

GFZ

POTSDAM

INFORMATION SYSTEMS AND DATA CENTER

Global Earth Science Data



IUGONET
Workshop



NiPR
National Institute of Polar Research



2015年8月17-19日

Current status of collaboration between the IUGONET and ESPAS projects

Bernd Ritschel (GFZ/京都大学),
Toshihiko Iyemori (京都大学),
Yukinobu Koyama (国立情報学研究所),
Guenther Neher (FHP),
Christoph Seelus (GFZ)



MoU between ESPAS and IOGONET

Mike Hapgood
RAL, STFC
Oxford, England



Toshihiko Iyemori
京都大学, WDC
Kyoto, Japan

Major Objectives and Scope

The aim of the cooperation is to promote and establish a research community to build the infrastructures to solve the global data issues.

1. Activities

The primary activities are agreed as follows:

- Exchange of information including data 0 X
- Exchange of researchers 0
- Promotion of cooperative projects in a common field of interest 0

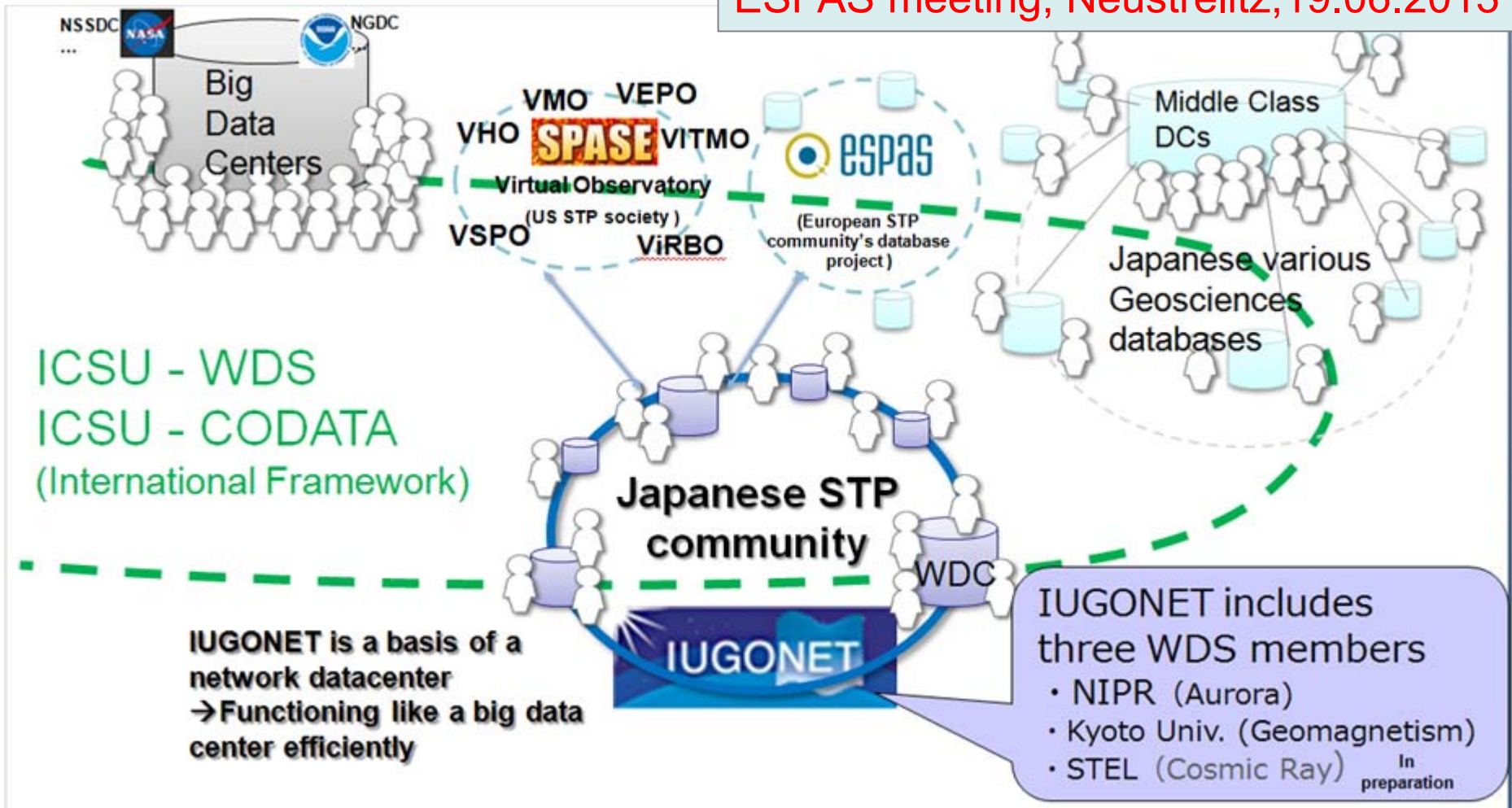
2. Cooperation field

The cooperation will focus on the following studies, and personnel of the other field of interest by either institute may be invited to participate in the joint work independently from other institutions with mutual agreement.

- Make global geophysical data accessible for other science domains X
- Make network of global observation data for integrated approach with the same metadata vocabulary 0 X
- Enhance usage of the common observational infrastructure 0 X
- Promote cooperation in the area of studies on solar, heliospherical, solar-terrestrial and geophysical activity 0
- Use the e-infrastructures for education and for capacity building 0

Future direction: To establish a close collaboration among various data systems

Source: Presentation from T. Iyemori at ESPAS meeting, Neustrelitz, 19.06.2013



Japanese e-science project

Project

- Inter-university Upper atmosphere Global Observation NETwork*

Objectives

- creation of a Metadata DataBase of ground-based observations of the upper atmosphere

