Enipedia.tudelft.nl & Moving Towards Linked Open Energy Data

Chris Davis
http://enipedia.tudelft.nl
c.b.davis@tudelft.nl

TUDelft

**Delft University of Technology**

Who am I?
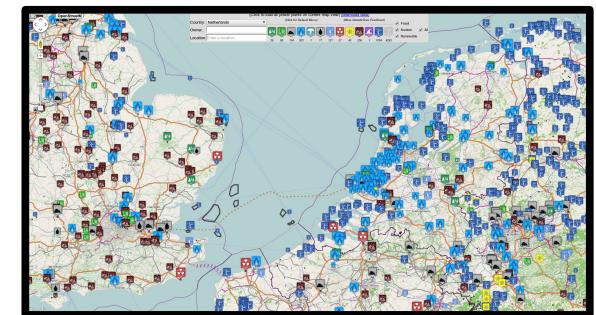
- Postdoc Energy & Industry, TBM, TU Delft
- Focus on Industrial Ecology, Open Data, Collaborative Software, Modeling, Visualization, Analytics, etc.

Christopher Bryan Davis
Making Sense of Open Data
From Raw Data to Actionable Insight

52

NEXT
GENERATION
INFRASTRUCTURES
FOUNDATION

TUDelft

Delft University of Technology

Motivations

- Energy and sustainability are some of the most important topics of the 21st century

- Need both aggregated and fine-grained data

- Research can be data intensive

- There's a lot out there, but connecting it is tedious

- Researchers often duplicate effort

- It would be great to revolutionize how we deal with this data

- The energy sector is only slowly embracing the ICT & Open Data revolutions

TU Delft

Delft University of Technology

Enipedia.tudelft.nl

Page   Discussion

Read   Edit   View history   [ ] [Search box] Go   Search

Enipedia

(Redirected from Main Page)

Enipedia is an active exploration into the applications of wikis and the semantic web for energy and industry issues. Through this we seek to create a collaborative environment for discussion, while also providing the tools that allow for data from different sources to be connected, queried, and visualized from different perspectives.

- Help
- Enipedia Blog
- SPARQL Endpoint
- Videos

- Recycling
- Integration in Energy Systems
- Power Plants
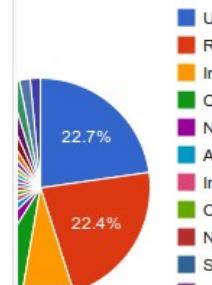
- Industry
- Environment
- Advanced Topics
- Feature Requests

Featured Content

Navigation

Main page
Enipedia Blog
Enipedia Mailing List
Contact Us
Add Reference
Report a Problem
Recent changes
Random page
Help

Portals

Power Plants
Direct Carbon Fuel Cells
Renewable Energy
Port of Rotterdam
Oil
Videos

Advanced

Sparql Endpoint
Advanced Topics

Toolbox

What links here
Related changes
Upload file
Special pages
Printable version
Permanent link
Browse properties

**Natural Gas Overview**

- UnitedSt...
- Russia
- Iran
- Canada
- Norway
- Algeria
- Indonesia
- China
- Netherla...
- SaudiAr...

22.7%
22.4%

Overview of natural gas production, consumption, reserves, infrastructure and trade networks worldwide.

Worldwide Flows of Natural Gas

Natural gas world trade network - dynamically generated from country level data.

Exploring Global Electricity Production
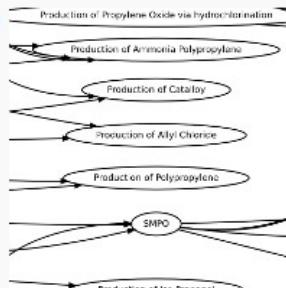
Learn about global electricity production, based on data linked together from sources such as Carma.org and eGRID. Browse the fuel sources and power outputs for 50,000 power plants using a KML file for Google Earth.

Natural Gas Infrastructure
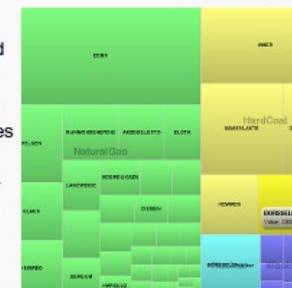
Compilation of natural gas infrastructure world wide: major pipelines, LNG terminals, cross-border points - all located on map.

Industrial Production Chains
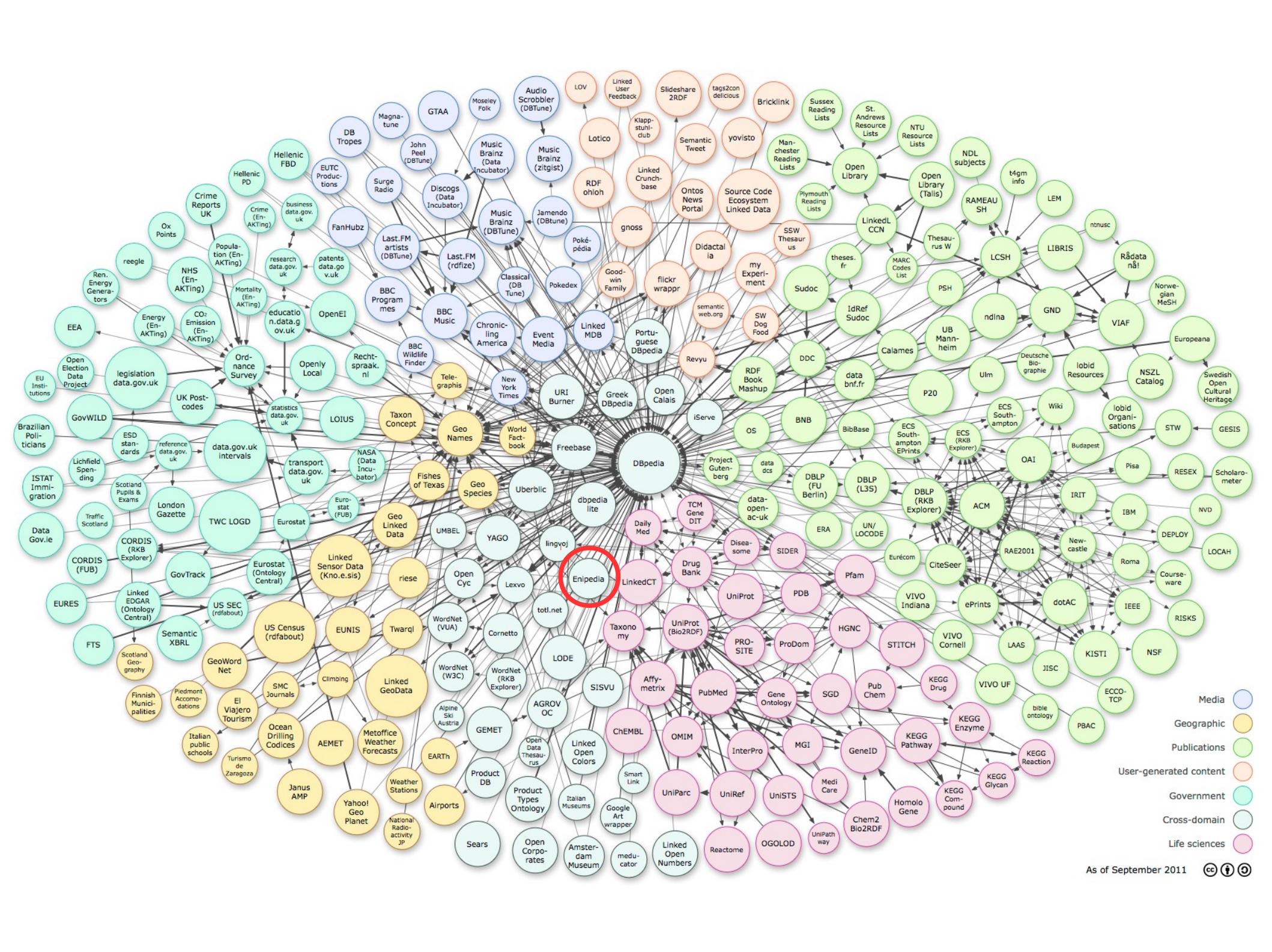
Semantic wiki technology is used to document the different chains that convert raw resources into products. This can set the basis for a collaborative approach in documenting processes for a Life Cycle Assessment.

