

Smart Spaces

Chapter 3:

Semantic Technology: Knowledge representation and reasoning

▶ 1 Dmitry G. Korzun, 2011-2021

Outline

- § 1. Web Evolution
- § 2. Resource Description Framework (RDF)
- § 3. Ontology representation model and OWL

▶ 2 Dmitry G. Korzun, 2011-2021

§1. Web Evolution

From <http://lifeboat.com/ex/web.3.0>

▶ 3 Dmitry G. Korzun, 2011-2021

The Web is Database

Beyond the Limits of Keyword Search

From Nova Spivack, 2009

▶ 4 Dmitry G. Korzun, 2011-2021

Web 1.0

- ▶ WWW: World Wide Web
- ▶ Web 1.0: Collection of multimedia human-readable material
- ▶ HTML: HyperText Markup Language
- ▶ HTTP: HyperText Transfer Protocol
- ▶ Web site: its users are passive readers

▶ 5 Dmitry G. Korzun, 2011-2021

Web 2.0

- ▶ Applications with information sharing, interoperability, user-centered design, and collaboration
 - ▶ Web services, eXtensible Markup Language (XML), and Service Oriented Architecture (SOA)
 - ▶ Collaborative self-publishing (blogs, wikis, ...)
 - ▶ Users interact and collaborate with each other in a social media dialogue as content creators in a virtual community
- ▶ A giant web of resources: Intelligence is in the connections

▶ 6 Dmitry G. Korzun, 2011-2021

Semantic Web

- ▶ Provision of machine readable information in order to allow automating many tasks that the web is currently used for manually
- ▶ **Semantic Web**
 - ▶ Web of data that can be processed directly and indirectly by machines
- ▶ Social graph connects people; Semantic graph connects everything
- ▶ Tim Berners-Lee
 - ▶ World Wide Web Consortium (W3C)

▶ 7

Dmitry G. Korzun, 2011-2021

The dream

- ▶ Each application in context tries to determine the meaning of the text or other data
- ▶ Then it creates connections for the user
- ▶ Users share and utilize computerized applications simultaneously in order to cross reference the time frame of activities with documentation and/or data
- ▶ The availability of machine-readable metadata would enable automated agents and other software to access the Web more intelligently
- ▶ The agents would be able to perform tasks automatically and locate related information on behalf of the user

▶ 8

Dmitry G. Korzun, 2011-2021

Examples

- ▶ Semantic Publishing
 - ▶ real-time publishing and sharing of scientific data on the Internet
- ▶ Semantic Blogging
 - ▶ changing the way blogs are read (search, ranking, clustering, aggregation, ...)
- ▶ Web 3.0
 - ▶ Covers semantic web (or equal)

▶ 9

Dmitry G. Korzun, 2011-2021

Smart Data

- ▶ Smart Data = Data that carries whatever is needed to make use of it
- ▶ The smartness moves into the data itself rather than being hard-coded into the software

▶ 10

Dmitry G. Korzun, 2011-2021

Five Approaches to Semantics

- ▶ Tagging (Flickr, Wikipedia)
- ▶ Statistics (Google)
- ▶ Linguistics
- ▶ Semantic Web
- ▶ Artificial Intelligence (WolframAlpha)

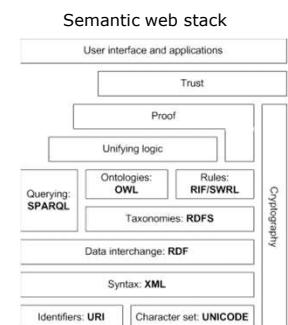
▶ 11

Dmitry G. Korzun, 2011-2021

W3C and Semantic Web

Methods/tools for formal description of concepts, terms, and relationships within a given knowledge domain

- ▶ Resource Description
- ▶ Data interchange formats
- ▶ Semantic rules
- ▶ ...

