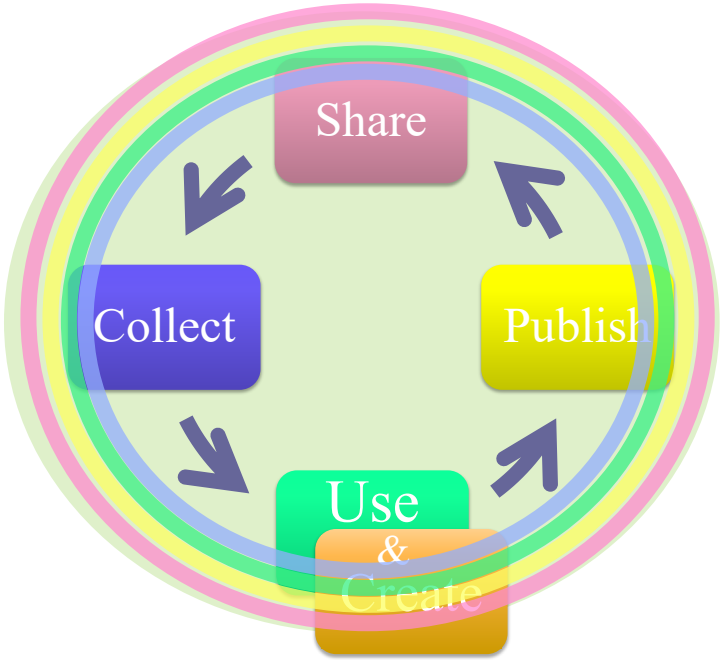
Semantic Web

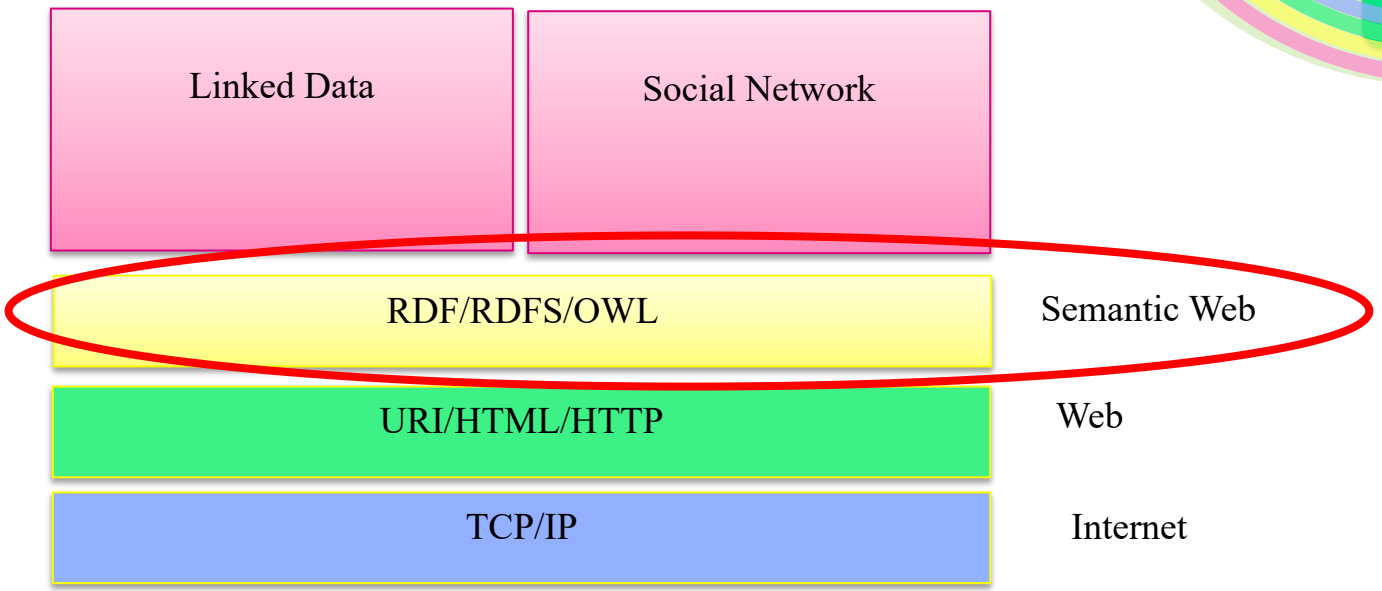
Web as a communication infrastructure between human and machines