Electricity Production by Fuel Type in NL

Data sourced from Carma.org is combined with information collected on fuel sources to show how the Netherlands generates most of its power.

Timeline of Investments in the Port of Rotterdam

Energy Calculations from CIA World Factbook

Facilities in the Port of Rotterdam

Legend:
- Media
- Geographic
- Publications
- User-generated content
- Government
- Cross-domain
- Life sciences

CARMA
Carbon Monitoring for Action

Start Here! Search CARMA by country, state, province, county, metro area, city, power company, power plant, or zip code.

[_____] Go

| Home | Power Plants | Power Companies | Geographic Regions | Blog |

Home > Plant overview

AMER

Company	RWE AG
Plant Location	Geertruidenberg, North Brabant
	Netherlands
	Europe

5,750,000 Tons CO2
6,671,600 MWh Energy
862 Intensity

Power Trends

For more about the terms or data used here, search the Glossary, learn All About Icons, or check out our FAQs. Information on plant specifics can be found here. If you use the data, please see our citation policy.

	Tons CO2	MWh Energy	Intensity
2004:	6,599,600	7,822,100	844
2009:	5,750,000	6,671,600	862
Future:	4,033,100	6,417,100	628

Top Power Producing Plants in the Region

Highest CO₂ Emitting Plants in the Region

See More
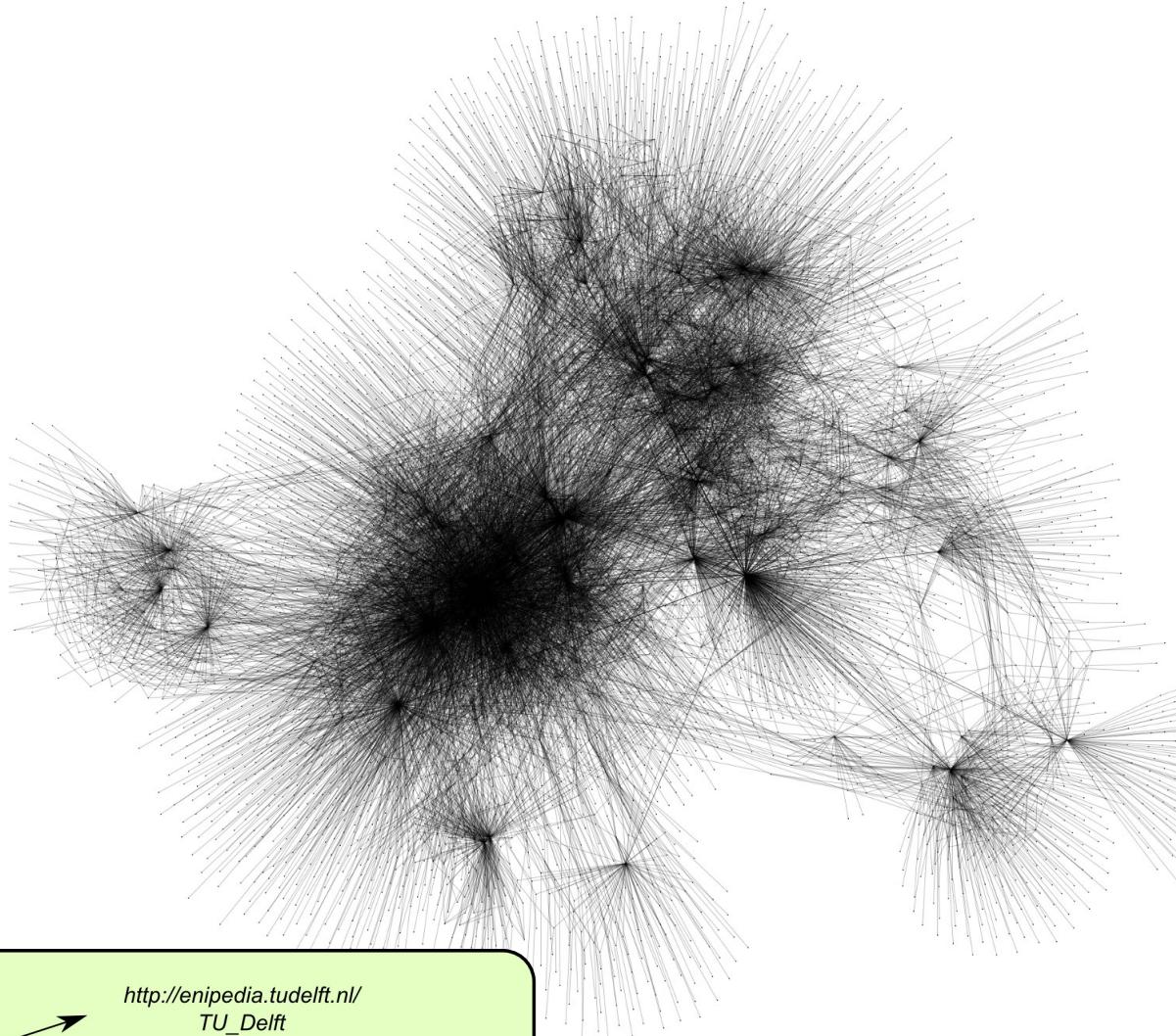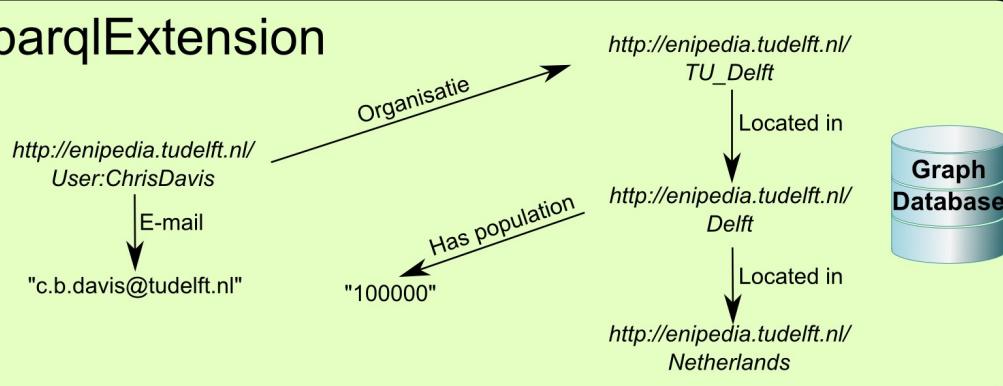
Carma.org