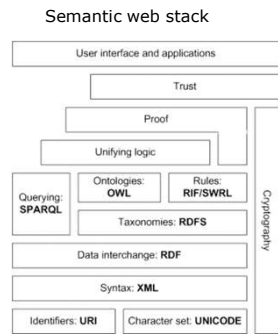


▶ 12

Dmitry G. Korzun, 2011-2021

Hypertext Web technologies

- ▶ **URI (Unified Resource Identifier)**
 - ▶ unique identification of resources
- ▶ **Unicode**
 - ▶ texts in many languages
- ▶ **XML**
 - ▶ documents composed of structured data

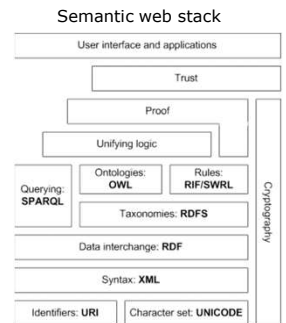


▶ 13

Dmitry G. Kozun, 2011-2021

Resources: One giant global graph

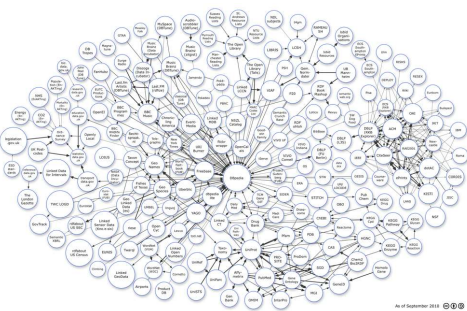
- ▶ **Resource Description Framework (RDF)**
 - ▶ Information is represented as a set of triples
- ▶ **RDF triple store**
- ▶ **One giant graph describes all resources of the web**



▶ 14

Dmitry G. Kozun, 2011-2021

http://richard.cyganiak.de/2007/10/lod/lod-datasets_2010-09-22.html



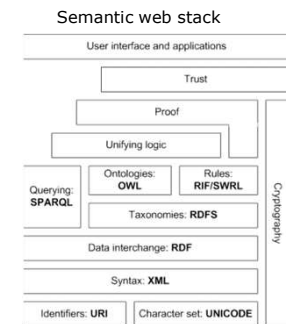
W3C SWEOW Linking Open Data project
The 203 data sets
• consist of over 25 billion RDF triples
• interlinked by around 395 million RDF links

▶ 15

Dmitry G. Kozun, 2011-2021

Ontological approach

- ▶ **Ontology describes shared vocabulary for modeling a particular domain (thesaurus, taxonomy)**
- ▶ **Ontology structures a part of the graph needed at the moment**

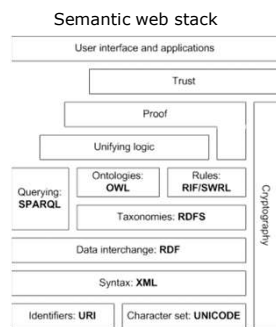


▶ 16

Dmitry G. Kozun, 2011-2021

Representation languages

- ▶ **RDF-Schema: RDFS**
 - ▶ basic vocabulary for RDF
 - ▶ hierarchies of classes and properties
- ▶ **Web ontology language: OWL**
 - ▶ advanced constructions to describe semantics of RDF statements
 - ▶ cardinality, restrictions of values, transitivity, ...
 - ▶ based on description logic
 - ▶ reasoning

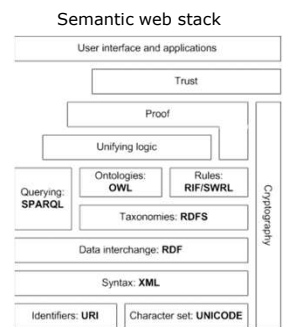


▶ 17

Dmitry G. Kozun, 2011-2021

Querying languages

- ▶ **Querying language is necessary to retrieve information for applications**
- ▶ **SPARQL is an RDF query language**
- ▶ **Simpler language: WQL (WilburQL) by Nokia (in original Smart-M3)**

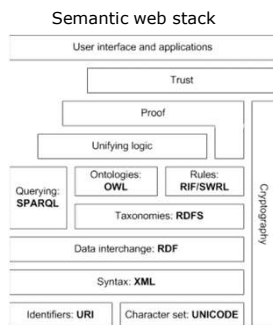


▶ 18

Dmitry G. Kozun, 2011-2021

Semi-structured information

- ▶ Common ontology is similar to standardization
 - ▶ Difference: possibility of leaving information only partially defined
- ▶ The web is not the best platform for sharing the rapidly changing, dynamic local information about the immediate environment of a device