Hideaki Takeda
National Institute of Informatics
takeda@nii.ac.jp

Layers for Information Cycle

- Internet
- Web
- Semantic Web
- Linked Data
- Social Network





Outline

- What is Semantic Web?
  - XML
- Realization of Semantic Web
  - Metadata
  - Dublin Core
  - RDF
  - RDF Schema
  - OWL

The Aim of The Semantic Web

- "The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation."

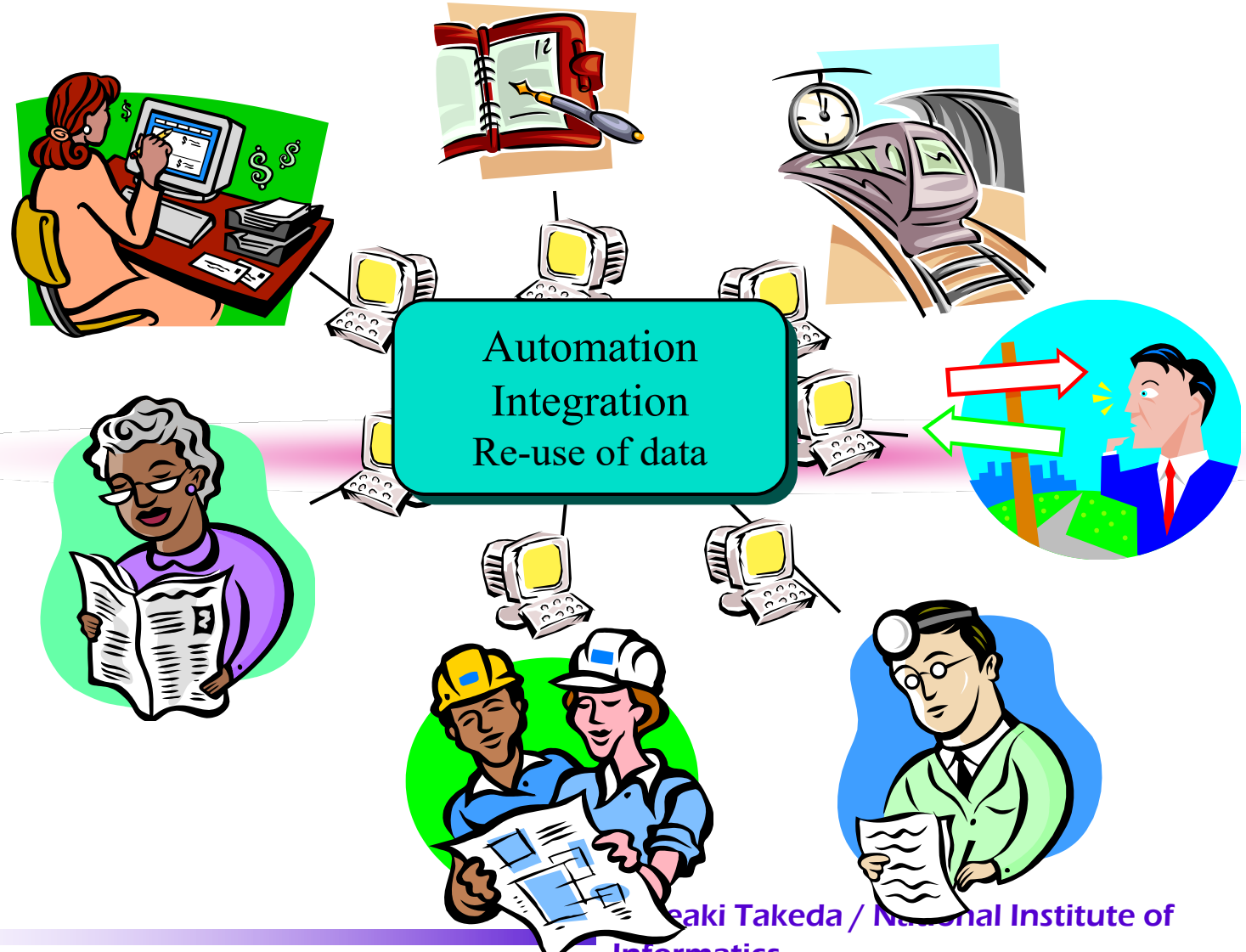
  *The **Semantic Web**, **Scientific American**, May 2001, Tim Berners-Lee, James Hendler and Ora Lassila*

