Semantics for Enterprise Architecture Workshop

This Presentation

- The presentation given at the Workshop **PLUS**
- The conclusions from the Workshop discussions
  - In added slides that are identified by 
- (Note that none of the backup slides at the end of the presentation were used in the workshop.)

Semantics for Enterprise Architecture Workshop

San Francisco

Wednesday 30th January 2008

THE Open GROUP

Dr Christopher J Harding
Forum Director

Tel +44 774 063 1520 (mobile)

c.harding@opengroup.org

Thames Tower
37-45 Station Road
Reading
RG1 1LX
UK

www.opengroup.org

Objectives

- Information Management and Semantics form the next big area in which middleware will take over functions traditionally performed by applications software.
- This will lead to big cost savings for enterprises - if they have the right architectures.
- The workshop will explore the incorporation of semantics into Enterprise Architecture in general, and TOGAF in particular.
- It will look at the different viewpoints involved, at some basic reference models, and at the use of ontologies and controlled vocabularies.
- The aim of this workshop, and of further work within The Open Group, will be to show how to produce enterprise architectures that make effective use of the semantic technologies that are now emerging.

Presentations to Follow the Workshop

- John Boddie: Deriving Services from Data Function Points
- Dave McComb: Semantic SOA
- Ron Schuldt: A SOA Controlled Vocabulary Case Study
- Arkady Kleyner: SOA in the world of Business Intelligence and Data Warehousing
- Arnold Van Overeem: Advances in Semantic Interoperability in Europe
- Steven van t Veld: Information as a Corporate Resource at the Demand Side of an Organisation

This Workshop is NOT About

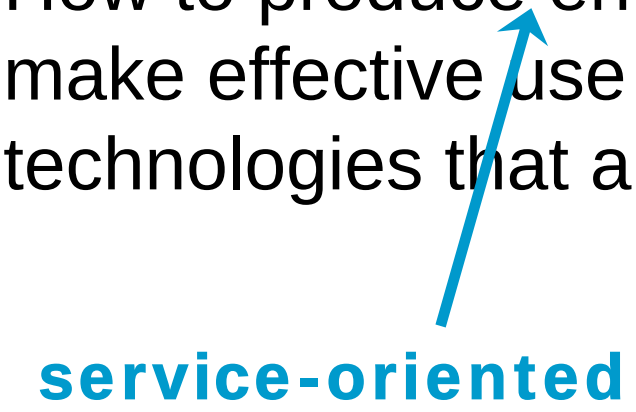
- Philosophy
- Advanced Research
- Detailed explanation of standards
  - ISO 11179
  - The Semantic Web
- How to develop an enterprise ontology

Important Unsatisfied Need

- How to develop an enterprise ontology
  - Would be subject of a very useful workshop!

This Workshop IS About

- How to produce enterprise architectures that make effective use of the semantic technologies that are now emerging



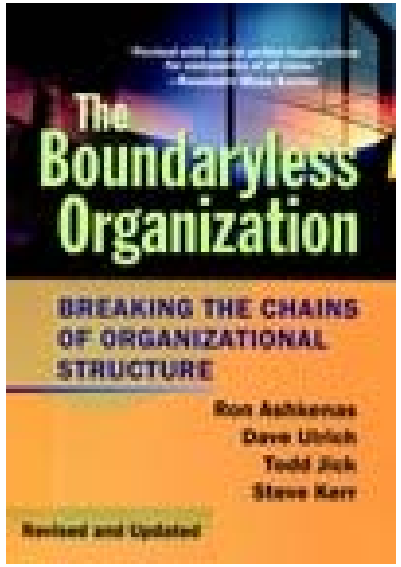
- What are the building blocks?

Agenda

11:00	Introductions
11:10	The Problem
11:20	Semantics v Syntax
11:30	Stakeholders and Requirements
11:45	Models and Building Blocks
12:15	Wrap-Up: Summary and Next Steps

The Problem

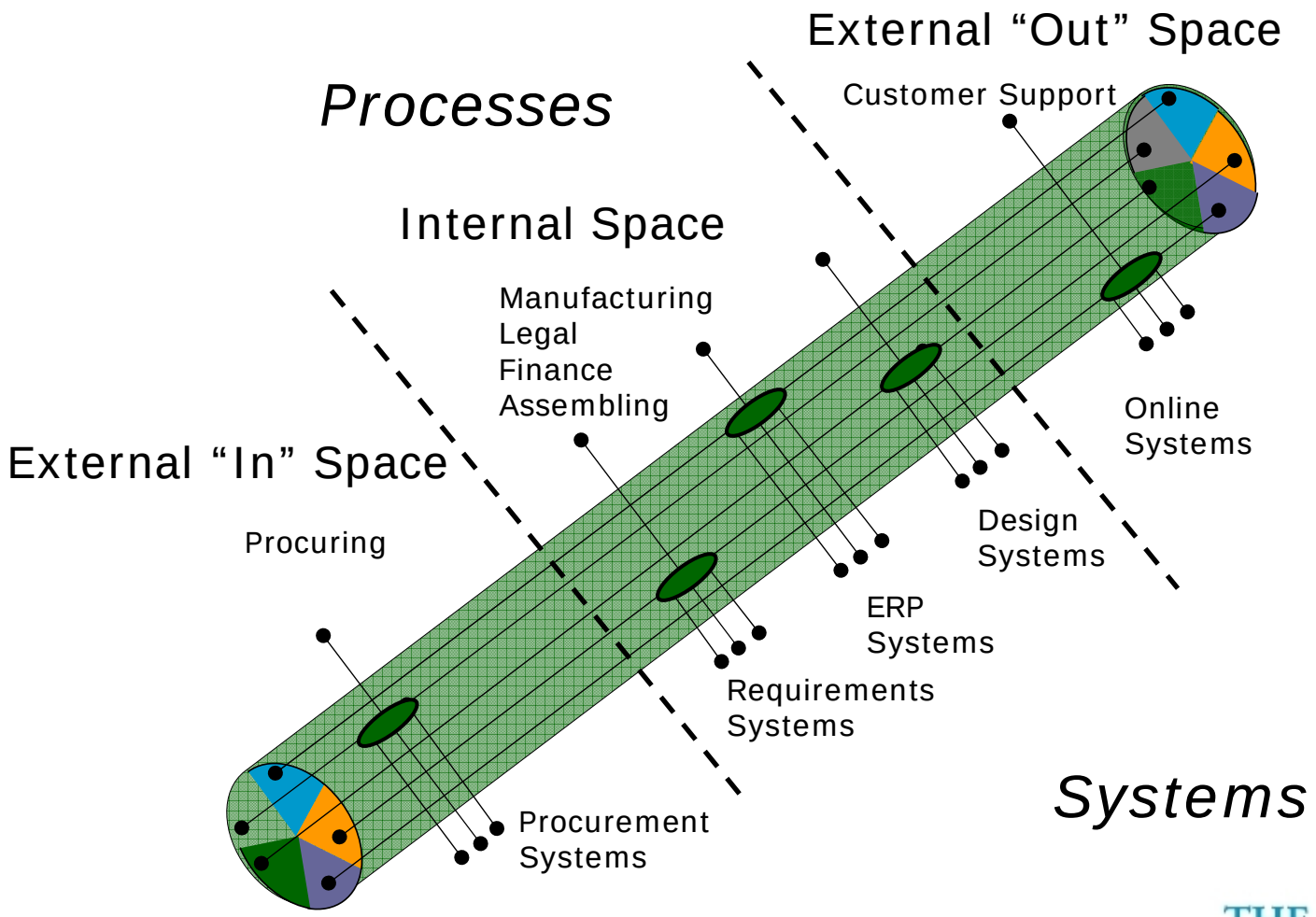
Boundaryless Information Flow



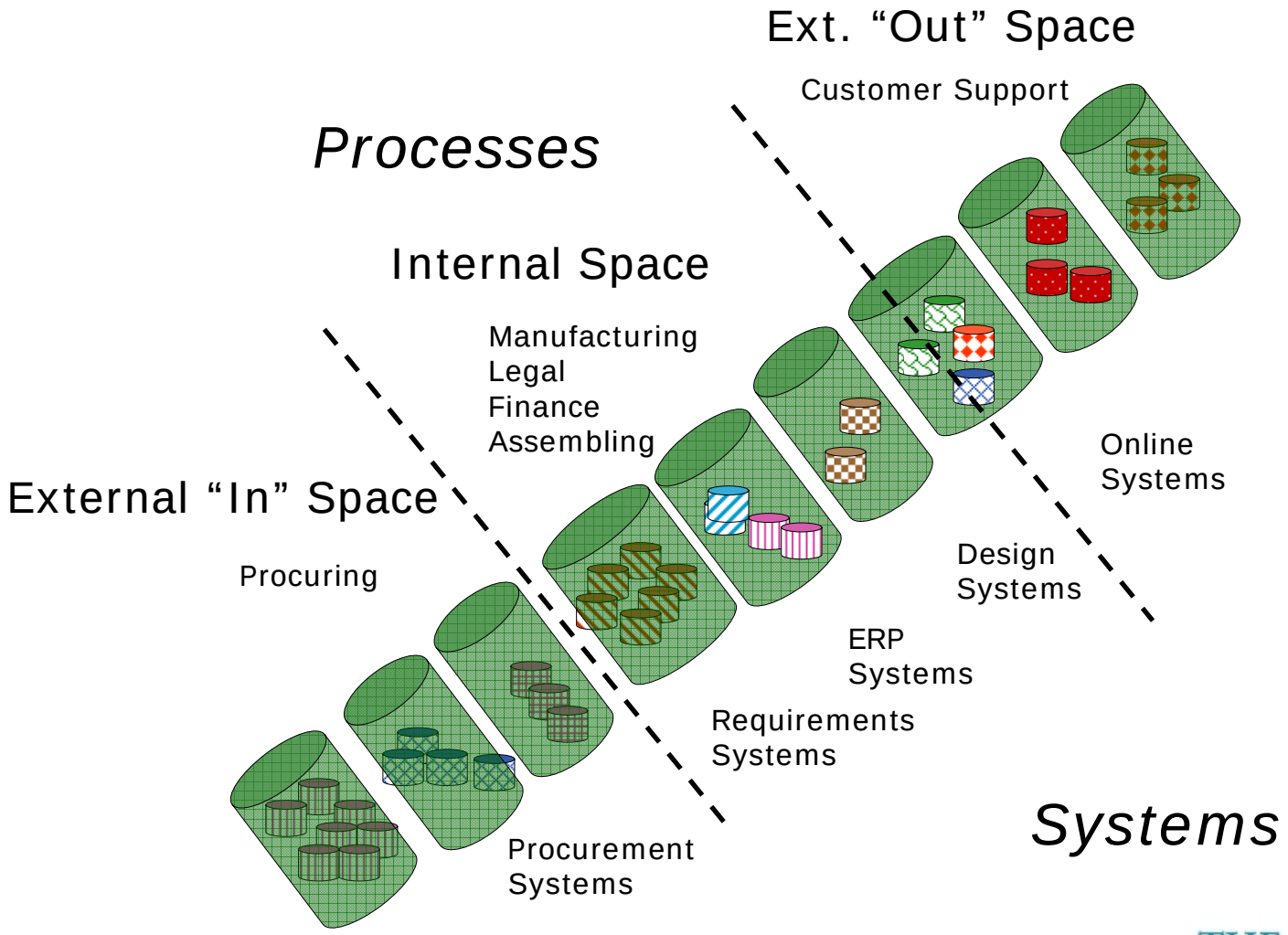
- Permeable boundaries between
  - Nations
  - Enterprises
  - Organizational levels
  - Departments
- Deliver
  - Productivity
  - Agility

But traditional IT architectures hinder this!