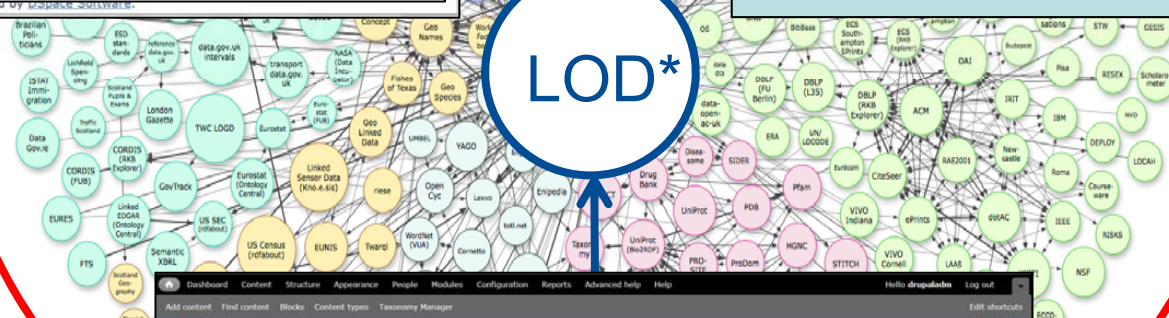
www.iugonet.org

ESPAS EU 7th framework project

Objectives

- Platform to integrate data from thermosphere, ionosphere, plasmasphere & magnetosphere
- Provide unique access to scientific near-earth space data
- Offer Web-based applications and services

www.espas-fp7.eu



Domain Ontology Merging

- FOAF
- BIBO
- GeoNames
- dbPedia
- SPASE

Terminological Ontology Merging

- GCMD Keywords
- GEMET
- SPASE
- PDS
- IVOA

Information model & resources

- ISDC ontology network
- Keyword vocabularies
- GFZ RDF+RDB data, LOD data

Technical implementation

- Drupal CMS, Virtuoso Triple Store
- Open Semantic Framework OSF

GFZ ISDC Drupal 7 proof of concept

Resources

- Geodetic data
- Magnetic field data
- Atmosphere data

*Linked open Data
<http://lod-cloud.net/>

Metadata, Data Models, Frameworks

Differences

- Metadata
 - SPASE (IUGONET)
 - ISO/OGC (ESPAS)
 - NASA DIF (ISDC)
- Metadata models
 - SPASE version 2.2.2
 - ISO/OGC 19xxx
 - ISDC ontology version 1.4
- Frameworks/Apps
 - DSpace
 - D-Net
 - OSF software stack (Drupal/Virtuoso/Solr/...)

Commons

- Data (scientific domain related)
- Metadata & metadata model entities
 - Data (granuals/products)
 - Catalog (classification)
 - Instruments
 - Platforms (observatories)
 - Persons and Institutions
 - Projects and Phenomena
- Vocabulary entities
 - Classifications (keywords)
 - Thesauri (keywords+links)

GFZ ISDC Semantic Web Project*



GFZ Department 1.1

GFZ ISDC

Github repositories

ISDC Drupal 7 Open Semantic Framework



We have a separate Virtual Machine for the Open Semantic Framework stack, including Drupal 7. The stack includes, amongst others:

- Apache Solr
- Open Link Virtuoso
- OWL API
- GATE
- PHP/Java Bridge
- Memcached

There is a [live](#) version and a [Git repository](#) (as always with further information inside the README) on-hand.

ISDC Drupal 7 proof of concept

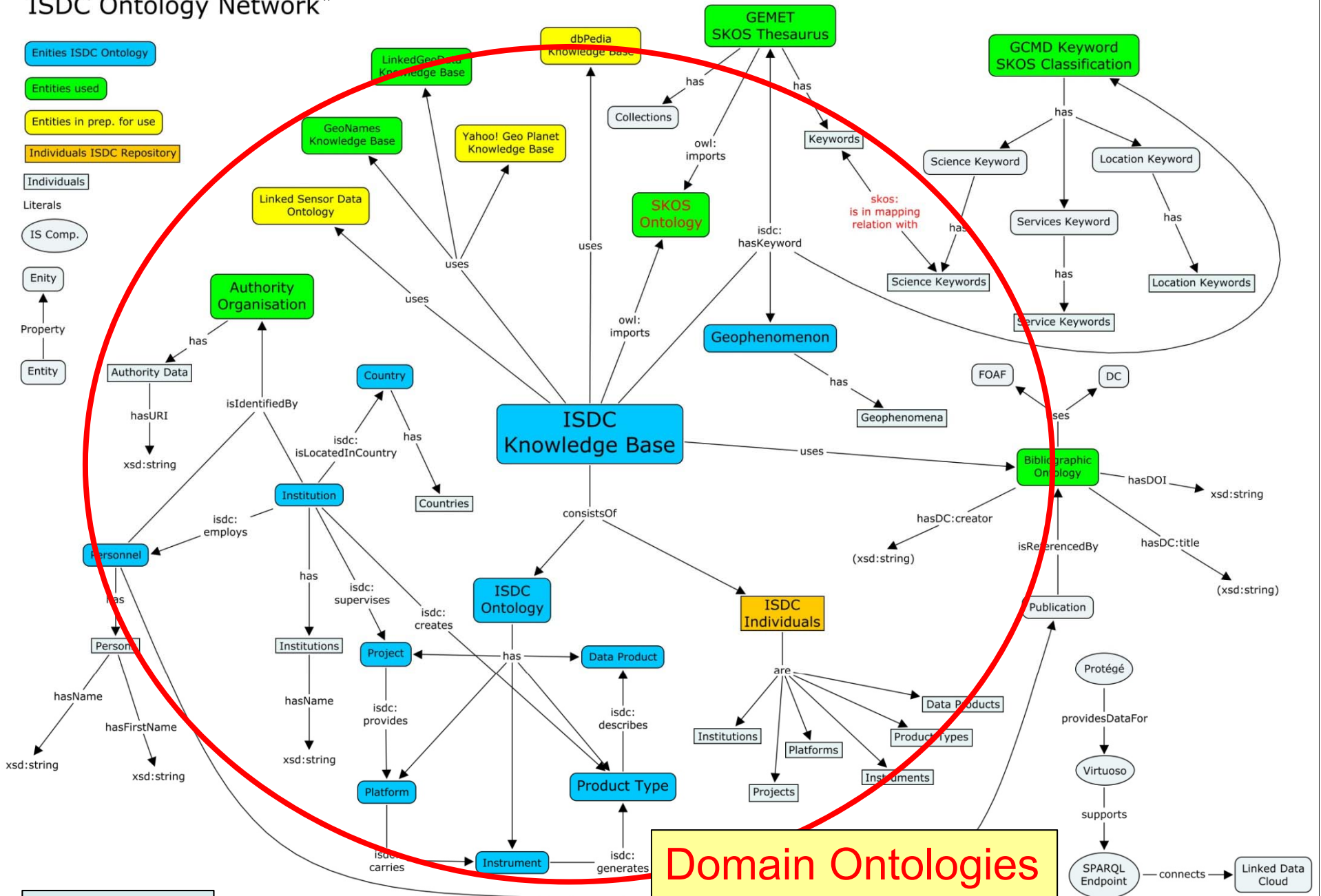


There is a [live](#) version, a [Git repository](#) and a [Github Wiki](#) available.