- Huge data set (70,000 plants)
- Diverse data - emissions, output, location, owner
- Largest power plant was in Finland (fixed now)

SparqlExtension

http://enipedia.tudelft.nl/
User:ChrisDavis

Organisatie → http://enipedia.tudelft.nl/
TU_Delft

E-mail → "c.b.davis@tudelft.nl"

Located in ↓

Has population → http://enipedia.tudelft.nl/
Delft

"100000"

Located in ↓

http://enipedia.tudelft.nl/
Netherlands

Graph Database
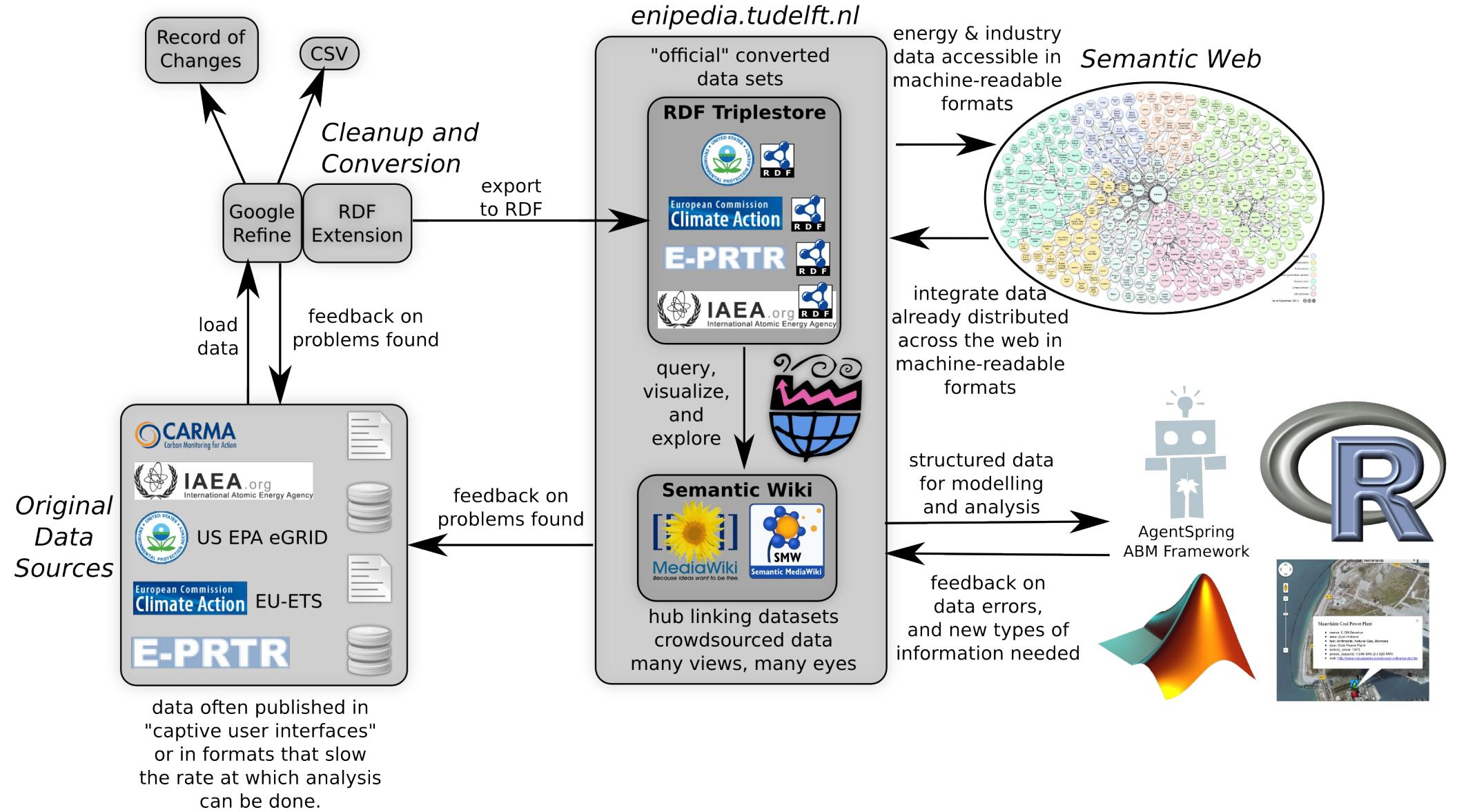
SPARQL Endpoint

(Query Structured Data from the Wiki) →

Structured data from the wiki

(RDF, XML, JSON, or CSV)

TUDelft

Delft University of Technology

enipedia.tudelft.nl

Record of Changes

CSV

Cleanup and Conversion

Google Refine

RDF Extension

export to RDF

"official" converted data sets

RDF Triplestore

European Commission
Climate Action

E-PRTR

IAEA.org
International Atomic Energy Agency

energy & industry data accessible in machine-readable formats

Semantic Web

load data

feedback on problems found

query, visualize, and explore

integrate data already distributed across the web in machine-readable formats

Original Data Sources

CARMA
Carbon Monitoring for Action

IAEA.org
International Atomic Energy Agency

US EPA eGRID

European Commission
Climate Action EU-ETS

E-PRTR

feedback on problems found

Semantic Wiki

MediaWiki
Because ideas want to be free.

SMW
Semantic MediaWiki

hub linking datasets
crowdsourced data
many views, many eyes

structured data for modelling and analysis

AgentSpring ABM Framework

feedback on data errors, and new types of information needed

data often published in "captive user interfaces" or in formats that slow the rate at which analysis can be done.

Navajo Powerplant

General

Name:	*(no data)*
Operator:	*(no data)*
Wikipedia page:	Navajo Generating Station
DBpedia page:	Navajo Generating Station
Year built:	1975
Owner company:	Salt River Project
Carma.org reference ID:	30650. (original data source)
eGRID reference ID:	4941

Location

City:	Page
Metro area:	Phoenix-mesa
State:	Arizona
Zip code:	85072
Country:	United States

Power Conversion Units

List of Power Conversion Units:

Add New Power Generating Unit:

- click here to create with auto-generated name
  - (creates incremental names - Unit 1, Unit 2, etc.)
- **or** below enter in a specific name (e.g. "Reactor Number 2")

[   ] [Add power conversion unit]

Map View

| Map | **Satellite** | Hybrid | Terrain |

Overlays

POWERED BY Google   2000 ft   500 m   Imagery ©2011 - Terms of Use

Energy

Fuel type:	Coal
Cooling method:	*(no data)*
Power plant type:	*(no data)*
Efficiency:	*(no data)* %
Power output (electrical): (nameplate capacity)	*(no data)*
Power output (thermal): (nameplate capacity)	*(no data)*
Operating cost:	*(no data)*