- The Semantic Web is a vision: the idea of having data on the web defined and linked in a way that it can be used by machines not just for display purposes, but for automation, integration and reuse of data across various applications.

  http://www.w3.org/2001/sw/

Semantic Web

- Realization of various information exchanging via Web



Next Generation Web?

- Evolution of Web
  - HTML: Web for Display
  - XML: Web with Syntax
  - ?? : Web with Semantics

- Why should we embed semantics into Web?

  From
  - Web for Human

  To
  - Web for human and machines

  cf. Web for machines

A brief introduction of XML

- Limitation of HTML
  - Chaos by mixture of displaying and text structures
    - e.g.,
      - <h3></h3> should be used for “the third-level heading”, but are often used just for bigger fonts
      - is specifying “bold” , not “emphasis”.
  - Fixed Structure
    - e.g.,
      - If you need <h7></h7>….
      - I need a structure just for my data

```
<h1> A list of lectures</h1>
<h2> Knowledge Sharing Systems</h2>
<h3> Lecturer : Hideaki Takeda</h3>
<h3>Wednesday 3rd</h3>
```

XML

- XML(eXtensible Markup Language)
  - Can define original tags
  - Represent logical structures of data
    - DTD
  - Do not include style information
    - XST

```
<lecturelist>
<lecture>
<title id=1234> Knowledge Sharing Systems</title>
<lecturer> Hideaki Takeda</lecturer>
<schedule>
  <week>Wednesday</week>
  <time>3rd</time>
</lecture>
...
</lecturelist>
```

XML

- XML(eXtensible Markup Language)
  - A standard by W3C(The World Wide Web Consortium) (1998)
  - A subset of SGML
  - Features
    - Fixed character set
    - DTD is not mandatory
    - Should not omit end-tags
    - Style can be specified
- SGML(Standard Generalized Markup Language)
  - ISO8879 (1986)
  - Features
    - Can specify character set
    - DTD is mandatory
    - Can omit end-tags
    - Style can be specified

DTD

- DTD(Document Type Definition):
  - Specify elements and attributes for XML logical structures
    - Defining a specific markup language
      - XHTML
    - Logical structure can be specified
  - Should share DTD between document authors and users

```
<!DOCTYPE lecturelist [
<!ELEMENT lecturelist (lecture*)>
<!ELEMENT lecture (title, lecturer*, schedule)>
<!ELEMENT title (#PCDATA)>
<!ATTLIST title id ID #REQUIRED >
<!ELEMENT lecturer (#PCDATA)>
<!ELEMENT schedule (week, time)>
<!ELEMENT week (#PCDATA)>
<!ELEMENT time (#PCDATA)>
]>
```

DTD

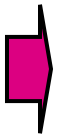
- Element type declaration
  <!ELEMENT element-name (element-content) >
  - Element name
  - Element content
    - Data type
    - Children
    - Order
    - Number of appearance: *(0-), +(1-, ?(0-1)
- Attribute-list declaration
  <!ATTLIST element-name attribute-name attribute-type default-value>
- Entity declaration
  <!ENTITY entity-name “string to be replaced">
- Notation declaration

