The role of semantics in the development from Web2.0 to Web3.0”

SDM – IMD New Gen Lib Workshop, 2008-02-25

Johannes Keizer
Information Systems Officer
FAO of the United Nations
Knowledge Exchange and Capacity Building Division

This presentation

1. Introduction
   - Knowledge to combate Hunger and poverty
2. Important Developments
   - The web becomes a mass activity
   - Global Public Goods
   - Going Digital
   - “Web20”
3. The next Internet Revolution
   - Requirements
   - Does semantics matter?
   - Bricks for the semantic web
4. The Agricultural Ontology Service
   - Application Profiles
   - AgroVoc
   - Domain Ontologies
   - Semantic Applications (Agropedia Indica)

Millenium Development Goals

Target year 2015

- Eradicate extreme poverty and hunger
- Achieve universal primary education
- Promote gender equality and empower women
- Reduce child mortality
- Improve maternal health
- Combat HIV/AIDS, malaria and other diseases
- Ensure environmental sustainability
- Develop a global partnership for development

http://www.un.org/millenniumgoals/

Cross Cutting Issues for our Future Work

- **Knowledge management;**
- Policy support and the development of policy and strategy;
- Capacity building;
- Basic data and statistics;
- Environment and natural resource management;
- Emergencies and rehabilitation; and
- Gender mainstreaming and women's empowerment.

http://www.fao.org

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS

helping to build a world without hunger

العربية | 中文 | français | español

Google™ Custom Search search

- What is FAO?
- From the Director-General
- Governing Bodies
- Agriculture and Consumer Protection
- Economic and Social Development
- Fisheries and Aquaculture
- Forestry
- Natural Resources Management and Environment
- Technical Cooperation
- Programme and Budget; Evaluation
- Legal Office
- Decentralized Offices
- Employment
- Procurement Service
- Interdisciplinary Activities: Trade, Biotechnology, Gender...
- Statistical Databases
- Country information
- Virtual Library
- Publications and Documents
- more...

Newsroom

Improved management of fishing's "last frontier" needed
Countries discuss how to better protect deep sea species and habitats

Tea prices to maintain upward trend in 2008
Output in Viet Nam increases 28 percent, another record crop in China

Significant increase in world cereal production forecast for 2008, but prices remain high
New FAO web portal tracks market trends

FAO unveils new bioenergy assessment tool
Weighs impact on food security

more news...

Webcasting from FAO

Jatropha nuts to make biodiesel in Thailand

more photos...

Knowledge Forum

- **World Food Situation**
- World Agricultural Information Centre
- **H5N1 Avian Influenza**
- Progress since the World Food Summit
- **World Food Day**
- **FAO Hunger Map**
- **TeleFood**
- Feeding Minds Fighting Hunger

Information Finder

All FAO & related sites search

Selected Key Programmes

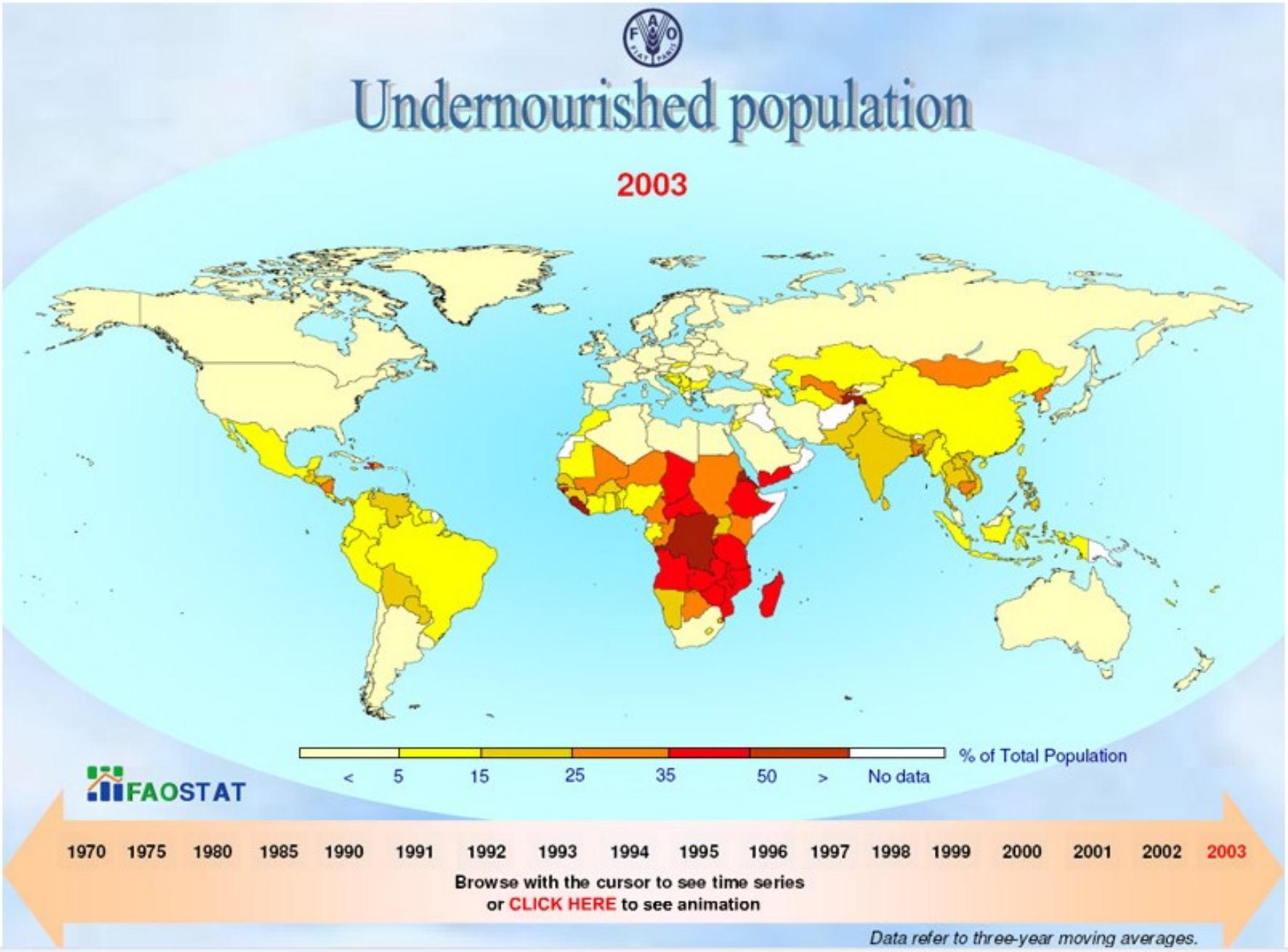
SPFS: Special Programme for Food Security

