**Towards an Integrated Information Space for the Environment in Europe**

September 10, 2008, Leuphana University of Lüneburg

# Semantic Network Service (SNS) Update and Future Vision


**Thomas Bandholtz**
Principal Consultant
innoQ Deutschland GmbH
thomas.bandholtz@innoq.com

**Maria Rüther**
FG II 1.2
Umweltbundesamt (DE)
maria.ruether@uba.de


## Outline

- Origins and history
- SNS Topic Map
- Services and clients today
- Semantic Data Integration (INSPIRE, SEIS/SISE)

## The problem: semantic noise

"daphnia" "direct discharger" "permit for the use of water" "deep water" "water aeration" "snow water" "terms of waste water" "agricultural effluent" "boiling water reactor" "water pollution load" "combined waste water" "waste water composition" "black water" "grey water" "river works" "water board law" "water supervision" "protection area for water regulation" "groundwater resources" "salamander" "agriculture" "tertiary purification of sewage" "sewage purification close to nature" "persistent chemical substance" "primary treatment" "flood" "chemical sewage purification" "secondary treatment of sewage" "industrial effluent" "lowering of groundwater level" "waste water purification" "bathing waters" "water protection" "waste water reclamation" "surface water" "water privilege" "underground disposal of waste water" "water rate statute" "water levy act" "hydro-isobath" "groundwater flow" "groundwater storey" "public goods" "hot water storage" "hot water heating system" "groundwater" "waste water legislation" "water quality management" "water heating" "seepage water disposal" "hot water" "saline water intrusion" "industrial water" "mineral water" "stocktaking" "municipal water management" "hydraulic and sanitary engineering" "waste water examination" "groundwater table" "void water" "seepage water treatment" "percolating water" "waste water reduction" "sewage flow" "deep sea" "sewage lagoon" "waste water statistics" "water protection policy" "well" "water quality directive" "salts" "surface water" "river water" "wastewater load" "indirect discharger" "back water" "waste water register" "river" "impregration (material)" "municipal sewage" "waste water sludge" "ordinance on parameters of noxiousness of waste water" "sludge" "harmfulness of wastewater" "aquifer" "impregnating agent" "sewer" "desalination of brackish water" "waste water decontamination" "brackish water" "Waste Water Origins Ordinance" "intertidal area" "feed water" "groundwater contour line" "ground water conservation" "soil moisture regime" "soil water" "small body of water" "waste water charge legislation" "waste water charge code" "human settlement" "stagnant water" "waste water charge fixation" "state water law" "waters (topographic)" "water sciences" "waste water charge" "environmental quality objective" "turtle" "residue" "water works" "waste water disposal" "proprietary right" "water course regulation" "sewage decontamination" "liquid manure" "industrial installation" "waste water disposal embargo" "wastewater discharge" "rinse water" "EU Water Protection Directive" "industry" "Framework Water Directive" "river filtrate" "waterfowl" "water pollution" "rhizosphere" "dump impounded water" "turbomachine" "water supply" "water pollution prevention" "raw water" "deep water" "sea water protection" "outfall" "water evaporation" "water consumption" "water board decree" "biological water testing" "sea water fish" "water analysis" "sea water desalination" "material insoluble in water" "shore belt bird" "waste water disposal embargo" "waste water disposal scheme" "sea water" "drinking water preparation regulation" "sewage treatment plant" "turbidity of water" "drinking water" "water management plant" "general planning on water resources development" "sewage disposal" "residual amount of water" "flowing waters" "water management" "wastewater quality" "tail water" "condenser" "under water coating" "planning permission" "aquatic animal" "water temperature" "water reuse" "tide" "waste water treatment plant" "physical sewage treatment" "mechanical sewage treatment" "water resources" "international convention" "electrochemical sewage treatment" "chemical sewage treatment" "rural area" "anaerobic sewage treatment" "aerobic sewage treatment" "water mike" "permit" "drinking water supply" "drainage" "drinking water regulation" "water quantity management" "water volume" "water market" "water statistics" "water level" "water sports" "reservoir" "waste water treatment" "drinking water examination" "drinking water protection area" "water shortage" "solubility in water" "drinking water quality" "water line" "water conductivity" "low water" "regulation on securing of enough water" "law on the securing of enough water" "securing enough water" "shallow water" "protected water catchment area" "Act Pertaining to Charges Levied for Discharging Waste Water into Waters" "wastewater levy" "water cooling system" "discharged water" "water cycle" "hydroelectric power plant" "water power" "water partition" "waterborne sound" "water pollution" "rainwater" "flood runoff" "drinking water treatment plant" "drinking water treatment" "pollution of waters" "sewerage system" "water contents" "vapour" "Water Hygiene Act" "water purification" "water act" "water association" "vadose water" "inland water way" "water hygiene" "water resources policy act" "biological water balance" "water hardness" "water pollution monitoring" "water sample" "water price" "water flow" "aquatic plant" "increasing water hardness" "water penny" "water surface" "surface runoff" "water utilization" "PWR-type reactor" "deep sea fishing" "inland waters" "algae bloom" "hydrologic balance" "water movement" "water bottom" "water quality model" "water deposit" "water protection directive" "water protection legislation" "water pollution control measure" "water pollution control act" "water pollution control deputy" "toilet" "pelagial" "water demand" "hydraulic construction" "water treatment" "water quality" "long distance water supply" "water catchment" "restoration of waters" "impounded water" "water act" "water company" "regulation of waters" "water content" "bilge water" "water ouzel" "water runoff" "water endangering" "regulation concerning water endangering matter" "wet-type cooling tower" "permit to exploit water" "reject water" "fresh water fish" "utilization of waters" "Framework Waste Water Administrative Regulation" "condition of water" "softening" "fresh water" "groundwater characteristics" "water fall" "administrative regulation on substances hazardous to water" "water prospecting" "water desalination" "lime water" "salt water" "black water" "drainage water" "Ground Water Ordinance" "Waste Water Ordinance" "Light water reactor" "washing water" "water sterilization" "ground water charges" "water softening" "quality of waters" "clarification basin planted with water plants" "endangering of water" "water saving" "water permeability" "cooling water" "available water supply" "steam" "water pool" "ecological assessment"