On the main VM a [clone of the ISDC Drupal proof of concept](#) is provided.

On the test VM another [clone of the ISDC Drupal proof of concept](#) is provided.

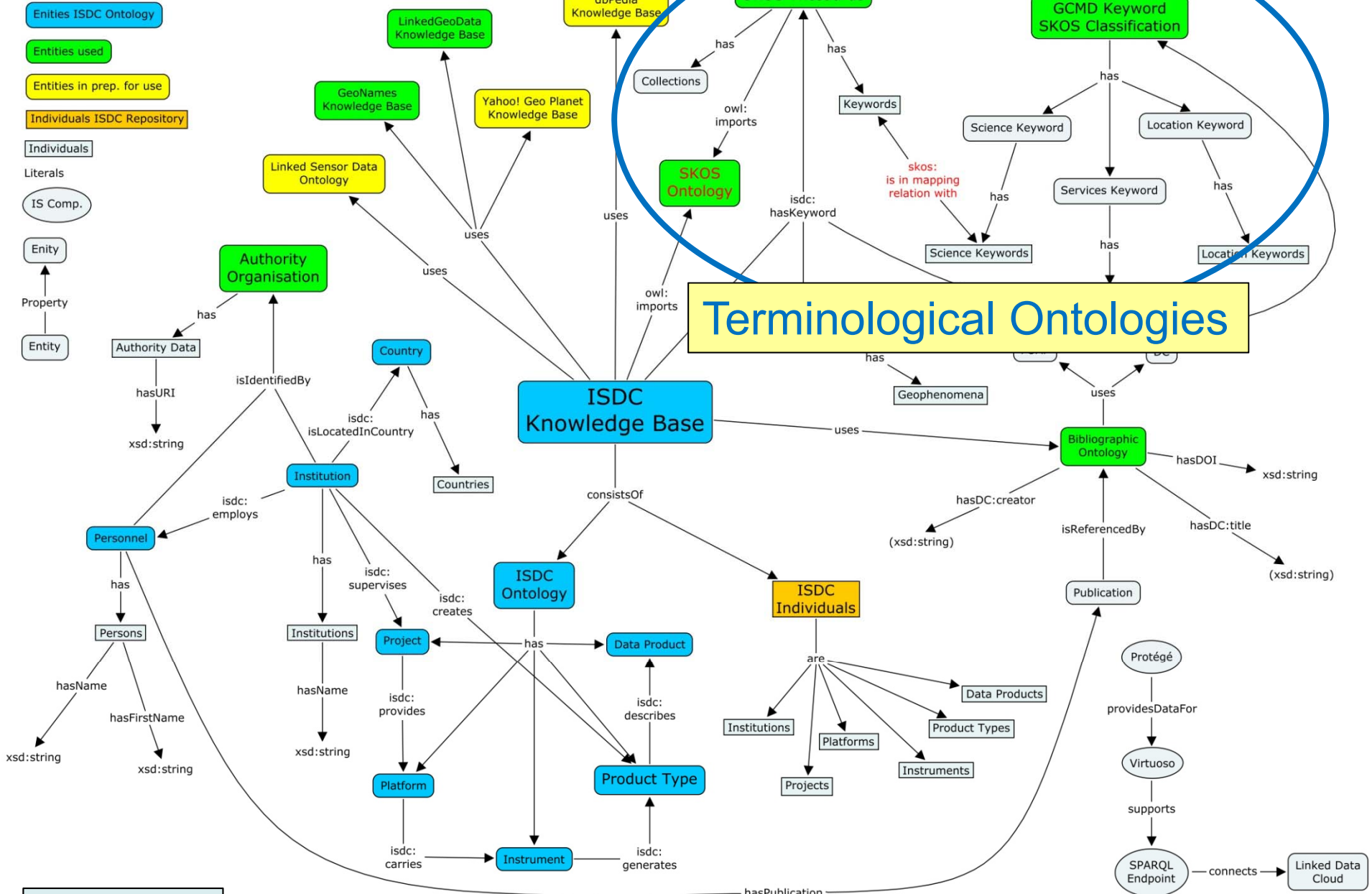
ISDC Ontology Network*



*December 2011

Domain Ontologies

ISDC Ontology Network*



*GCMD Science Keywords

Global Change Master Directory (GCMD) Science Keywords

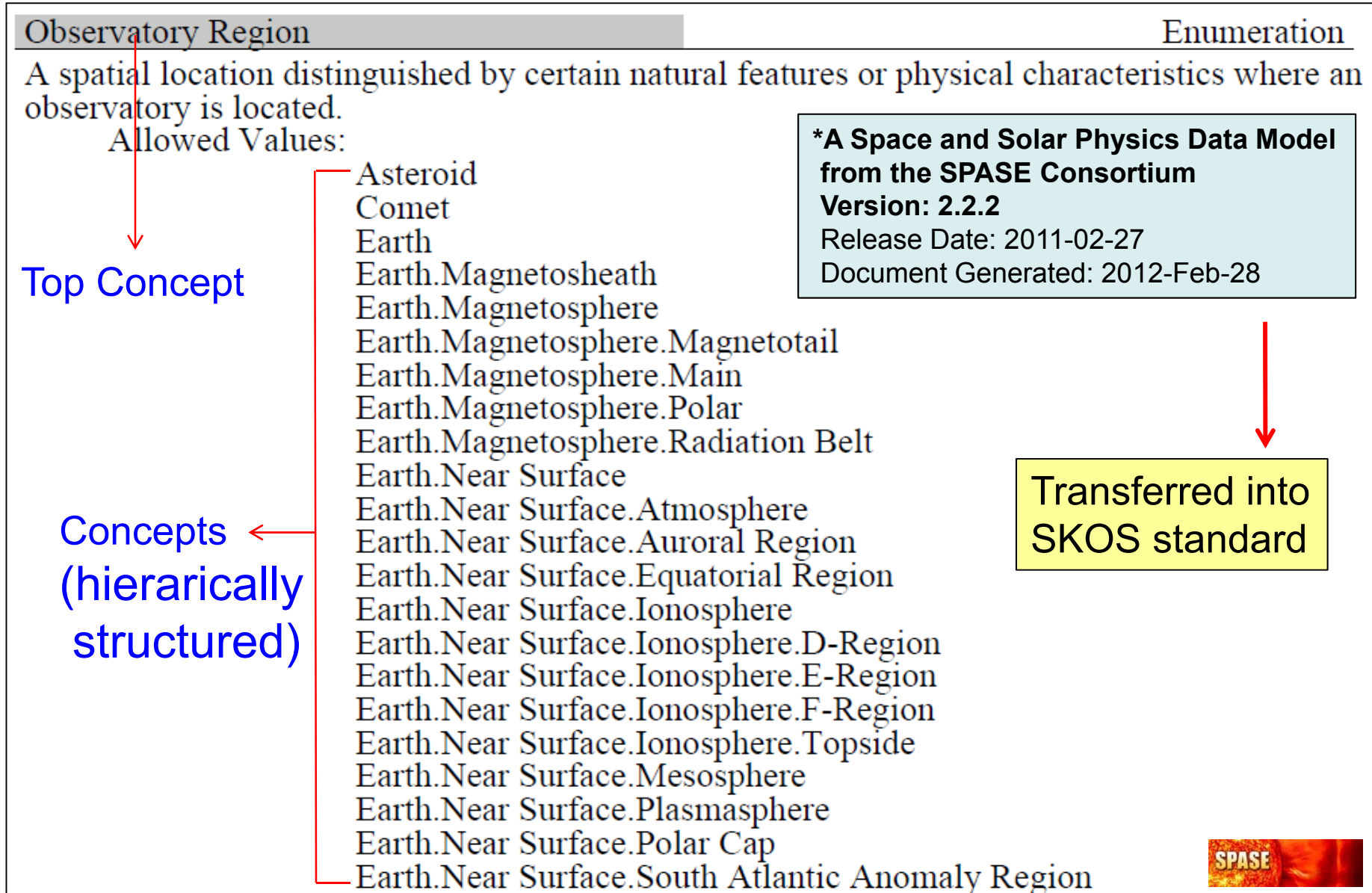
TOPIC > TERM > VARIABLE_LEVEL_1 > VARIABLE_LEVEL_2 > VARIABLE_LEVEL_3

Agriculture > Agricultural Aquatic Sciences > Aquaculture
Agriculture > Agricultural Aquatic Sciences > Fisheries
Agriculture > Agricultural Chemicals > Fertilizers
Agriculture > Agricultural Chemicals > Pesticides
Agriculture > Agricultural Engineering > Agricultural Equipment
Agriculture > Agricultural Engineering > Farm Structures