Enterprises Want This…



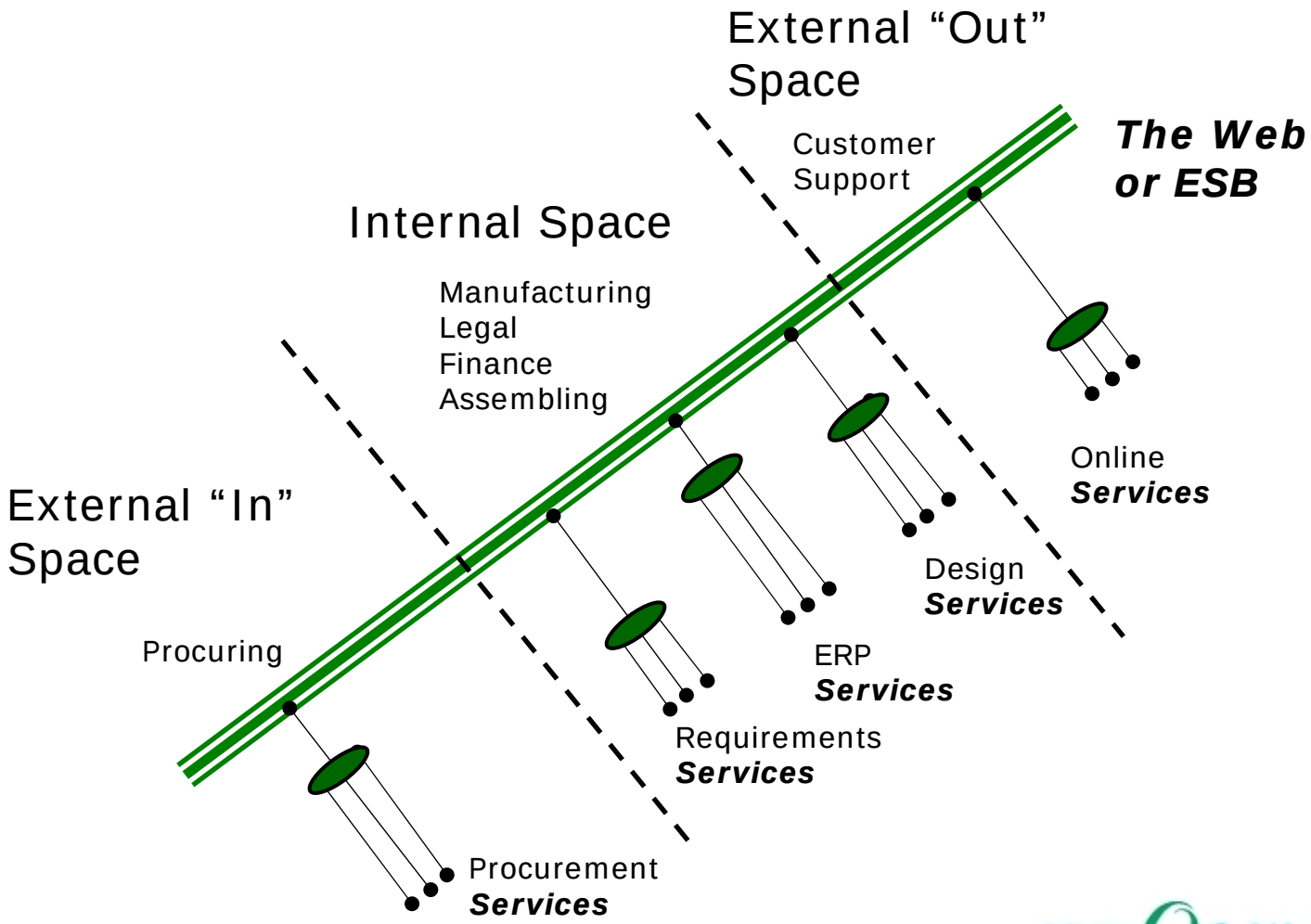
But Have This



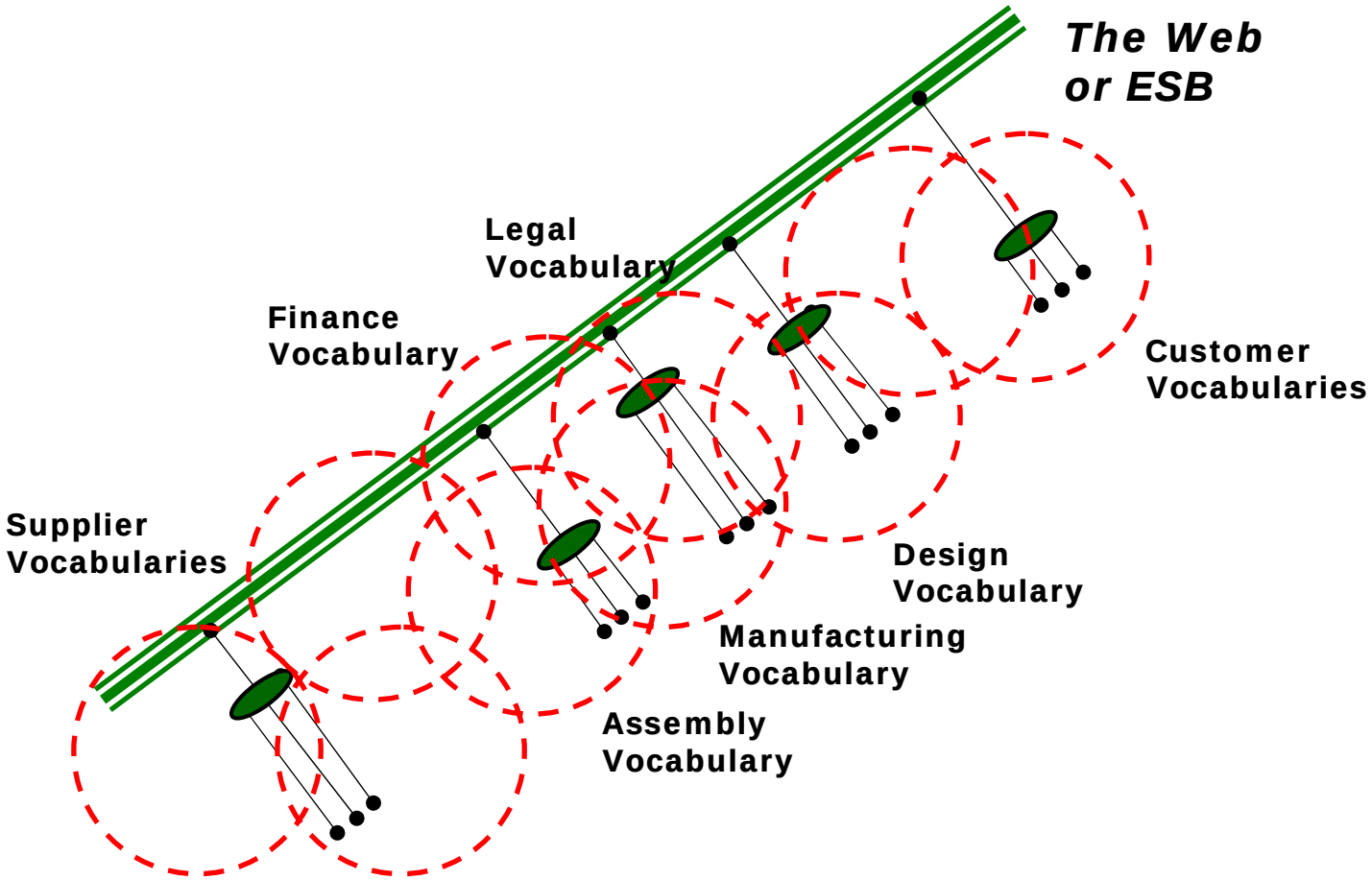
SOA

- SOA is an **architectural style** that supports **service orientation**
  - A way of thinking in terms of services and service based development and the outcomes that services bring
- SOA re-structures applications as loosely-coupled, modular services
- And provides for data flow between them