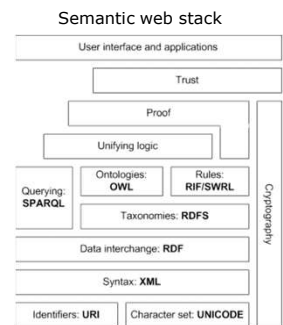


▶ 19

Dmitry G. Korzun, 2011-2021

Unrealized Technologies

- ▶ Top layers contain technologies that are not yet standardized or contain just ideas
- ▶ RIF/SWRL
 - ▶ Rule Interchange Format
 - ▶ Semantic Web Rule Language
 - ▶ describing relations that cannot be directly described using OWL
- ▶ Cryptography, Trust
- ▶ User interface
 - ▶ enable humans to use semantic web applications



▶ 20

Dmitry G. Korzun, 2011-2021

Literature

- ▶ Tim Berners-Lee, James Hendler, and Ora Lassila. The Semantic Web. Scientific American Magazine, 2001
- ▶ Web Evolution by Nova Spivack (2009) <http://www.slideshare.net/novaspivack/web-evolution-nova-spivack-twine>

▶ 21

Dmitry G. Korzun, 2011-2021

§2. Resource Description Framework (RDF)

- ▶ Knowledge representation model
 1. Data structures (memory cells, pointers): no a priori semantics
 2. **Logical:** formal semantics in terms of relations among objects
- ▶ Given a problem domain
 - ▶ Shared vocabulary
 - ▶ Entities: objects and their properties
 - ▶ Relations

▶ 22

Dmitry G. Korzun, 2011-2021

Ideas from Theory

- ▶ Artificial intelligence
- ▶ Logical facts as n-tuples
 - ▶ Predicates in descriptive logic
$$P(X_1, X_2, \dots, X_n)$$
- ▶ n-ary relations

$$(X_1, X_2, \dots, X_n)$$
- ▶ Simpler? n=2

$$P(X_1, X_2)$$

▶ 23

Dmitry G. Korzun, 2011-2021

Ideas from Practice

- ▶ Web: pages and links
 - ▶ $X_1 \rightarrow X_2$
 - ▶ ... or $P(X_1, X_2)$
 - ▶ Triple:

$$(X_1, P, X_2)$$

Subject – Predicate – Object
- Subject: <http://dig.csail.mit.edu/data#DIG>
 Predicate: <http://xmlns.com/foaf/0.1/member>
 Object: <http://www.w3.org/People/Berners-Lee/card#i>
- Subject: <http://data.linkedmdb.org/resource/film/77>
 Predicate: <http://www.w3.org/2002/07/owl#sameAs>
 Object: http://dbpedia.org/resource/Pulp_Fiction_%28film%29

▶ 24

Dmitry G. Korzun, 2011-2021

RDF triple

- ▶ Encoding model for knowledge:
Subject – Predicate – Object
- ▶ URI – URI – URI
- ▶ URI – URI – String

- ▶ A, B are people:
 - ▶ A knows B
- ▶ C is a person, D is a book:
 - ▶ C is the author of D

▶ 25

Dmitry G. Korzun, 2011-2021

RDF store

- ▶ Description of a domain:
a list of many triples
- ▶ Knowledge base
 - ▶ Actual data (facts)
 - ▶ Deduction rules

▶ 26

Dmitry G. Korzun, 2011-2021

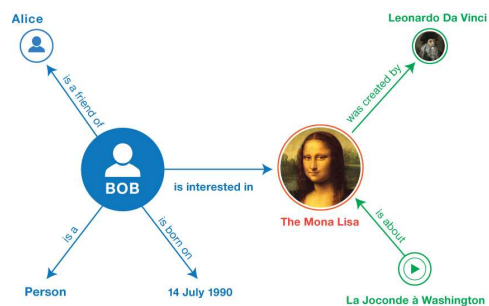
RDF graph

- ▶ Composite descriptions:
chains of knowledge
- ▶ Some objects, predicates, and subjects concise

▶ 27

Dmitry G. Korzun, 2011-2021

Example



<http://milicicvuk.com/blog/2014/09/06/what-semantic-web-can-learn-from-javascript/>

▶ 28

Dmitry G. Korzun, 2011-2021

Example

<Bob> <is a> <person>.
<Bob> <is a friend of> <Alice>.
<Bob> <is born on> <the 4th of July 1990>.
<Bob> <is interested in> <the Mona Lisa>.
<the Mona Lisa> <was created by> <Leonardo da Vinci>.
<the video 'La Joconde à Washington'> <is about> <the Mona Lisa>

