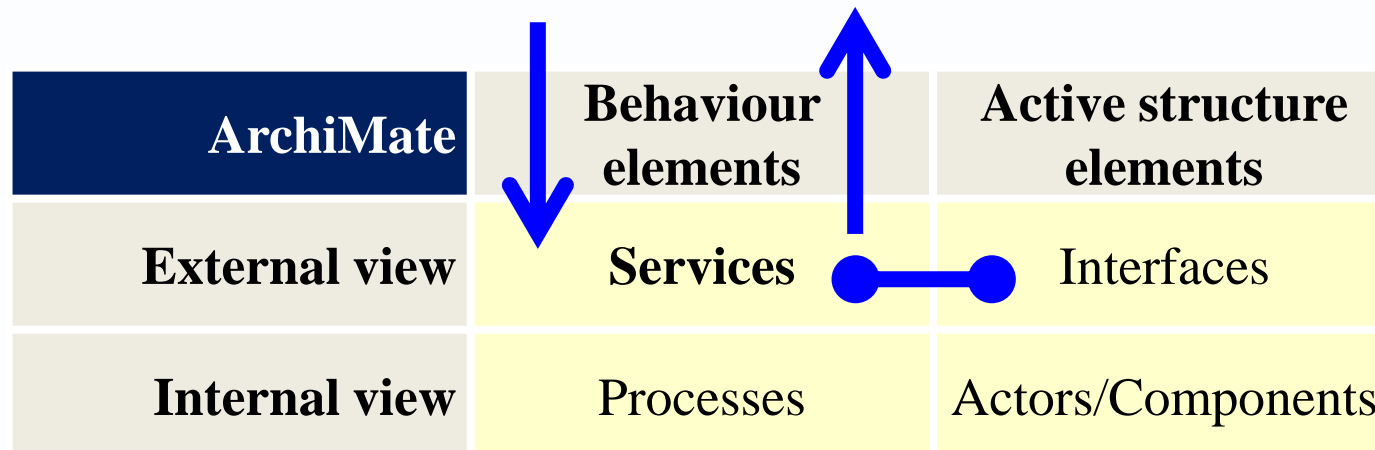


▶ Client devices and user interfaces



Defining data flows in service contracts

- ▶ A system, subsystem, building block, role or component is definable by its interface(s).
- ▶ An interface lists requestable services
- ▶ A service contract
 - encapsulates (hides) internal process flows and actors/components
 - should identify input/output data flows



Design target data architecture (AM level 3 and 4)

1. Define the business context for data creation and use
2. **Define data flows (I/O messages, displays, forms and reports)**
 - **Data structures used to perform activities in roles or processes**
 - **Data structures created in performing activities in roles or processes**
 - **Data structures passed between applications**
3. Define data dictionary or canonical data model
4. Define data store(s): relational and document stores
5. Address data quality issues

- ▶ Leading bank is urgently seeking a proven EA to engage and lead IT projects including
 - the Enterprise Information Architecture
 - **information models and flows,**
 - data dictionaries, data standards
 - data quality standards and processes

- ▶ develop and maintain the logical Enterprise Information Architecture that enables seamless information interoperability of all Bank systems for efficiency and cost-effectiveness.

- ▶ eutopiaonline.com

Define documents needed/used to perform activities

Salesman wants

Customer Order History
Customer id
Customer name and address
Orders Placed
Order id
Order value
Products Ordered
Product type
Product amount
Products Ordered End
Order Placed End
Customer Order History END

Product management

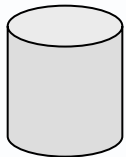
Product Demand
Product type
Amount on hand
Products ordered
Product amount
Order id
Products Ordered End
Product Demand End

Define documents created in the course of performing activities

- ▶ Customer creates

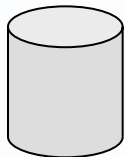
Shopping Basket
Customer id
Order id
Order value
Products Ordered
Product type
Product amount
Products Ordered End
Shopping Basket