Facts about Navajo Powerplant ⓘ RDF feed ⚡

Carbonemissions	18,325,131.93 t (18,325,131,929.61 kg) + 🔍
Carbonemissions2000	18,234,413.455 t (18,234,413,454.711 kg) + 🔍
Carbonemissionsnextdecade	18,415,850.405 t (18,415,850,404.509 kg) + 🔍
CarmaId	30650 + 🔍
City	Page + 🔍
Congdist	John Shadegg + 🔍
Continent	North America + 🔍
Country	United States + 🔍
County	Maricopa + 🔍
DBpedia Page	http://dbpedia.org/resource/Navajo_Generating_Station 🔗 + 🔍
EGRID ID	4,941 + 🔍
Energyoutput	19,100,000 MWh (29,895,652,170.924 BigMacs) + 🔍
Energyoutput2000	18,100,000 MWh + 🔍
Energyoutputnextdecade	19,300,000 MWh + 🔍
Fuel type	Coal + 🔍
Intensity	957.08 kg (0.957 t) + 🔍
Intensity2000	1,009.697 kg (1.01 t) + 🔍
Intensitynextdecade	957.08 kg (0.957 t) + 🔍
Isocountry	USA + 🔍
Latitude	36.913 + 🔍
Longitude	-111.392 + 🔍
Metroarea	Phoenix-mesa + 🔍
Name	NAVAJO + 🔍
Owl:sameAs	http://dbpedia.org/resource/Navajo_Generating_Station 🔗 + 🔍, and http://enipedia.tudelft.nl/data/eGRID/Plant/4941 🔗 + 🔍
Ownercompany	Salt River Project + 🔍
Point	36.9125 N, 111.3917 W + 🔍
State	Arizona + 🔍
Wikipedia page	http://en.wikipedia.org/wiki/Navajo_Generating_Station 🔗 + 🔍
Year built	1,975 + 🔍
Zipcode	85072 + 🔍

Page | Discussion

Read | Edit with form | Edit | View history | ▼ | [_____] | Go | Search

Edit Powerplant: Amer Powerplant

| General | Advanced |

Power plant

General

Operator:	Essent N.V.
Wikipedia page:	http://en.wikipedia.org/wiki/Amercentrale
Year first built:	1980
Owner company:	Essent
Status:	Deutsche **Essent** Gmbh
CarmaID	**Essent** Milieu
(id used on carma.org ⧉)	**Essent** N.V.
EU ETS ID	868

Location

City:	Geertruidenberg
Metro area:	
State:	Noord-brabant
County:	
Zip code:	
Country:	Netherlands
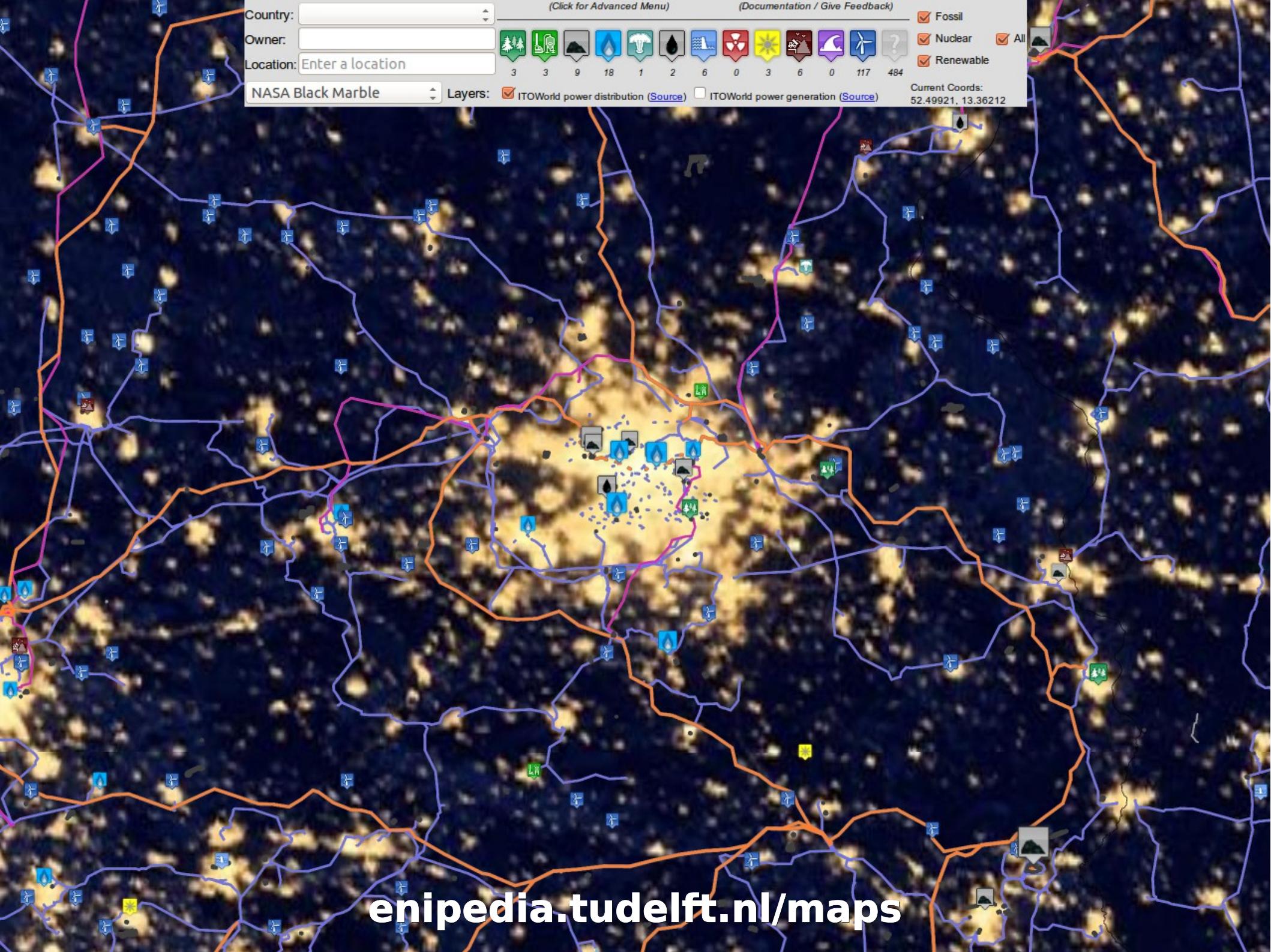
Power Conversion Units

Details about individual power conversion units (e.g. units composed of their own set of boilers, turbines, generators) can be specified in more detail by clicking on save, and then following the instructions on the "Power Conversion Units" section of the wiki page.

