HTW Chur
Swiss Institute for
Information Research

Coping with Evolving Knowledge

Dynamic Domain Ontologies for Web Intelligence

Powered by webLyzard technology

5 September 2012

albert.weichselbraun@htwchur.ch

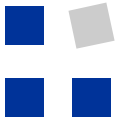
Agenda

Part I: Ontologies for Web Intelligence

- Web Intelligence
- NOAA Media Watch
- Applications of ontologies for Web Intelligence tasks

Part II: Dynamic Lightweight Domain Ontologies

- Ontology learning from heterogeneous sources
- Coping with change
- Outlook and conclusions



HTW Chur

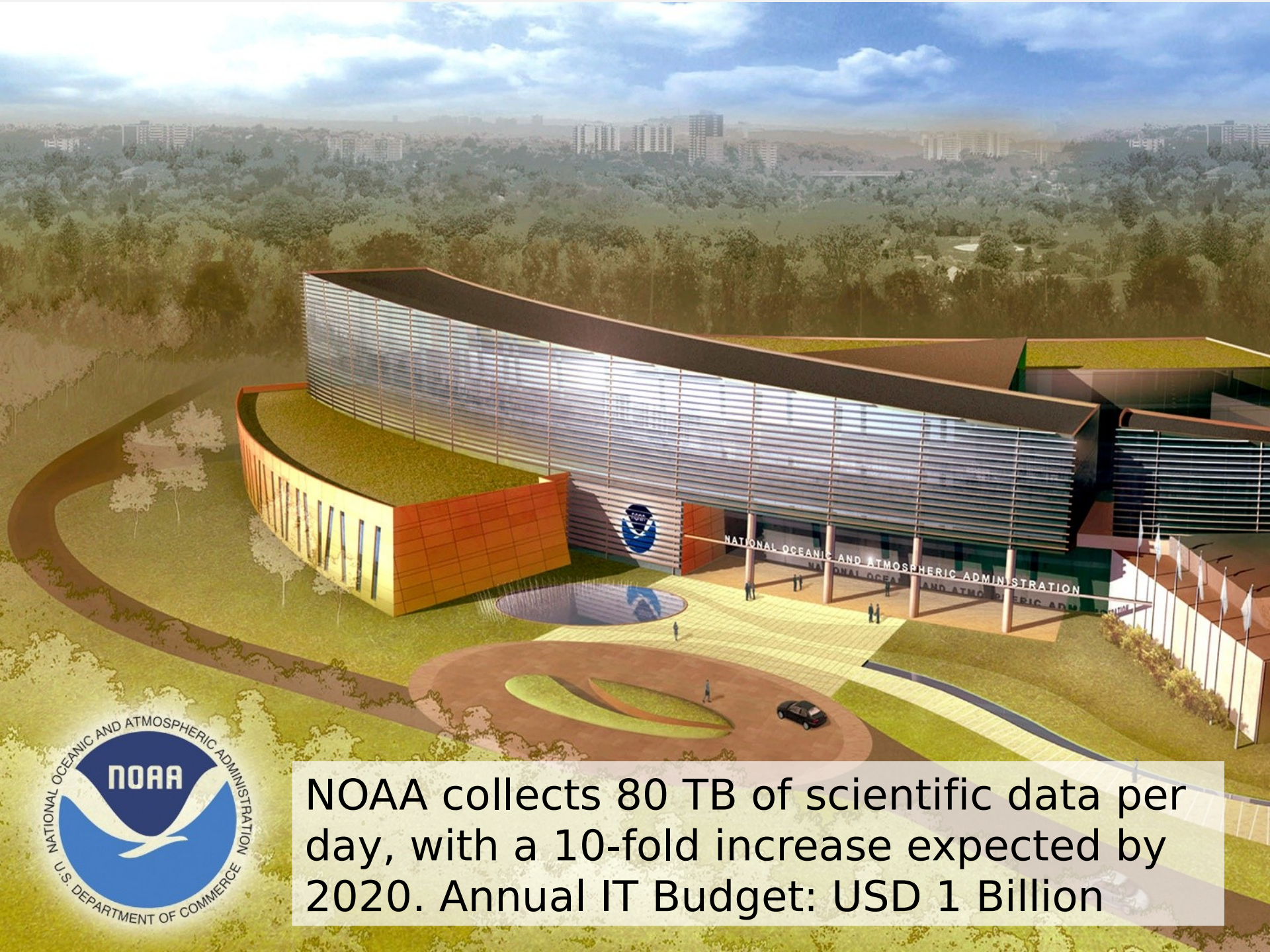
Swiss Institute for
Information Research

Part I
Ontologies for Web Intelligence

Powered by webLyzard technology

Web Intelligence

- **Business Intelligence** has been used as an umbrella term to describe concepts and methods for improving business decision making by using fact-based support systems (Chen 2010).
- Luhn (1958): “A business intelligence system”
- Structured data sources (databases, warehouses, …)
- Case studies
  - Continental Airlines (Anderson-Lehman et al. 2004)
  - Parkway (Negash and Gray 2008)
  - 1-800-CONTACTS (Watson and Wixom 2007)



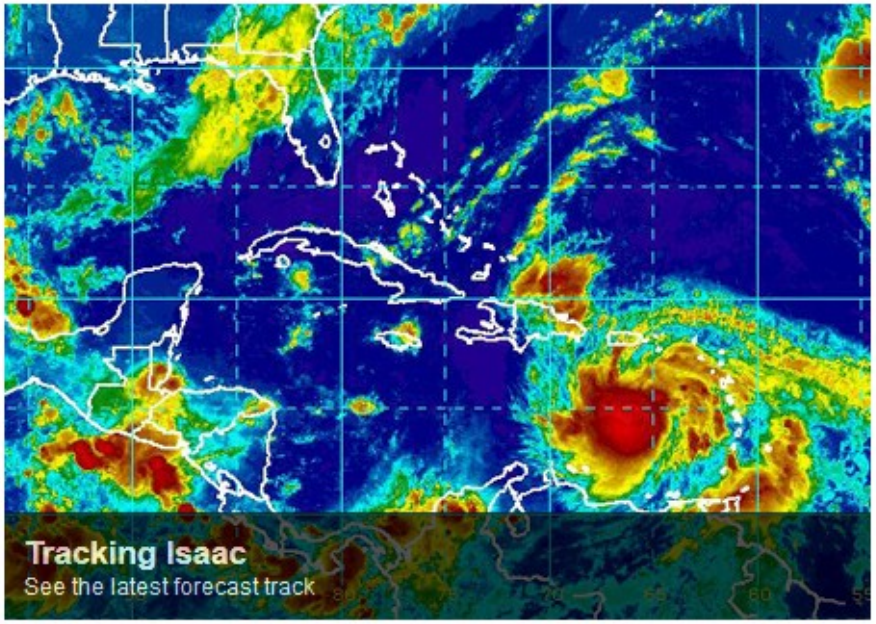
NOAA collects 80 TB of scientific data per day, with a 10-fold increase expected by 2020. Annual IT Budget: USD 1 Billion

NOAA NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

UNITED STATES DEPARTMENT OF COMMERCE

Weather.gov Forecast
City, ST » GO

» Active Weather Alerts
» Our Organization
» Working With Us
» Media & Constituents
» NOAA In Your State
» Budget Information
» Emergency Information for NOAA Employees
» Weather Ready Nation



EXPLORE OUR AGENCY

FROM THE NEWSROOM

All Releases

» New committee will advise federal leaders on integrating ocean observation systems

» Underwater noise decreases whale communications in Stellwagen Bank sanctuary

» New director for NOAA Corps and Office of Marine and Aviation Operations appointed

» Glang named as nation's hydrographer, director of Coast Survey

» Record depth for Lophelia coral discovered on Gulf of Mexico energy platforms

Subscribe to NOAA News [RSS]

NOAAWATCH
ALL-HAZARDS MONITOR

Climate.gov
science and services for society

RECOVERY at NOAA

Law Enforcement Updates

» July global temperatures fourth highest on record; Arctic sea ice is second lowest July extent on record

According to NOAA NCDC scientists, the globally-averaged temperature for July 2012 marked the fourth warmest July since record keeping began in 1880. It also marked the 36th consecutive July and 329th consecutive month with a global temperature above the 20th century average.

» Updated Atlantic hurricane outlook calls for near- or above-normal season

This year's Atlantic hurricane season got off to a busy start, with 6 named storms to date, and may have a busy second half, according to the updated hurricane season outlook issued today by NOAA's Climate Prediction Center, a division of the National Weather Service.