- ▶ The HR department maintains a spreadsheet of all employees



Human resources
Employee Number, Name, Role, Grade

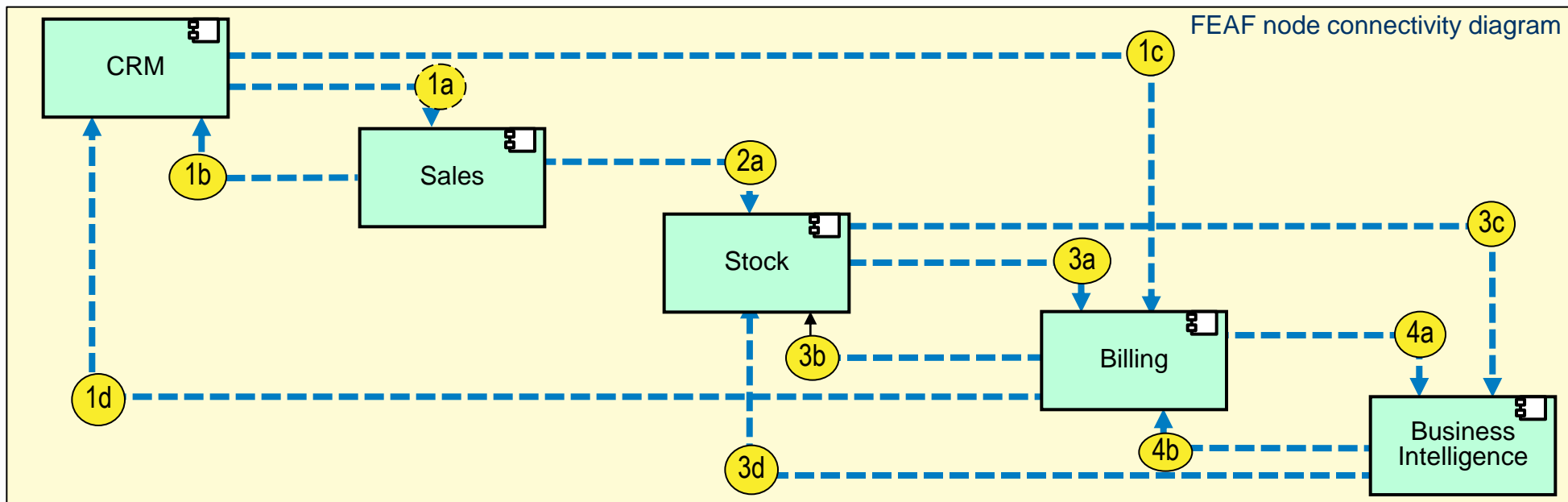
- ▶ The sales manager has a card file with all salesmen in it



Salesman card file
Employee Number, Name, Commission Rate, Sales Area

Define inter-application data flows

Data Flow (Application Communication) Diagram



Data Flow Catalogue

Id	Trigger event	Source App	Destination App	Data content	Technology
1a	Order entry	CRM	Sales	Sales order request	ESB
1b	Order accepted	Sales	CRM	Sales order confirmation	ESB
2a	Notification	Sales	Stock	Requisition	ETL

Define finer-grained inter-application data flows?