**e.g. some 500 terms about "water" in use in the information corpus**

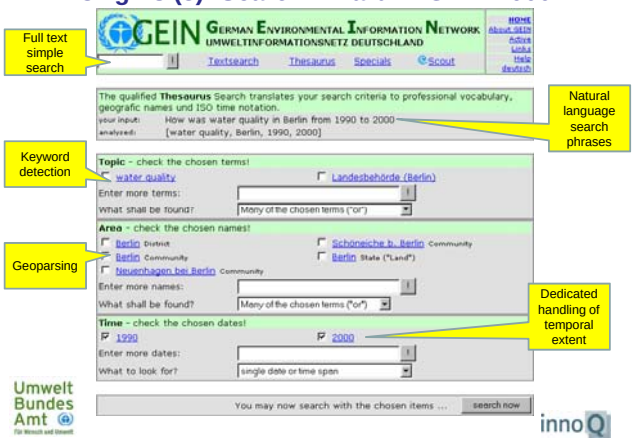
## Origins (1): UMTHES in GEIN 2000

## Origins (2): Gazetteer in GEIN 2000


Umwelt Bundes Amt

innoQ

## Origins (3): Search wizard in GEIN 2000




Umwelt Bundes Amt

innoQ

## 2001-2003 SNS Research Project

Environmental Research Programme Of the Federal Ministry for Environment, Nature Conservation and Nuclear Safety, UFOPLAN 201 11 612

"Implementation of a Semantic Network Service (SNS) in the context of the German Environmental Information Network (gein®)"

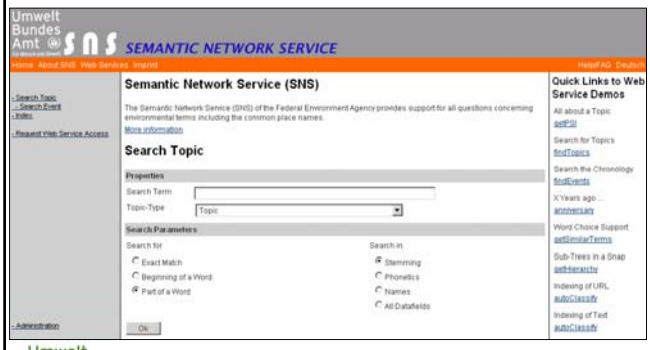
- Integration of the gein® Taxonomy in a Topic Map (ISO 13250)
- Topic Map Browser and Navigator in the Web
- Editorial Staff Services
- Enhanced Automatic Indexing
- Semantic Web Services
- Getting integrated