» July 2012 marked the hottest month on record for the contiguous United States

Drought expands to cover nearly 63 percent of the Lower 48; wildfires consume two

A CLOSER LOOK

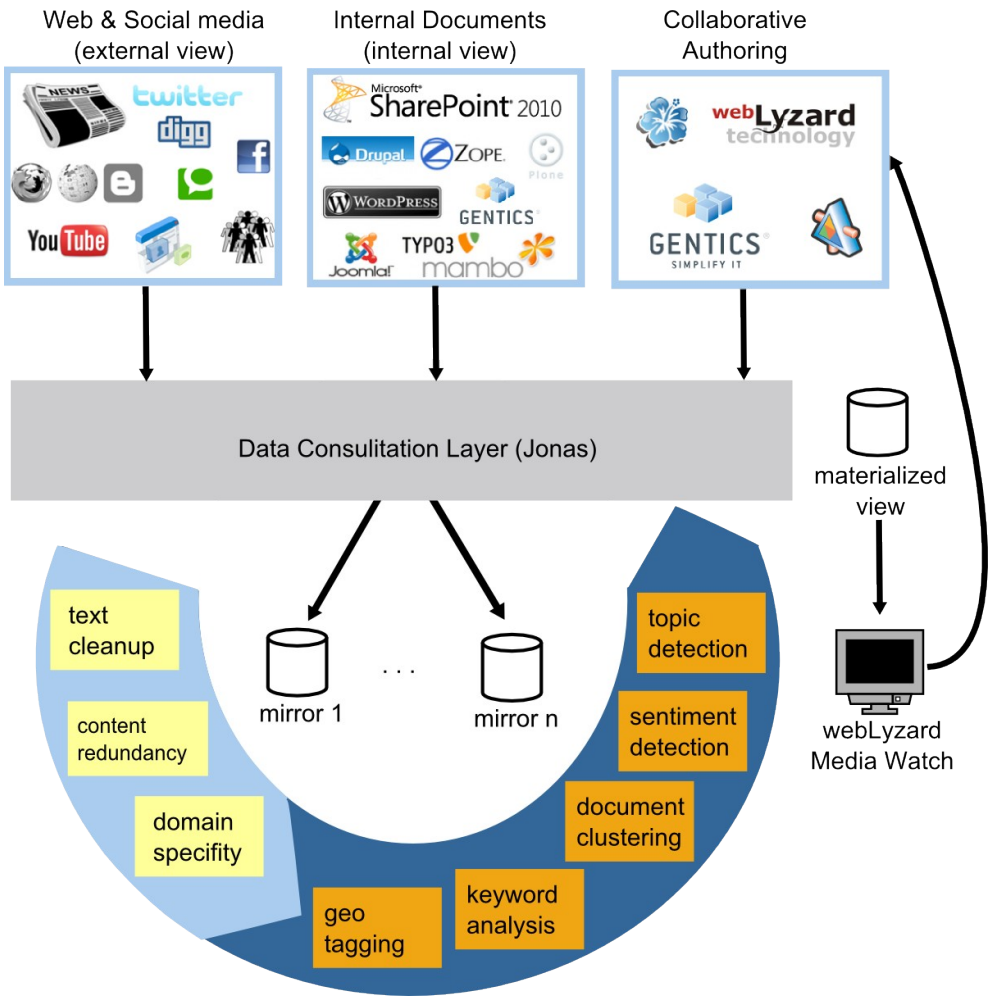
NOAA Education
Science Resources for Students and Teachers

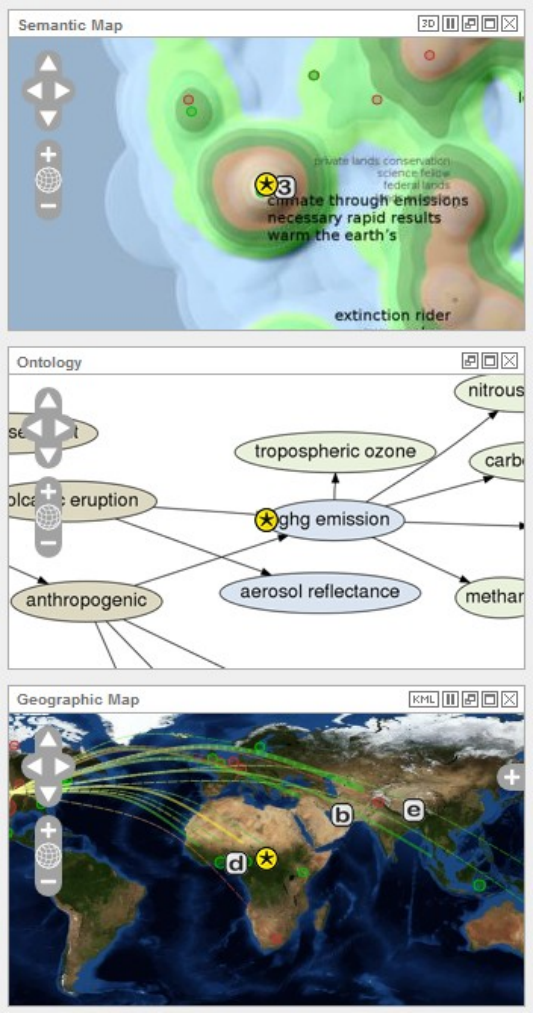
NOAA Corps
The New Corps of Discovery [see video]

STAY CONNECTED

View all of NOAA's social media sites

Data Acquisition & Contextualization





webLyzard

ECOresearch.net

Sources and Settings

Climate Change Media Watch

17 Jan 2012 — 17 Mar 2012 | News Media | Unfiltered | About | Help | hurricane

TOPICS

Topics

General

Climate Change	4112
Climate Sciences	287
Climate Policy	95
Mitigation	146
Adaptation	166

World Summits

Qatar (COP18)	57
Durban (COP17)	394
Cancun (COP16)	78
Copenhagen (COP15)	225

Own Topics

Trend Charts

Frequency | Sentiment | Disagreement

Content View

Document Summary

-4.0

LIVES AND HUNDREDS | HAVE SAVED LIVES | AUDITOR'S REPORT

Energy Efficiency - Environment - smh.com.au

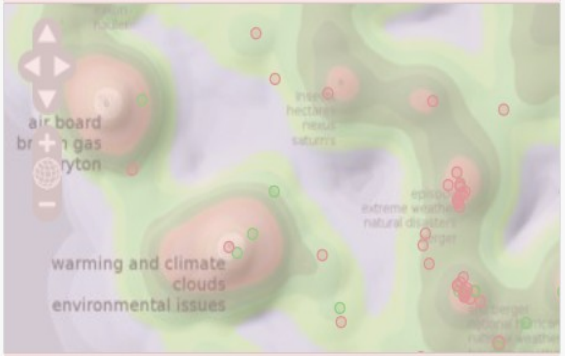
www.smh.com.au/environment/resources/energy-efficiency

Published in: AU • Text refers to: Sydney • Date: 2012-03-12

« Tom Arup ENVIRONMENT CORRESPONDENT A [illegible] released accidentally by the federal government shows [illegible] brown coal power plant will resist energy efficiency efforts because... ... Jason Dowling A SECRET report from an expert advisory panel might have saved lives and hundreds of millions of dollars had the federal government commissioned it before launching its disastrous $2. Jason Dowling Asecret report from an expert advisory panel might have saved lives and hundreds of millions of dollars had the federal government

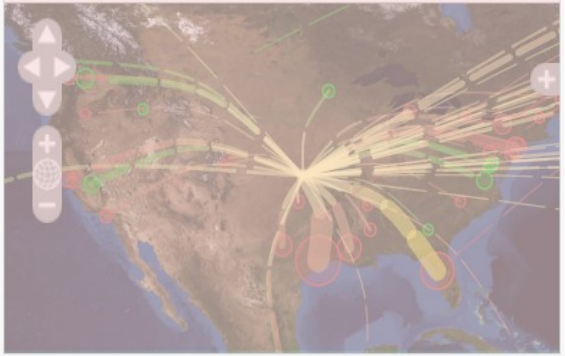
Visualizations

Semantic Map



Tag Cloud

Geographic Map



ASSOCIATIONS

Chart Terms

warming and climate	
ben cubby	
activism	
heartland institute	

hurricane

- hurricanes
- eric berger
- hurricane irene
- national weather service
- hurricane season
- national hurricane center
- storms
- high tide
- tornado

Search Results

Similar Documents | Nearby Locations | Search Results | Keyword Graph

Results: 182 out of 182 | Average Sentiment: -7.9

03/16	-10	rise even more. Flooding from hurricane Irene last year illustrated how vulnerab
03/16	-9	less that 40mpg, uphill in a hurricane headwind with 4 sumo wrestlers inside he
03/15	-7	lew York City last summer from hurricane Irene. Scientists are predicting a possi
03/15	-34	youtube.com Rina becomes 17th hurricane of the season theextinctionprotocol.worc
03/14	+18	m/Sling_Shot_Orlando_Florida. hurricane Wilma Video - Miami Beach, Florida. www.
		teChase.com captured on video hurricane Wilma's strong Category 2 force winds or
		barometric pressure of 959mb, hurricane Wilma lost some of her punch before hitt
03/14		Irene last year illustrated how vulnerab
03/14		Irene's landfall. New York City, often r
03/14	-18	storm surge from a Category II hurricane, combined with a projected sea level ris
		Grenada - Grenada was hit by hurricane Ivan six years ago and has been struggli
		now apart from another deadly hurricane, like Ivan is rising sea levels and sea
03/14	-28	in numerous sections, during hurricane Katrina , in the city proper and in east
03/13	-12	in the past including after hurricane Katrina, as my colleague Michael Grunwal
03/12	-4	ously.… More ». Live-Blogging hurricane Irene as It Moves Up the Eastern Coast.
		hquake jolted the East Coast, hurricane Irene is gathering steam for a possible

NOAA NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

25 Jun 2012 - 25 Aug 2012 ▼ Unfiltered ▾ noaa|"national oceanic

Help | News | Blogs | Social Media | Press Releases | NOAA Sites | Partners | NAPA Dialogue | Results: 100 ▾ out of 1084 | Semantic Map | Geo Map | Ontology | Tag Cloud | Keywords

TOPICS

NOAA Overview

Climate	724
Lubchenco	67
NOAA	1084
Oceans	954
Satellites	254
Weather	2410

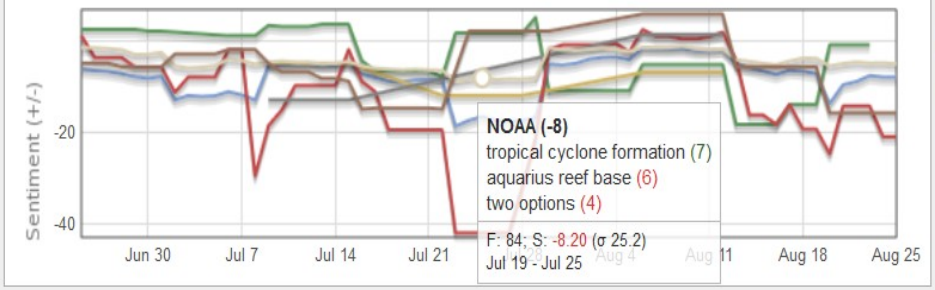
NOAA Services

adaptation	36
climate change	522
climate impact	3
climate science	59
climate service	5
climatewatch	1
copenhagen	24
mitigation	62