EMPRES: Emergency Prevention

http://www.fao.org/waicent

Dr.Johannes Keizer, Knowledge Exchange and Capacity Building Division

http://www.fao.org/es/ess/faostat/foodsecurity/FSMap/map14.htm

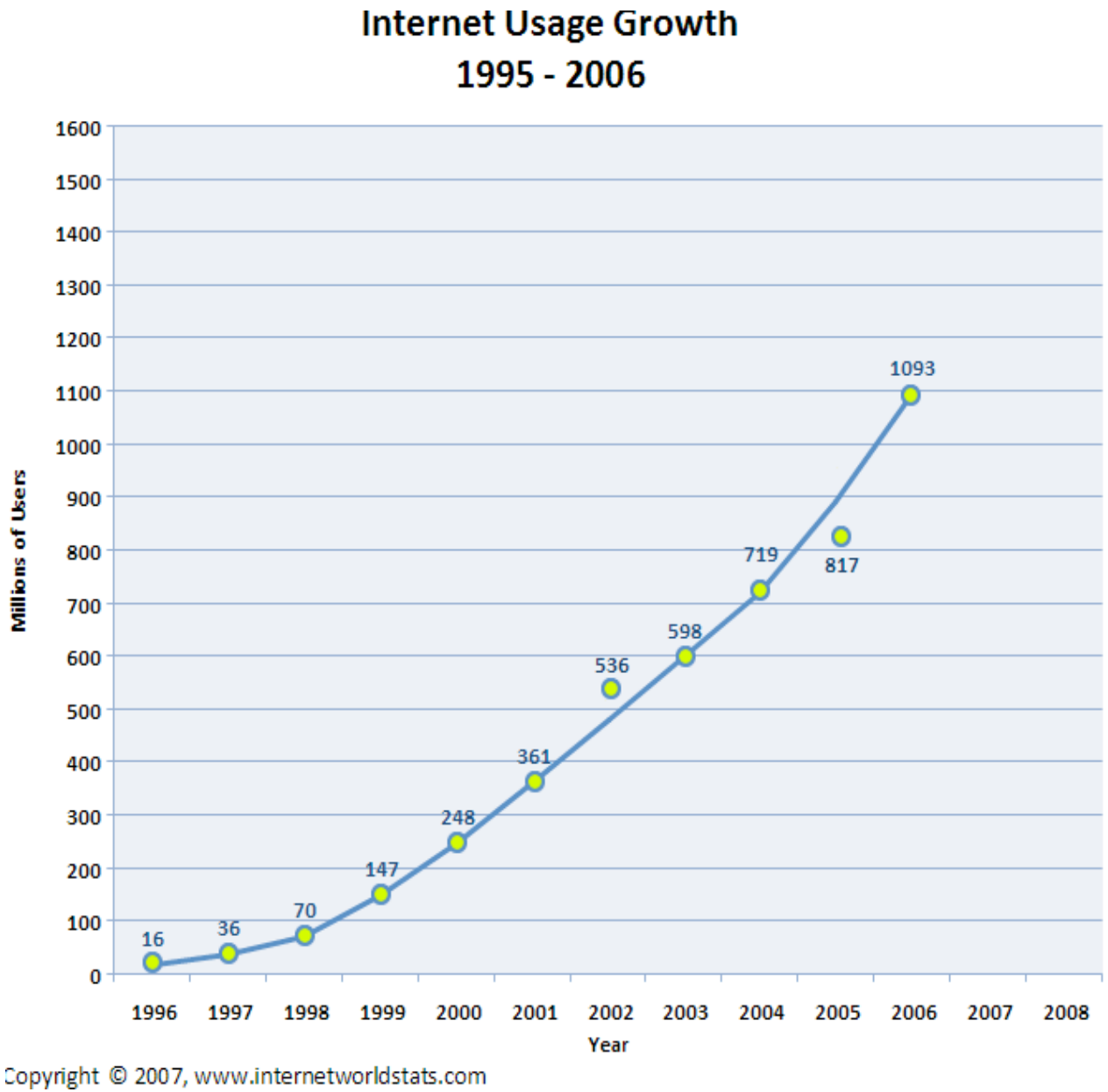


Dr.Johannes Keizer, Knowledge Exchange and Capacity Building Division

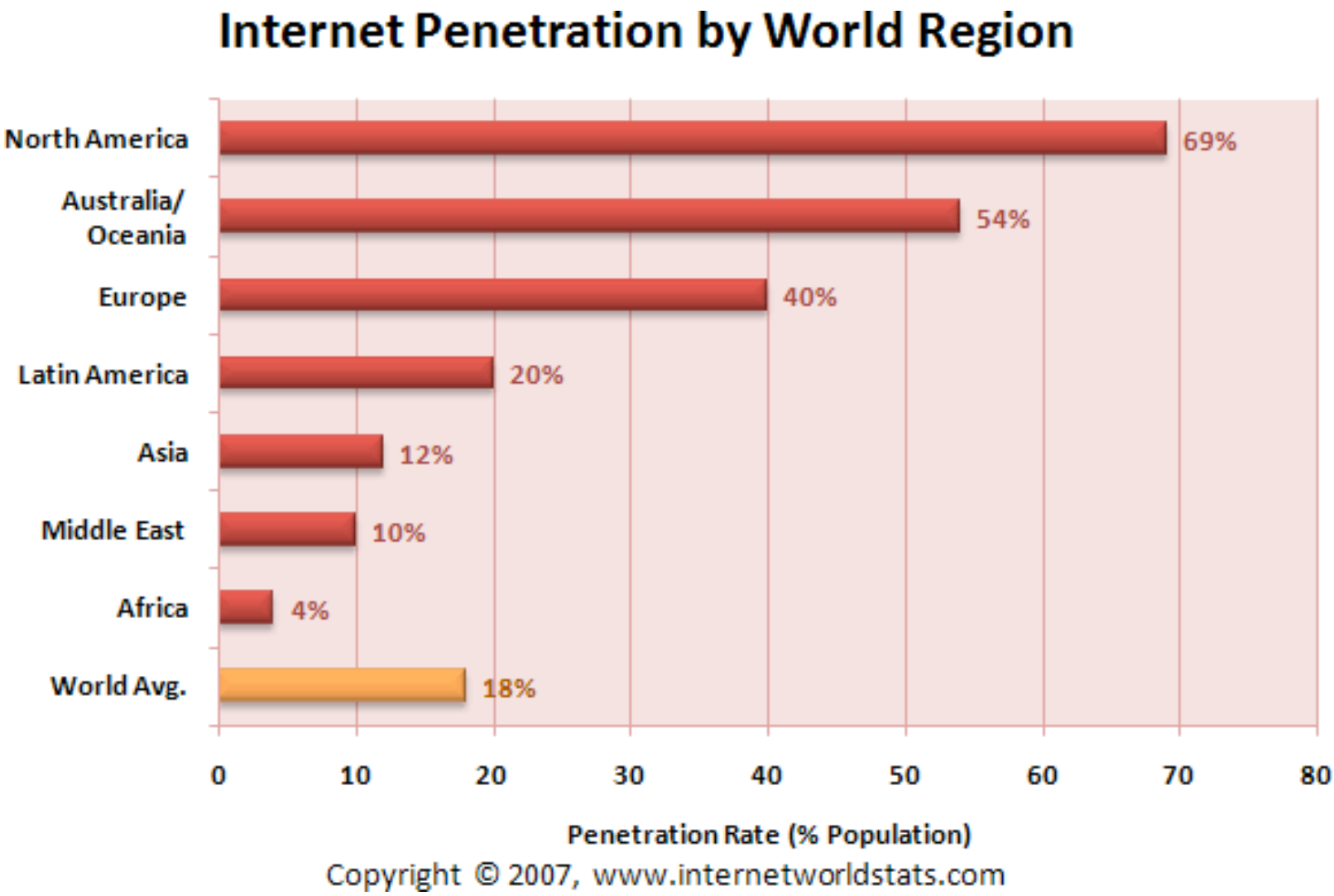
Important Developments

1. The web becomes a mass activity
2. Global Public Goods
3. Going Digital
4. “Web20”

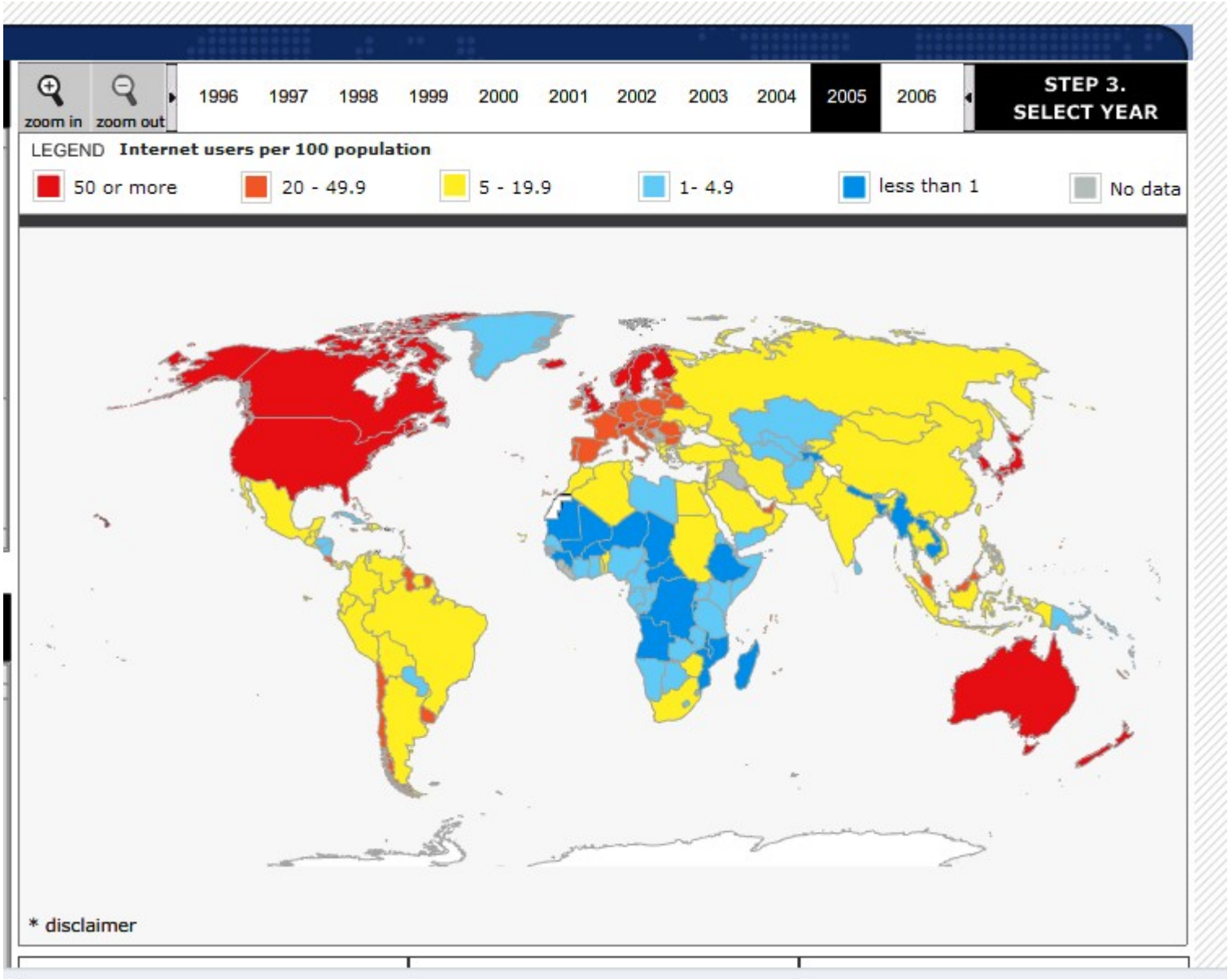
The web becomes a mass activity



The web becomes a mass activity



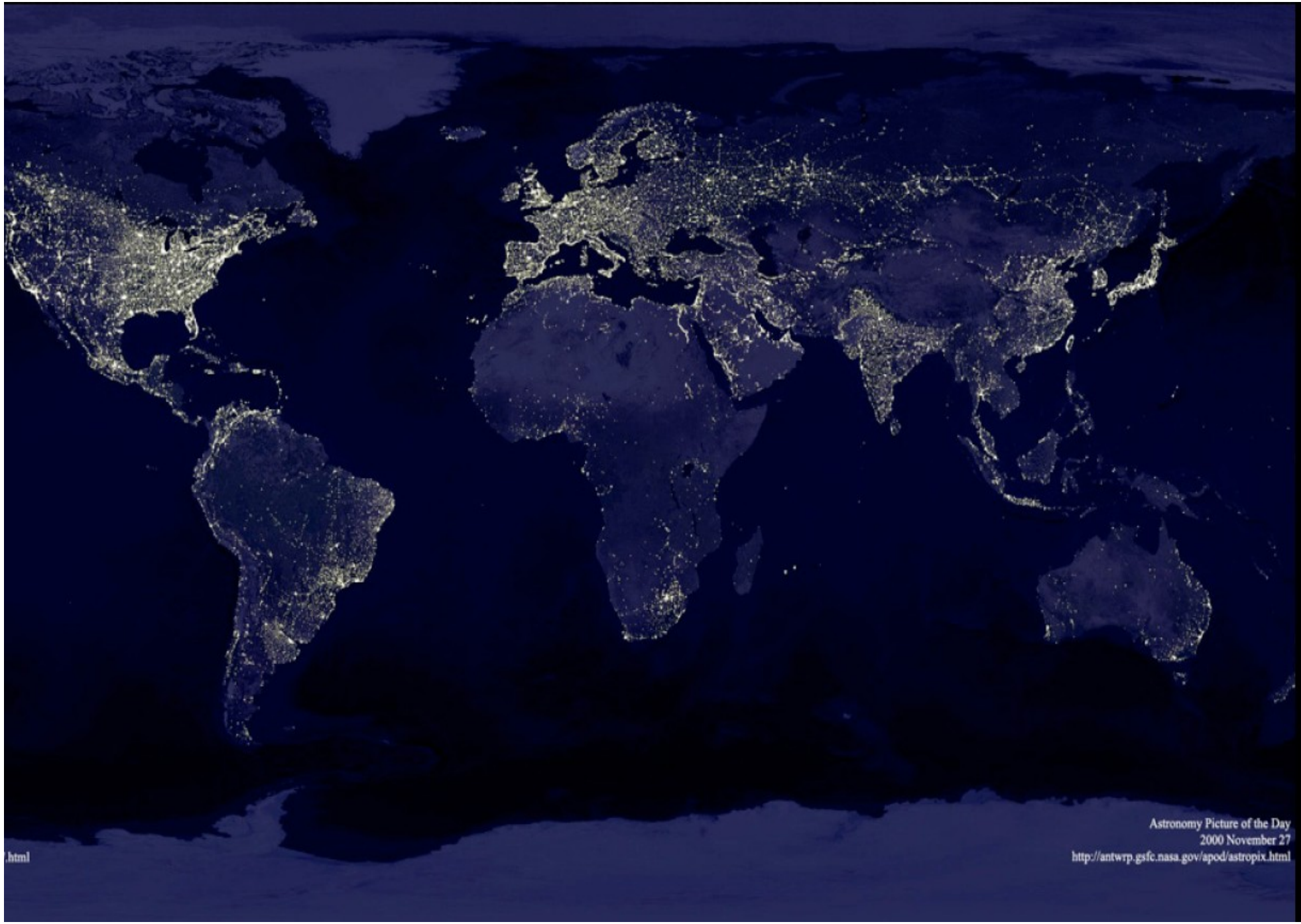
Internet Usage in 2005



Internet Usage the tendencies

Internet Users/100 population	2000	2005
India	0.54	5.44 (61 Mln)
China	1.77	8.44 (112 Mln)
USA	44.06	66.33 (200 Mln)
Germany	30.15	43.17
Ghana	0.15	1.81
Sweden	45.08	76.21

Source: Millennium Development Goals Monitor


Astronomy Picture of the Day
2000 November 27
http://antwrp.gsfc.nasa.gov/apod/astropix.html

Global Public Goods

- The Open Access publishing Paradigm gets mainstream
- Research Results available as global public goods (Gene Ontology, CGIAR Programs)
- Substitution of proprietary software by open Source

Open Access Publishing

- Started in 90s in Los Alamos
- Now mainstream for all bigger research organizations
- Mandatory in US and UK for research done with public funding
- No change in scholarly publishing system, but in addition institutionally and subject repositories
- Most scholarly publishing available in open access with OAI-MPH protocol in the next 20 years?