▶ All the arrows indicate data flows

- Human to human
- Human to computer
- Computer to computer

1. Sender

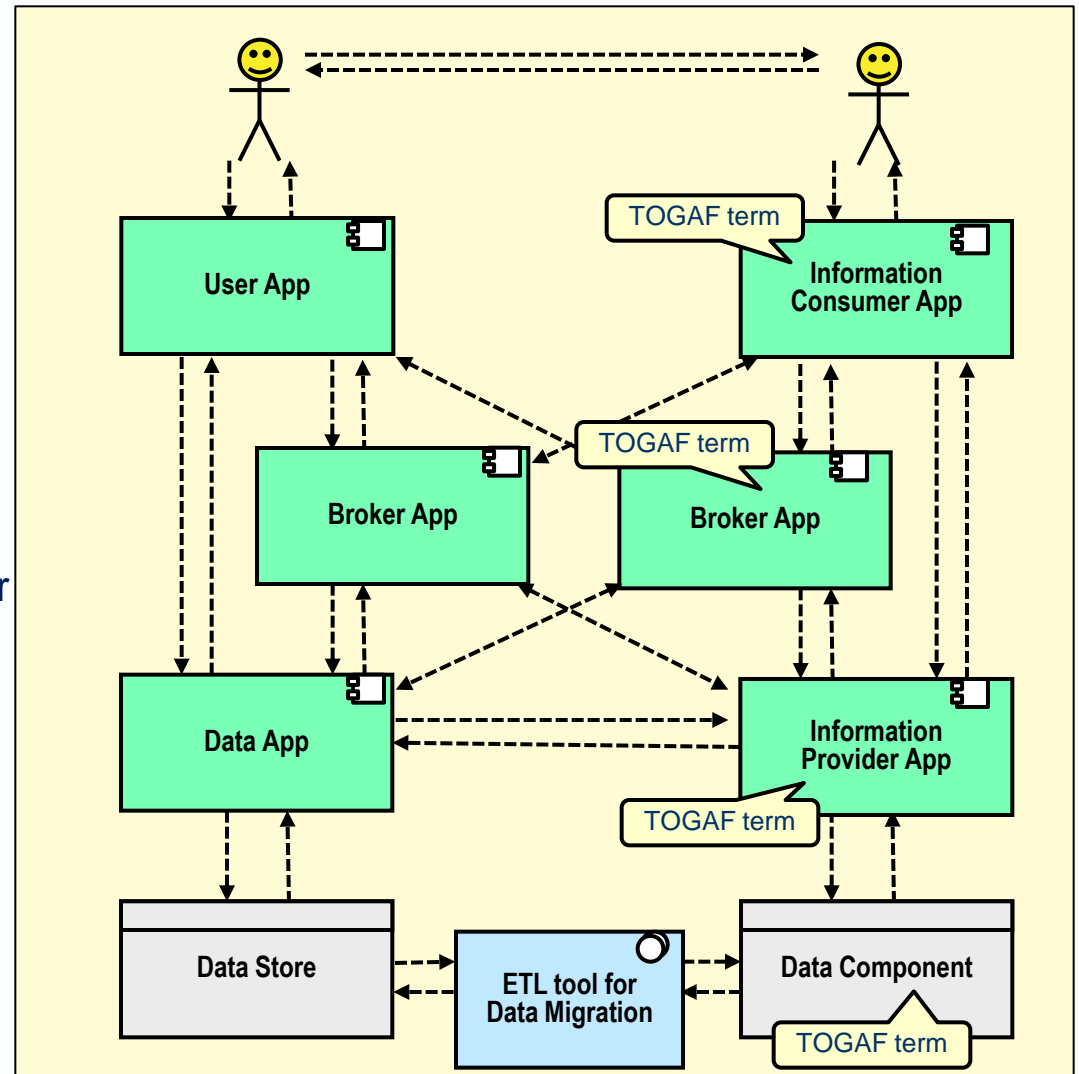
- marshals facts to be sent

2. Transport

- conveys data from sender to receiver via a physical medium

3. Receiver

- detects facts in the flow and manipulates them to some purpose



Catalogue the most significant data flows

A Data Flow Catalogue (Interface Catalogue in TOGAF)				
	Flow name	Enquiry	Response	Order
Functional attributes	Trigger		Enquiry	
	Source	Customer	Sales	Customer
	Destination	Sales	Customer	Sales
	Information	Unstructured	Unstructured	Order details (tbd)
	Frequency	1,000/day	1,000/day	30/day
Non-functional attributes	Volume			500K
	Confidentiality	High	High	High
	Integrity	Medium	Medium	High
	Availability	24/7	09.00-18.00	24/7
Transport mechanisms	Technology	Web	Telephone	Web
	Protocol	HTTP		HTTPS

Like many such illustrations, this shows what could be documented rather than what most actually do. But understanding what is possible in theory is a precursor to deciding what to do in practice.

Design target data architecture (AM level 3)

1. Define the business context for data creation and use
2. Define data flows (I/O messages, displays, forms and reports)
3. **Define data dictionary or canonical data model**
4. Define data store(s): relational and document stores
5. Address data quality issues

Define data items in a data dictionary

- ▶ Leading bank is urgently seeking a proven EA to engage and lead IT projects including
 - the Enterprise Information Architecture
 - information models and flows,
 - **data dictionaries, data standards**
 - data quality standards and processes

- ▶ develop and maintain the logical Enterprise Information Architecture that enables seamless information interoperability of all Bank systems for efficiency and cost-effectiveness.