Type in an address, or click on the map to move the marker

Coordinates:

| 51.710730368168, 4.8433399200439 | Update map |
| Enter address here | Look up coordinates |

Energy

Fuel type: **Cooling method:** **Power plant use:**

Navigation

- Main page
- Enipedia Blog
- Enipedia Mailing List
- Contact Us
- Add Reference
- Report a Problem
- Recent changes
- Random page
- Help

Portals

- Power Plants
- Direct Carbon Fuel Cells
- Renewable Energy
- Port of Rotterdam
- Oil
- Videos

Advanced

- Sparql Endpoint
- Advanced Topics

Toolbox

- What links here
- Related changes
- Upload file
- Special pages

Annual Electricity Production (Source: eGRID, measurements in MWh)

name	trend	1996	1997	1998	1999	2000	2004	2005	2007
MWh		13,297,129	15,102,096	16,484,808	16,874,933	18,096,243	17,734,190	17,030,674	17,616,339

Generation Percent (Source: eGRID, measurements in percent)

name	trend	1996	1997	1998	1999	2000	2004	2005	2007
Coal		99.8537	99.8179	99.8670	99.8974	99.8885	99.9239	99.8742	99.8908
Oil		0.1463	0.1821	0.1330	0.1026	0.1115	0.0761	0.1258	0.1092

Net Generation by Fuel Type (Source: eGRID, measurements in MWh)

name	trend	1996	1997	1998	1999	2000	2004	2005	2007
Coal		13,277,675	15,074,601	16,462,883	16,857,616	18,076,057	17,720,688	17,009,250	17,597,111
Oil		19,454.0	27,495.0	21,925.0	17,317.0	20,186.0	13,502.0	21,424.0	19,228.4

Emissions (Source: carma.org)
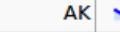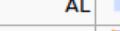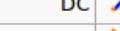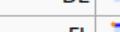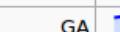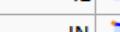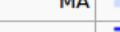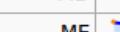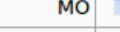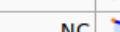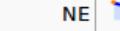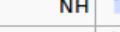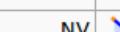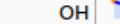
| Name	Amount	Year	Units
Carbon Dioxide	18,234,413,455	2000	kg
Carbon Dioxide	18,325,131,930	2007	kg
Carbon Dioxide	18,415,850,405	2020	kg

Annual Emissions (Source: eGRID, measurements in tons)

name	trend	1996	1997	1998	1999	2000	2004	2005	2007
CH4		-	-	-	-	-	-	446,094.3	458,255.8
CO2		14,341,710	17,156,377	19,800,996	19,499,180	20,137,721	20,237,545	19,677,241	20,178,992
N2O		-	-	-	-	-	-	669,141.5	686,400.3
NOx		25,245.00	34,296.00	39,157.80	35,275.24	37,267.01	33,985.68	33,221.26	35,252.94
SO2		63,878.00	66,230.00	39,845.60	9,162.60	4,837.10	3,918.39	3,943.91	4,436.80

ton CO2/MWh per state

name	trend	1996	199
AK		0.7866408	0.73
AL		1.0729445	1.09
AR		1.1320019	1.12
AZ		1.1225671	1.11
CA		0.7461570	0.73
CO		1.1289670	1.24
CT		1.0013961	0.97
DC		0.5638427	1.43
DE		1.4206414	1.08
FL		0.9641245	0.96
GA		1.1233446	1.14
HI		0.8856505	0.92
IA		1.272581	1.2
ID		1.3054466	1.30
IL		1.150678	1.1
IN		1.215465	1.2
KS		1.220004	1.1
KY		1.173276	1.1
LA		0.9492622	0.90
MA		0.9228521	0.91
MD		1.1156948	1.12
ME		1.0609585	0.99
MI		1.0205069	1.02
MN		1.219044	1.2
MO		1.152154	1.1
MS		1.0828998	1.22
MT		1.141739	1.2
NC		1.0985647	1.07
ND		1.343287	1.3
NE		1.230613	1.2
NH		1.1161865	1.16
NJ		1.2516628	0.85
NM		1.1560940	1.16
NV		1.4983544	1.03
NY		0.7283189	0.77
OH		1.1205714	1.09

A tale of one (or four?) power stations and seven data sets

Wilton Power Stations

LCPD
- Thermal Output (MW)
- Fuel Consumption (TJ)

EU ETS
- CO$_2$ Emissions
- Emissions Allowances

Wilton

DUKES (DECC)
- Capacity (MW)
- Fuel Type
- Year Begin

How the European Commission manages data

Year	Plant Number	Plant name	Plant location	MWth	Biomass (TJ)	Other solid fuels (TJ)	Liquid fuels (TJ)	Natural gas (TJ)	Other gases (TJ)	SO2 (t)	NOx (t)	Dust (t)
2007	94	Wilton	SembCorp Utilities, Wilton P Stn		418.598	5576.306	143.159	606.001	0	5303.3	3446.1	129.4
2008	60	Sembcorp Utilities U.K Ltd Wilton	England	714	0	8302.55	10.214	1161.335	0	2570	1456.2	211.7
2008	204	Sembcorp Utilities U.K Ltd Wilton	England	100	2139	0	0	88.492	0	1.1	135.1	4.6
2008	205	Sembcorp Utilities U.K Ltd Wilton	England	100	0	0	0	50.672	0	0.01	0.1	0.02
2009	74	Sembcorp Utilities U.K Ltd Wilton	England	714	0	4246.646	3.562	5647.128	0	1164.2495795525	992.0418719788	87.9914050096
2009	268	Sembcorp Utilities U.K Ltd Wilton 2	England	100	2669	0	0	100.07	0	2.01	193.6	1.07
2009	269	Sembcorp Utilities U.K Ltd Wilton 3	England	100	0	0	0	204.3	0	0.17	0.17	0.29

Large Combustion Plants Directive
http://ec.europa.eu/environment/air/pollutants/stationary/lcp/legislation.htm

TUDelft

Delft University of Technology

Entity	Data Sources						
(based on most commonly encountered name)	SembCorp Website	UK Department of Energy and Climate Change	EU ETS	E-PRTR	Large Combustion Plant Directive	Carma.org (subset of WEPP)	Wikipedia
(entire site, data for all units aggregated together)							
Wilton Power Stations			• SembCorp Utlilities Teesside Power Station *(This is likely without Wilton 10 as it burns biomass. Without the inclusion of the owner name, this could be confused with the other Teesside Power Station)*	• Sembcorp Utilities (uk) Ltd • Sembcorp Utilities (uk) Ltd Wilton 10 Power Station • Sembcorp Utilities (uk) Ltd, Wilton Power Station *(There's only one entry for a facility named Wilton that is owned by SembCorp. The labeling of this as Wilton 10 is likely wrong as mentioned in the discussion below)*		• Wilton Cogen *(aggregation not clear, are other units included?)*	• Wilton power stations • (a.k.a. SembCorp power station)
(Power stations within the site)							
Wilton Power Station *(main station)*	• Wilton Power Station	• Wilton Power Station			• Sembcorp Utilities U.K Ltd Wilton		• Wilton Power Station
Wilton GT2	• Wilton GT2	• Wilton GT2			• Sembcorp Utilities U.K Ltd Wilton • Sembcorp Utilities U.K Ltd Wilton 3		
Wilton 10	• Sembcorp Biomass Power Station • Wilton 10	• Wilton 10			• Sembcorp Utilities U.K Ltd Wilton • Sembcorp Utilities U.K Ltd Wilton 2		• Wilton 10
Wilton 11 *(planned)*							• Wilton 11 *(planned)*

How to manage this?