Style Sheet

- XSLT（eXtensible Stylesheet Language Transformations）
  - A language to describe rules to translate structures in XML
  - Originally a part of XML (eXtensible Stylesheet Language) which is a language to display or print XML data
  - Procedure
    - A template for each node in the source tree
    - Specify which node should be applied to a template with *match* attribute of *xsl:template* element
    - XSLT processor analyzes from the top-most node to down
    - Find a pattern of appearance of elements and attributes which is specified with *match* attribute of *xsl:template* element
    - Apply a rule of a template applied

```
<?xml version="1.0" encoding="shift_jis"?><xsl:stylesheet version="1.0"
xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
 <xsl:output method="html" encoding="shift_jis"/>
 <xsl:template match="/">
  <html><body>
   <xsl:apply-templates/>
  </body></html>
 </xsl:template>
 <xsl:template match="lecturelist">
  <table border="1">
   <tr><td>Lecture name</td><td>Lecturer</td><td>Time</td></tr>
   <xsl:apply-templates /></table>
 </xsl:template>
<xsl:template match="lecture">
  <tr><xsl:apply-templates/></tr>
 </xsl:template>
<xsl:template match="title">
<td><xsl:value-of select="." /></td>
 </xsl:template>
<xsl:template match="lecturer">
  <td><xsl:value-of select="." /></td>
 </xsl:template>
 <xsl:template match="schedule">
   <td> <xsl:apply-templates /> </td>
 </xsl:template>
<xsl:template match="week">
   <xsl:value-of select="." />/
 </xsl:template>
<xsl:template match="time">
   <xsl:value-of select="." />
 </xsl:template> </xsl:stylesheet>
```

```
<?xml version="1.0" encoding="shift_jis"?>
<?xml-stylesheet href="lecturelist.xsl" type="text/xsl" ?>
<!DOCTYPE lecturelist [
<!ELEMENT lecturelist (lecture*)>
<!ELEMENT lecture (title, lecturer*, schedule)>
<!ELEMENT title (#PCDATA)>
<!ELEMENT lecturer (#PCDATA)>
<!ELEMENT schedule (week, time)>
<!ELEMENT week (#PCDATA)>
<!ELEMENT time (#PCDATA)>
]>
<lecturelist>
<lecture>
<title id=1234> Knowledge Sharing
Systems</title>
<lecturer> Hideaki Takeda</lecturer>
<schedule>
  <week>Wednesday</week>
  <time>3rd</time>
</lecture>
</lecturelist>
```



```
<html>
<body>
<table border="1">
<tr>
<td>Lecture name</td>
<td>Lecturer</td>
<td>Time</td>
</tr>
<tr>
<td> Knowledge Sharing Systems </td>
<td> Hideaki Takeda</td>
<td>
  Wednesday/3rd
</td>
</tr>
</table>
</body>
</html>
```

Whey is XML not sufficient?