Going Digital

- Most formal knowledge will be digitally accessible
  - Institutional Repositories
  - Googlebooks and related projects
  - eGovernment?
  - Everything what you want to buy in e-commmerce(Amazonization?)
- Also tacit knowledge becomes explicit
  - Wikis
  - Blogs
  - Question and Answer Services

Web 20

- People on the web
  - Wikipedia (becomes mainstream information source)
  - Blogs for (nearly) everyone
  - Social Networking (MySpace, Facebook)

- More structured data
  - CMS Inflation and
  - Tagging - Folksonomies
  - Much more semistructured data and semi-defined semantics?? With a lot of metadata

Web 20 is more about people than about technology, some Web20 technology is even old(Wiki, Blogs) - the difference is the mass participation of people in the web

An what is to come

- 2 billion Internet users in the world in 2010 !?
- All relevant content will be digitally (and mostly freely) available!?
- 10 Million internet users with their own blogs?
- 100,000s of Folksonomies!?
- Most of Internet Continent in CMS or semistructured databases!?
- ICT Literacy of Internet Users will increase by order of magnitudes (today a majority f users of Internet ICT tools are “illiterate”)!?

Some requirements to a global network

1. Does the knowledge/ information/ resource I need exist ?
2. Where is it?

(Questions for reference librarians)

4. How can I use this knowledge for my own business
5. How can I collaborate with others over the network

Where is it

Up to know a satisfactory solution has been found only for Question 2

“Where is it?”

If you know precisely for what you are looking, and if it is present on the web, Google will most probably find it for you

Question 2

1. Does the knowledge/ information/ resource I need exist?

here is a bias: (some people think they can investigate this)

telling a story from the preparation of my last summer holidays

The problem of meaning

- Input to the Internet is mostly given by humans, but the information is processed by machines (computers)
- Humans can interprete Information
  - Semantic and Linguistic Agreements
  - Inferences through known relations (a cow is a mammal)
- Computers see only bit streams and transfer them into strings and other visible objects or they see strings and other visible objects and transfer them into bit streams
- What can be done to teach them meaning?

The problem of meaning (2)

cow, mucca, vacca, Kuh, vache, بقرات , गाय ,
……..

…in a certain environment the same concept…

..you know, a Computer does not know

Questions 3 + 4

- doing business and collaborative work on the web
  - very, very poor
    - How often do you need to register to access information
    - How easy is it for you to personalize information streams?
    - email dominates communication leading to wrong communication structures
    - social networking exists on proprietary platforms only, no social working exists, it is even complicated to link to your peers
  - the problem is also here, computer cannot interprete information

Why metadata tags were a failure

- DC 1995
  - tried to copy library cataloguing to the web
  - but catalogues were interpreted by humans, not by machines
  - computer could not understand the meaning of the metadata
- Semantic Identification of knowledge objects and concepts within knowledge objects

Knowledge Objects

- it's a Hotel
- it's a CV
- it's a Recipe
- it's a political journal

Concepts

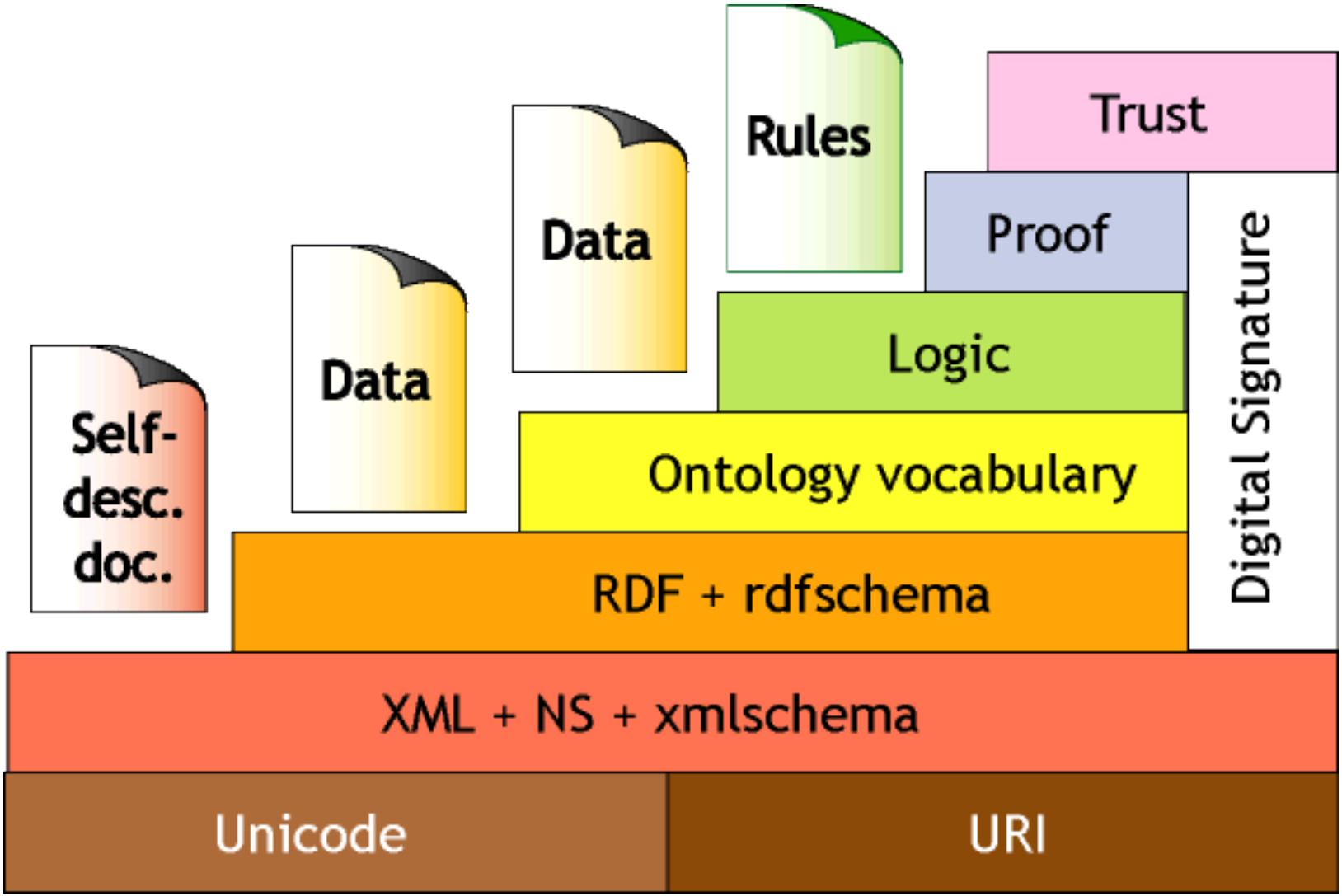
- it's a cow
- it's a Person
- it's Johannes Keizer
- it's .....