My Topics

+ Add new Category

Frequency | Sentiment | Disagreement



Document Summary | Full Text — PREVIEW DELETE NEW LINK

RAIN BAND | NIÑO EVENTS | EXTREME FLOODS

-1.0

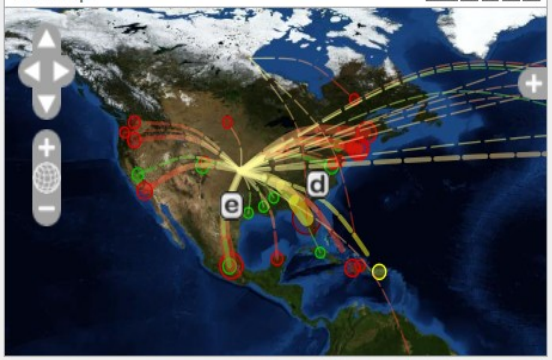
Warming causes more extreme shifts | Scoop News

www.scoop.co.nz/stories/WO1208/S00288/warming-causes-more-extreme-shifts.htm

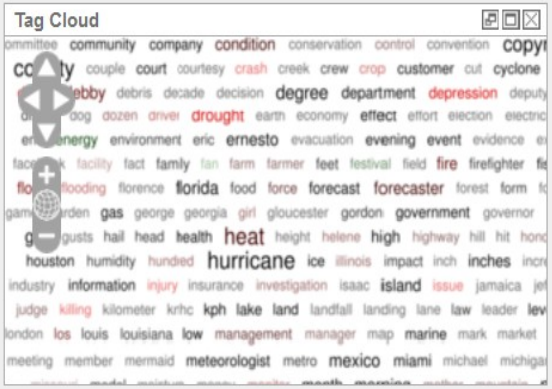
Published in: NZ • Text refers to: Oceania • Date: 2012-08-17

« Warming causes more extreme shifts of the Southern Hemisphere's largest rain band. ... Occasionally, the rain band moves northwards towards the Equator by 1000 kilometres, inducing extreme climate events. ... In turn, and in spite of disagreement about the future of El Niño events, this warming leads to the increased frequency of extreme excursions of the rain band. During moderate El Niño events with warming in the equatorial eastern Pacific, the rain band moves north-eastward by 300 kilometres. »

Geo Map



Tag Cloud



ASSOCIATIONS

NOAA

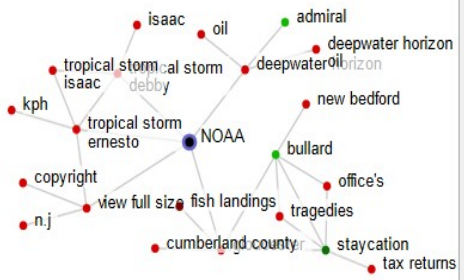
deepwater horizon	75
tropical storm ernesto	154
gloucester	117
tropical storm debby	256
view full size	100
oil	441
fish landings	14
mermaids	68
bullard	17
low chance	58
admiral	26
habitat	117
dolphins	58
size national hurricane	22
kerry	43

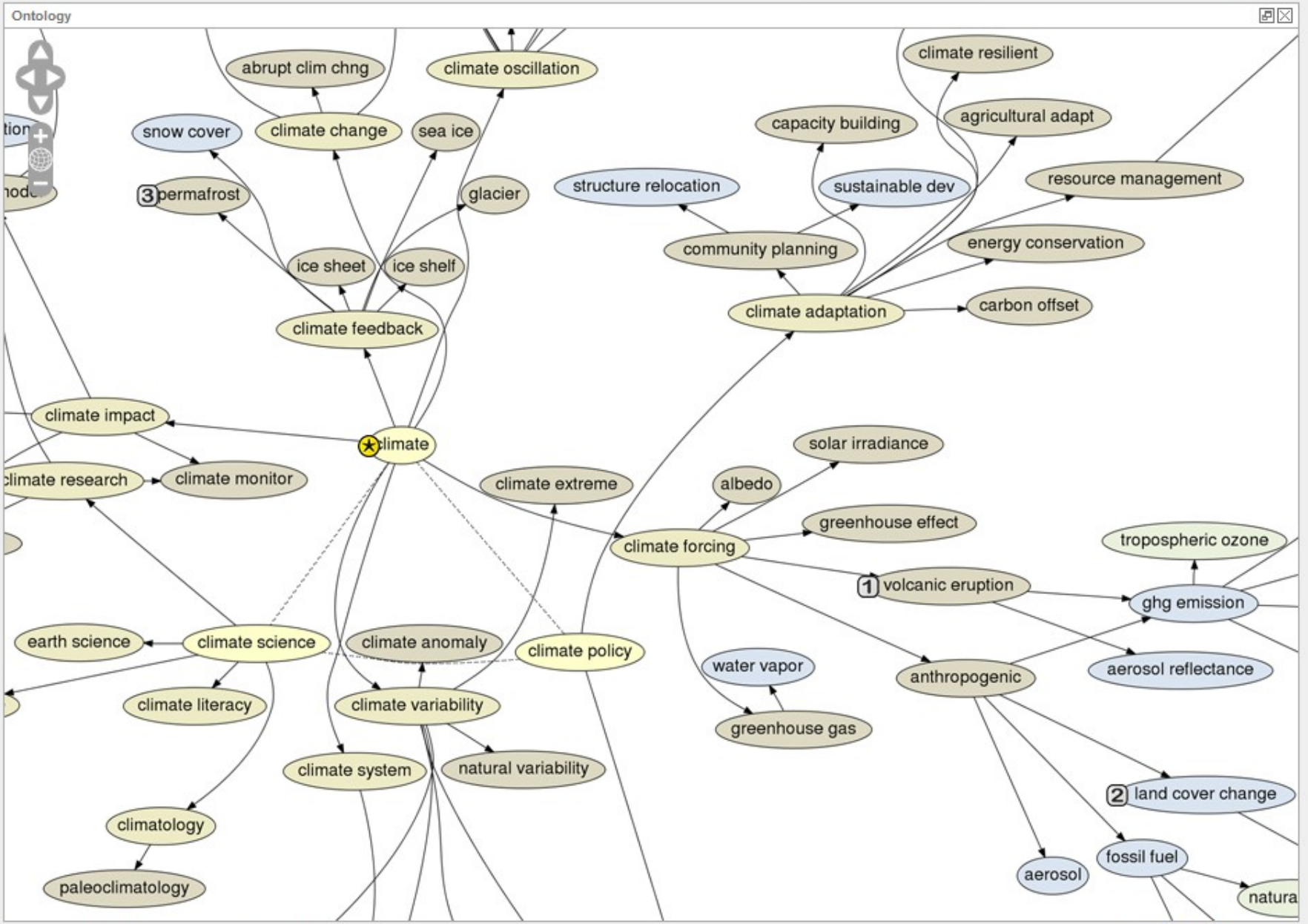
Similar Documents | Search Results | News Flow

Average Sentiment: -4.4

DATE	+/−		+/−
		2012. July 2012 weather events.	noaa. Last month was a record setter. July was
		hic Atmospheric Administration (	noaa) reported today (Aug. 8). The previous re
08/25	+5.0	ropical cyclone heat potential. (	noaa). In looking at this map you can see that
08/06	+6.0	o federal waters. Since mid-June	noaa data has shown no oil in the reopened area
		spill event," said Eric Schwaab,	noaa assistant administrator for NOAA's Fisheri
		catches being abandoned at sea.	noaa 's first Notice of Violation and Assessment
08/23	+7.0	, explains that according to the	noaa National Climatic Data Center, the first h
07/09	+11.0	30.6 million pounds. Last year,	noaa predicted a Louisiana brown shrimp catch o
		n for Louisiana since 2007, when	noaa predicted brown shrimp catch at 32.9 milli
07/09	+11.0	inistrator for fisheries, out of	noaa headquarters in Silver Spring, Md. In his
		Office as he strives to balance	noaa 's multiple missions in a historically sign
08/24	+11.0	e Caribbean Sea this afternoon. (	noaa). This weakening has occurred because some
		er , national hurricane center ,	noaa , tropical storm isaac , tropical storm jo
08/20	+13.0	ds the depression at this time. (	noaa). The wind shear, now light, is forecast t
		gs: johnson space center , nasa.	noaa increases seasonal forecast as tropics com
08/20	+14.0	ews stories, social media feeds,	noaa ... Bright House Internet customers gain 50

Keywords — Edges: 5 ▾




Ontology
abrupt clim chng
climate oscillation
climate resilient
snow cover
climate change
sea ice
capacity building
agricultural adapt
3 permafrost
glacier
structure relocation
sustainable dev
resource management
community planning
energy conservation
ice sheet
ice shelf
climate adaptation
carbon offset
climate feedback
climate impact
climate
solar irradiance
climate research
climate monitor
climate extreme
albedo
greenhouse effect
climate forcing
tropospheric ozone
1 volcanic eruption
ghg emission
earth science
climate science
climate anomaly
climate policy
water vapor
anthropogenic
aerosol reflectance
climate literacy
climate variability
greenhouse gas
climate system
natural variability
2 land cover change
climatology
fossil fuel
aerosol
paleoclimatology
natura