Agriculture > Agricultural Plant Science > Crop/Plant Yields
Agriculture > Agricultural Plant Science > Cropping Systems
Agriculture > Agricultural Plant Science > Irrigation
Agriculture > Agricultural Plant Science > Plant Breeding and Genetics
Agriculture > Agricultural Plant Science > Plant Diseases/Disorders/Pests
Agriculture > Agricultural Plant Science > Reclamation/Revegetation/Restoration
Agriculture > Agricultural Plant Science > Weeds, Noxious Plants Or Invasive Plants
Agriculture > Animal Commodities > Dairy Products
Agriculture > Animal Commodities > Livestock Products
Agriculture > Animal Commodities > Poultry Products

Transferred into
SKOS standard**

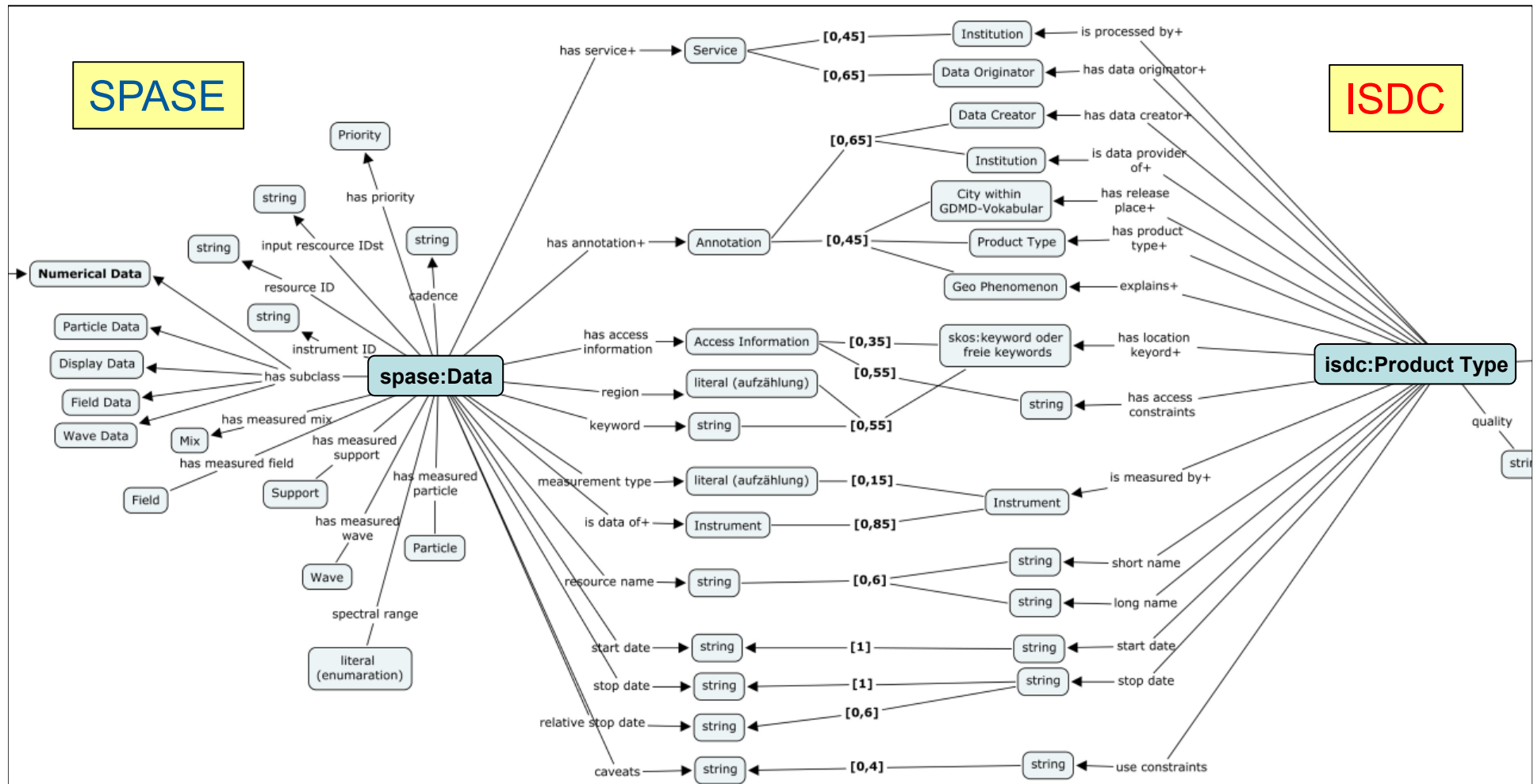
**Juliane Pilgrim,
University of Applied Sciences Potsdam
Faculty of Information Science