- Unvalidated Metadata are useless

Semantic Identification of Metadata/data

- How can computer validate Information
  - it must find a machine interpretable information
- Trusted Namespaces with standard URIs (Ontology Vocabulary Registries)
  - Persons, Organizations, Projects, Conferences
  - Concepts, thesauri, ontologies
- To be established by standard bodies or communities of practice

Sir Tim's Vision of the semantic Web



Agricultural Ontology Service

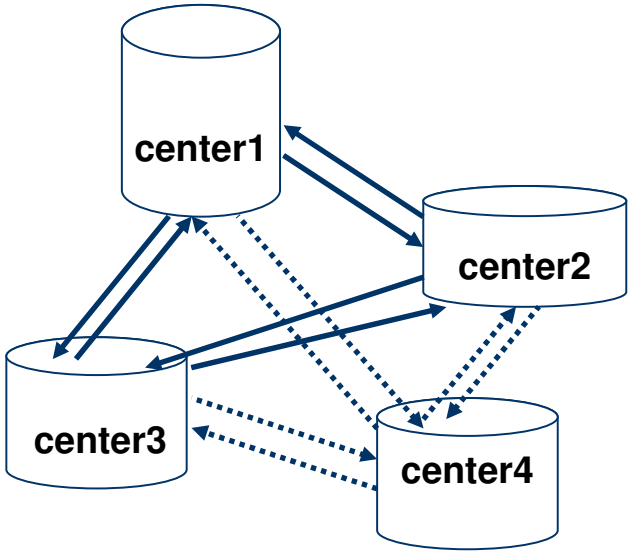
http;//www.fao.org/aims

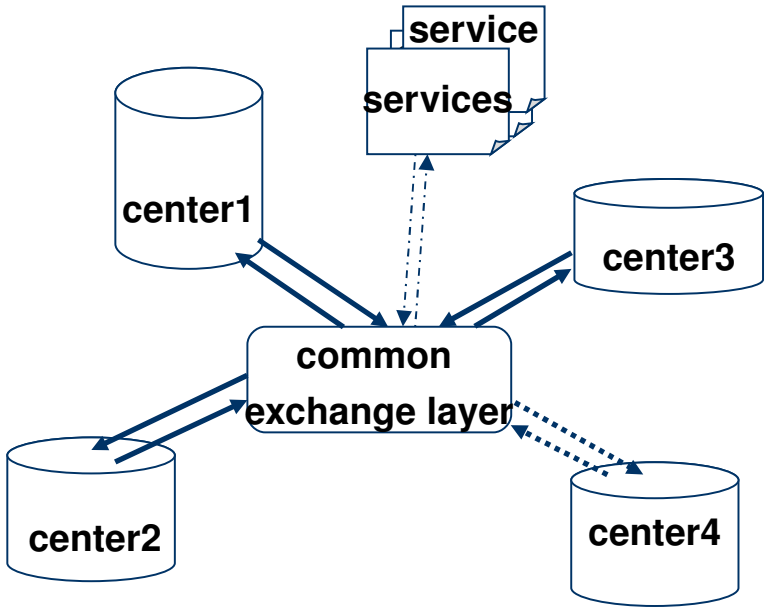
- Application Profiles/Ontologies
- AGROVOC
- Domain Ontologies
- Knowledge Based Applications (Agropedia Indica)

Application Profiles/ Application Ontologies

The AGRIS problem

$$n_P_2 = \frac{n!}{(n-2)!}$$

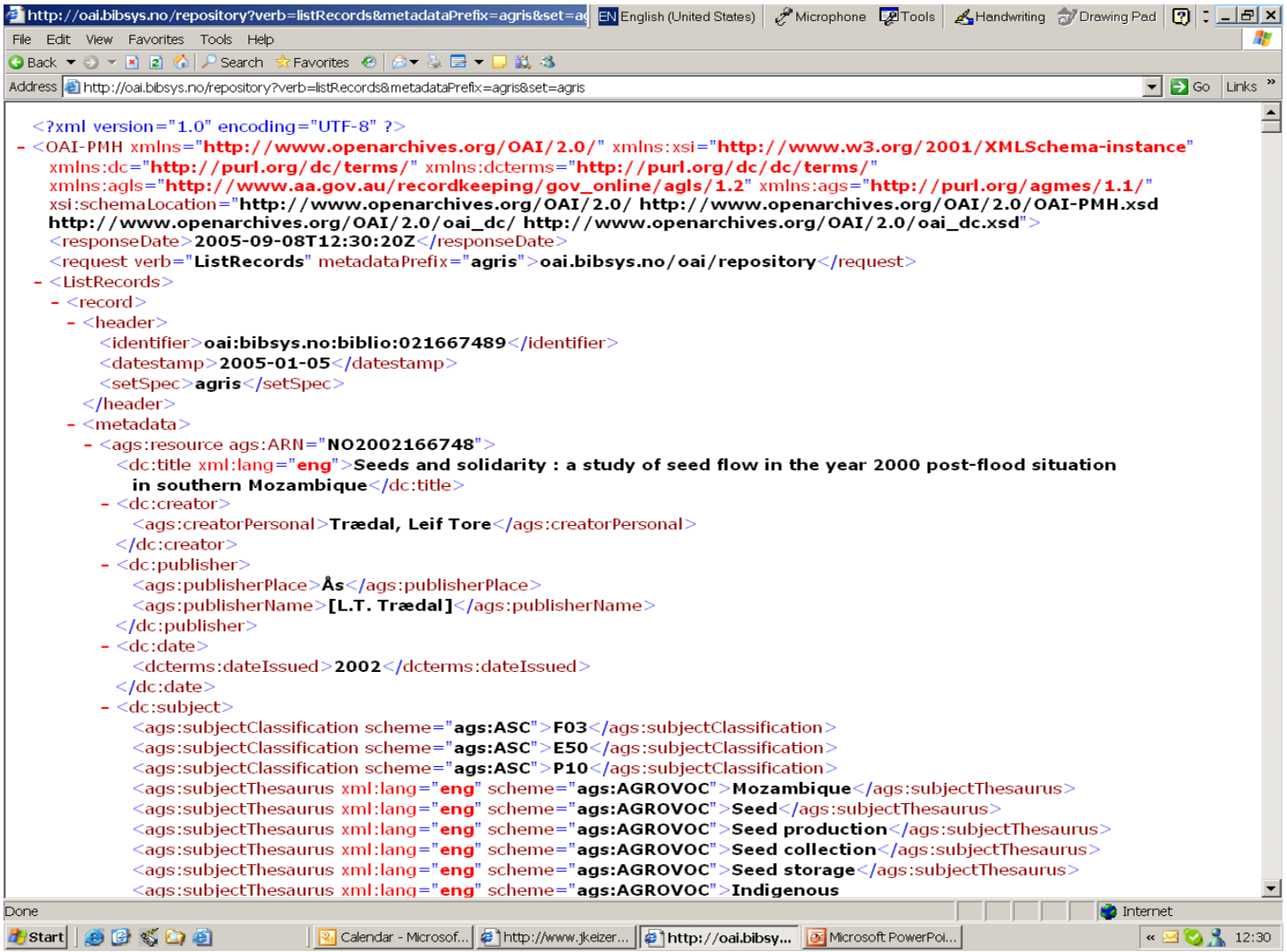




How to teach computers meaning

Section 1: Elements	Section 2: Refinements	Section 3: Schemes
(DC) Title	(DCTERMS) Alternative (AGS) titleSupplement	-
(DC) Creator	(AGS) creatorCorporate (AGS) creatorPersonal (AGS) creatorConference	-
(DC) Subject	(AGS) subjectClassification	(AGS) ASC (AGS) CABC (AGS) GFDC (DCTERMS) DDC (DCTERMS) LCC (DCTERMS) UDC
	(AGS) subjectThesaurus	(AGS) AGROVOC (AGS) CABT (AGS) ASFAT (AGS) NALT (AGS) BT (DCTERMS) MeSH (DCTERMS) LCSH
(DC) Description	(DCTERMS) Abstract (AGS) descriptionNotes (AGS) descriptionEdition	-
(DC) Publisher	(AGS) publisherName (AGS) publisherPlace	-
(DC) Date	(DCTERMS) Created (DCTERMS) Valid (DCTERMS) Available (DCTERMS) Modified (DCTERMS) Issued	(DCTERMS) W3CDTF
(DC) Type	-	(DCTERMS) DCMIType
(DC) Format	(DCTERMS) Extent	-