Umwelt Bundes Amt

innoQ

## Semantic Network Service (SNS) Today

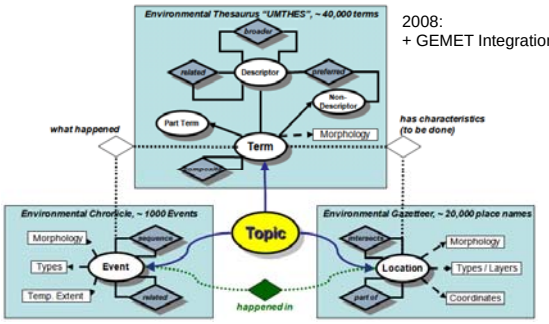
http://www.semantic-network.de/home.html?lang=en




Umwelt Bundes Amt

innoQ

## SNS Topic Map




2008:
+ GEMET Integration

Umwelt Bundes Amt

innoQ

## SNS Web Services

- **getPSI**
  - description and relations of a Topic (“published subject indicator”).
- **findTopics**
  - search for Topics by terms, types, etc.
- **findEvents**
  - adds temporal search conditions
- **anniversary**
  - what happened X years ago?
- **getSimilarTerms**
  - a “cloud” of semantically similar terms
- **getHierarchy**
  - (sub-)trees of broader / narrower terms
- **autoClassify**
  - automatic indexing of text documents

Umwelt Bundes Amt

innoQ

## SNS Web Services: getPSI

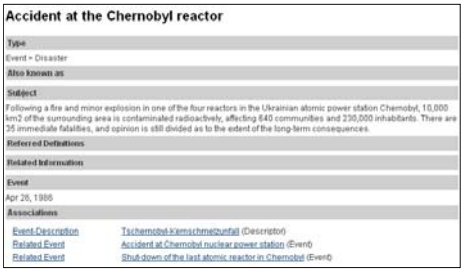
- PSI = ""Published Subjects Indicator" (OASIS)
- Direct Topic access by Topic-ID
- Returns a Topic representation with all attributes and associations
- URL (HTTP-GET) is a resolvable Topic reference

Umwelt Bundes Amt

innoQ

## getPSI example

http://www.semantic-network.de/displayTopic.html?lang=en&tid=calendarEvent_1



**Accident at the Chernobyl reactor**

Type

Event > Disaster

Also known as

Subject

Following a fire and minor explosion in one of the four reactors in the Ukrainian atomic power station Chernobyl, 10,000 km2 of the surrounding area is contaminated radioactively, affecting 640 communities and 230,000 inhabitants. There are 35 immediate fatalities, and opinion is still divided as to the extent of the long-term consequences.

Referred Definitions

Related Information

Event

Apr 26, 1986

Associations

| | |
|---|---|
| Event-Description | Tschernobyl-Kernschmelzunfall (Descriptor) |
| Related Event | Accident at Chernobyl nuclear power station (Event) |
| Related Event | Shut-down of the last atomic reactor in Chernobyl (Event) |

Umwelt Bundes Amt

innoQ

## SNS Web Services: findTopics / findEvents

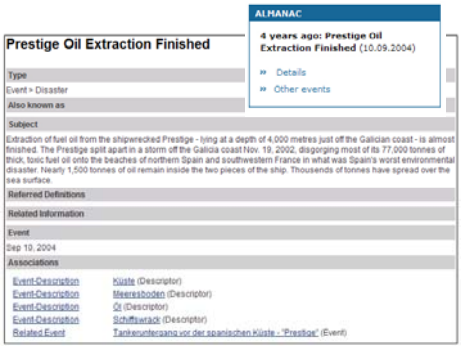
- Search for Topics by terms, Topic type, etc.
- Adjustable matching tolerance
- Returns a list of matching topics
- findEvents adds temporal search conditions

## findTopics example


## SNS Web Services: anniversary

- What happened X years ago?
- Pass any arbitrary date to this request (default is “today”).
- Returns events in the past matching the given date as close as possible.