So Enterprises Can Have This



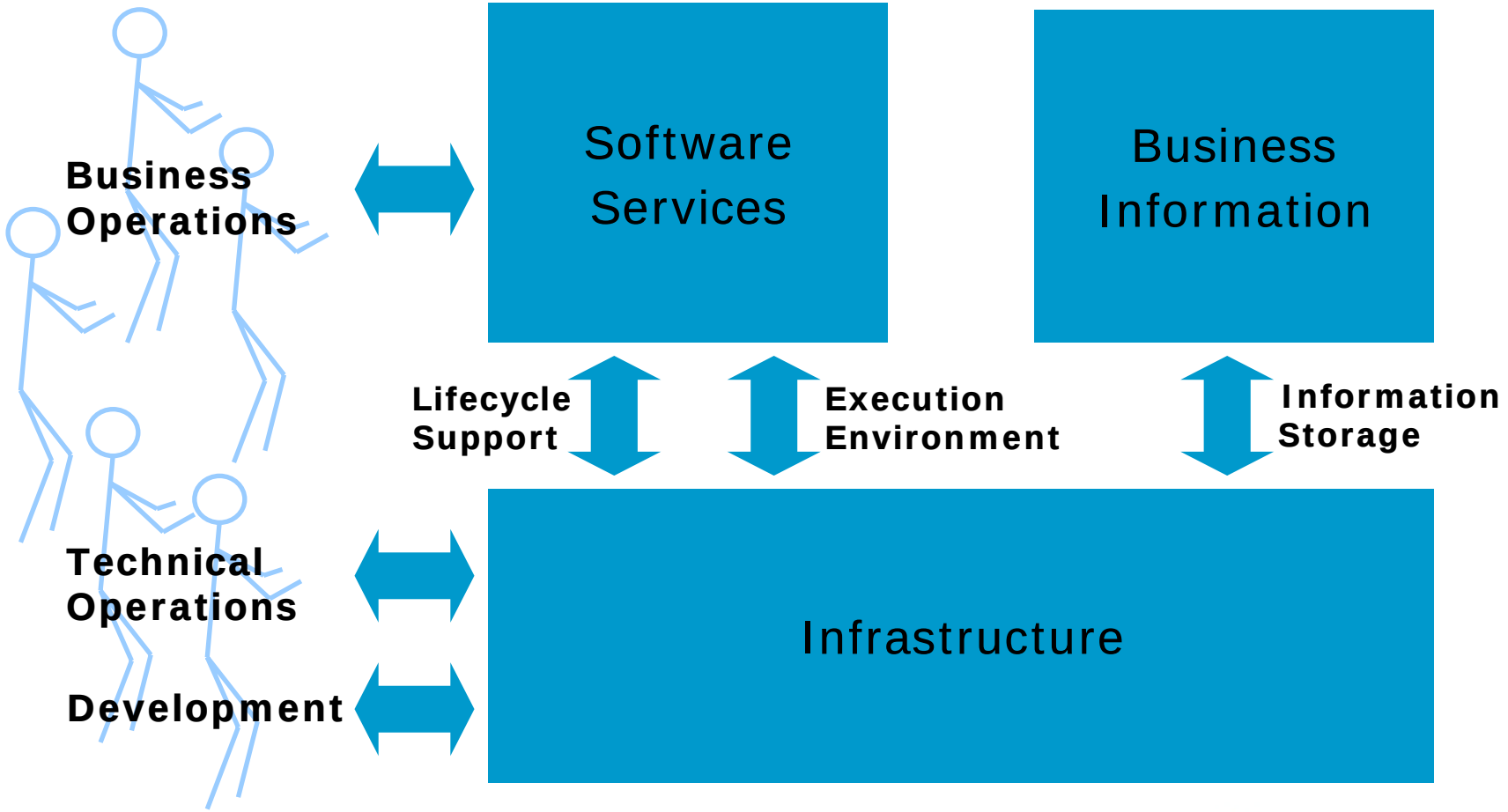
The Problem of Semantic Interoperability



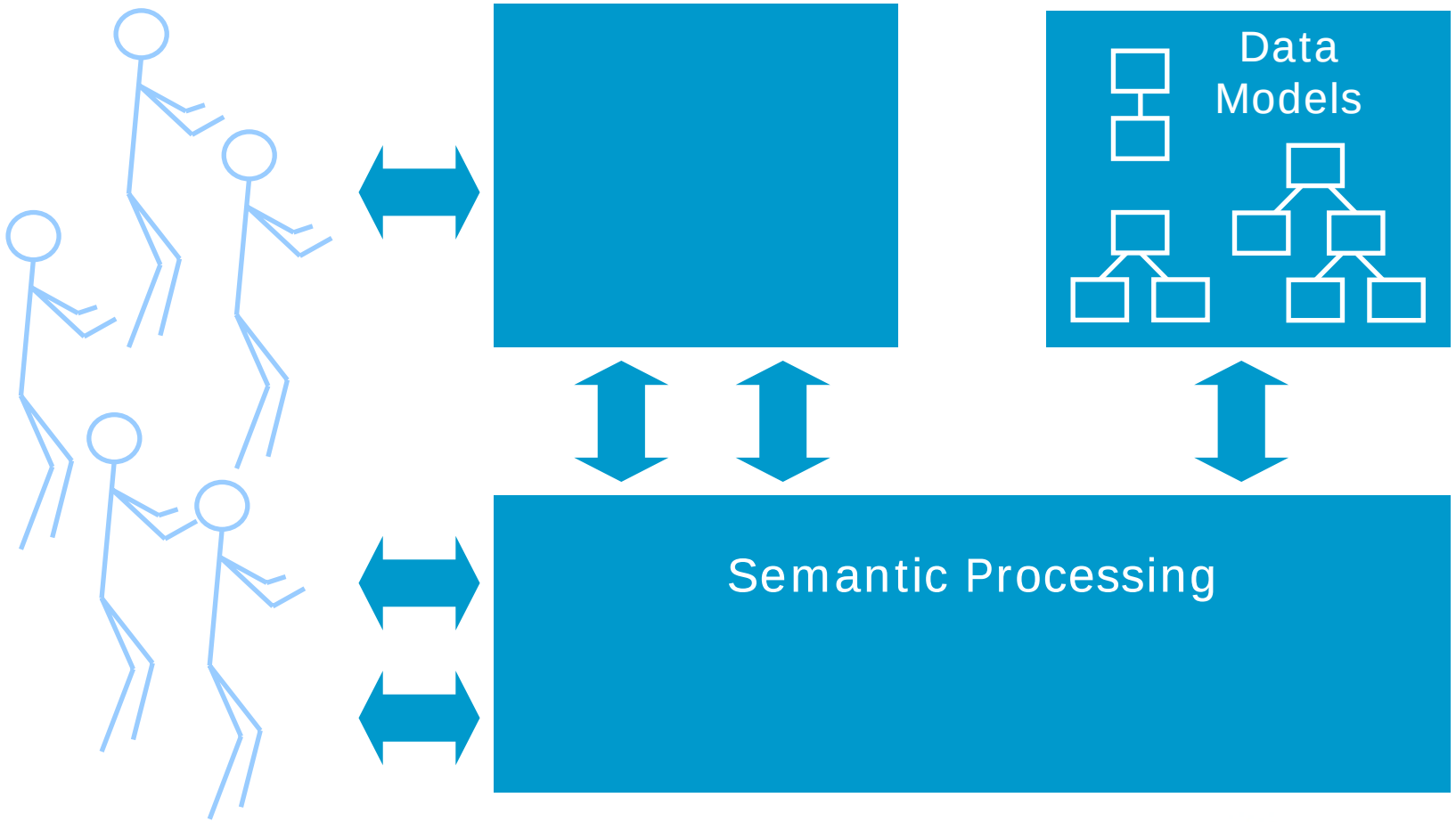
Boundaryless Information Flow

- If semantic interoperability can be achieved
- Then SOA will deliver Boundaryless Information Flow

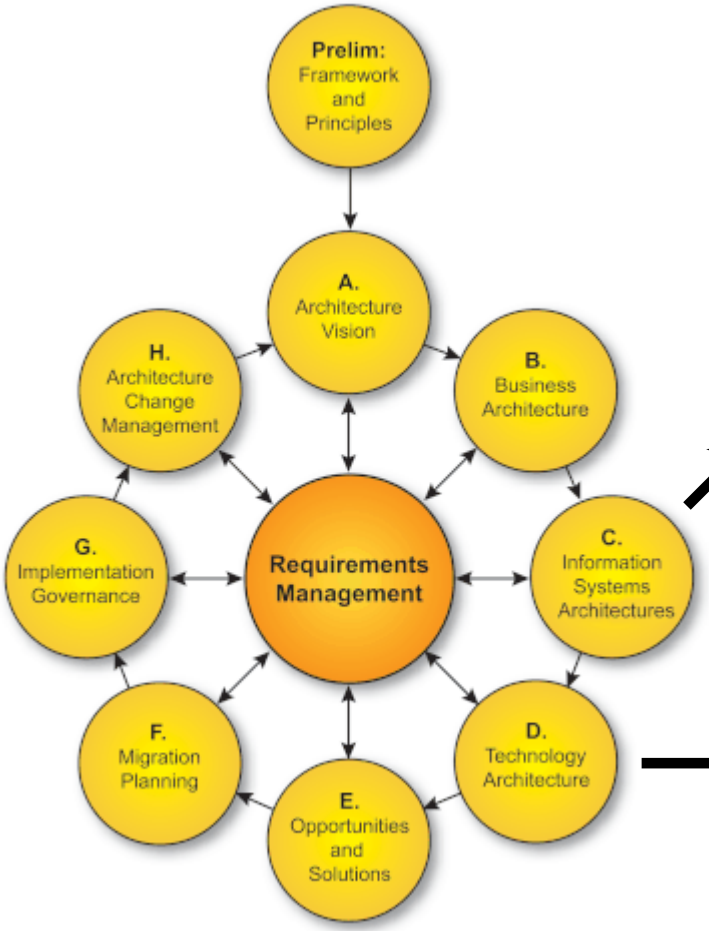
SOA Overview



SOA and Semantics



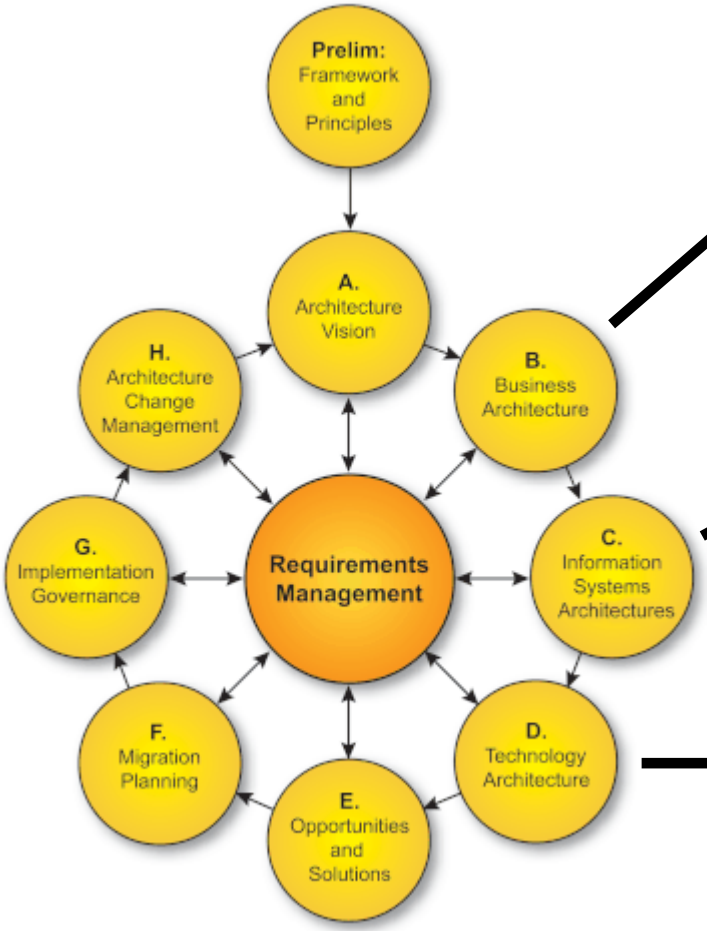
TOGAF



The Data Architecture part of Phase C defines the major ***types*** and sources of data necessary to support the business