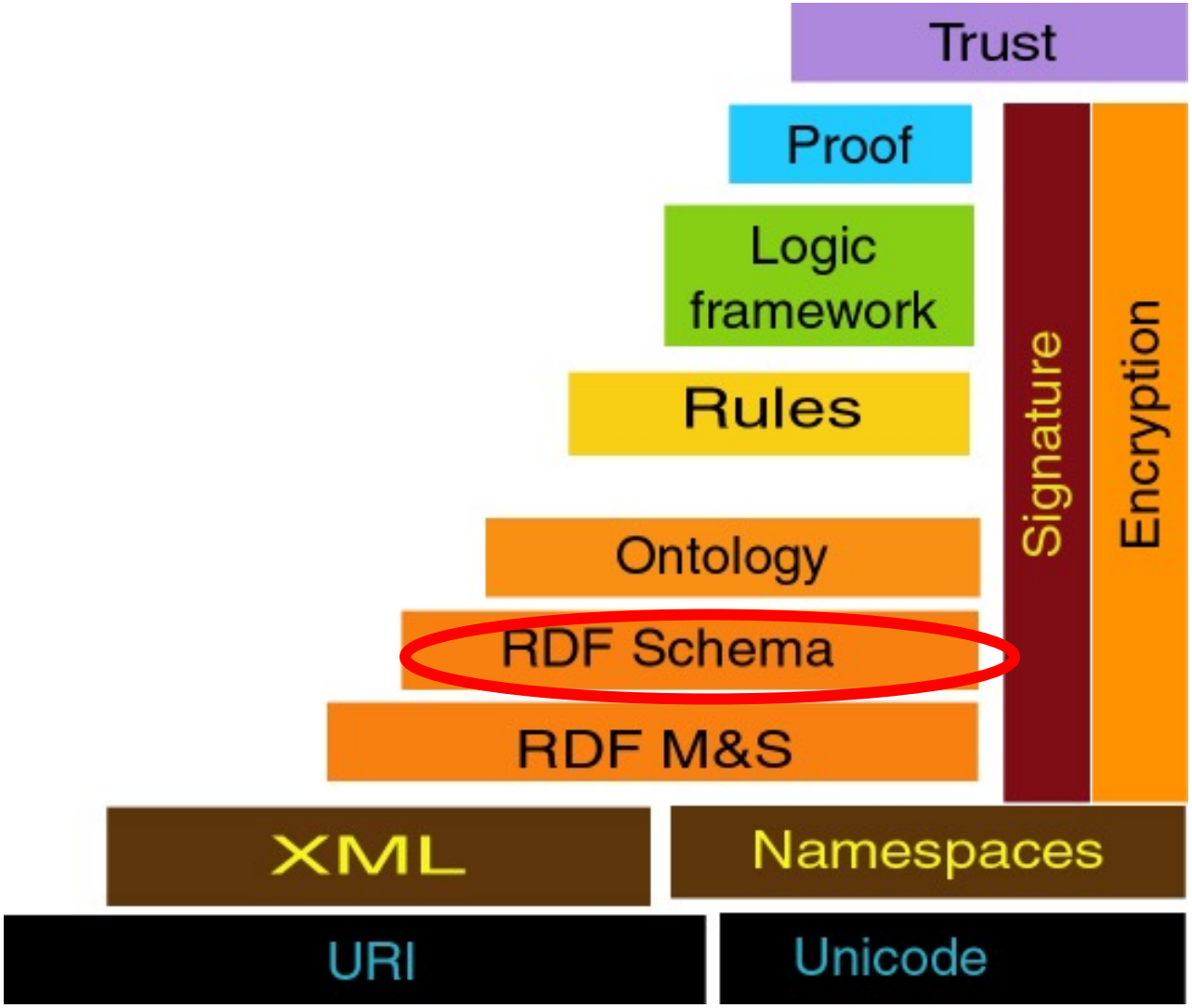
```
<person>
   <name> Hideaki Takeda</name>
   <age> 20</age>
</person>
```

```
<個人>
   <名前>Hideaki Takeda</名前>
   <年齢> 20</年齢>
</個人>
```

- What are specified by “person” and “name” ?
- Is “name” and “名前” the same?
- Is this description sufficient as a description for “person”?
- …