http://www.flickr.com/photos/maxbraun/98688824/

http://www.flickr.com/photos/acme/229065626/

TUDelft

Delft University of Technology
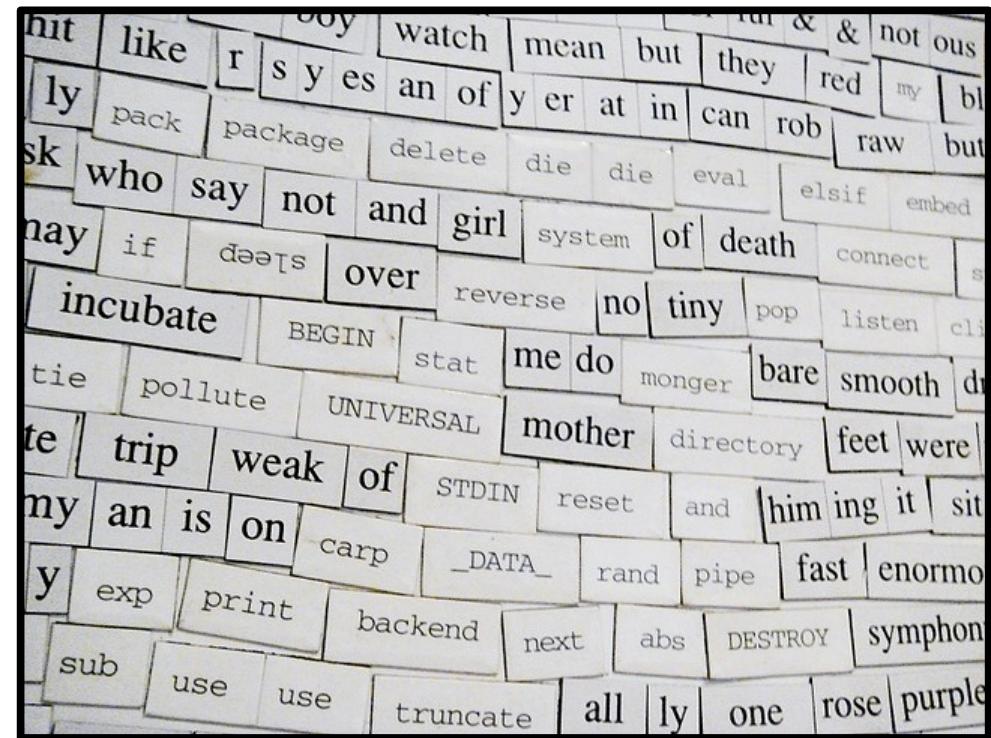
Further documentation and development notes can be found here.
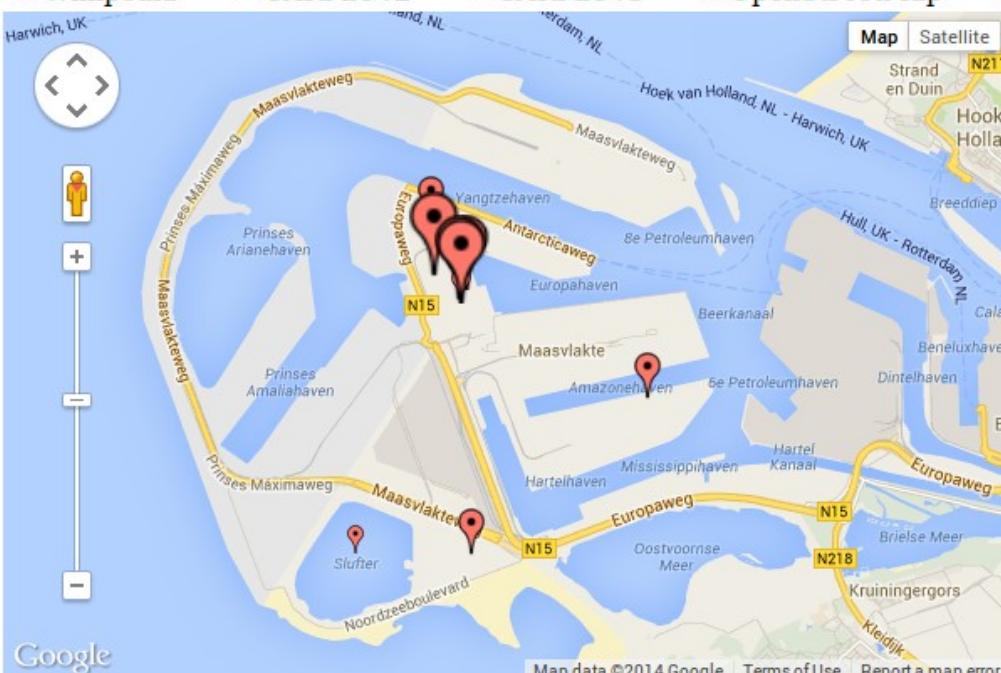
☑Wikipedia ☐CARMA v2 ☐CARMA v3 ☑OpenStreetMap ☑IndustryAbout ☑EU ETS ☑E-PRTR ☑LCPD

10 of 13 hits returned in 0.073 seconds

Score	Source	ID	building	changeset	lat	uid	power	generator:type	sour
3.1615386	OpenStreetMap	way/54115557	industrial	15925124	51.958459123076935	36080	generator	steam_turbine	3dSha
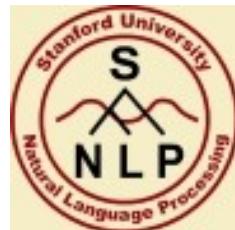
Score	Source	ID	building	changeset	lat	uid	power	generator:type	source	tir
3.140545	OpenStreetMap	way/54115421	industrial	15925124	51.959003	36080	generator	steam_turbine	3dShapes	2013-04

Score	Source	ID	building	changeset	lat	uid	power	generator:type	t
3.1240933	OpenStreetMap	way/171988323	industrial	15925124	51.96205817777778	36080	generator	steam_turbine	2013-0

Score	Source	ID	x	name	facilityID	country	lat	long	street
2.8561406	eprtr	6306	http://prtr.ec.europa.eu	E.On	6306	http://prtr.ec.europa.eu	51.960030	4.02681000	Colora

Next Steps?

- Named graphs preserving original data, provenance
- SKOS + Crowdsourcing to the rescue?
  - owl:sameAs doesn't work
  - skos:mappingRelation
  - skos:broadMatch
  - skos:relatedMatch
  - skos:linkTheseAndFigureOutWhatsGoingOnLater
- Annotation graphs to facilitate linking
  - Matching scores, verification, votes, notes, etc.
- Systems that get smarter the more that they're used
- Tools that support both soup and structure, humans & computers