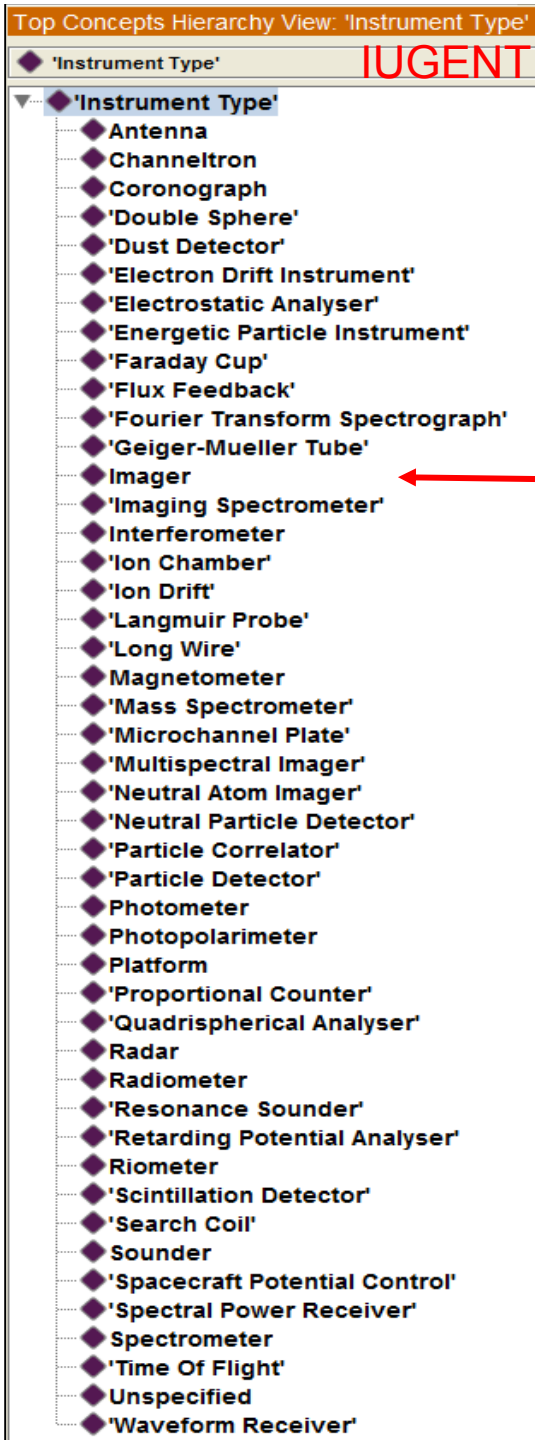
*SPASE „allowed values“ Classification



*Merging ISDC and SPASE|VMO/VHO Ontologies

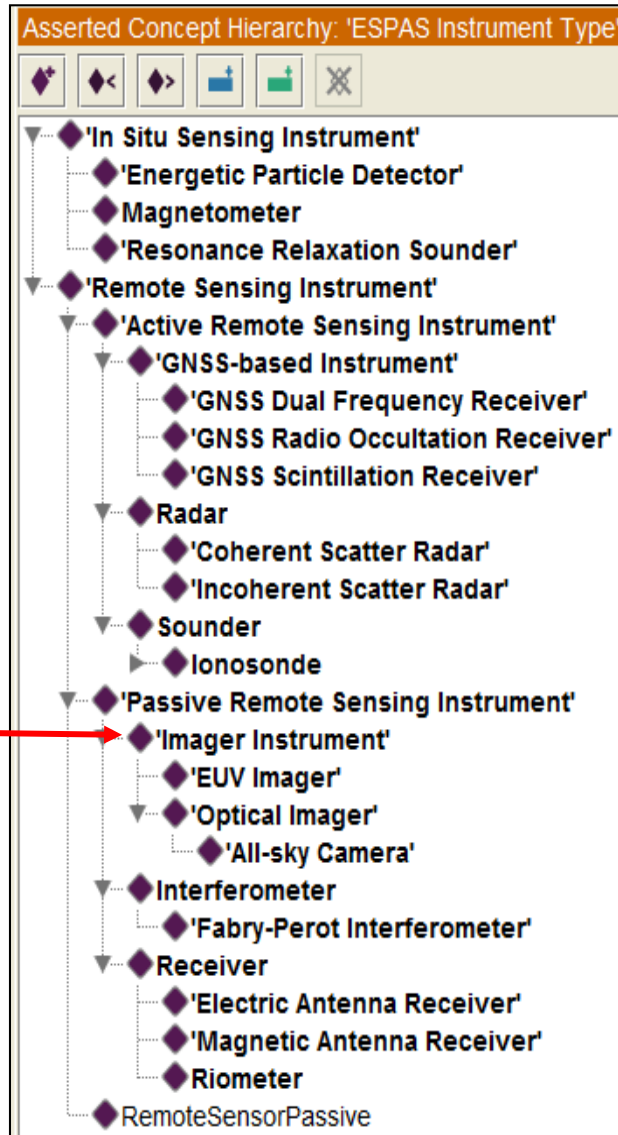
Example: spase:Data <=> isdc:Product Type





SPASE "allowed values" \Leftrightarrow ESPAS ontology
 (Collaboration project between ESPAS and IUGONET)