NOAA NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

27 Apr 2010 - 27 Jun 2010 ▼ Unfiltered ▼ oil spill

Help | News | Blogs | Social Media | Press Releases | NOAA Sites | Partners | NAPA Dialogue | Results: 250 ▼ out of 1562 | Semantic Map | Geo Map | Ontology | Tag Cloud | Keywords

TOPICS

NOAA Overview

Climate	513
Lubchenco	175
NOAA	1433
Oceans	988
Satellites	249
Weather	1239

NOAA Services

adaptation	19
climate change	427
climate impact	1
climate science	58
climate service	1
climatewatch	0
copenhagen	103
mitigation	30

My Topics

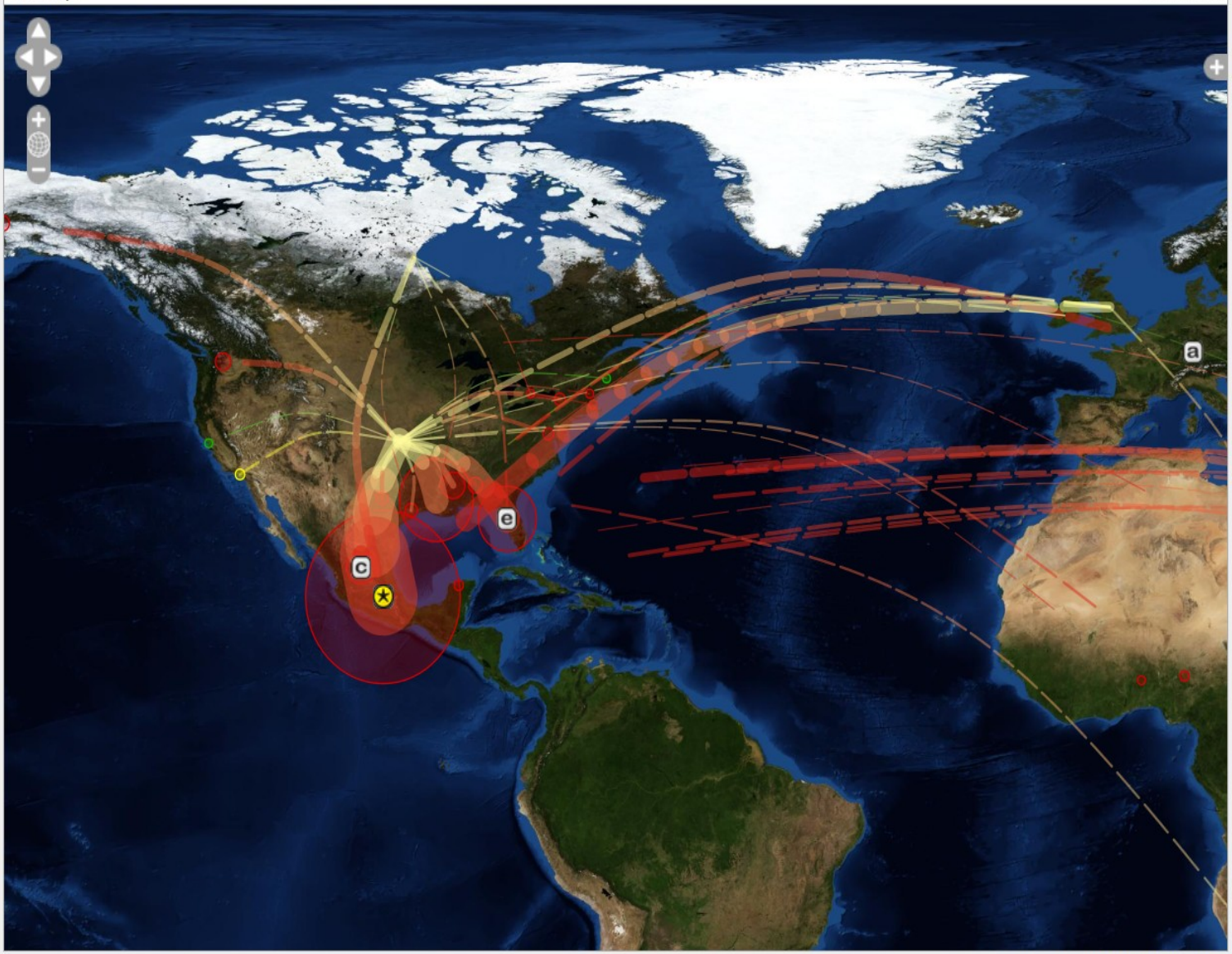
Add new Category

ASSOCIATIONS

oil spill

gulf	1474
oil	1535
rig	774
leak	765
oil slick	464
louisiana	887
spill	1365
barrels	489
horizon	714
containment	408
into the gulf	1474
mexico	1379
hayward	222
pipe	347
venice	253

Geo Map



TOPICS

NOAA Overview

Climate	724
Lubchenco	67
NOAA	1084
Oceans	954
Satellites	254
Weather	2410

NOAA Services

adaptation	36
climate change	522
climate impact	3
climate science	59
climate service	5
climatewatch	1
copenhagen	24
mitigation	62

My Topics

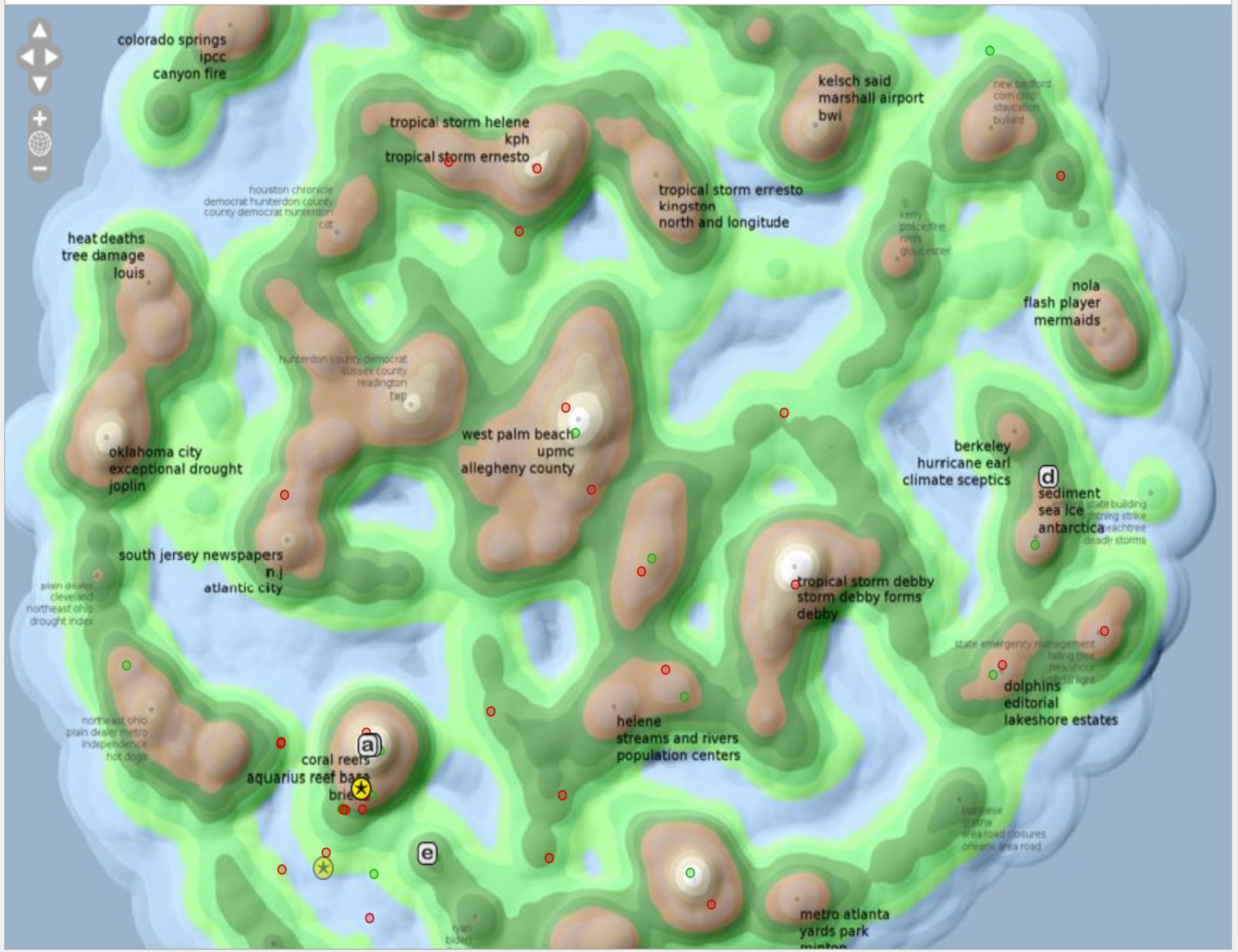
Add new Category

ASSOCIATIONS

coral reefs

reefs	57
hawaiian islands	17
coral	73
world's	122
aquarius reef base	10
lubchenco said	17
acid	40
acidity	15
ecosystems	55
key largo	8
two options	8
ocean	686
boynton beach	62
surface waters	12
csiro	12

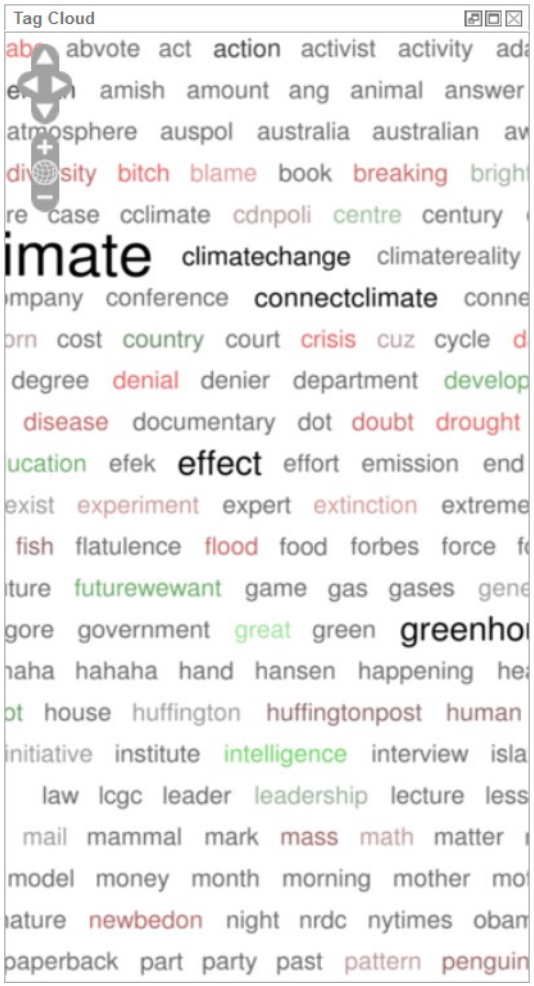
Semantic Map



Tag Cloud on Climate Change

News Media

Social Media



Fortune 1000

Analytical Questions & Applications

- How can we detect, describe and potentially predict patterns in online coverage (external events versus resonance patterns)?