- ▶ eutopiaonline.com

Define data items in a data dictionary

1. Define data items appearing in documents
2. Define items that persist to make outputs possible

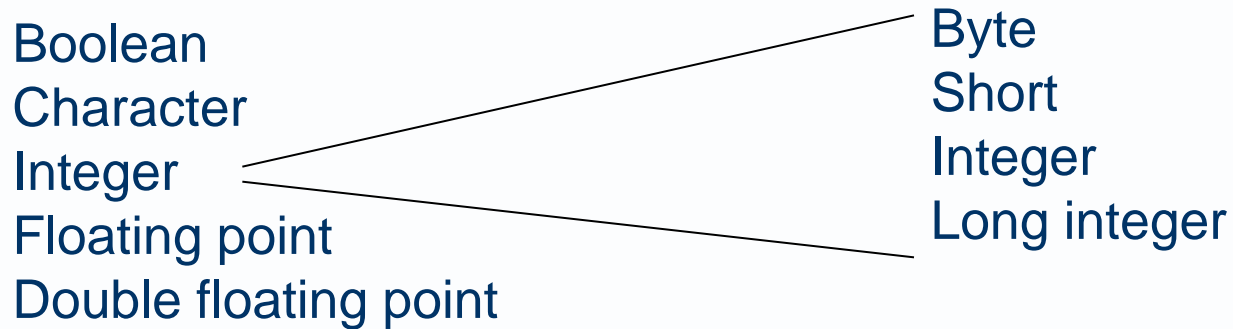
Object	Description (edited from Salesforce.com)
<u>Account</u>	An individual account, which is an organization involved with your business (such as customers, competitors, and partners).
<u>AccountContactRole</u>	The role that a given Contact plays on an Account.
<u>AccountPartner</u>	A relationship between two Account objects, such as partnerships or subsidiaries.
<u>AccountShare</u>	A sharing entry on an Account.
<u>AccountTag</u>	Associates a word or short phrase with an Account.
<u>AccountTeamMember</u>	A User who is a member of an Account team.
<u>AdditionalNumber</u>	An additional phone number for a CallCenter.
<u>AuthSession</u>	The AuthSession object represents an individual user session in your organization.
<u>Bookmark</u>	A link between two opportunities.

A more formal data dictionary

1. Define data items in input and output data flows
2. Define data items that must persist to make outputs possible
3. **Define business rules associated with data items**

Term	Currency Code
Facts	Currency Code [abbreviates] Currency [which denominates a] Value
Constraints	Currency [is a] three letter String [in the range] defined at
Derivation	
Term	Item Value
Facts	Item Value [is an attribute of an] Order Item Item Value [is associated with a] Currency
Constraints	Item Value [is a] Decimal Number [in the range] 000.00 to 999.99
Derivation	Item Value = Product Amount Ordered * Unit Price
Term	Order Value
Facts	Order Value [is an attribute of] Order Order Value [is calculated from] Item Values
Constraints	Order Value [is a] Decimal Number [in the range] 0000.00 to 9999.99
Derivation	Order Value = sum of (Item Values for an Order) - Discount

Primitive data types, for example in Java



Type	Contains	Default	Size	Range
boolean	true or false	false	1 bit	NA
char	Unicode character	\u0000	16 bits	\u0000 to \uFFFF
byte	Signed integer	0	8 bits	-128 to 127
short	Signed integer	0	16 bits	-32768 to 32767
int	Signed integer	0	32 bits	-2147483648 to 2147483647
long	Signed integer	0	64 bits	-9223372036854775808 to 9223372036854775807
float	IEEE 754 floating point	0.0	32 bits	$\pm 1.4E-45$ to $\pm 3.4028235E+38$
double	IEEE 754 floating point	0.0	64 bits	$\pm 4.9E-324$ to $\pm 1.7976931348623157E+308$

Primitive data types

- ▶ Boolean
- ▶ Text
- ▶ Number

User defined data types

- ▶ Name (Text)
- ▶ City (Text)
- ▶ Order value (Number)


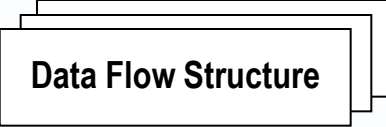

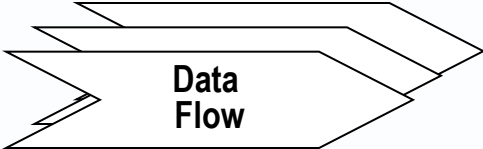
Complex data types (simple data structures)

- ▶ Date
 - DD
 - MM
 - YYYY
- ▶ Person
 - Title
 - First name
 - Last name
- ▶ Address
 - Address Line 1
 - Address Line 2
 - Address Line 3
 - City
 - County/State
 - Postcode



A source of many system integration problems

Canonical data model

Conceptual	Defines the ideal or common types for data items in data flows	
Logical		
Physical		
Real		

- ▶ **A canonical data model** can define
 - common data types and structures used in data flows
 - Staff [Staff name, NiNo, DoB, StartDate]
 - the compression or codification of data for transmission in messages.

- ▶ A data architect might
 - Define canonical data model at a logical level
 - In some kind of data dictionary

- ▶ An applications architect might
 - Have to design multi-point application integration
 - Encode the CDM in an XML schema
 - Use a message bus to transform all data flows into and out of the canonical data model form

CDM
Logical

CDM
XSD

Design target data architecture – TO FOLLOW

1. Define the business context for data creation and use
2. Define data flows (I/O messages, displays, forms and reports)
3. Define data dictionary or canonical data model
4. **Define data store(s): relational and document stores**
5. **Address data quality issues**