Phase D defines the technical solution at the level necessary to support implementation

TOGAF



The Business Architecture addresses business information

The Data Architecture part of Phase C defines the major ***types*** and sources of data necessary to support the business

Phase D defines the technical solution at the level necessary to support implementation

The Role of Semantic Technology

- Semantic Technology enables semantic processing to be done cheaply in the infrastructure
- Instead of by expensive application-specific programs
- Used at run-time to select services and interpret information
- Used at design/development time to assist service composition and realization

The Problem

- How do semantic products fit into service-oriented enterprise architectures to perform this role?

Semantics v Syntax

11:00	Introductions
11:10	The Problem
11:20	**Semantics v Syntax**
11:30	Stakeholders and Requirements
11:45	Models and Building Blocks
12:15	Wrap-Up: Summary and Next Steps

Syntax and Semantics

Syntax

- Same data definitions
- Reproduceable interpretation rules
- Apply to dataelements
- Are language independant (although defaults may different from country to country)
- The characteristics that are compressible (Shannon)
- Syntax errors can be enumerated

Semantics

- Different data definitions
- Context dependant interpretation rules
- Apply to data element concepts
- More often then not convey language specific connotations .
- The characteristics that are the essence (Kolmogorov)
- Semantic errors can be classified

Semantic Views

- Semantic views for the "CUSTOMER" master data might be
  - CustomerAsThePayer
  - CustomerAsTheConsumer
  - CustomerAsTheBeneficiary

Claude Blouin

This is merely Syntax …

Field name	Address stored	Address requested
First name	Johann Wolfgang	Johann-Wolfgang
Last name	Goethe	Göthe
Name extension	von	
Birth day	00/00/1913	13/12/1913
Place of residence	Frankfurt (Main)	Frankfurt/Main
ZIP code	60313	
Street	Schiller Straße	Schillerstrasse
Street number	24 A	24a

- Different syntax, different characterset (code page)
- Same concept, same business object, same classification
- No relevent understanding of business context required to enable correct mapping

But this is Semantics

Field name	Address stored	Field name	Address requested
First name	Johann Wolfgang	Name	Dr J.W. Göthe
Last name	Goethe	Address	Schiller Str 24A
Name extension	von	Place of Residence	60313 Frankfurt/Main
Birth day	00/00/1913	Telephone	(+49) 69 87651234
Place of residence	Frankfurt (Main)	E-mail	goethe@boehringer.de
ZIP code	60313		
Street	Schiller Straße		
Street number	24 A		

- Different concept classificatation. Different metadata
- Understanding of business context may be required for correct mapping
- External knowledge can be applied to validate the match

Boundaryless Information Flow Requires Interoperability

Procedural layer (process layer)
ensuring that information that has to be shared and is correctly understood can be properly acted upon (legislation, objectives, priorities, resources, retribution, deeds)

Semantic layer
ensuring that information that has to be shared and is correctly interpreted can be properly understood (context specific metadata, ontologies, taxonomies, masterdata)

Technical layer (syntactic layer)
ensuring that information that has to be shared and is correctly received can be properly interpreted (characterset, syntax, data formats, currency, units, presentation language, etc)

Trivial layer (exchange layer)
ensuring that information that has to be shared can be properly received (connectivity, security, etc)

What Distinguishes Semantic from Syntactic Processing?

- Information is “data in context”
- Semantic processing requires understanding of context
- Which typically means that semantic processing of a message requires information that is not included in the message

Stakeholders and Requirements

11:00	Introductions
11:10	The Problem
11:20	Semantics v Syntax
11:30	**Stakeholders and Requirements**
11:45	Models and Building Blocks
12:15	Wrap-Up: Summary and Next Steps

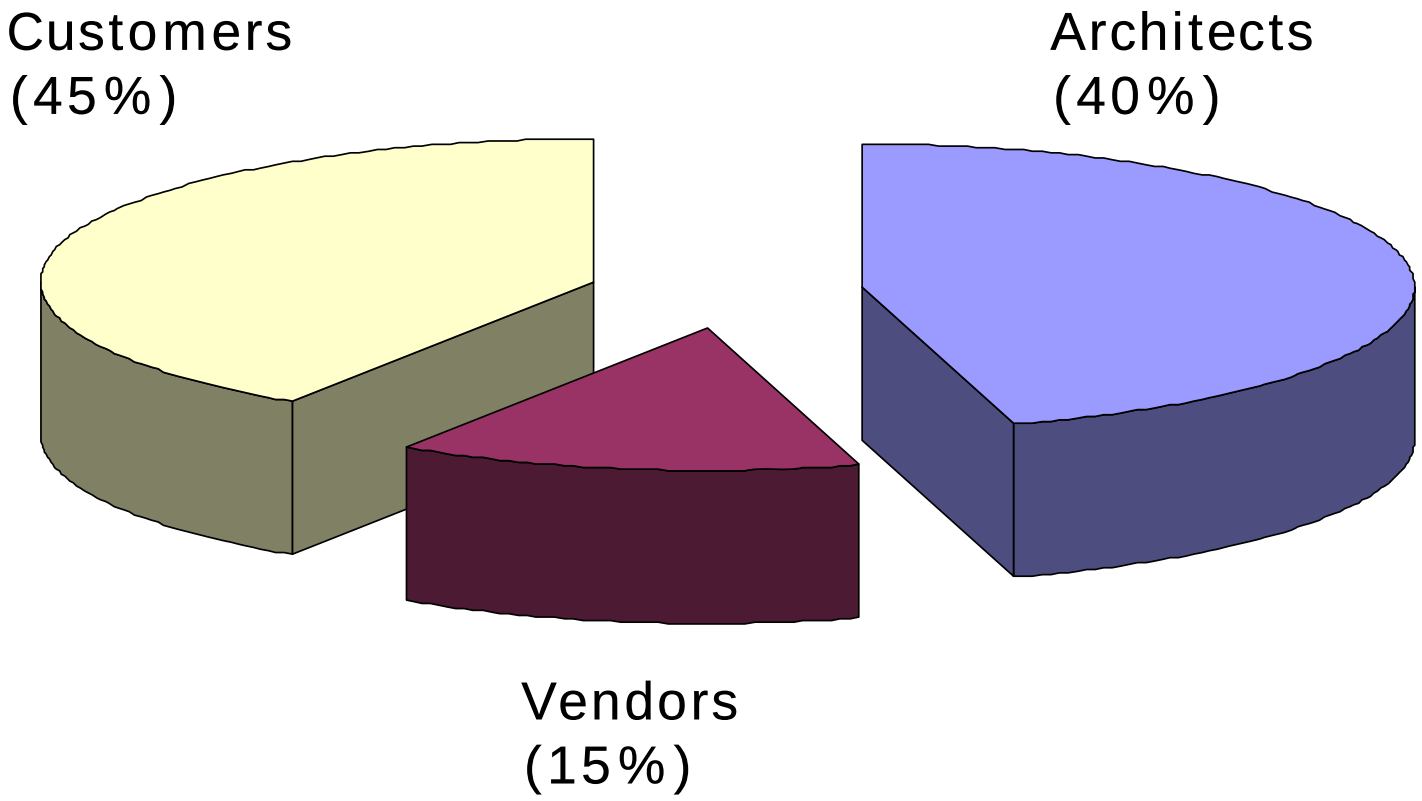
Who Are The Stakeholders?

- **Business owners**
- **Product Vendors**
- **Architects**
- Developers
- Operations staff
- . . .

Additional Stakeholders

- Auditors, including for legal compliance, such as to operational health and safety legislation
- Business analysts
- Architects of specific kinds that should be distinguished
  - Information architects
  - Business Architects
- Government and regulators
- Industry organizations
- Service creators (composers and developers)
- Service providers
- Service consumers
- Information managers (creators, collectors, custodians)

This Workshop Includes



Semantic Requirements for Enterprise Architecture

- Semantic architecture should
  - Include business object semantics
  - Include semantic standards, at least for the master data objects
  - Support model-based product development
  - Enable product data integration of systems
  - Be vendor-neutral.
- Semantic frameworks must fit with existing organization/industry frameworks

Requirements (continued)

- . .