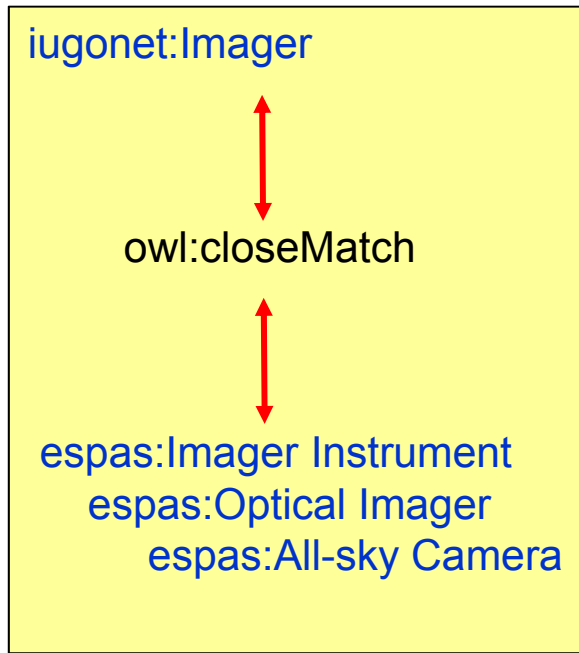
ESPAS Instrument Type



ESPAS SKOS
 resource browser:
espas.spaceweatherservices.com

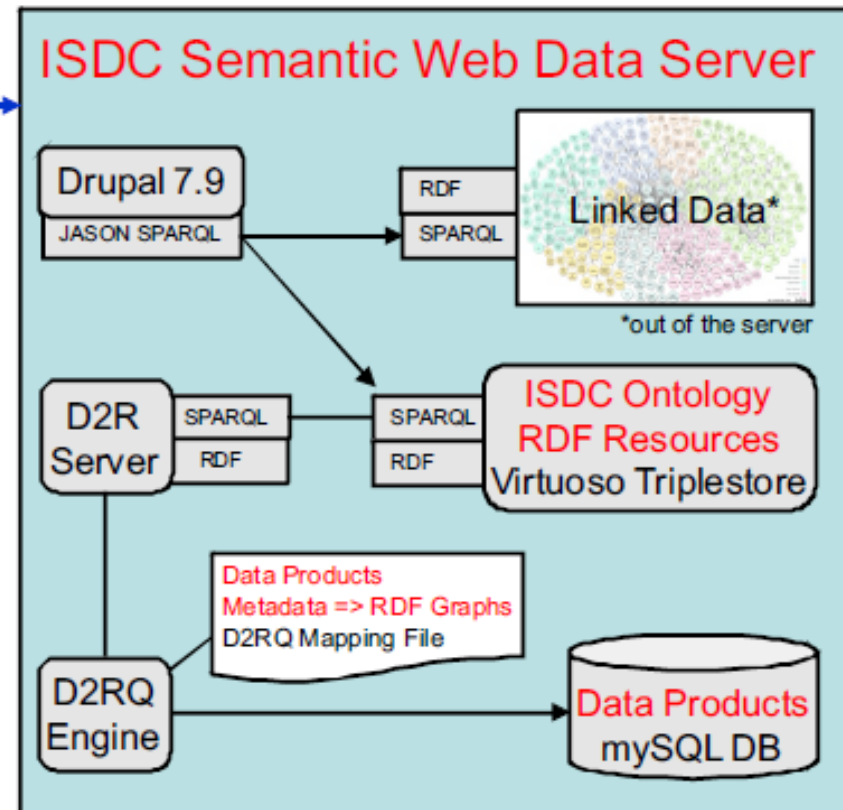
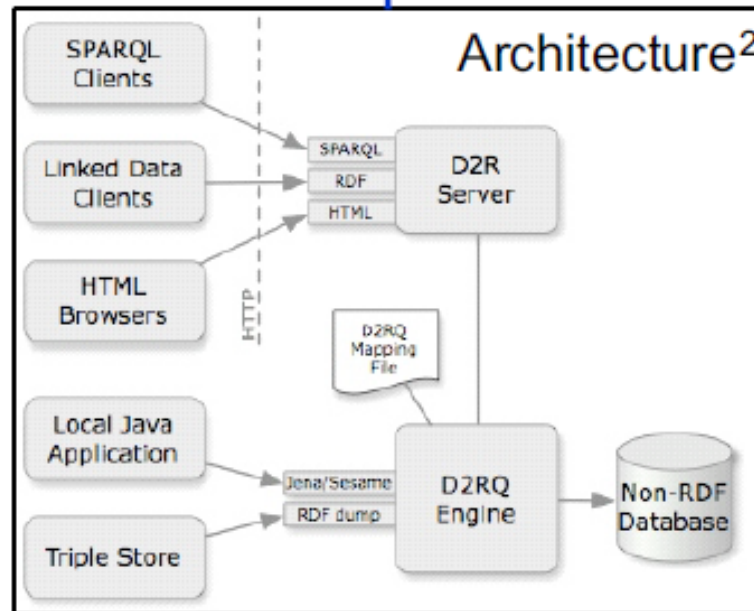


RDF + OWL + SKOS



D2RQ for mashup of RDBMS with Triple Stores

- D2RQ Platform: accessing relational databases as virtual, read only RDF graphs¹
- D2R Server: tool for publishing content of relational databases on the [Semantic Web](#)



¹<http://d2rq.org/>

²<http://d2rq.org/images/architecture.png>



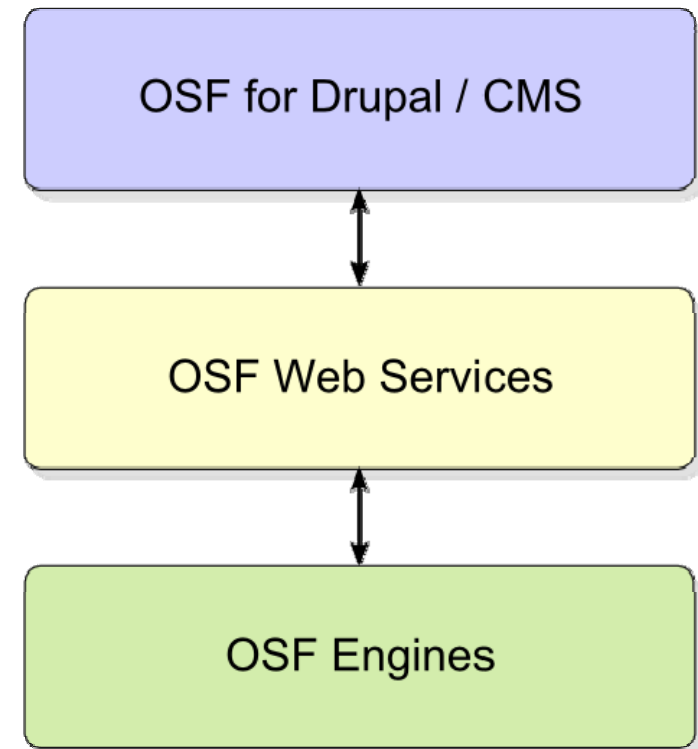
Open Semantic Framework (OSF)