  MARKET RESEARCH, TREND SCOUTING

- How widespread is content redundancy, and what influences content replication within and across networks (e.g. network and community characteristics)? Who are the most influential opinion leaders driving this process?

  OPINION LEADER ANALYSIS, IMPACT MEASURES, VIRAL MARKETING STRATEGIES

- How do macroscopic information flows shape public opinion? What are appropriate methods to measure and model the extent, dynamics and latency of this process?

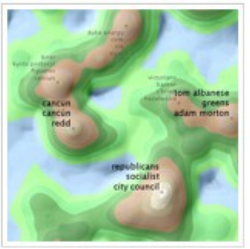
  FINANCIAL MARKETS, POLITICAL CAMPAIGNS, SCIENCE COMMUNICATION

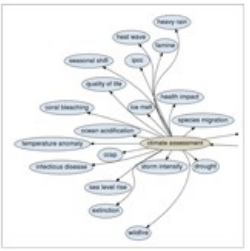
Ontologies for Sentiment Analysis

- Problem: Ambiguities, Context
- Data Sources: ConceptNet, Wordnet

- **Workflow**:

  1. Identify ambiguous terms based on a training corpus
  2. Determine the context for ambiguous terms
     → contextualized sentiment dictionary
  3. Use the contextualized sentiment dictionary for ambiguous terms



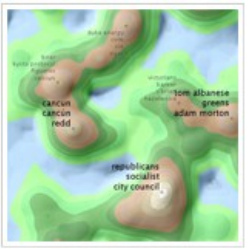


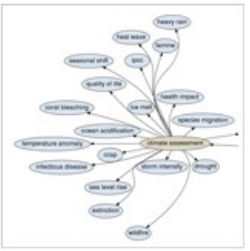
Ontologies for Geotagging

- Problem: Ambiguous Locations, Focus
- Data Sources: GeoNames

- **Workflow**:

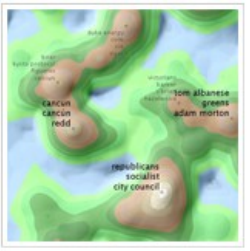
  1. Build a data model based on the GeoNames ontology and the corresponding instance data.
     → nearbyFeatures, parentFeature, parentCountry, population
  2. Use this context information for disambiguation and focus selection.

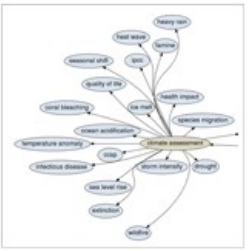




Further Application Areas

- Navigation
- Search query expansion
- Pre-structuring of the semantic map
- Named entity detection





HTW Chur
Swiss Institute for
Information Research

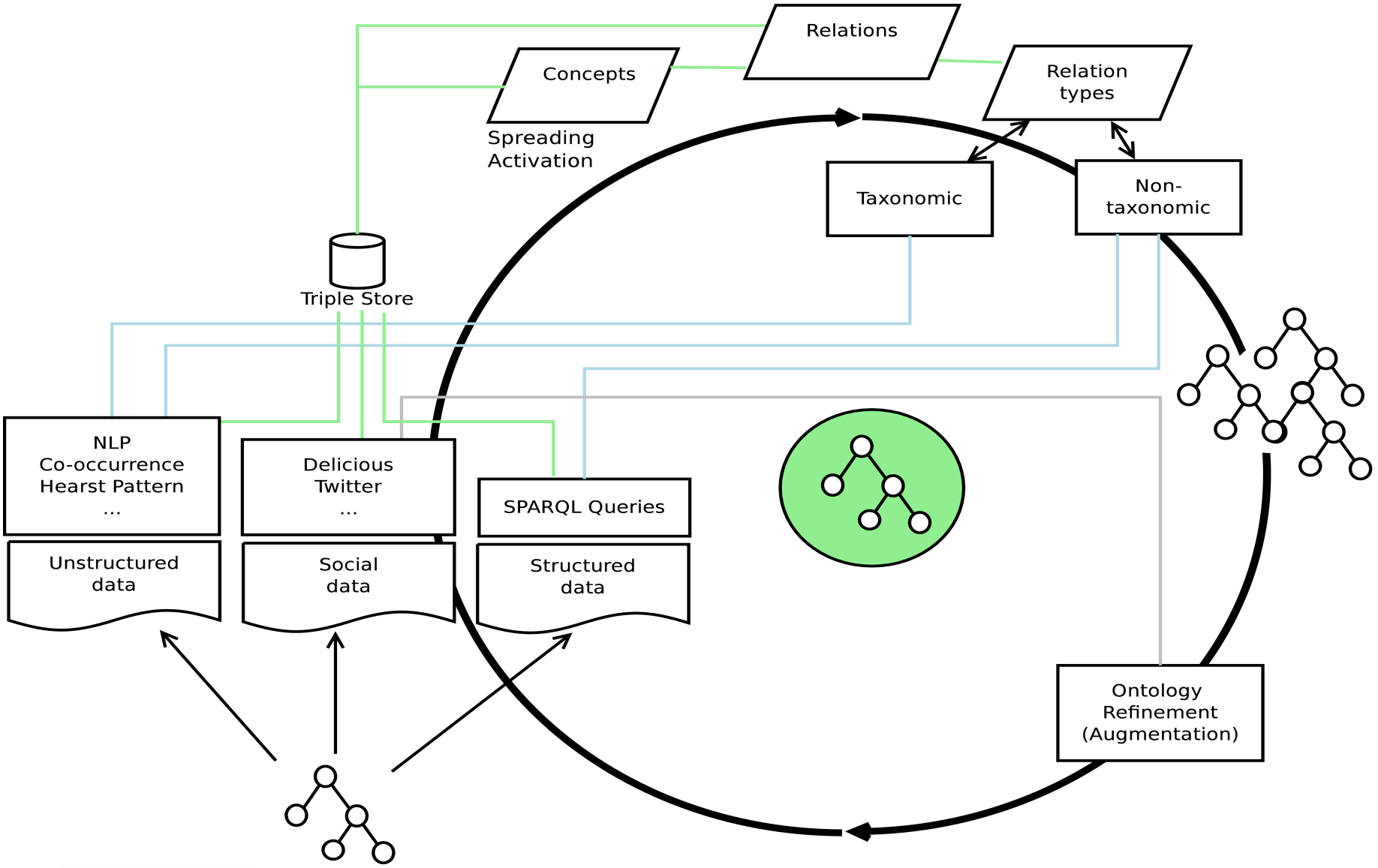
Part II
Dynamic Domain Ontologies

Powered by webLyzard technology

Ontology Learning

- ontology learning framework for lightweight domain ontologies (Hendler 2009, Alani et al. 2008)
- based on a seed ontology and domain documents
  - extract relevant terms
  - integrate them into the ontology
- benefits of integrating social sources
  - potential of providing background knowledge
  - contain the latest terminology (Angeletou et al. 2007)
  - (evolve at much a higher pace than domain documents)