Additional Semantic Requirement for Enterprise Architecture

- Facilitate compliance (enterprise conformance to regulation such as SOX and HIPAA)

Requirements for an Information Architecture Framework

- Enable consistent description
- Facilitate stakeholder communication
- Include reference model
- Show roles of relevant standards
- Provide a methodology for architecting solutions
- Separate infrastructure support from business need
- Distinguish *information* from *data*
- Address both man-machine and machine-machine communication
- Incorporate cross-cutting concerns, including privacy, data security, and lifecycle management
- Enable interoperability
- Basis for metrics
- Vendor-Neutral

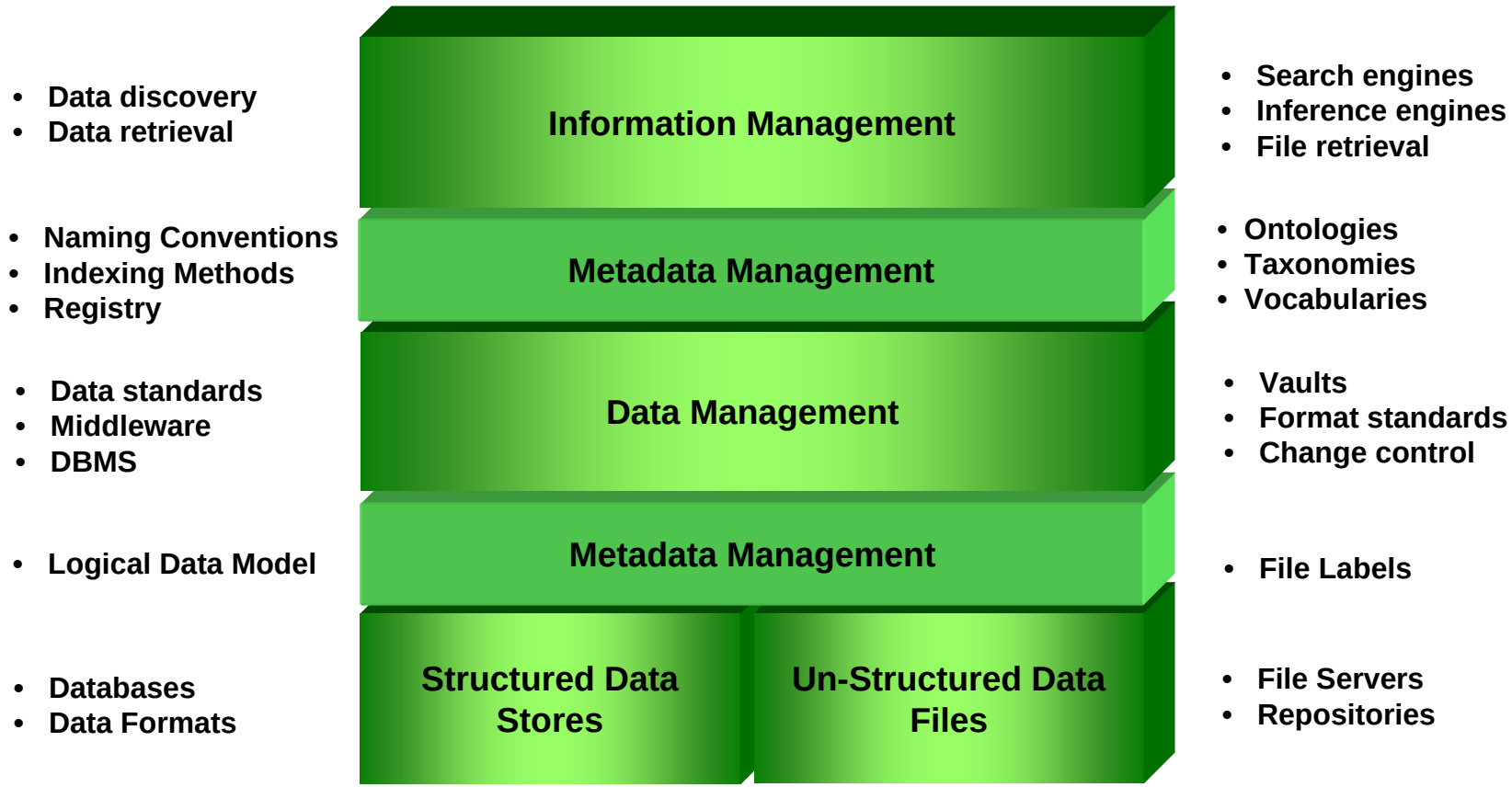
Requirement for Semantic Technology

- Increase automation of information processing and management

Models and Building Blocks

11:00	Introductions
11:10	The Problem
11:20	Semantics v Syntax
11:30	Stakeholders and Requirements
11:45	**Models and Building Blocks**
12:15	Wrap-Up: Summary and Next Steps