Frédéric Giasson

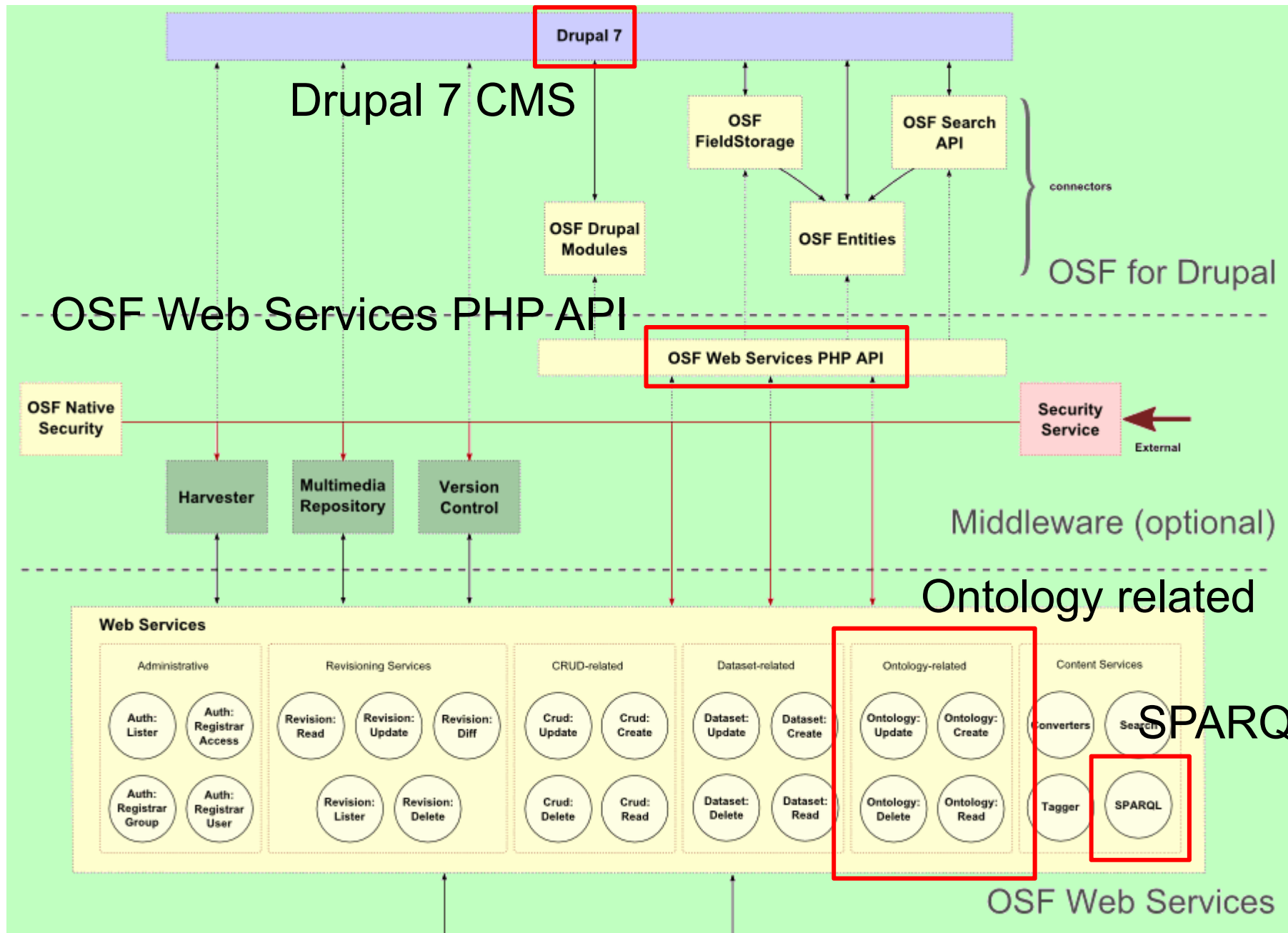


Web-oriented architecture

- Data is generally exposed/open and universally available as [linked data](#)
- SPARQL endpoints and APIs are generally [RESTful](#) in design
- The overall architecture is modular, with inherent decentralized and distributed aspects
- All display and visualization aspects are cross-browser ready and capable.



