- Learning Goals
- Development of Web
- Limitation of current Web
- How to improve current Web?
- Vision of Semantic Web
- What is the Semantic Web?
- Ontology definition

LEARNING GOALS

- Basic understanding of semantic technologies and their evolution.

- Knowledge of:
  Core semantic languages like RDF, RDFS, OWL,…
  working principles and technologies for handling semantic systems
  e.g. annotation, reasoning, relation to semantic web services…

- Knowledge of typical / most prominent recent developments, trends, applications in the field.

DEVELOPMENT OF WEB

- Internet - 1st generation
- Web 1.0 - 2nd generation
- Web 2.0 – 3rd generation

Internet - 1st generation

"The **Internet** is a global system of interconnected computer networks that use the standard Internet Protocol Suite (TCP/IP) to serve billions of users worldwide. It is a network of networks that consists of millions of private and public, academic, business, and government networks of local to global scope that are linked by a broad array of electronic and optical networking technologies."

- Internet evolution started from 1945 and existed till 1995.
- It enabled disparate machines to exchange data.

Web 1.0 - 2nd generation

"The **World Wide Web** ("**WWW**" or simply the "**Web**") is a system of interlinked, hypertext documents that runs over the Internet. With a Web browser, a user views Web pages that may contain text, images, and other multimedia and navigates between them using hyperlinks". Started 1989

- It enabled new applications on top of growing internet, making enormous amounts of information available, in human-readable form, allowing revolution in new applications B2C.

Netscape

It is associated with the breakthrough of the Web.
It connects user community making attractive to present their information on the Web.

Google

It is the incarnation of Web 1.0 mega grows
It's indexed already in 2008 more than 1 trillion pages [*]
Google and other search engines can fetch information faster on Web.

WEB 1.0 PRINCIPLES

The success of Web 1.0 is based on three simple principles:

- A simple and uniform addressing schema to identify information chunks i.e. Uniform Resource Identifiers (URIs)
- A simple and uniform representation formalism to structure information chunks allowing browsers to render them i.e. Hyper Text Markup Language (HTML)
- A simple and uniform protocol to access information chunks i.e. Hyper Text Transfer Protocol (HTTP)

1. Uniform Resource Identifiers (URIs)

- Uniform Resource Identifiers (URIs) are used to name/identify resources on the Web
- URIs are pointers to resources to which request methods can be applied to generate potentially different responses
- Resource can reside anywhere on the Internet
- Most popular form of a URI is the Uniform Resource Locator (URL)

2. Hyper-Text Markup Language

- Hyper-Text Markup Language:
  - A subset of Standardized General Markup Language (SGML)
  - Facilitates a hyper-media environment

- Documents use elements to “mark up” or identify sections of text for different purposes or display characteristics

- HTML markup consists of several types of entities, including: elements, attributes, data types and character references

- Markup elements are not seen by the user when page is displayed

- Documents are rendered by browsers

3. Hyper-Text Transfer Protocol (HTTP)

- Protocol for client/server communication
- The heart of the Web
- Very simple request/response protocol
- Client sends request message, server replies with response message
- Provide a way to publish and retrieve HTML pages
- Stateless
- Relies on URI naming mechanism

Web 2.0 – 3rd generation

- “The term "Web 2.0" (2004–present) is commonly associated with web applications that facilitate interactive information sharing, interoperability, user-centered design, and collaboration on the World Wide Web”.
- Web 2.0 is a vaguely defined phrase referring to various topics such as social networking sites, wikis and communication tools.
- With Web 1.0 technology a significant amount of software skills and investment in software was necessary to publish information.
- Web 2.0 technology changed this dramatically.

Web 2.0

The four major breakthroughs of Web 2.0 are:

- Blurring the distinction between content consumers and content providers.
  - Wiki, Blogs, and Twiter turned the publication of text in mass phenomena, as flickr and youtube did for multimedia
- Moving from media for individuals towards media for communities.
  - Social web sites such as del.icio.us, facebook, FOAF, linkedin, myspace and Xing allow communities of users to smoothly interweave their information and activities
- Blurring the distinction between service consumers and service providers
  - Mashups allow web users to easy integrate services in their web site that were implemented by third parties
- Integrating human and machine computing in a new and innovative way.
  - Amazon Mechanical Turk - allows to access human services through a web service interface blurring the distinction between manually and automatically provided services