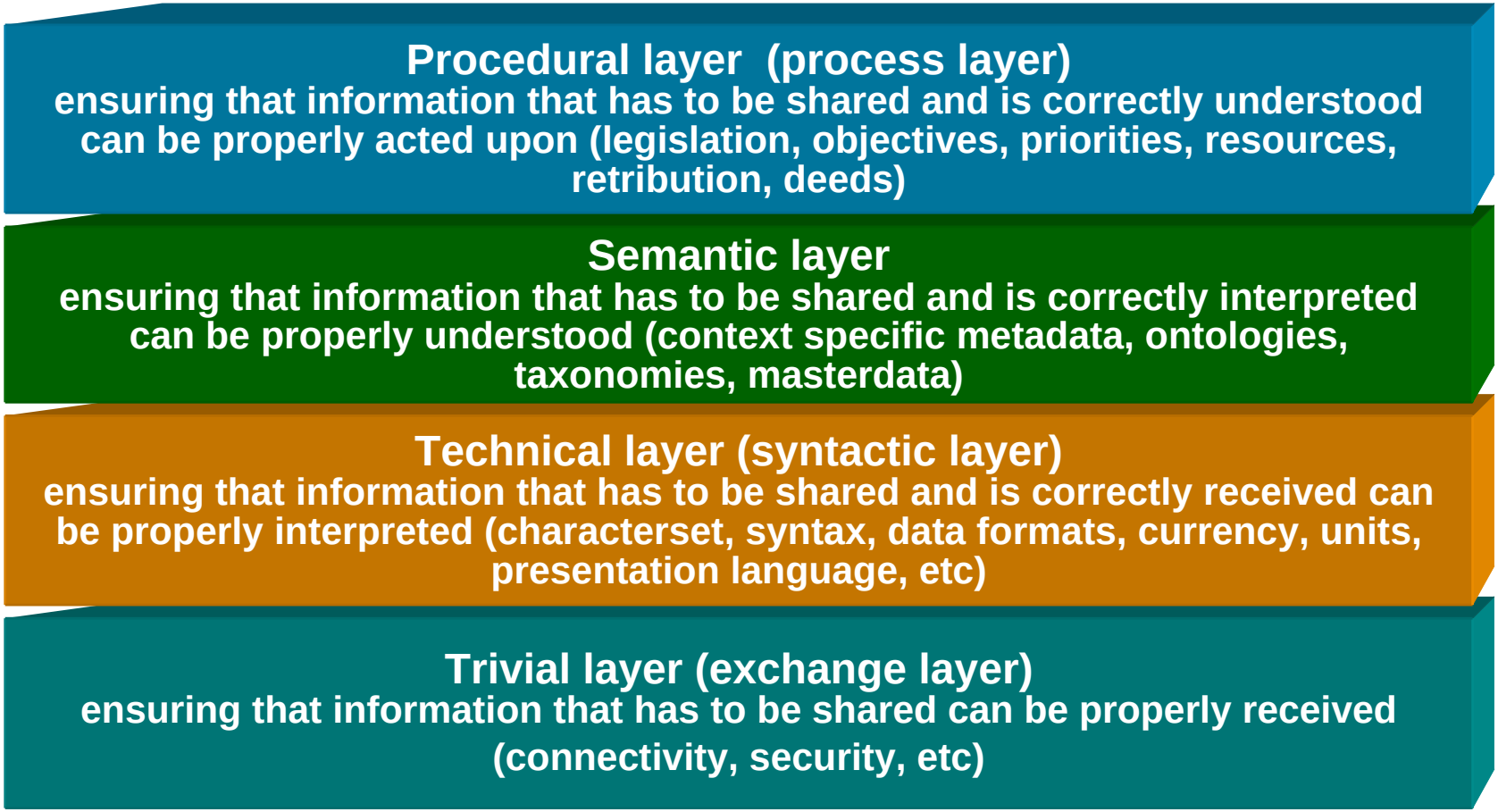
Draft GEIA Reference Model



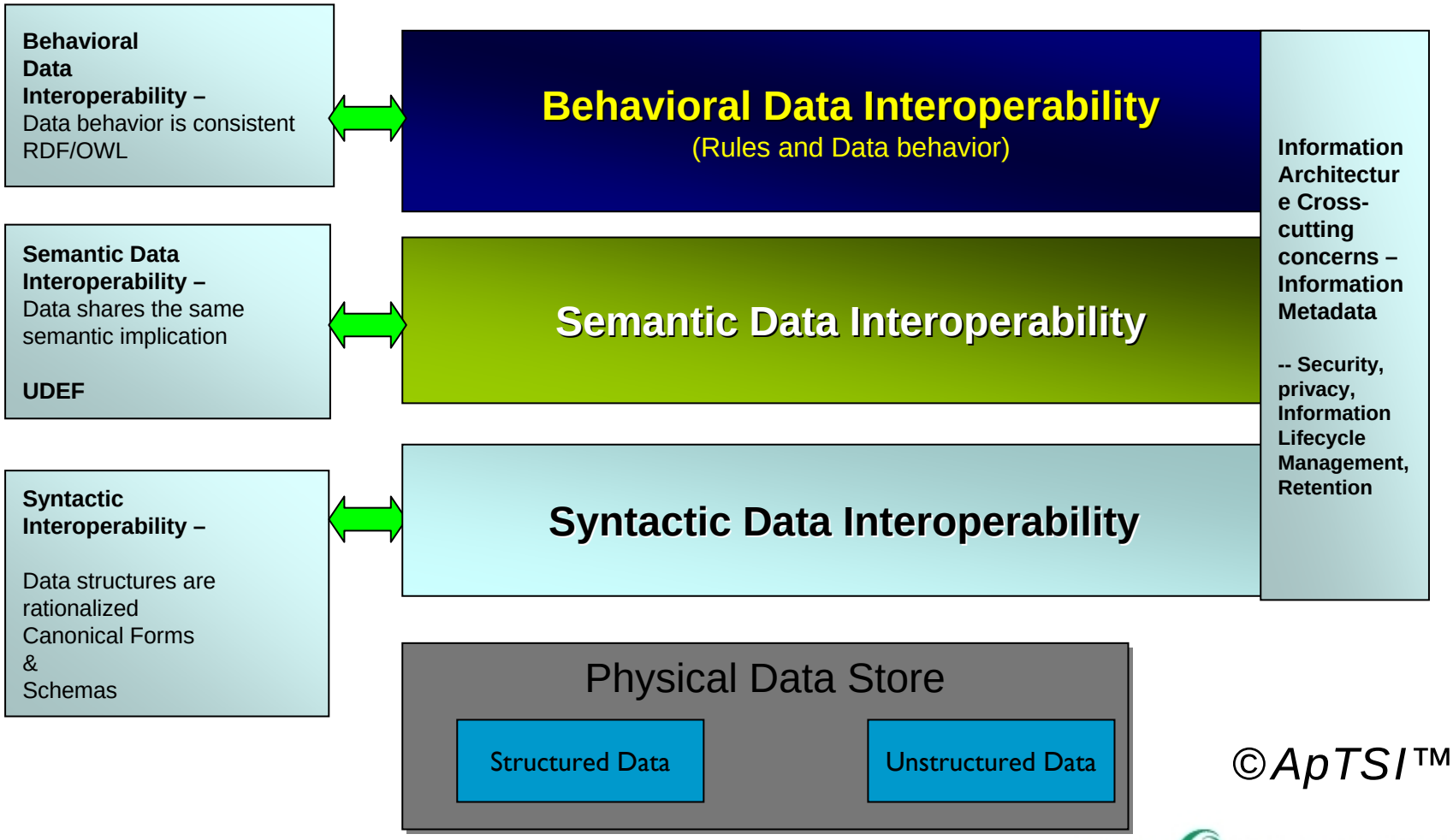
The Draft GEIA Model

- Is a reference model for information management rather than information processing
- Doesn't specifically show distinguish syntax and semantics but
  - The lower part relates to syntax, and
  - The upper part relates to semantics

Boundaryless Information Flow Requires Interoperability



Interoperability Reference Model



©ApTSI™

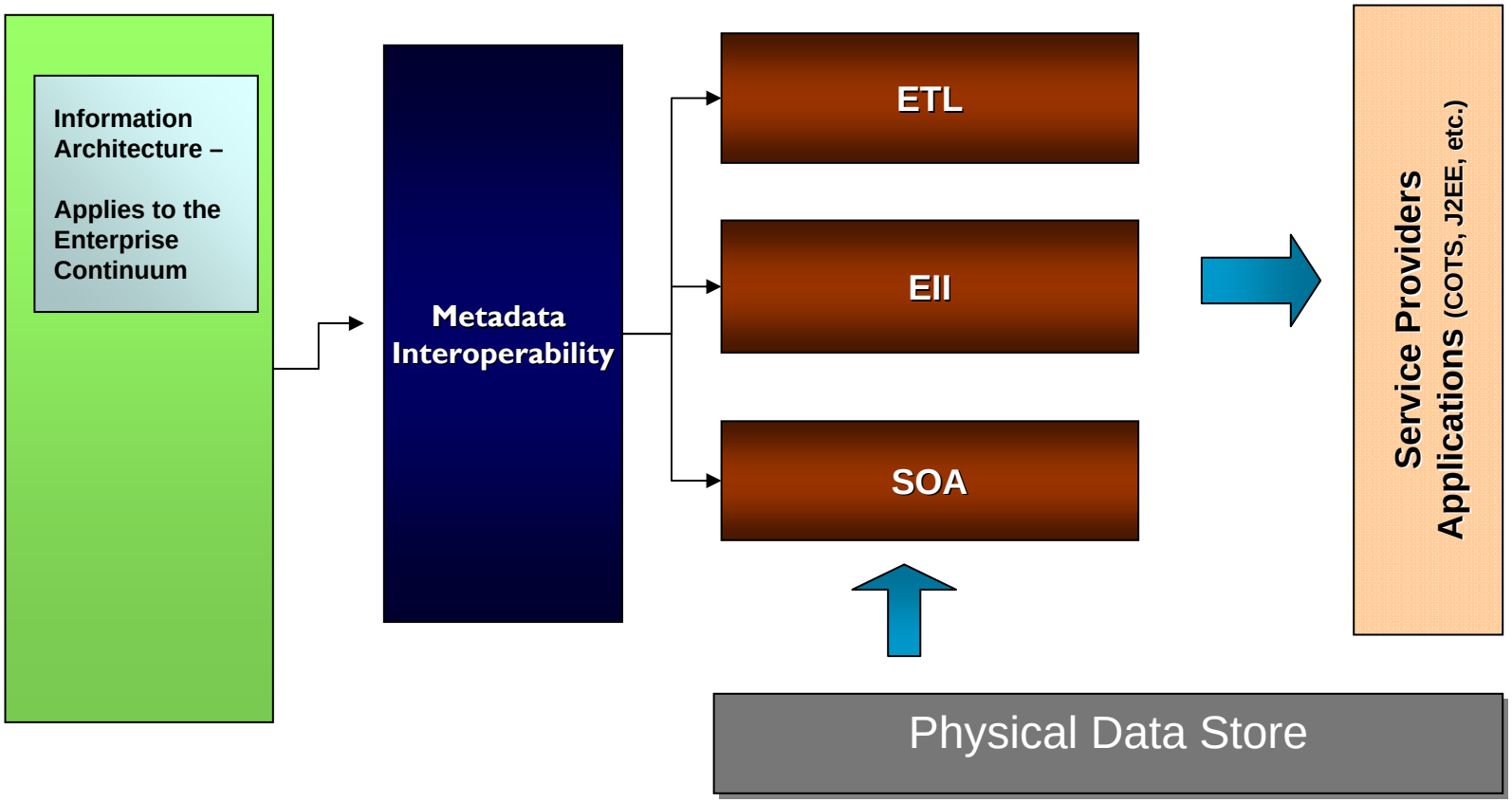
Reference Model

- There are striking similarities between previous 2 slides
- Indicating that there are four layers that are a good starting point for a model:
  - Procedural/Behavioral
  - Semantic
  - Syntactic
  - Trivial/Physical

Semiotic Reference Model

- There is a well-respected semiotic reference model with 6 layers
- Two of these are *Semantic* and *Syntactic*
- It has two layers above the Semantic layer:
  - Social
  - Pragmatic
- Machine processing can address part of the pragmatic layer, but arguably not the social layer
- It could be useful to include the social/pragmatic distinction to help the model to illustrate the limits of machine processing

Information Architecture Across the Enterprise Continuum



©ApTSI™

Metadata Interoperability

- Metadata interoperability is vital
- It needs a semantic repository that
  - Will be maintained
  - Will produce output in forms that ETL, EII and SOA can process

Semantic Building Blocks

Metadata Repository		
Semantic Web		
Metadata Bridge		
Ontology (Enterprise, Vertical Domain . .)		
Search/Discovery Engine		
Inference Engine		
Program that resolves a semantic conflict		
Product Configurator		

Additional Semantic Building Blocks

- Metadata Management (MDM) repository
- MDM Engine
- Rules Engine
  - Eg. For SWRL
- Semantic gateway
  - Translates messages
- Semantic maintenance tools
  - Eg. repository tool for UDEF
- Semantic language processing tools
- Analysis tools
  - Help model semantics from enterprise repository

Standards

- Semantic languages
  - Eg. OWL, RDF, WSML
- Rules languages
  - Eg. SWRL

Summary and Next Steps

11:00	Introductions
11:10	The Problem
11:20	Semantics v Syntax
11:30	Stakeholders and Requirements
11:45	Models and Building Blocks
12:15	**Wrap-Up: Summary and Next Steps**

Summary and Next Steps

- TBA

Summary

- A framework and reference model will meet a number of needs, especially:
  - Better communication between stakeholders
  - Methodology for architecture/development
  - Basis for automation
- The four-layer model (trivial, technical, semantic, procedural) is an excellent starting point, but requires further consideration of
  - Additional layers, per semiotic model
  - Cross-cutting concerns (security, etc.)

Next Steps

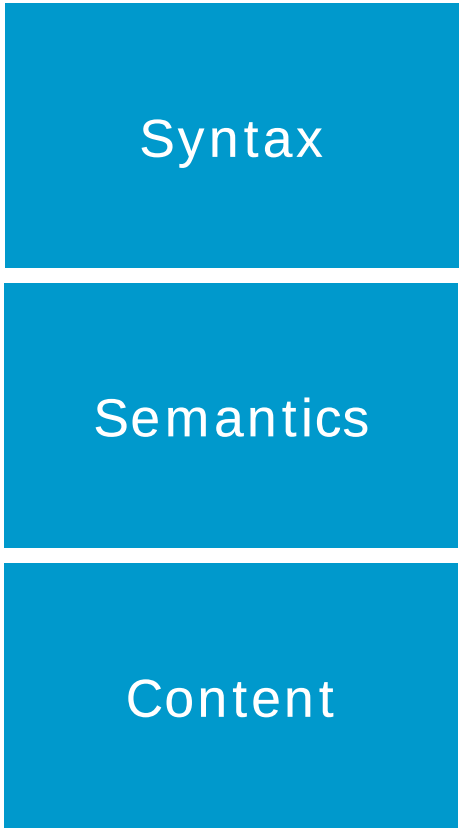
- Add the workshop conclusions to the input presentation
  - This presentation is the result of that process
- Discuss in Semantic Interoperability Working Group
  - http://www.opengroup.org/projects/si/

Semantics for Enterprise Architecture Workshop

Thank you!