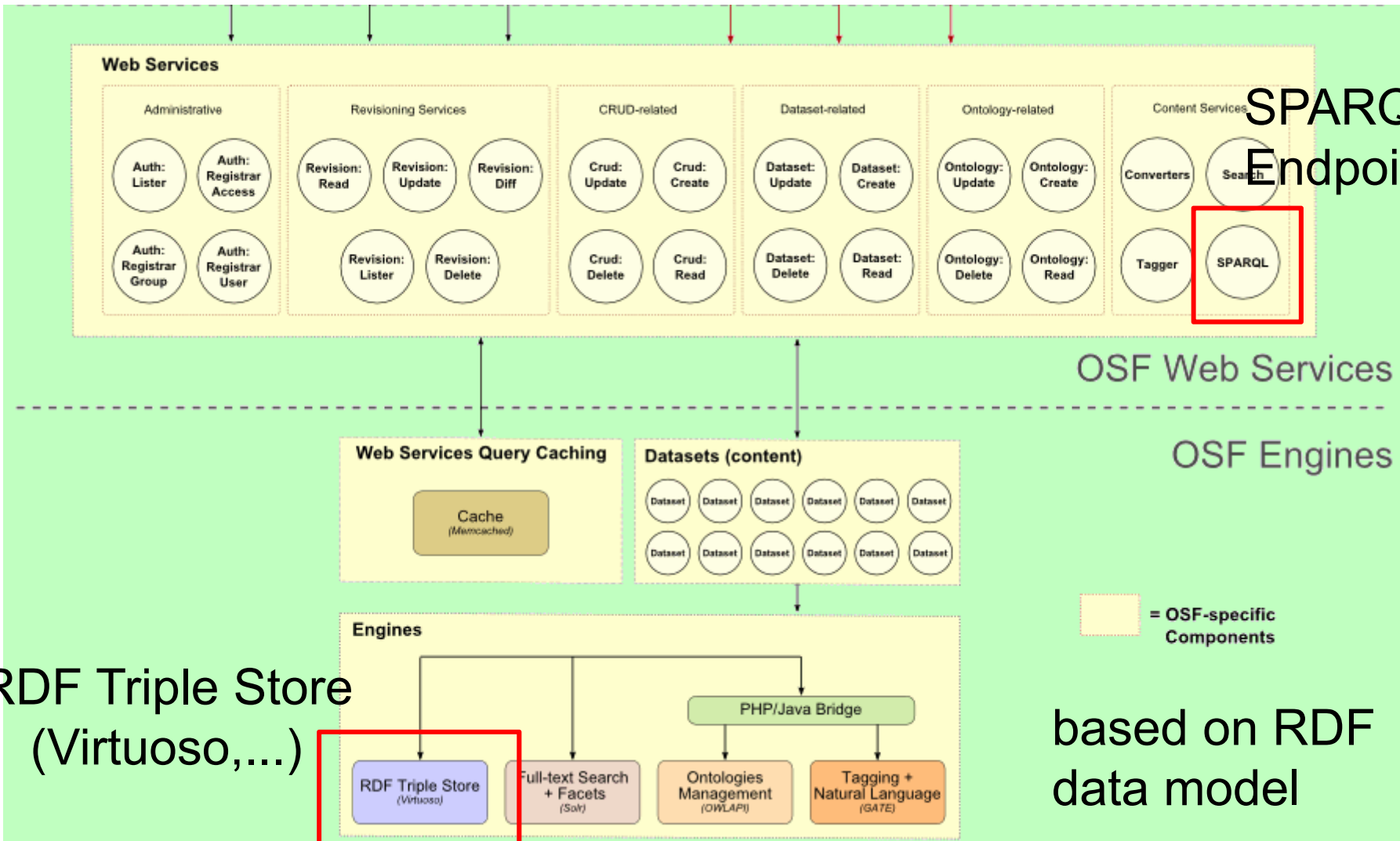
<http://opensemanticframework.org/>
<http://structureddynamics.com/index.php>
<http://fgiasson.com/blog/>



http://wiki.opensemanticframework.org/index.php/Category:Complete_Overview_of_OSF

OSF Web Services PHP API

SPARQL
Endpoint



RDF Triple Store
(Virtuoso,...)

Mashup of IUGONET, ESPAS and ISDC data server

Mapping of terminological ontologies

- IUGONET: SPASE, GCMD science keywords
- ESPAS: ESPAS ontology
- ISDC: GCMD keywords, SPASE, GEMET

and/or mapping of domain models using DCAT domain ontology

Proof of concept (keyword based mashed-up catalog queries):

- Open Semantic Framework for application
- Ontology based (mashup of particular catalog entities)
- IUGONET API: DSpace OpenSearch Query
- ESPAS API: under construction
- ISDC API: SPARQL
- Integration of context data via LOD: SPARQL

Creation of a Semantic Knowledge Network

- WDS is establishing a Knowledge Network (KN) for scientific metadata (similar to Web of Science) *

Whether looking at data, books, journals, proceedings or patents, [Web of Science](#) provides a single destination to access the most reliable, integrated, multidisciplinary research. Quality, curated content delivered alongside information on emerging trends, subject specific content and analysis tools make it easy for students, faculty, researchers, analysts, and program managers to pinpoint the most relevant research to inform their work.



- KN implementation working group in WDS
 - D2RQ Framework (e.g. using GFZ/FHP experiences)
 - Mashup of semantic/vocabulary resources (e.g. IUGONET, ESPAS, ISDC)
 - ...

Research Data Management with Semantic Web Technology

- Collaboration in education and project based research since 2010
- Education and research in information science => **data scientist**
- Current topics: **4 work packages**

Transformation of GFZ XML data to RDF for GFZ semantic Web based ISDC prototype

Merging of terminological ontologies for integration of further structured LOD data

Data mining and named entity recognition/relation extraction for integration of **unstructured** web data

Implementation of semantic Web application based on OSF stack (Drupal 7, Virtuoso, Solr, OWL API)

Future Developments and Capabilities – Internet of Things

- Big player activities (Google, Microsoft, Apple, Amazon, ..)
 - **Big Data**: collecting more and more data ("*Datenkrake*")
 - **Smart data**: mashing-up Big Data (e.g. social networks)
 - **Smart services**: maps & navigation, education, ... (e.g. Google Maps/Earth, Scholars, ...)

Artificial Intelligence + Deep Learning => **Knowledge Graph**

Google: 40 billion facts, 570 million things, 15.000 types

- (Behshad Behzadi, Director of Conversional Search, Google, SMX March 2015)

- **Smart sensors/Big Data**: data output in RDF format

Google Search Will Be Your Next Brain, Back- channel, Medium:

<https://medium.com/back-channel/google-search-will-be-your-next-brain-5207c26e4523>

resources in LOD (via triple stores & SPARQL + reasoning)

- **Internet of things** (scientific platforms and instruments)



We have to trace the developments very carefully, or even better, should become one of the active drivers.



<http://spitfire-project.eu>

<http://www.pcgames.de/Google-Firma-97880/>

Call for design of scientific vocabularies

CAWSES-II Nagoya 2013, AGU/EGU/JpGU/AOGS 2013/2014/2015

Please help to create a well agreed keyword vocabulary for space weather and climate including neighbor disciplines such as e.g. earth magnetic field or solar-terrestrial physics.

- Report the keyword vocabulary you are using
- Find common agreement in the use => standard
- Discuss and agree about concordances and mash-ups in your own domain and cross-domain



Please contact IUGONET or ESPAS:
iyemori@kugi.kyoto-u.ac.jp, rit@gfz-potsdam.de

Can we trust the Web any longer?

- internet spying
- censorship
- manipulation
- erasing
- destruction

ありがとうございます。ございました。しづもんが
ありますか。

rit@gfz-potsdam.de

ritschel@kugi.kyoto-u.ac.jp

Thank you, Questions?