Architecture - Overview



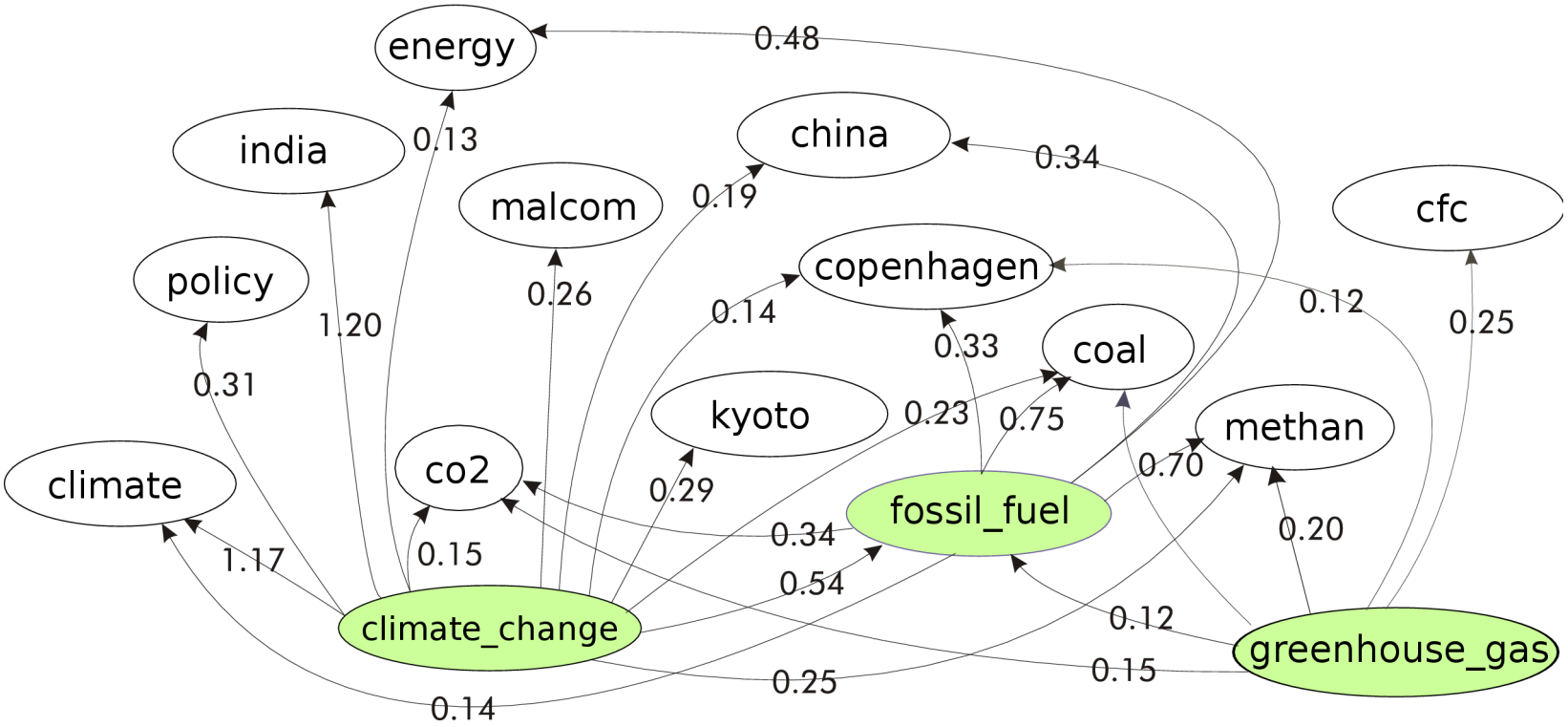
Suggested Terms

Unstructured data sources	Social	
	delicious	flickr
targets	animalcare	architecture
building	architects	art
coal	atmosphere	auckland
levels	award	beach
climate change policy	carbonfoodprint	bicycle
pact	...	...
reduce greenhouse gas	**technorati**	**twitter**
pollution	agile	aces
firm	apple	afghan
carbon dioxide emissions	art.	afghanistan
etcs	automotice	africa

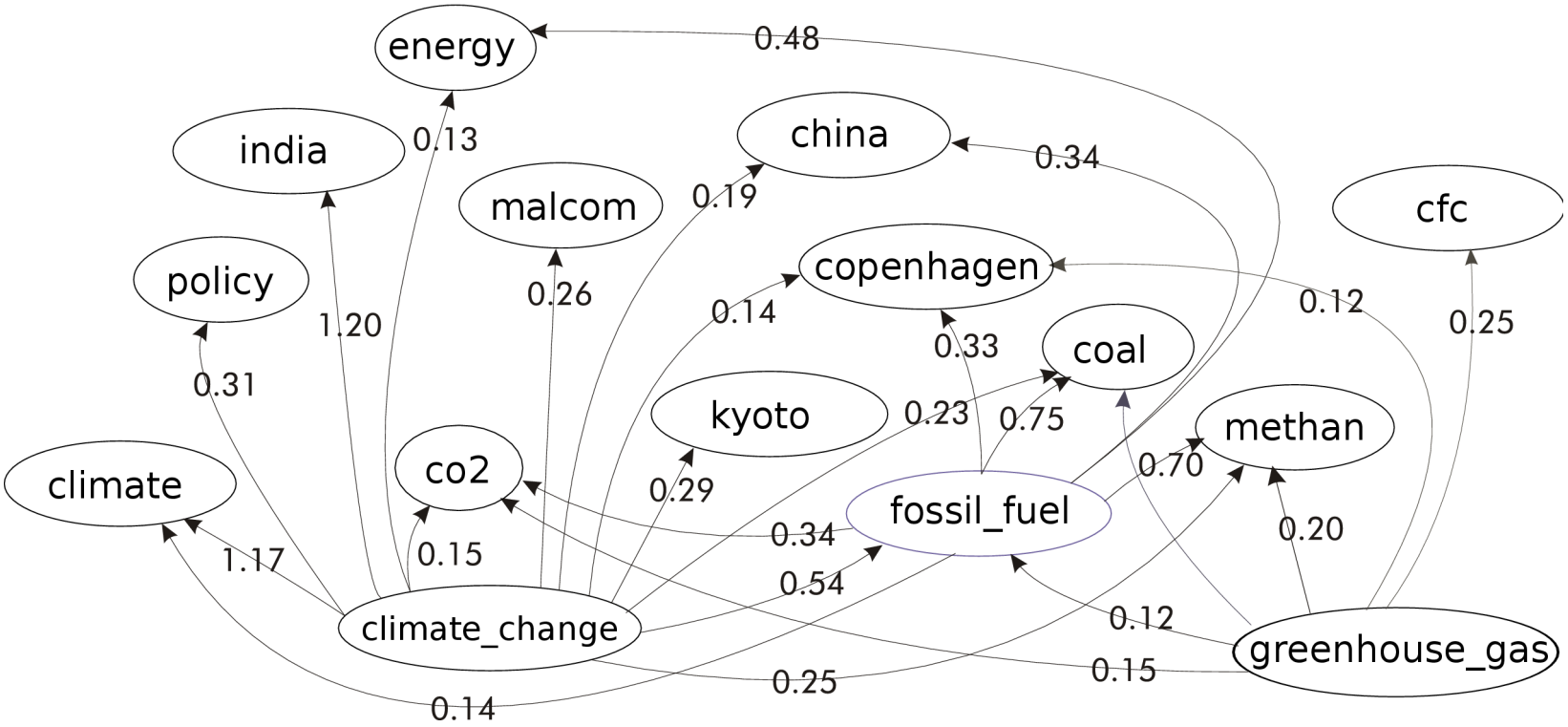
Suggested Terms

Seed Concept (C_s)	Evidence-Source (e)	Candidate Concept (C_c)
climate_change	oe:coOccurs	“CO2”
_:1	rdf:subject	climate_change
_:1	rdf:predicate	oe:coOccurs
_:1	rdf:object	“CO2”
_:1	rdf:type	rdf:Statement
_:1	oe:significance	“3.20”
climate_change	ow:twitter	“CO2”
	...	
climate_change	wn:hyponym	temperature_change