## anniversary example 10-sept-2008



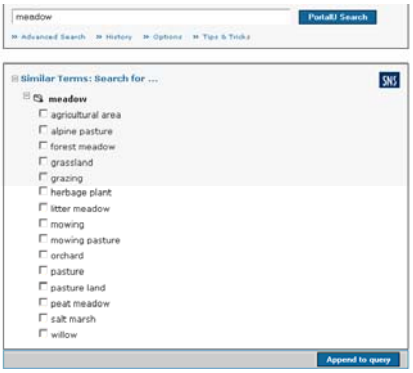
## SNS Web Services: getSimilarTerms

- Pass any term to this Service. If the term is known:
- Returns a “cloud” of semantically similar terms
- Meant as suggestions for a full text search or for a word choice while writing documents

Umwelt Bundes Amt

innoQ

## getSimilarTerms example




Umwelt Bundes Amt

innoQ

## SNS Web Services: getHierarchy

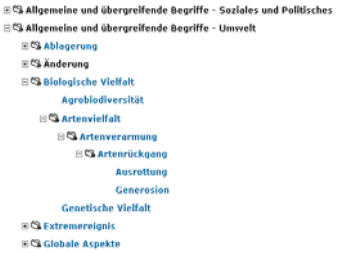
- Expects the Topic ID of a Thesaurus preferred term
- returns the full term hierarchy (broader/narrower terms).
- may be directed "upwards" or "downwards" or restricted to a given depths.

*Due to the current structure of the vocabulary, this service is only applicable to thesaurus descriptors in German language today.*

Umwelt Bundes Amt

innoQ

## getHierarchy example (German)



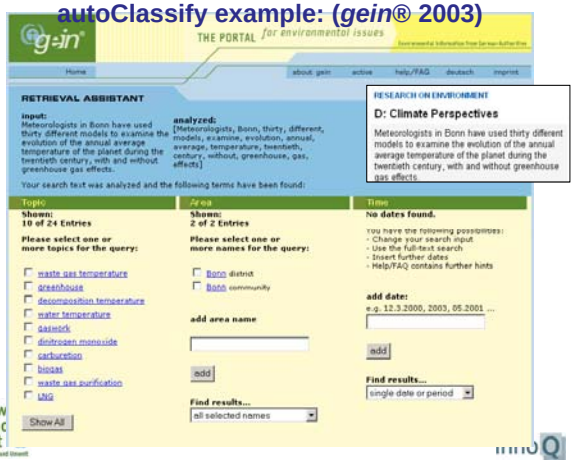

Umwelt Bundes Amt

innoQ

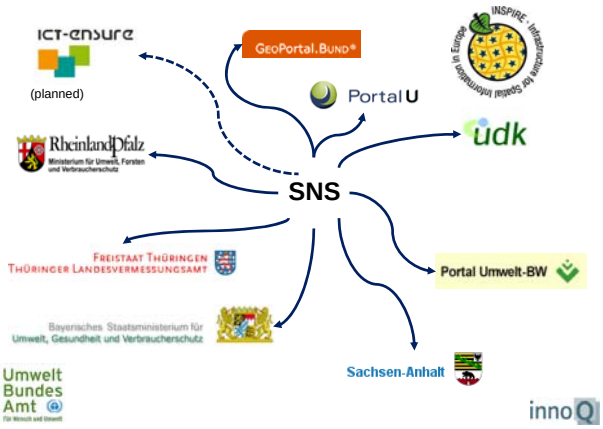
## SNS Web Services: autoClassify

- takes any kind of text and identifies significant keywords (Topics) from the controlled vocabulary
- fully automated text analysis
- text may be passed as a request parameter or by URL
- Supports plain Text, HTML, XHTML, PDF

## autoClassify example: (*gein*® 2003)




## Sharing Concepts by SNS 2008




## INSPIRE: ISO 19115 *descriptiveKeywords*

Metadata Implementing Rules: At least one keyword from GEMET (General Multilingual Environmental Thesaurus)

A rather poor support of non-spatial topics. Example:

```
<gmd:descriptiveKeywords>
  <gmd:MD_Keywords>
    <!-- Part B 3.1 Keyword Value -->
    <gmd:keyword>
      <gco:CharacterString>Hydrography</gco:CharacterString>
    </gmd:keyword>
    <!-- Part B 3.2 Originating Controlled Vocabulary -->
    <gmd:thesaurusName>
      <gmd:CI_Citation>
        <gmd:title>
          <gco:CharacterString>GEMET Thesaurus version 1.0</gco:CharacterString>
        </gmd:title>
```

## SNS can support INSPIRE

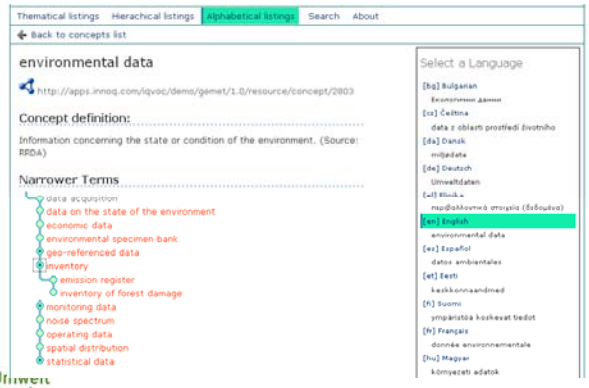
- UMTHES is the German Source of GEMET
- SNS can provide GEMET keywords for PortalU Catalogue Data