Next Steps?
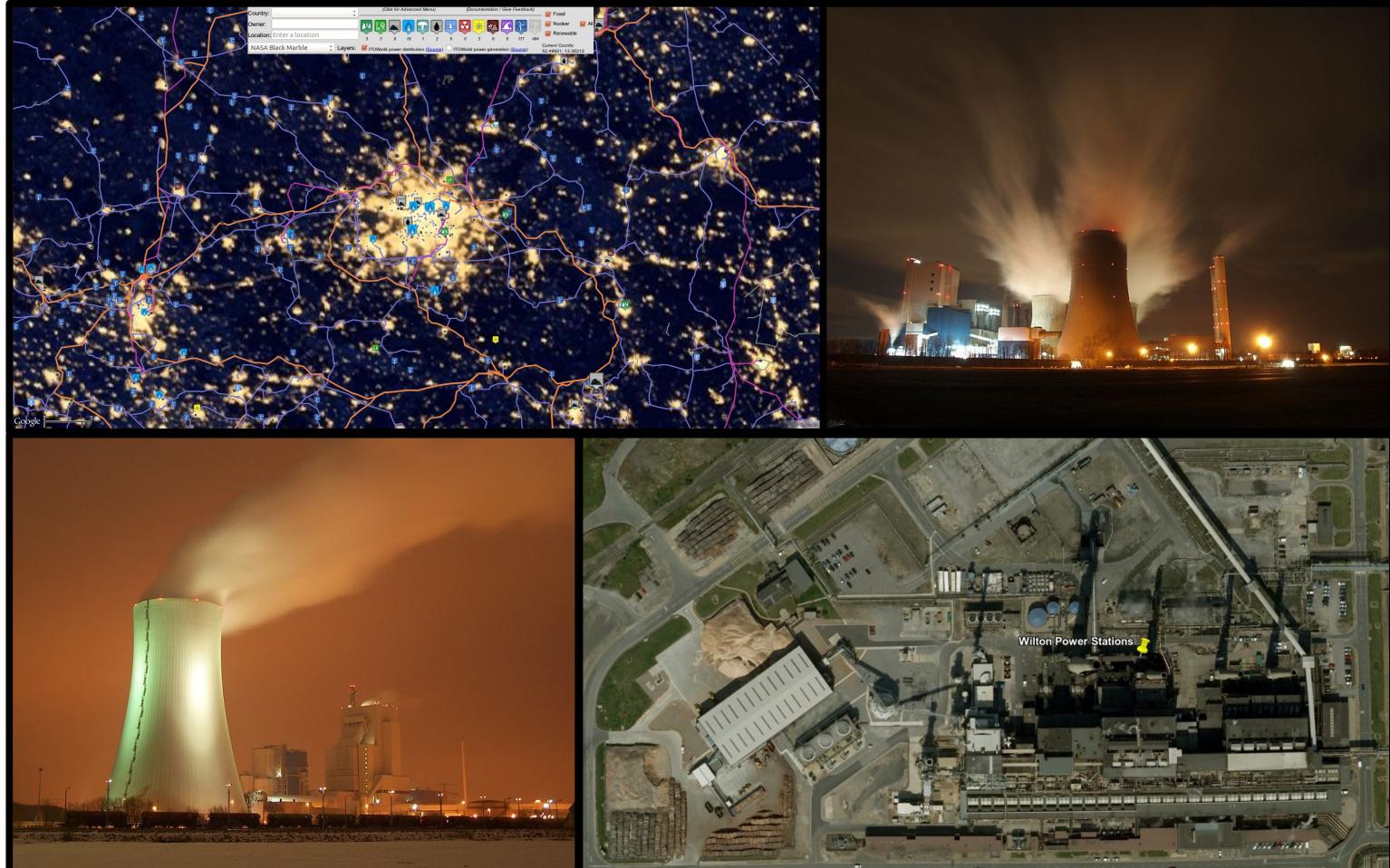
Big Data?

|	Big Data	Tedious Data
Size	>TBytes	<GBytes
Collection	Automated	Manual, difficult to automate
Formats	Machine readable	Human readable
Sources	Centralized data	Widely distributed data
Tools	Hadoop, MapReduce, NoSQL	Excel, MS Access, PDF
Challenges	Software, Hardware	Social, Institutional
Information	Simple Facts	Complex Knowledge
Processing Speed	>MBytes/sec	<5 words/sec

Conclusions
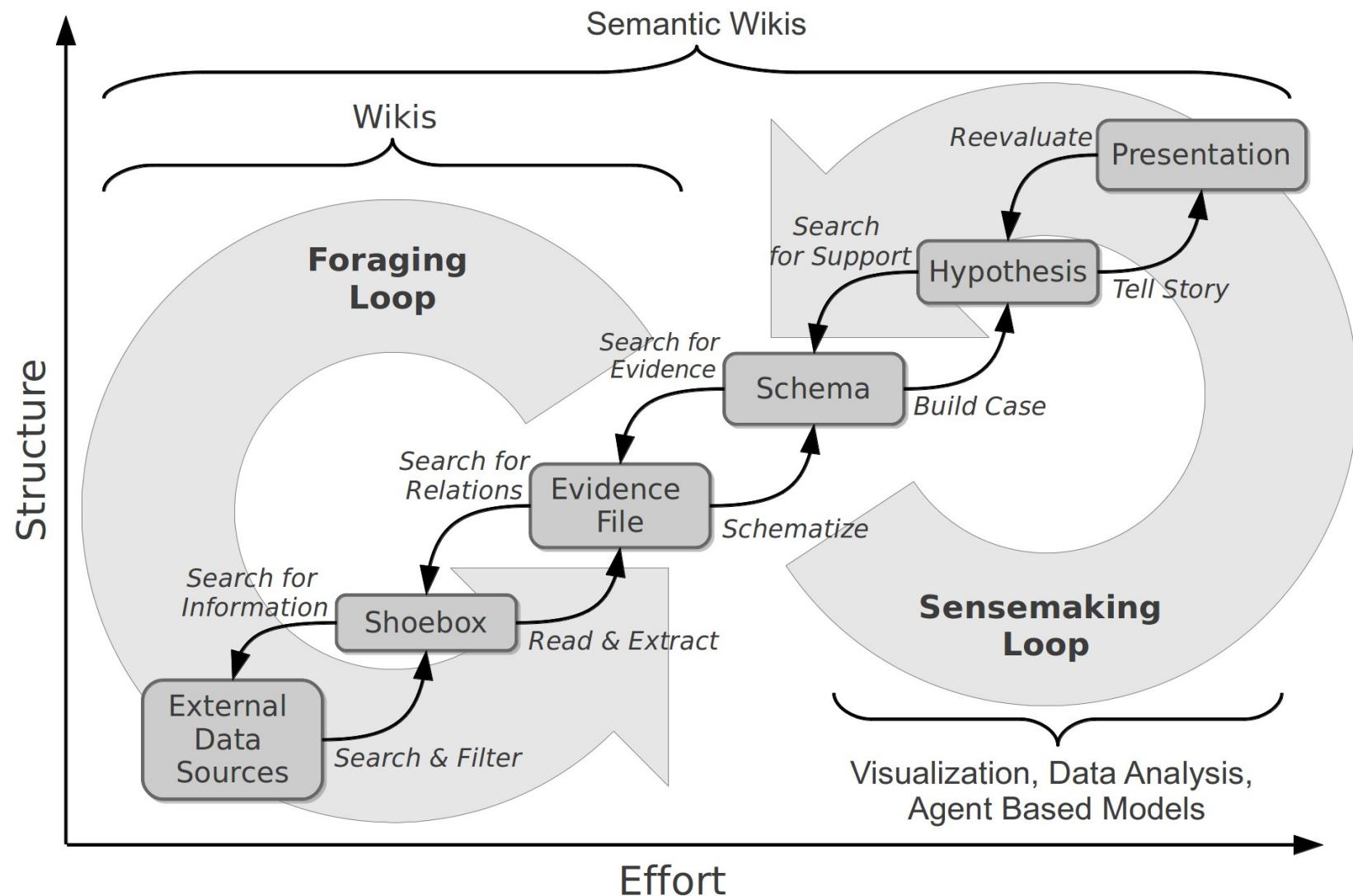
- Graph databases work great as a backend
- The issues (I face) are social
  - Lack of consistent data publishing standards
  - Lack of standard identifiers
  - Lack of technical skills
- Hoping Semantic CMS, NER, etc offer a better way forward

Questions?