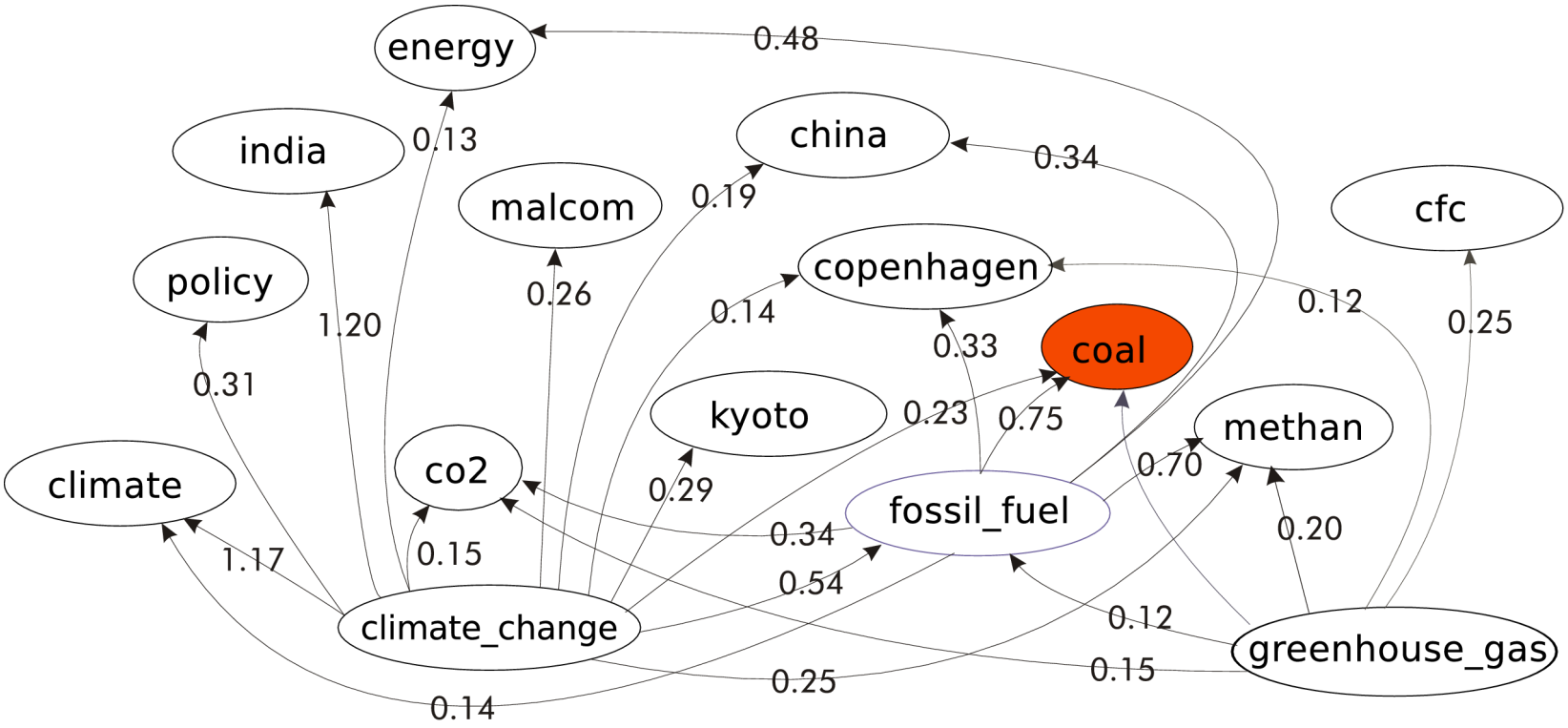
Concept Selection and Positioning



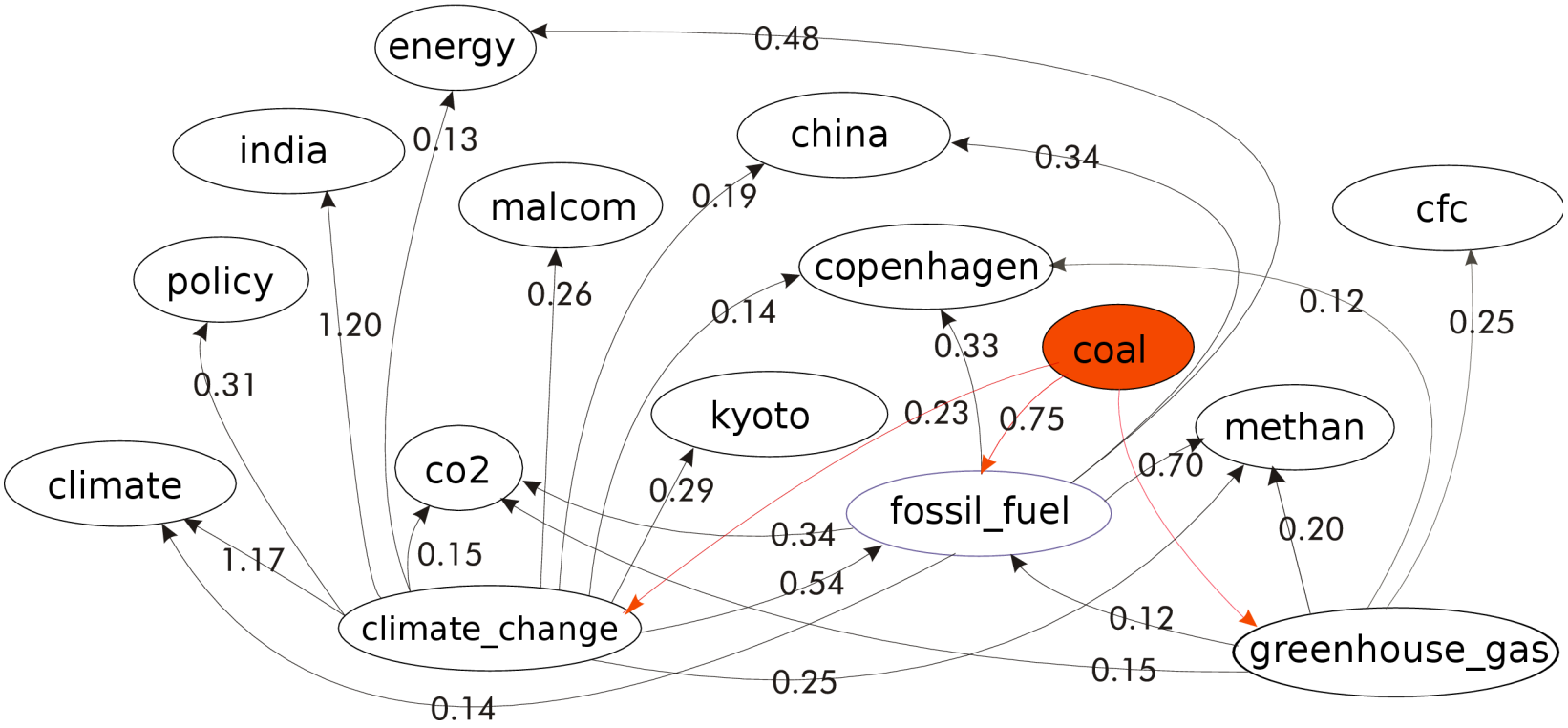
Concept Selection and Positioning



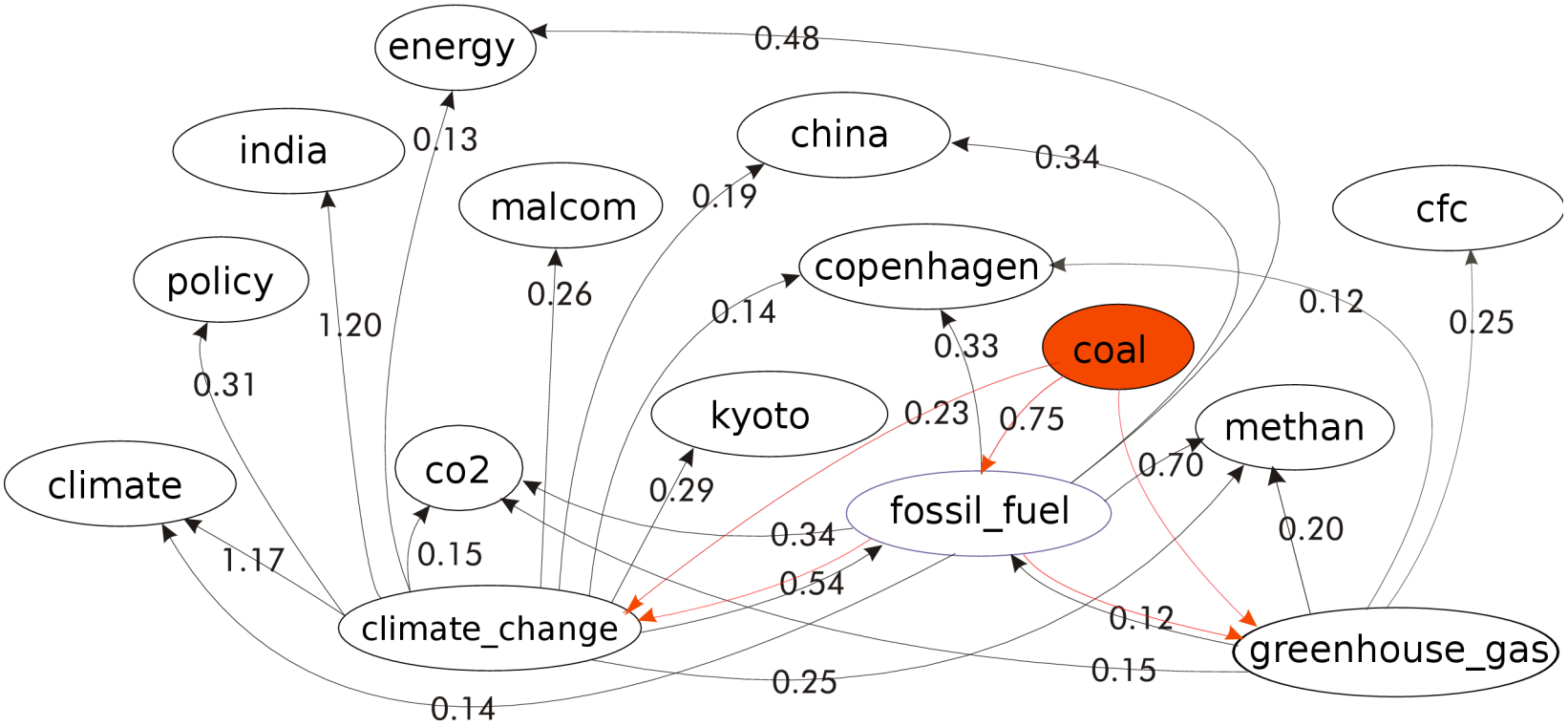
Concept Selection and Positioning



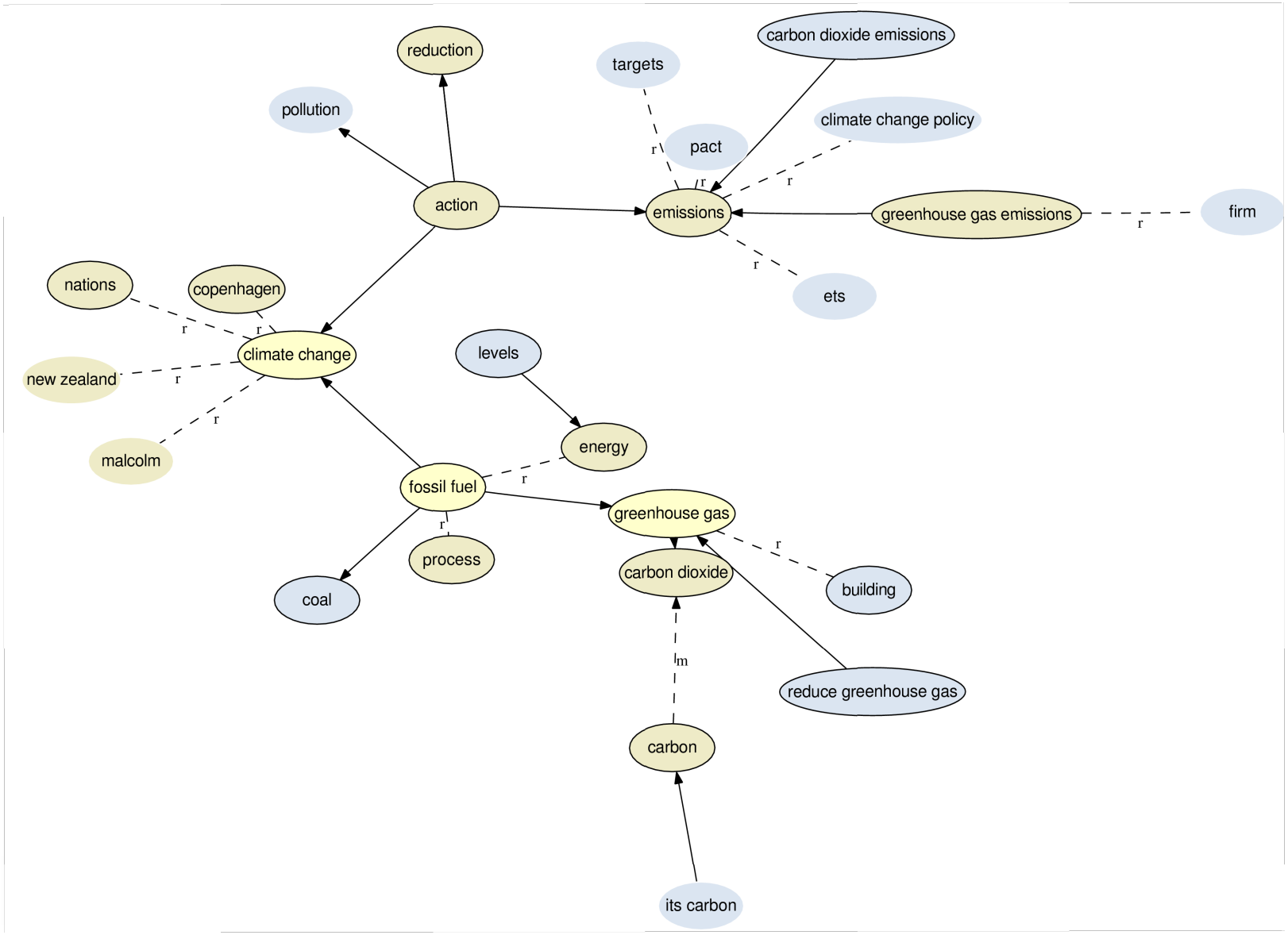
Concept Selection and Positioning



Concept Selection and Positioning



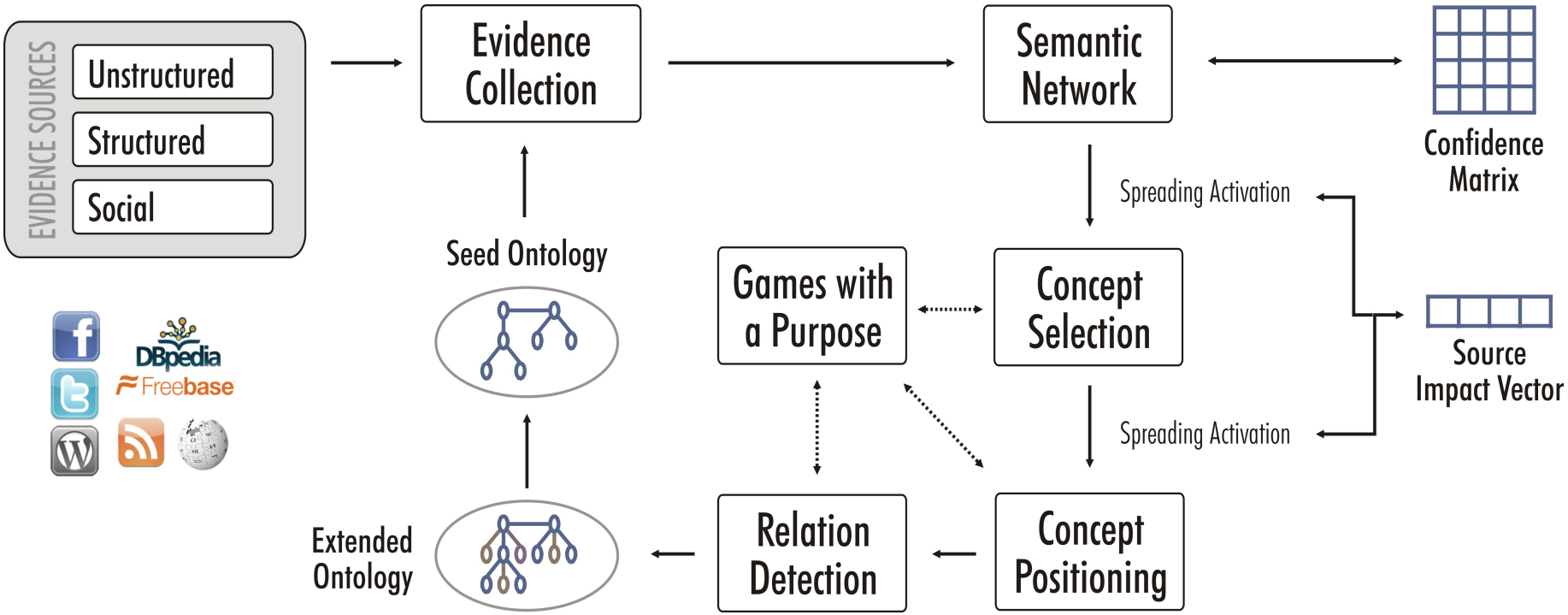
Extended Ontology



Domains Evolve

- Static Approach
  - Fixed transformation functions
    (evidence sources → spreading activation weights)
  - Determined by domain experts
- Problems
  - Domains evolve
  - Importance and reliability of evidence sources changes
- “Gold Standard”
  - Users
  - Games with a Purpose

Learning Evolving Ontologies



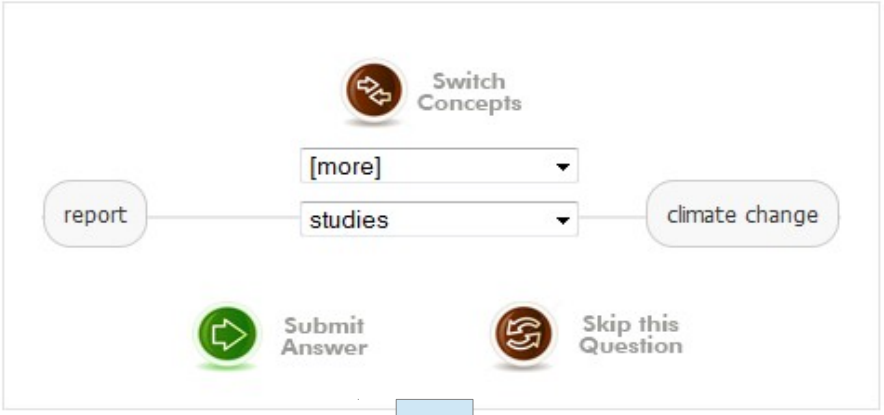
Relation Strength Matrix

- A formalized representation of the ontology
- Records the relation strength (rs) for any relation r_{ij} between seed concept “i” and candidate concept “j”.

$$RSM = \begin{bmatrix} rs_{r_{11}t_1} & rs_{r_{11}t_2} & \cdots & rs_{r_{11}t_n} \\ rs_{r_{12}t_1} & rs_{r_{12}t_2} & \cdots & rs_{r_{12}t_n} \\ \vdots & & & \vdots \\ rs_{r_{21}t_1} & rs_{r_{21}t_2} & \cdots & rs_{r_{21}t_n} \\ \vdots & \vdots & & \vdots \\ rs_{r_{mn}t_1} & rs_{r_{mn}t_2} & \cdots & rs_{r_{mn}t_n} \end{bmatrix}$$

- Most direct visualization of changes in the domain
- Distinguish: changes in the domain versus changes in the user's perception.

Games with a Purpose



$$\begin{bmatrix} rs^*_{r_{11}t_1} & rs^*_{r_{11}t_2} & \cdots & rs^*_{r_{11}t_n} \\ rs^*_{r_{12}t_1} & rs^*_{r_{12}t_2} & \cdots & rs^*_{r_{12}t_n} \\ \vdots & & & \vdots \\ rs^*_{r_{21}t_1} & rs^*_{r_{21}t_2} & \cdots & rs^*_{r_{21}t_n} \\ \vdots & \vdots & & \vdots \\ rs^*_{r_{mn}t_1} & rs^*_{r_{mn}t_2} & \cdots & rs^*_{r_{mn}t_n} \end{bmatrix}$$