▶ 29

Dmitry G. Korzun, 2011-2021

Linked data

- ▶ Uniform Resource Identifiers (URI)
 - ▶ Initially: addresses of documents located on the Web
 - ▶ Generally: identification of any entity that exists in the world
- ▶ Data sets and their namespaces
 - ▶ Hierarchy and ID readability
- ▶ HyperText Transfer Protocol (HTTP)
 - ▶ Dereferencing a URI
 - ▶ Retrieving resources serialized as a stream of bytes
 - ▶ Retrieving descriptions of entities that cannot be sent across the network

▶ 30

Dmitry G. Korzun, 2011-2021

Uniform Resource Identifier

- ▶ URL + semantics
`<scheme name> : <hierarchical part> [? <query>] [# <fragment>]`
- ▶ Scheme name
"ftp", "mailto" or "file "
- ▶ Hierarchical part
cs.karelia.ru
- ▶ Query
type=animal
- ▶ Fragment
http://en.wikipedia.org/wiki/URI#Examples_of_URI_references

▶ 31

Dmitry G. Korzun, 2011-2021

Knowledge Encoding

- ▶ Structured formalism (graph)
- ▶ Natural for web resources (links)
- ▶ Understandable to experts of the problem domain
- ▶ Understandable to software agents searching for information

▶ 32

Dmitry G. Korzun, 2011-2021

§3. Ontology representation model and OWL

Terminological clarifications

- ▶ Conceptualization
 - ▶ A system of categories
 - ▶ Independent of specific language
- ▶ Engineering artifact
 - ▶ Specific vocabulary to describe a certain reality
 - ▶ Assumptions for intended meaning of the vocabulary words

▶ 33

Dmitry G. Korzun, 2011-2021

Basic Idea

- ▶ Extending RDF
 - ▶ Triples are natural for web resources
- ▶ OO-approach
 - ▶ Classes, objects, properties, hierarchies
 - ▶ OO class: operational properties (methods)
 - ▶ Ontology: structural properties
- ▶ OO-model != Ontology model

▶ 34

Dmitry G. Korzun, 2011-2021

Ontology

A formal explicit specialization of a conceptualization

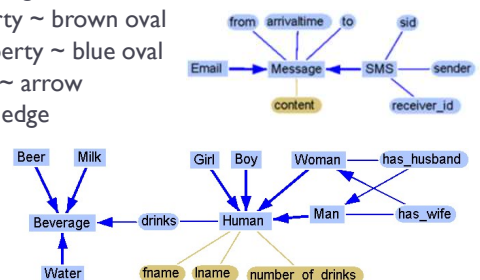
- ▶ Classes ~ Concepts
 - ▶ A class ~ a set of individuals (instances)
- ▶ Properties of each concept
 - ▶ Features, attributes, roles
- ▶ Restrictions on properties

▶ 35

Dmitry G. Korzun, 2011-2021

Example

- ▶ Volunteer?
 - ▶ class ~ rectangle
 - ▶ data property ~ brown oval
 - ▶ object property ~ blue oval
 - ▶ subclass-of ~ arrow
 - ▶ property ~ edge

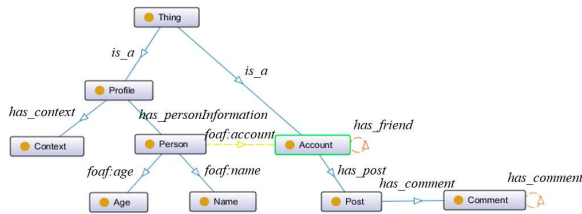


▶ 36

Dmitry G. Korzun, 2011-2021

Example

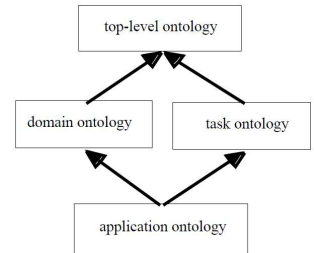
- ▶ SmartScribo system
 - ▶ User representation model and related semantics



▶ 37 Dmitry G. Korzun, 2011-2021

Granularity

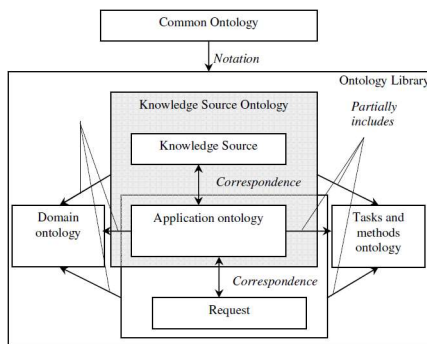
- ▶ Accuracy or generality levels
- ▶ Top-level: general concepts
 - ▶ space, time, event, human, ...
- ▶ Domain vocabularies
 - ▶ medicine, automobiles
- ▶ Task or activity vocabularies
 - ▶ diagnosing, messaging
- ▶ Application: specialization of both a domain ontology and a task ontology (ontology library)
 - ▶ Blogging scenarios in a smart car



▶ 38 Dmitry G. Korzun, 2011-2021

More advanced ...