Limitation of current Web

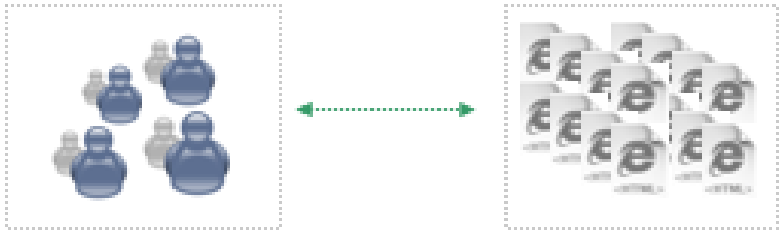
The current Web has its limitations when it comes to:

- finding relevant information
  - Finding information on the current Web is based on keyword search
- extracting relevant information
  - Current search engines provide no means to specify the relation between a resource and a term
- combining and reusing information
  - Searching for the same information in different digital libraries
  - Information may come from different web sites and needs to be combined

How to improve current Web?

- Increasing automatic linking among data
- Increasing recall and precision in search
- Increasing automation in data integration
- Increasing automation in the service life cycle
- Adding semantics to data and services is the solution!

Vision of Semantic Web



More than 3 billion users,
more than a trillion pages (2016)

Static

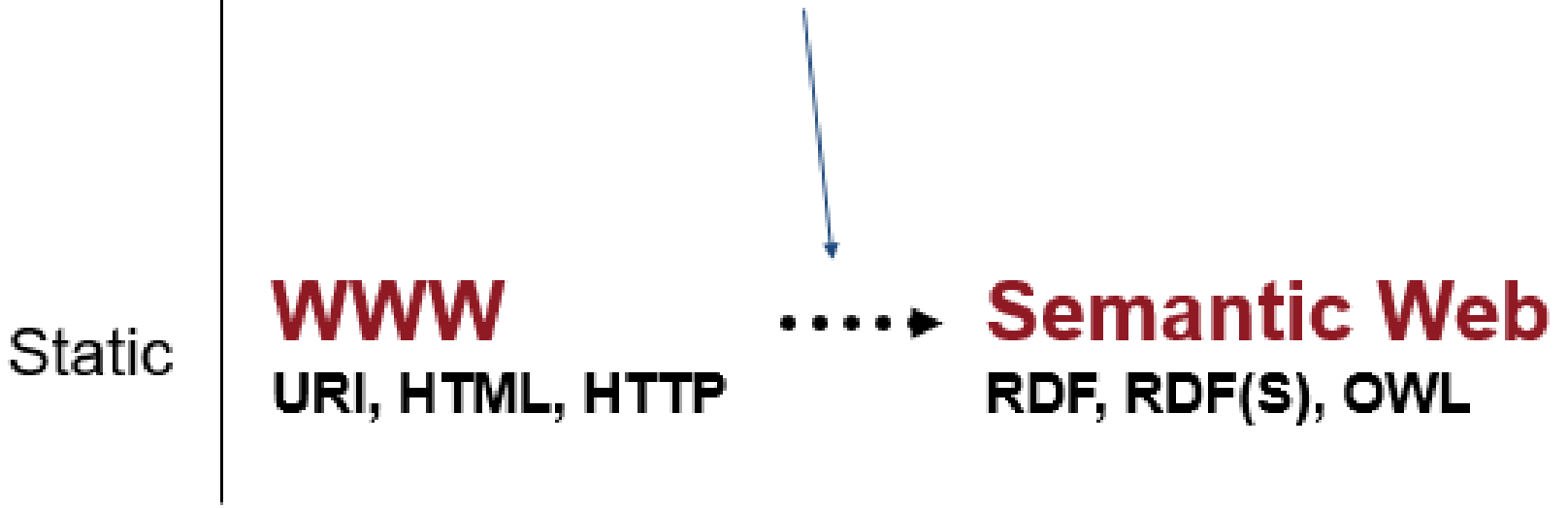
WWW

URI, HTML, HTTP

Vision of Semantic Web

Serious problems in

- information finding,
- information extracting,
- information representing,
- information interpreting and
- and information maintaining.



What is 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."