Chris Davis
http://enipedia.tudelft.nl
c.b.davis@tudelft.nl

TUDelft

Delft University of Technology
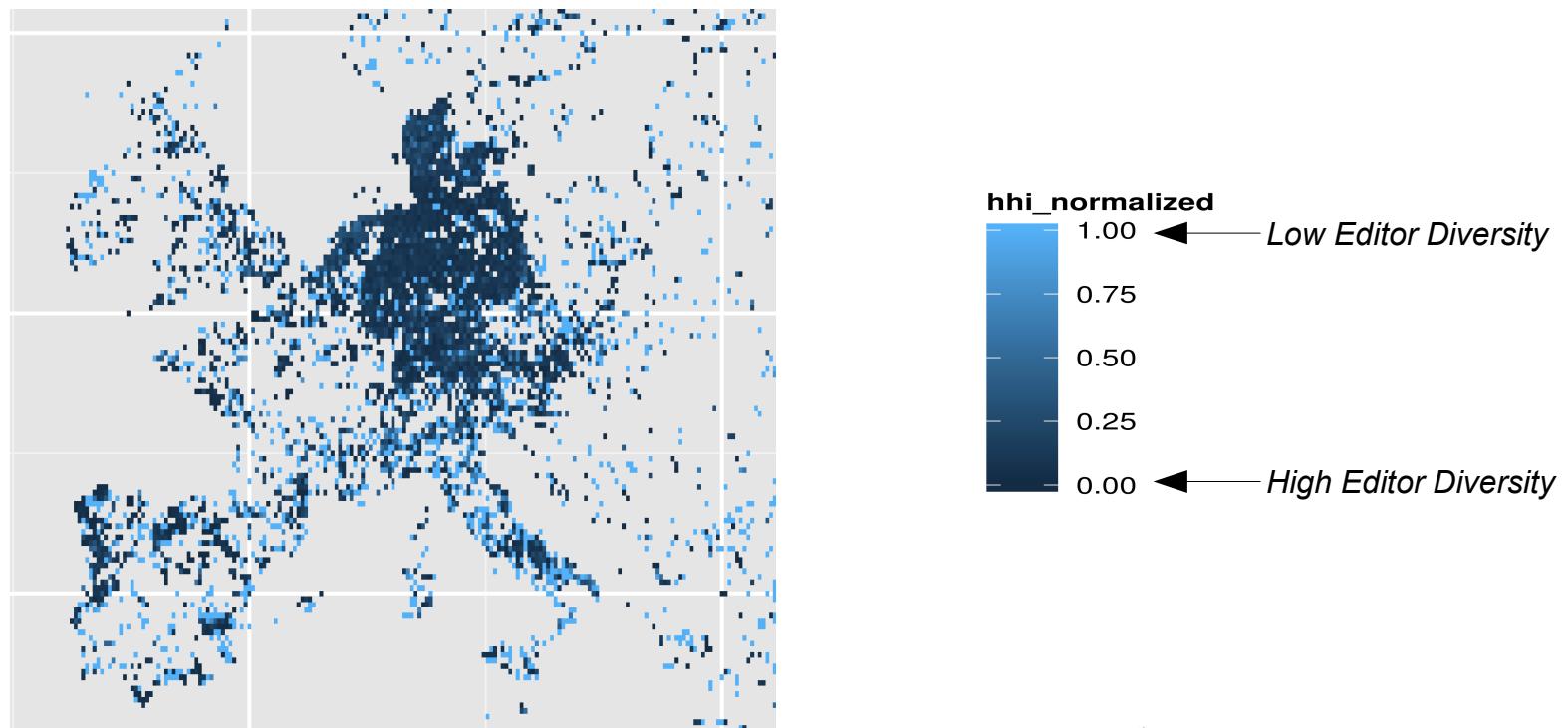
Information Supply Chains

Pirolli & Card (2005) *The Sensemaking Process and Leverage Points for Analyst Technology as Identified Through Cognitive Task Analysis*
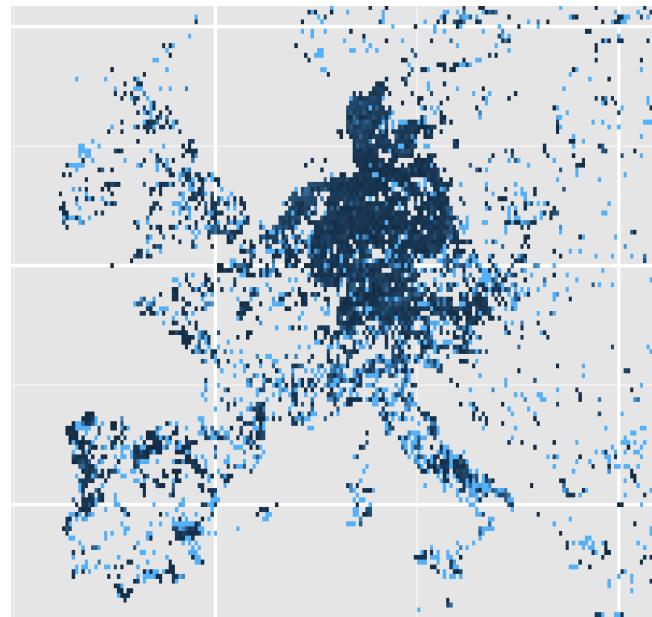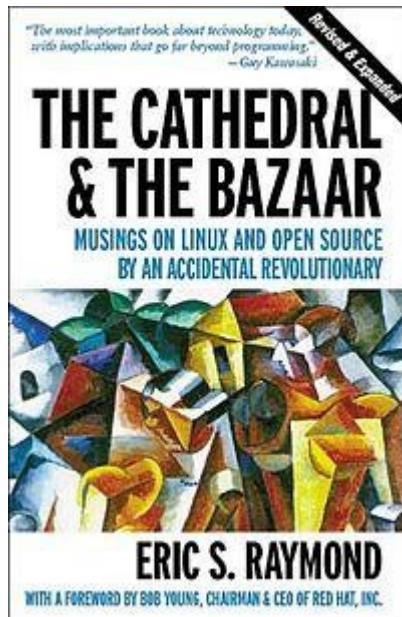
# How to Measure Data Quality?

Data Quality $=$ Researcher Skill/Experience $\times$ # Viewers/ Editors $\times$ Ease of Independent Verification

hhi_normalized

1.00 ← *Low Editor Diversity*

0.75

0.50

0.25

0.00 ← *High Editor Diversity*

TUDelft

Delft University of Technology

How to Measure Data Quality?

- Eric Raymond – "With many eyes all bugs are shallow"
- But… not all eyes are evenly distributed

Issues

- Regulation Driven Conceptualization
- Knowledge Reengineering Bottleneck
  - "The difficulty of the correct and continuous use of preexisting knowledge for a new task"

http://www.slideshare.net/rinkehoekstra/the-knowledge-reengineering-bottleneck

TU Delft

Delft University of Technology