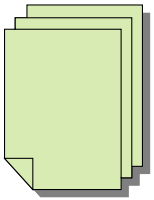
http://oai.bibsys.no/repository?verb=listRecords&metadataPrefix=agris&set=agris



```
<?xml version="1.0" encoding="UTF-8" ?>
- <OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:dc="http://purl.org/dc/terms/" xmlns:dcterms="http://purl.org/dc/dc/terms/"
  xmlns:agls="http://www.aa.gov.au/recordkeeping/gov_online/agls/1.2" xmlns:ags="http://purl.org/agmes/1.1/"
  xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd
  http://www.openarchives.org/OAI/2.0/oai_dc/ http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
  <responseDate>2005-09-08T12:30:20Z</responseDate>
  <request verb="ListRecords" metadataPrefix="agris">oai.bibsys.no/oai/repository</request>
- <ListRecords>
  - <record>
    - <header>
      <identifier>oai:bibsys.no:biblio:021667489</identifier>
      <datestamp>2005-01-05</datestamp>
      <setSpec>agris</setSpec>
    </header>
    - <metadata>
      - <ags:resource ags:ARN="NO2002166748">
        <dc:title xml:lang="eng">Seeds and solidarity : a study of seed flow in the year 2000 post-flood situation
        in southern Mozambique</dc:title>
        - <dc:creator>
          <ags:creatorPersonal>Trædal, Leif Tore</ags:creatorPersonal>
        </dc:creator>
        - <dc:publisher>
          <ags:publisherPlace>Ås</ags:publisherPlace>
          <ags:publisherName>[L.T. Trædal]</ags:publisherName>
        </dc:publisher>
        - <dc:date>
          <dcterms:dateIssued>2002</dcterms:dateIssued>
        </dc:date>
        - <dc:subject>
          <ags:subjectClassification scheme="ags:ASC">F03</ags:subjectClassification>
          <ags:subjectClassification scheme="ags:ASC">E50</ags:subjectClassification>
          <ags:subjectClassification scheme="ags:ASC">P10</ags:subjectClassification>
          <ags:subjectThesaurus xml:lang="eng" scheme="ags:AGROVOC">Mozambique</ags:subjectThesaurus>
          <ags:subjectThesaurus xml:lang="eng" scheme="ags:AGROVOC">Seed</ags:subjectThesaurus>
          <ags:subjectThesaurus xml:lang="eng" scheme="ags:AGROVOC">Seed production</ags:subjectThesaurus>
          <ags:subjectThesaurus xml:lang="eng" scheme="ags:AGROVOC">Seed collection</ags:subjectThesaurus>
          <ags:subjectThesaurus xml:lang="eng" scheme="ags:AGROVOC">Seed storage</ags:subjectThesaurus>
          <ags:subjectThesaurus xml:lang="eng" scheme="ags:AGROVOC">Indigenous
```

Application Profiles

- Document – Like Information Objects (DLIOs)



- Learning resources

- Organizations

- Projects

- News and Events

AGROVOC

- Thesauri, were traditionally the tool of librarians to describe the content of a document

- They are now basis to describe knowledge and meaning in a machine readable way

- AGROVOC was maintained by FAO since the 80s

AGROVOC

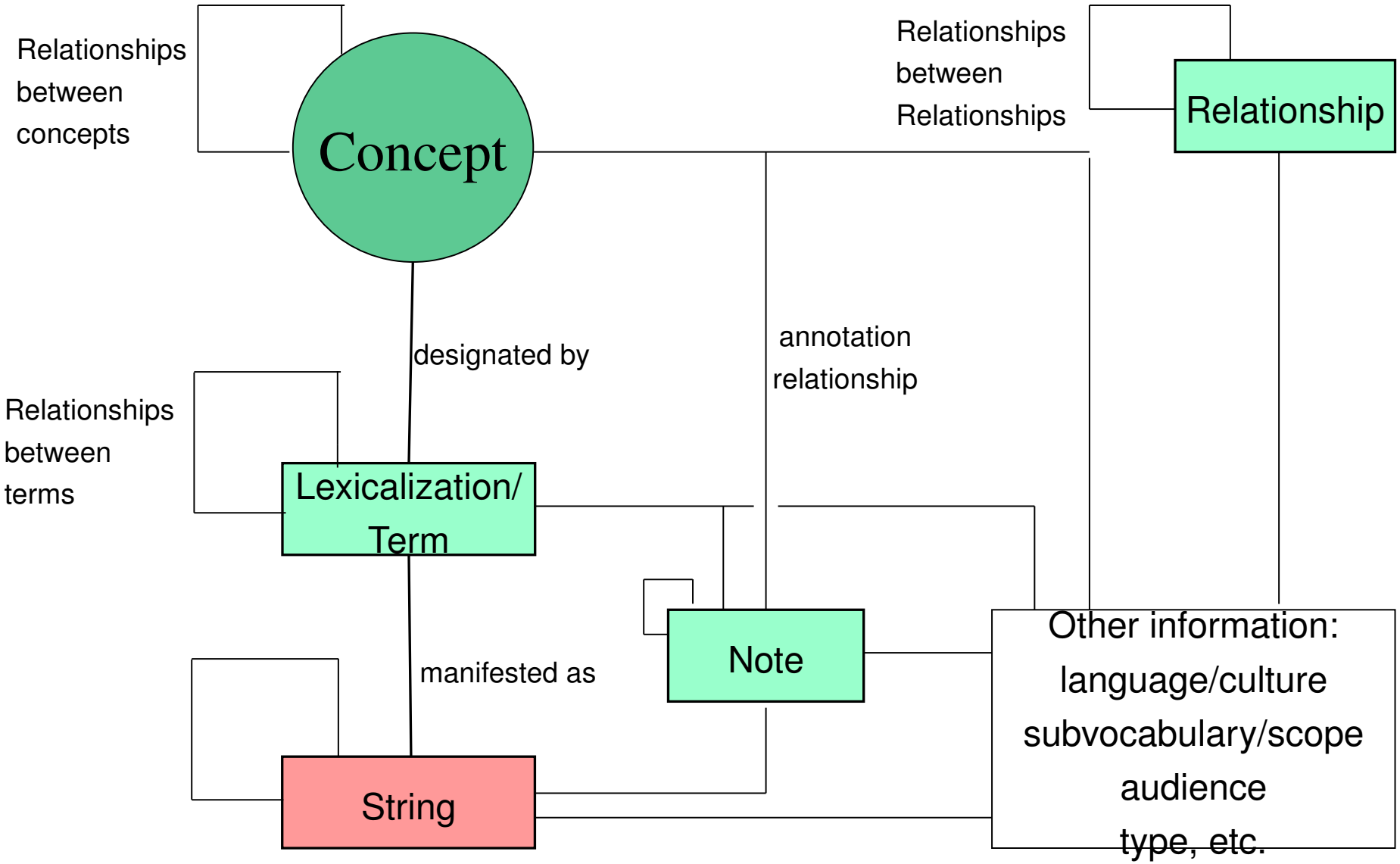
Learn more about AGROVOC by browsing: AGROVOC Flyer