- Some open issues before we heard „INSPIRE-ing GEMET – Enhancing Metadata Creation and Discovery“ by *Nicole Ostländer, Michael Lutz* :
  1. Keyword selection strictly restricted by INSPIRE themes?
  2. Will *all* GEMET keywords be linked to a theme?
  3. (When) will there be an INSPIRE keyword service?
  4. Will this service support automatic keyword assignment?

Umwelt Bundes Amt

innoQ

## GEMET in SNS (pre-production)




Umwelt Bundes Amt

innoQ

## SEIS: Basic Principles

**Streamlining and simplification of existing reporting**

- information should be managed as close as possible to its source
- information is provided once and shared with others for many purposes

That’s how SNS works.

Umwelt Bundes Amt

innoQ

## SISE: Priorities (excerpt)

**Flexible chaining of distributed environmental services**

- service discovery and chaining
- automated data fusion
- semantics, thesauri, ontology services and standardisation
- chaining of models, predictive tools
- interactive and contextualised user interfaces

(Workshop Brussels Feb 2008)

Will SISE make use of **Semantic Web experience?**

Umwelt Bundes Amt

innoQ

## Semantic Web today: Linking Open Data (LOD)

LINKINGOPENDATA

**LOD Design Issues**

1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information.
4. Include links to other URIs, so that they can discover more things.

(Tim Berners-Lee, 2007)

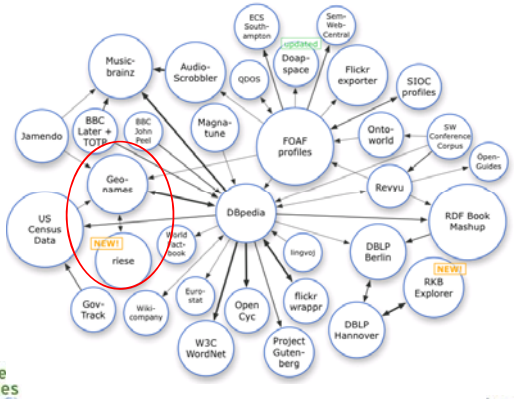
http://www.w3.org/DesignIssues/LinkedData.html

Umwelt Bundes Amt

http://esw.w3.org/topic/SweoIG/TaskForces/CommunityProjects/LinkingOpenData

innoQ

## Linking Open Data dataset cloud




Umwe Bundes Amt

innoQ

## http://www.geonames.org/2875115/lueneburg.html


Umwelt Bundes Amt

innoQ

## LOD: EuroStat Data (Riese*)


Umwelt Bundes Amt

*RDFizing & Interlinking the EuroStat Dataset Effort

innoQ

## SNS supporting Semantic Web Technology

- **HTTP URIs for "things" (= concepts = topics) …**
  - provide definitions and relations of these concepts
  - provide links to related data and information
  - in human and machine readable formats ("content negotiation").
- **… should be used as references to environmental concepts in**
  - datasets
  - metadata collections
  - reports in the Web
- **Weaving a Semantic Web of environmental information**
  - ubiquitous access
  - seamless navigation

Umwelt Bundes Amt
Für Mensch und Umwelt

innoQ

## SNS supporting SEIS/SISE

- **German Thesaurus, place names, and chronicle of events**
- **European GEMET Thesaurus**
- **To be extended:**
  - European place names ⇨ EuroGeoNames (EGN Content+)
  - European chronicle of events
  - Indicators
  - Substances
  - Methods
  - Units of measurement
  - Sensor types
  - …

Umwelt Bundes Amt
Für Mensch und Umwelt

innoQ

## Thank You for listening!
## Questions & opposition welcome!


thomas.bandholtz@innoq.com
+49 178 4049387

maria.ruether@uba.de
+49 30 8903 1503


innoQ


innoQ Schweiz GmbH
CH-6330 Cham
Tel +41 41 743 01 11

innoQ Deutschland GmbH
D-40880 Ratingen
Tel +49 2102 77 1620


Umwelt Bundes Amt
Für Mensch und Umwelt


Wörlitzer Platz 1
D-06844 Dessau-Roßlau
Telefon: 0340 2103-0