T. Berners-Lee, J. Hendler, O. Lassila, "The Semantic Web", Scientific American, May 2001

- The next generation of the WWW
- Information has machine-process-able and machine-understandable semantics
- Not a separate Web but an augmentation of the current one
- The backbone of Semantic Web are ontologies

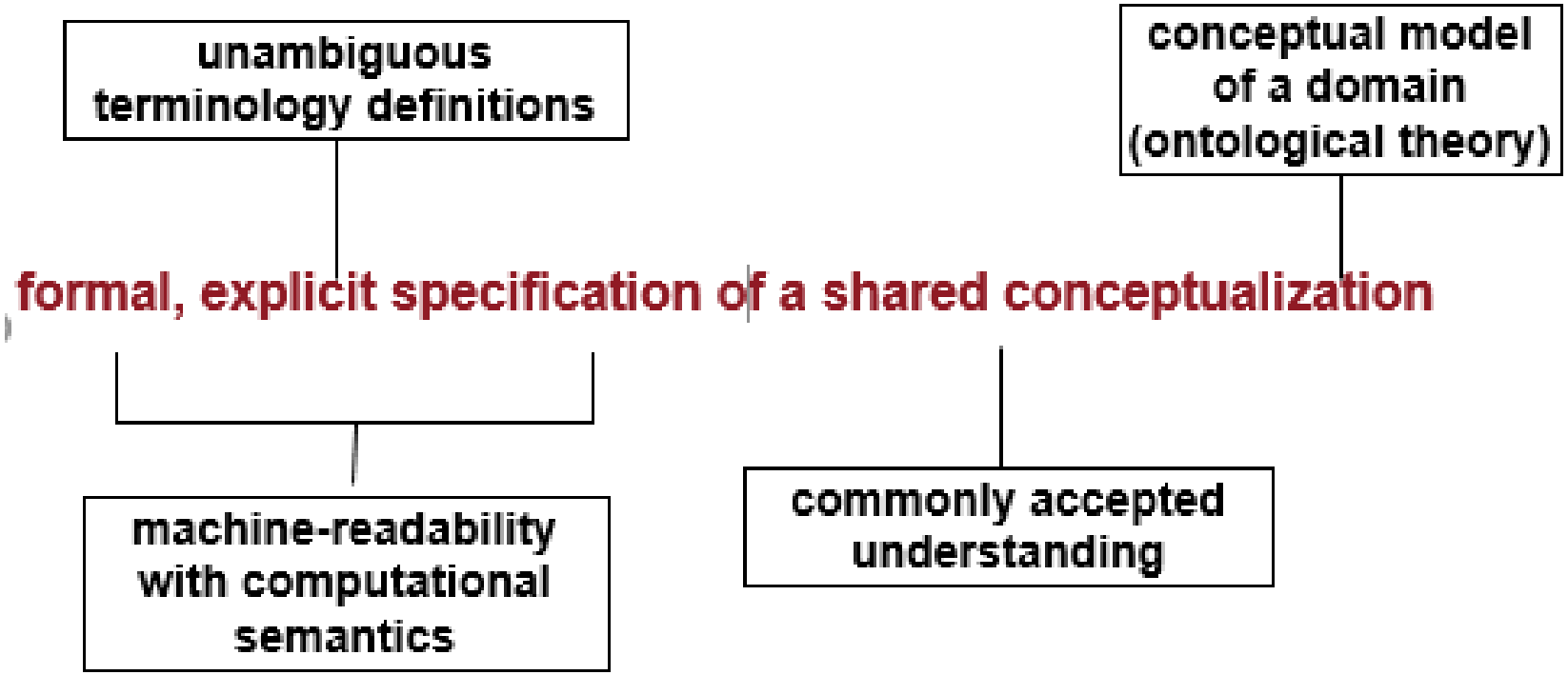
Ontology definition

"An ontology is an explicit specification of a conceptualization" Gruber, "Toward principles for the design of ontologies used for knowledge sharing?"

Ontologies are the modelling foundations to Semantic Web.

- They provide the well-defined meaning for information.

Ontology definition



An ontology is:

- A conceptualization

  An ontology is a model of the most relevant concepts of a phenomenon from the real world
- Explicit

  The model explicitly states the type of the concepts, the relationships between them and the constraints on their use
- Formal

  The ontology has to be machine readable (the use of the natural language is excluded)
- Shared

  The knowledge contained in the ontology is consensual, i.e. it has been accepted by a group of people.