Search term: [ ] Search in AGROVOC

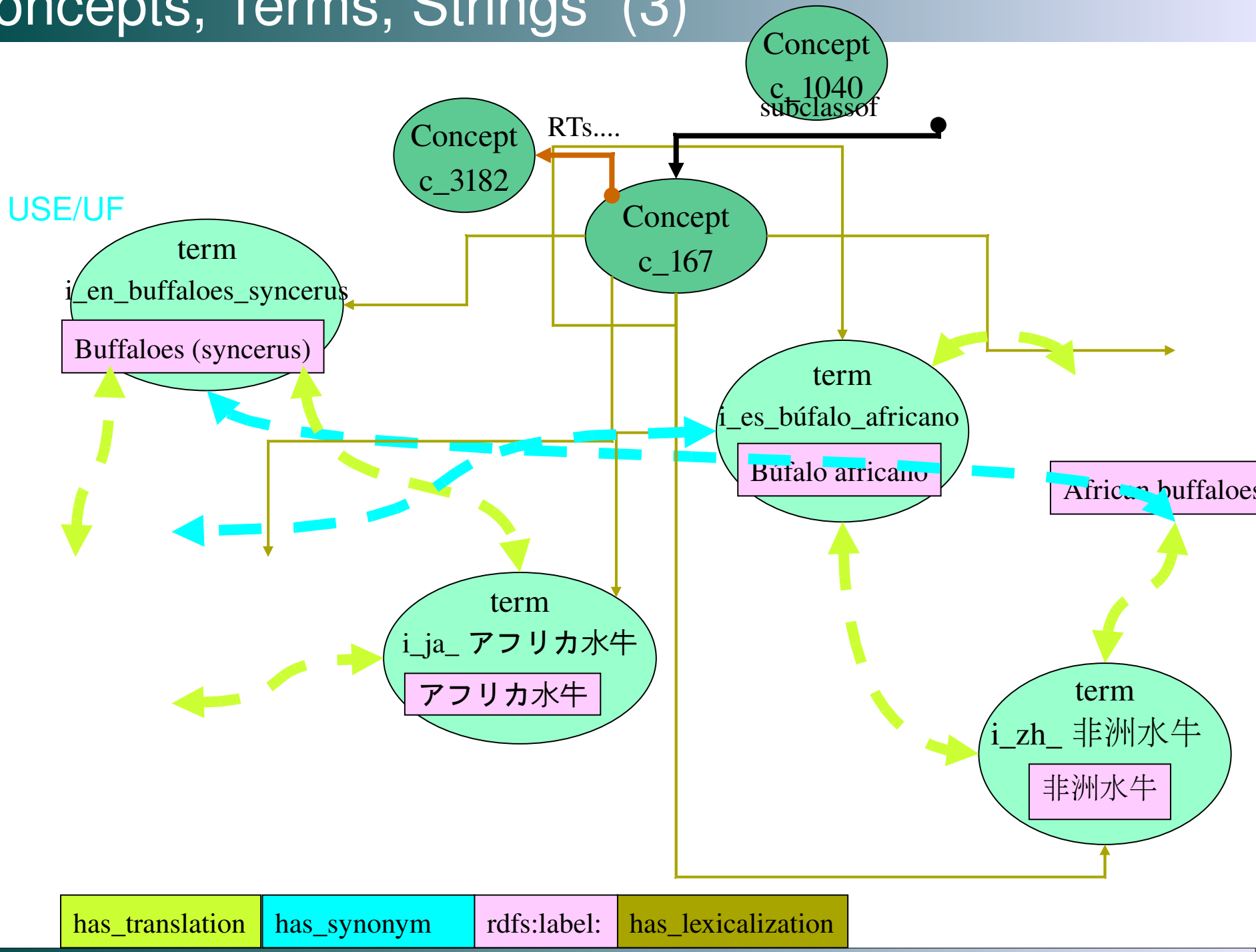
(•) starting with ( ) containing text ( ) exact match

EN : Rice	BT : Cereals
FR : Riz	RT : Rice flour
ES : Arroz	RT : Oryza
AR : أرز	RT : Rice straw
ZH : 稻米	UF : Paddy
PT : Arroz	
CS : rýže	
JA : 米	
TH : ข้าว	
SK : ryža	
DE : REIS	
HU : rizs	
PL : Ryż (ziarno)	
FA : برنج	
IT : Riso	
HI : चावल/धान	
LO : ?????	

Concepts, Terms, Strings (2)



Concepts, Terms, Strings (3)

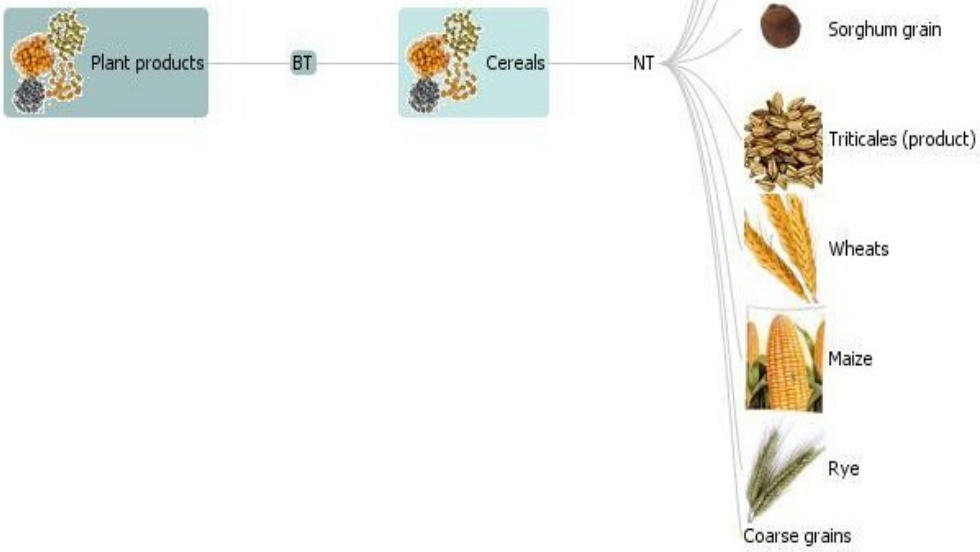


Different Views on the Same knowledge

Plant products
- NT Cereals
  - NT Barley
    - NT Malting barley
  - NT Coarse grains
  - NT Maize
    - NT Dent maize
    - NT Flint maize
    - NT Popcorn

C Plant products(en)Prodotti di origine
उत्पाद(hi)ထးဘ ိဝးမ ်ဘမ ိဝ(lo)

C धान्य(hi)Cereali(it)ธัญพืช(th)Cerea
- C Riso(it)ເ ်ာ(lo)Rice(en)ข้าว(t
- C ข้าวโพด(th)मक्का /मक्की(hi)ဘာဝ ိ(
- C Sorgo(it)ဃ ်ဝເ ်ာမ ်ာ(lo)So
- C Avena (prodotto)(it)ข้าวโอ๊ต(th)
- C Triticale (prodotto)(it)ทริติเคล (
(product)(en)ເ ်ာပ_ပးဘ ်ဃ(
- C Millets(en)ເ ်ာမ ်ာ(lo)Miglio
- C ข้าวไรย์(th)राई(hi)Segale(it)Rye
- C Orzo(it)ข้าวบาร์เลย์(th)ເ ်ာပား
- C ธัญพืชเมล็ดโต(th)Coarse grains(en)Cereali per l'alimentazione animale(it)मोटे दाने(hi)ဖ ိဝဃ ်ဝใတย ' (lo)
- C गेहूँ(hi)ເ ်ာဘာဝ ိ(lo)ข้าวสาลี(th)Grani (cereali)(it)Wheats(en)



Creating Ontologies (an example)