Back-up Slides

Simple Information-Processing Reference Model

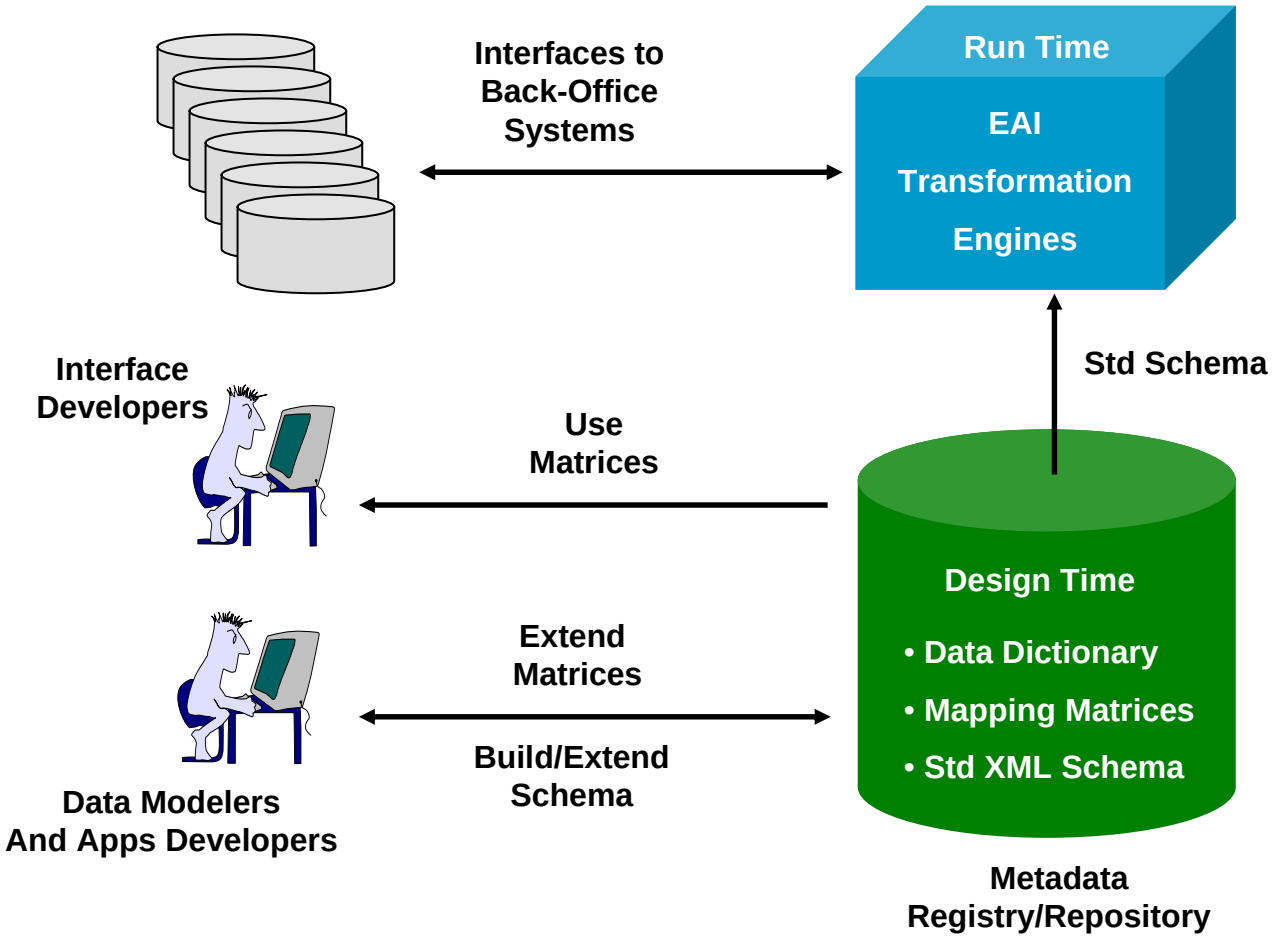


Identifying fields and describing their representations, properties and inter-relations

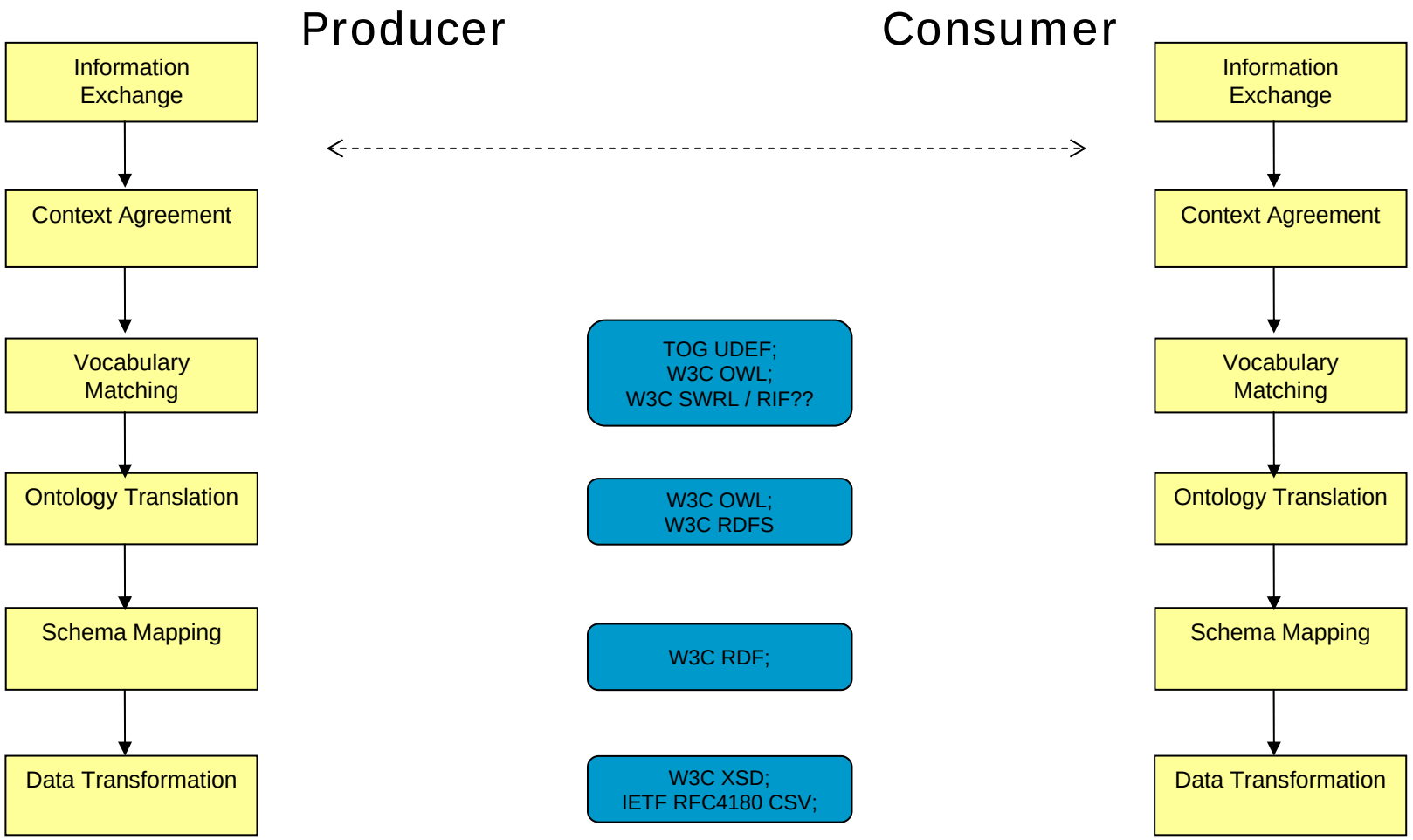
Identifying resources and describing their types, properties and inter-relations