Evidence Confidence Matrix

- Records the per evidence source confidence between source and candidate terms and its temporal evolution
  → base for optimization and learning of the weights

$$ECM_{C_s,C_c} = \begin{bmatrix} c_{es_1,t_1} & c_{es_1,t_2} & \cdots & c_{es_1,t_m} \\ c_{es_2,t_1} & c_{es_2,t_2} & \cdots & c_{es_2,t_m} \\ & & \vdots & \\ c_{es_n,t_1} & c_{es_n,t_2} & \cdots & c_{es_n,t_m} \end{bmatrix}$$

- Supports trend detection and perception experiments
- Perception != Truth; example: food for kids (Kinder Milk Slice)

Source Impact Vector

- Contributes the per source impact weights for translating evidence sources into spreading activation weights.
  → Determines the impact of a single source.

$$SIV_{t_i} = \left[\begin{array}{cccc} I_{es_1} & I_{es_2} & \cdots & I_{es_n} \end{array} \right]$$

- Initial settings
  - Domain experts
  - heuristics and metrics such as Google Page Rank, etc.
- Gets adjusted based on user feedback
  → better alignment with the optimal relation strength matrix

Summary

- Ontologies and the corresponding instance information play a crucial role in improving Web Intelligence
- Timely adaptation of ontologies to changes in the domain are still challenging
- Presented an approach for aligning ontologies to user perception that allows
  - Keeping ontologies up to date
  - Analyzing the **sources** of changes (evidence confidence matrix)
  - Detecting trends and patterns (evidence confidence matrix, source impact vector)
  - Making the reasons for change more traceable (evidence confidence matrix, source impact vector)

Interface Description
www.weblyzard.com/interface

Media Watch on Climate Change
www.ecoresearch.net/climate

Contact + Login Requests
albert.weichselbraun@htwchur.ch
info@weblyzard.com