- In short, syntax alone cannot solve these problems

Architecture for the Semantic Web



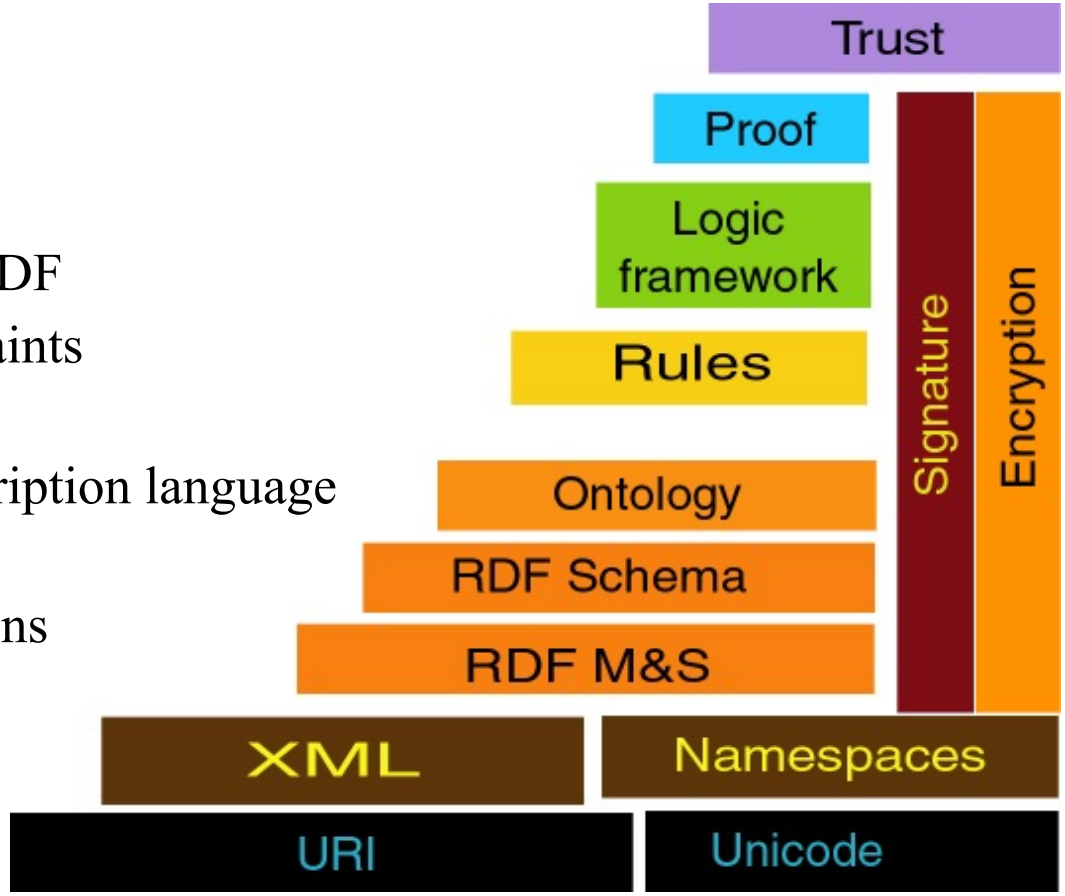
Tim Berners-Lee http://www.w3.org/2002/Talks/09-lcs-sweb-tbl/

How to describe “meaning”?

- Need to describe “information on information”
  - “Meaning of something” is a description (“meaning”) to a description (“something”) in computers
  - Metadata
    - Data about data
- Need to architecture for common understanding
  - Syntax (language or scheme)
  - Vocabulary (ontology)

A Layer model for Semantic Web

- RDF (Resource Description Framework)
  - The most primitive model for metadata description
    - SVO model
    - Entity-Relation Model
    - Semantic net
- RDF Schema
  - Addition of "concept" to RDF
    - class-subclass, constraints
- OWL
  - More general concept description language
    - Logical consistency
    - Various class expressions
    - Various constraints



Tim Berners-Lee http://www.w3.org/2002/Talks/09-lcs-sweb-tbl/

Two origin of Semantic Web

- From the viewpoint of Web Community
  - More meaning, more knowledge
  - HTML -> XML -> RDF -> RDF Schema -> OWL
    ->XML schema ->

- From the viewpoint of Knowledge Sharing Community (AI)
  - From closed knowledge presentation to open KR
    - Not for depth but for width
    - Not for completeness, but for possibility
  - KIF, Ontolinga -> OKBC -> OWL