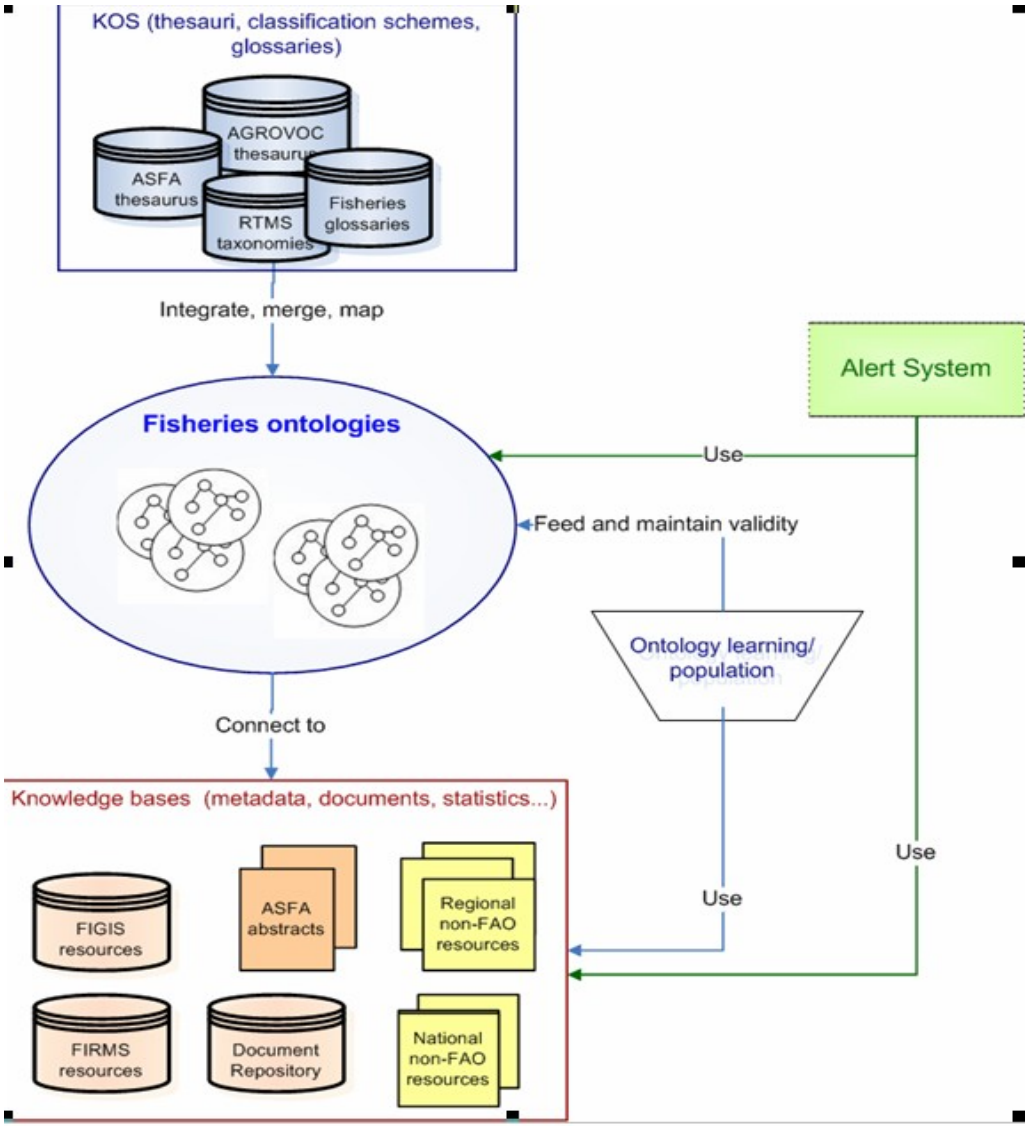


Figure 4 – Overview of the ontology-driven FSDAS

Creating Ontologies (2)

How it looks like

- http://www.fao.org/aims/aos/fi/fishing_areas

And how we did it

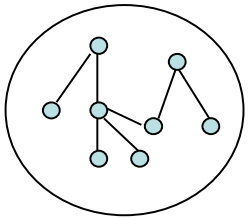
- http://www.fao.org/aims/neon.jsp

Building applications (ex. Agropedia Indica)

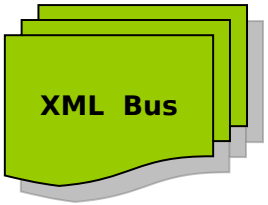
Value added Information Services

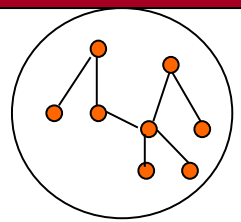
Portal 1 | Portal 2 | Portal 3 | Portal 4 | Portal 5 | Portal 6 | Portal 7

Specific Knowledge Views



Metadata ontologies (Application Profiles)





Subject ontologies

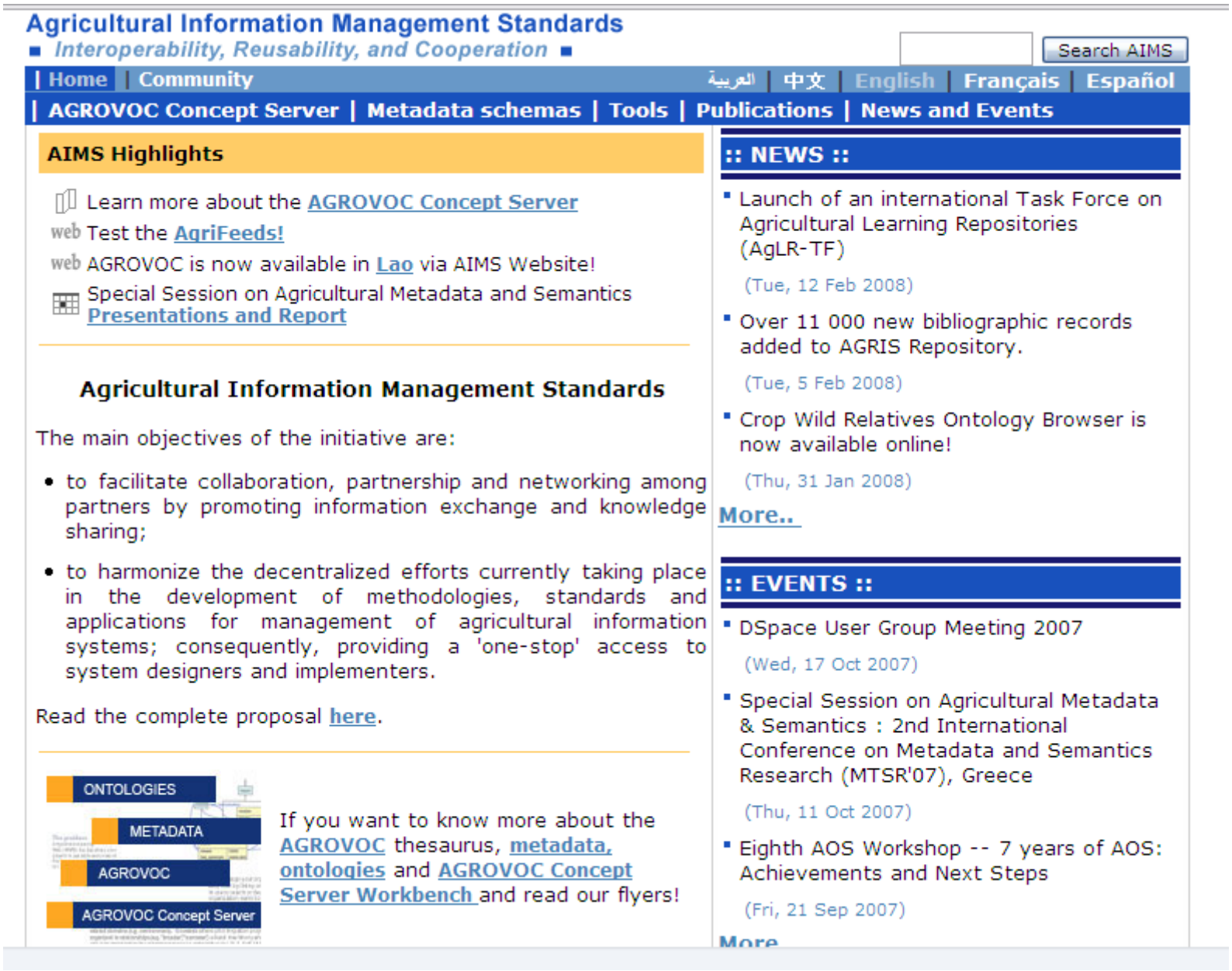
Repository Specific Wrappers

OAI-MHP | CMS | eGovApps | PlantOnto | BioMoby | Blogs | WIKIs | QA

Digitization: Unicode and URI

Agricultural Knowledge

http://www.fao.org/aims



Thank You

.....for your attention

johannes keizer, MSc, PhD

Informations systems officer

FAO of the United nations

johannes.keizer@fao.org

http://www.fao.org

http://www.fao.org/aims

special thanks to the great team in Rome, Bangkok, Kanpur, Washington, Djakarta, Bangalore and Hyderabad