Ontology modeling



▶ 39 Dmitry G. Korzun, 2011-2021

Knowledge Base

- ▶ Ontology describes state-independent information (the logical component)
 - ▶ Concept model (with particular syntax)
 - ▶ Ontology class graph
- ▶ “Core knowledge base” contains state-dependent information (the actual data component)
 - ▶ All individual instances
 - ▶ Ontology instance graph (~ RDF graph)
- ▶ Inference mechanism

▶ 40 Dmitry G. Korzun, 2011-2021

OWL to RDF

OWL is a particular language to write ontologies

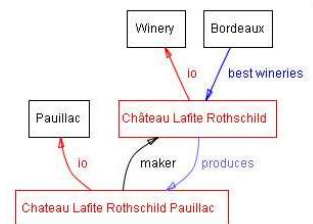
- ▶ OWL ontology models a structure of a given problem domain
 - ▶ Ontology class graph
- ▶ High-level knowledge encoding by the ontology (individuals and their properties)
 - ▶ Ontology instance graph (~ RDF graph)
- ▶ At the lowest level, RDF triples are used
 - ▶ RDF store

▶ 41 Dmitry G. Korzun, 2011-2021

Class + Instance graph

Wine domain

- ▶ Classes are in black
- ▶ Individuals are in red
- ▶ Links
 - ▶ instance-of (io)
 - ▶ subclass-of
 - ▶ data properties
 - ▶ sugar level
 - ▶ flavor
 - ▶ ...
 - ▶ object properties
 - ▶ maker
 - ▶ produces



▶ 42 Dmitry G. Korzun, 2011-2021

Properties

- ▶ Data properties
 - ▶ Datatype values
- ▶ Object properties
 - ▶ Relationships to other individuals
- ▶ Restrictions
 - ▶ Cardinality
 - ▶ Value type (string, number, Boolean, enumerated, instance)
 - ▶ Range (allowed classes for type "instance")
 - ▶ Domain (classes a property is attached to)
 - ▶ ...

▶ 43

Dmitry G. Korzun, 2011-2021

Development of a Knowledge Base

The basic scheme:

- ▶ Classes in the ontology
- ▶ Class structure as a subclass-superclass hierarchy (composition of several tree-like structures, multiple inheritance)
- ▶ Properties and allowed values
- ▶ Individual instances (individuals) and values of their properties

▶ 44

Dmitry G. Korzun, 2011-2021

Remarks

- ▶ Reusing existing ontologies
- ▶ Tools for constructing ontologies
 - ▶ Protégé
 - ▶ ...

▶ 45

Dmitry G. Korzun, 2011-2021

Часть 3 проекта

Онтологическое моделирование

- ▶ Интеллектуальное пространство как база знаний (БЗ).
- ▶ Какая информация становится доступной всем (описать в сценариях).
- ▶ Онтология как логическая структура БЗ (граф онтологических классов).
- ▶ Семантическая сеть как фактическая информация, хранимая в БЗ в некоторый момент времени (граф онтологических индивидов).

▶ 46

Dmitry G. Korzun, 2011-2021

Literature

- ▶ C.Bizer, T.Heath, T.Berners-Lee. Linked Data - The Story So Far (2009)
- ▶ N.Guarino. The Ontological Level- Revisiting 30 Years of Knowledge Representation (2009)
- ▶ N.F.Noy, D.L. McGuinness. Ontology Development 101: A Guide to Creating Your First Ontology
- ▶ Кашевник А.М., Корзун Д.Ж., Баландин С.И., Пономарев А.В. Разработка рекомендующих систем на основе интеллектуальных пространств. Уч.пос. Изд-во ПетрГУ, 2015. <http://elibrary.karelia.ru/book.shtml?id=22842>

▶ 47

Dmitry G. Korzun, 2011-2021