Creating and transforming information items about the identified resources

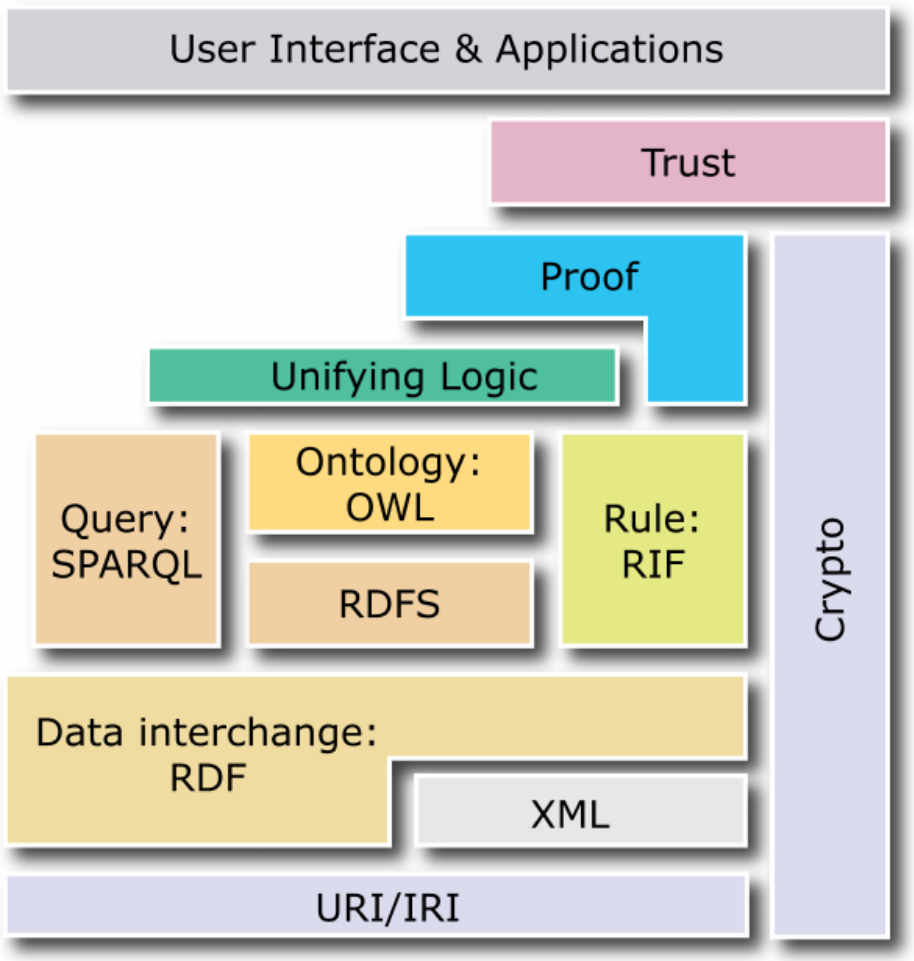
Metadata Management



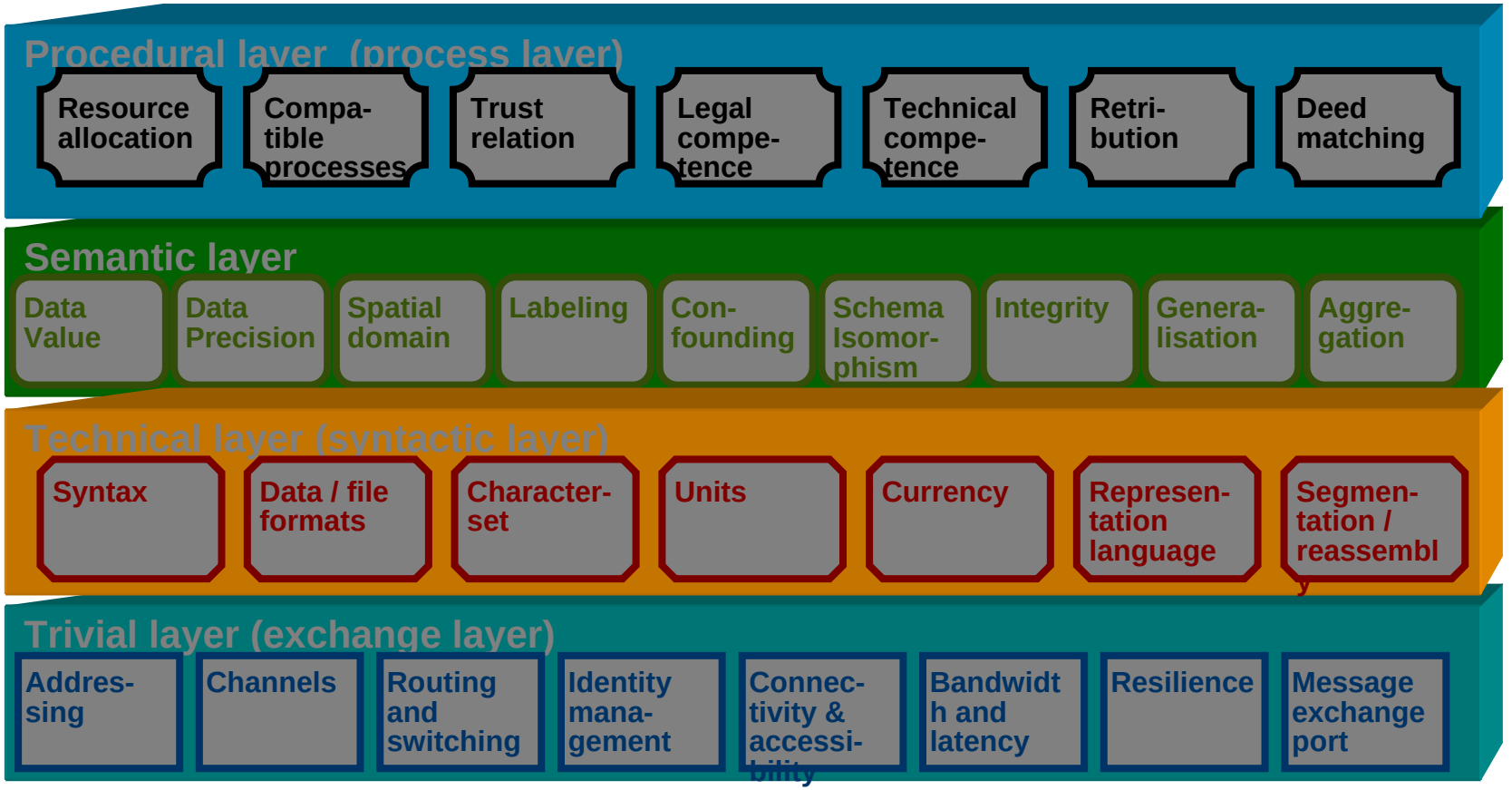
Information Exchange Model



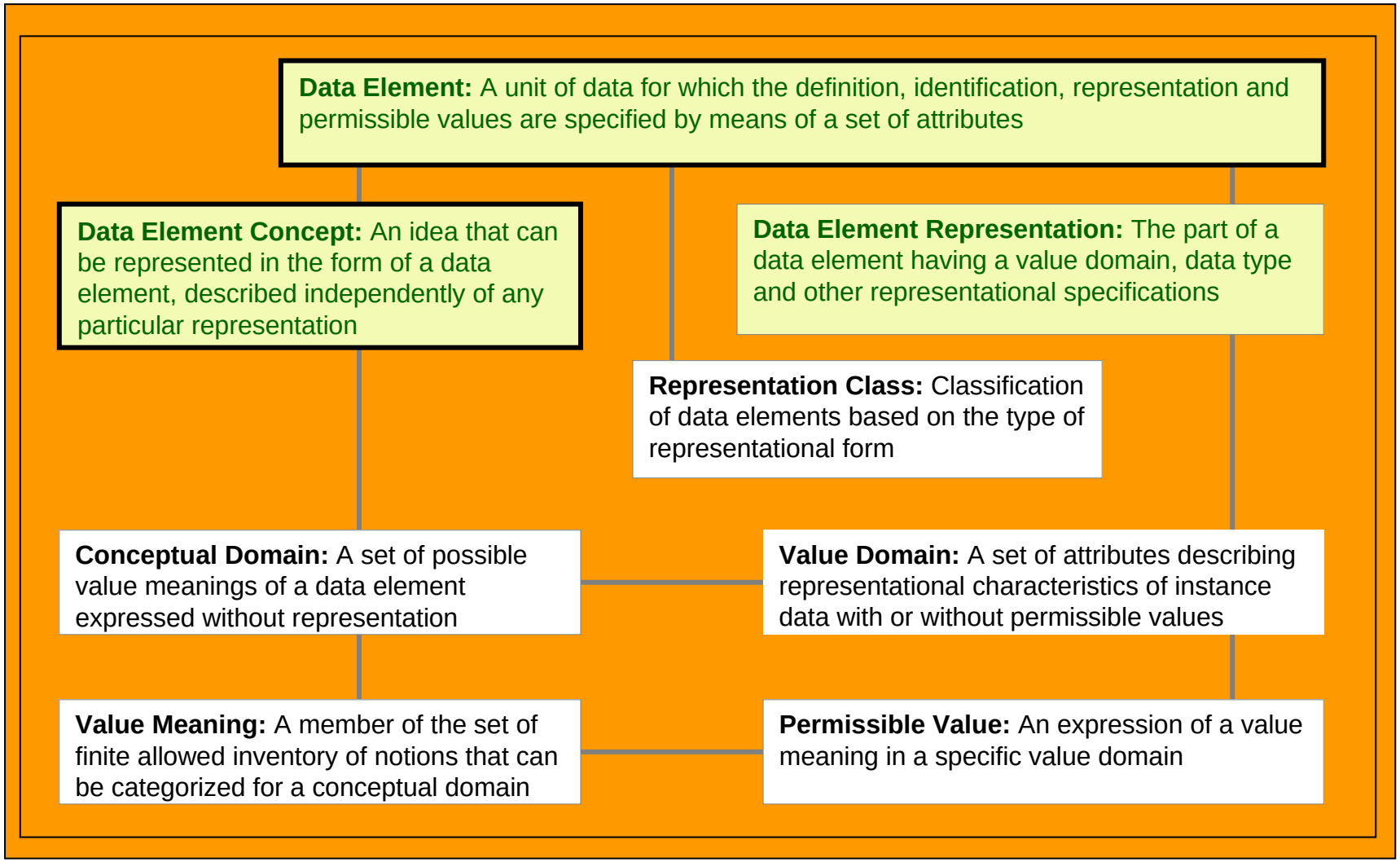
Semantic Web “Layer Cake”



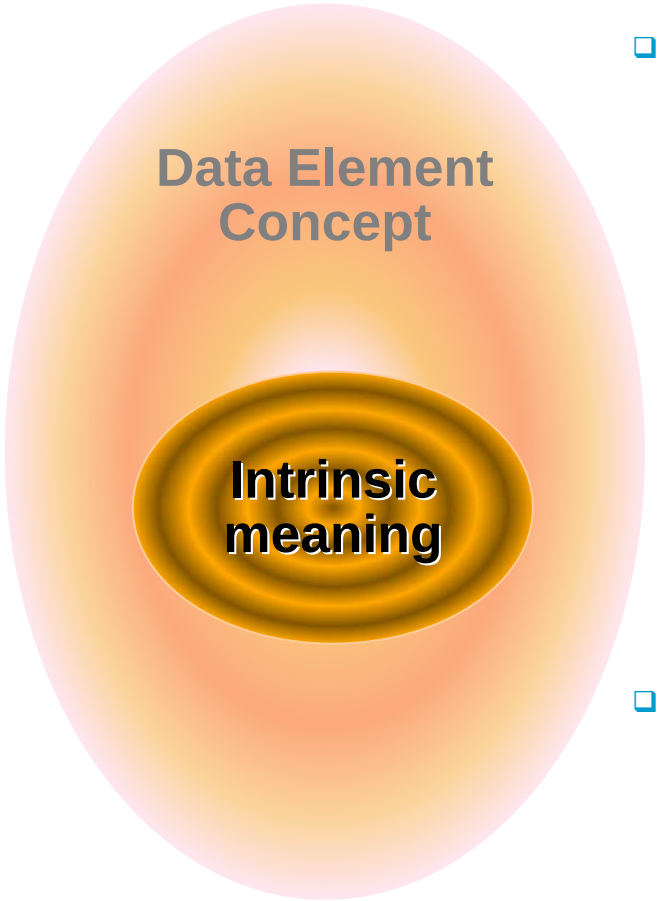
Functional overview of issues by layer



ISO/IEC 11179 Key Terms & Definitions



Meaning of information



- Kolmogorov complexity, named after the famous mathematician Kolmogorov (1903-1987), who introduced this theory in 1965, can be summarized by the following four mathematical principles:
  - Incompressibility: For each n an arbitrary word exists of length n.
  - Set compression: Each member of a computable set A has a description of length $log|A|$
  - Source compression: Each word x with a positive probability under a computable distribution P has a description of length $-\log P(x)$
  - Information symmetry: The quantity of information in the word x about the word y is equal to the quantity of information in the word y about x.
- Shannon has defined information entropy as the redundant quantity of the information that does not convey any meaning. A message so to say consists of an intrinsic meaning and an arbitrary amount of entropy, due to which it has a bigger size then the minimum that conveys its meaning. In daily life speech this becomes apparent when people who cannot completely hear each other still can understand a message.