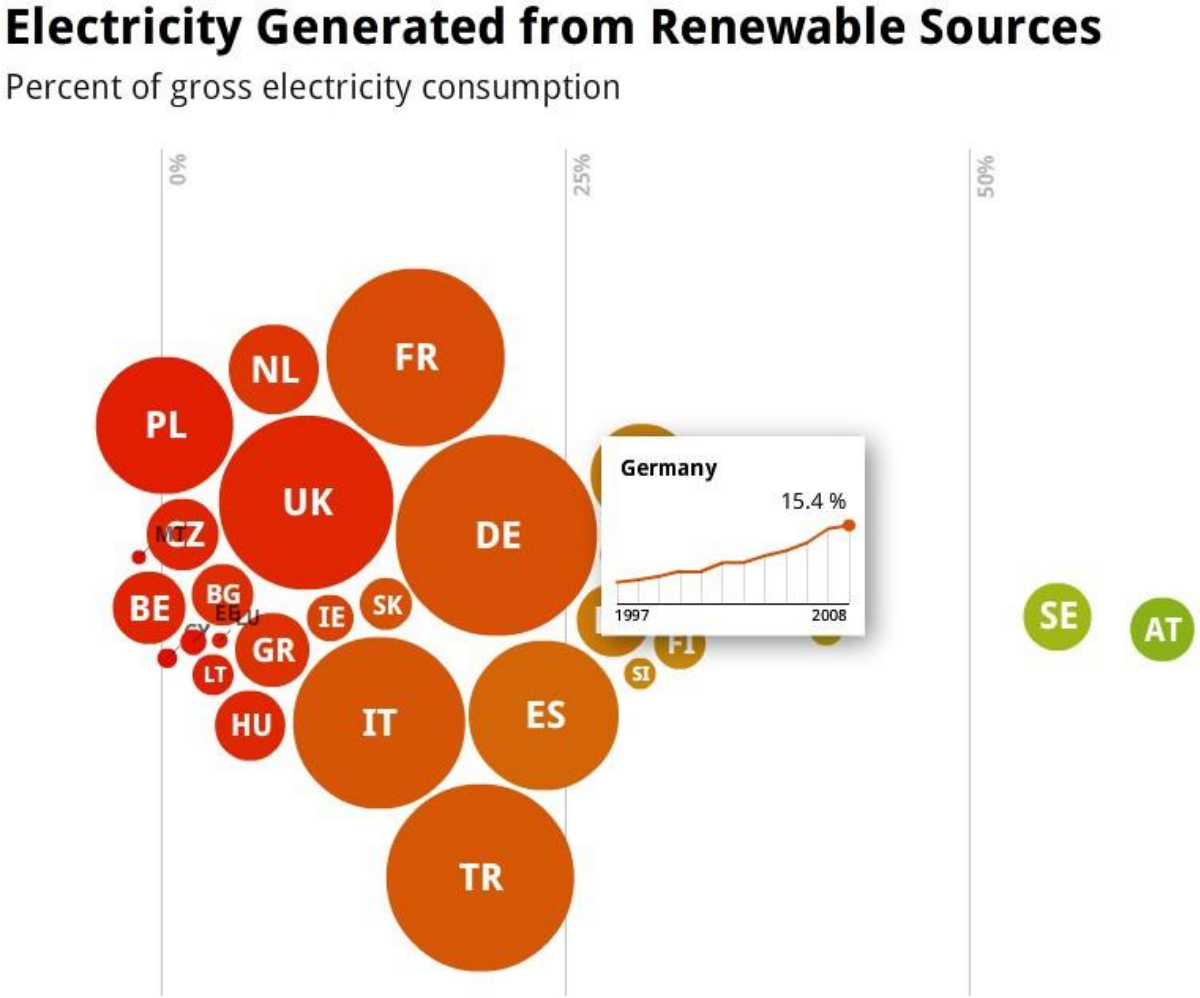
Unleashing climate and energy knowledge with Linked Open Data

Clean Energy Solution Center Webinar, 12.03.2013

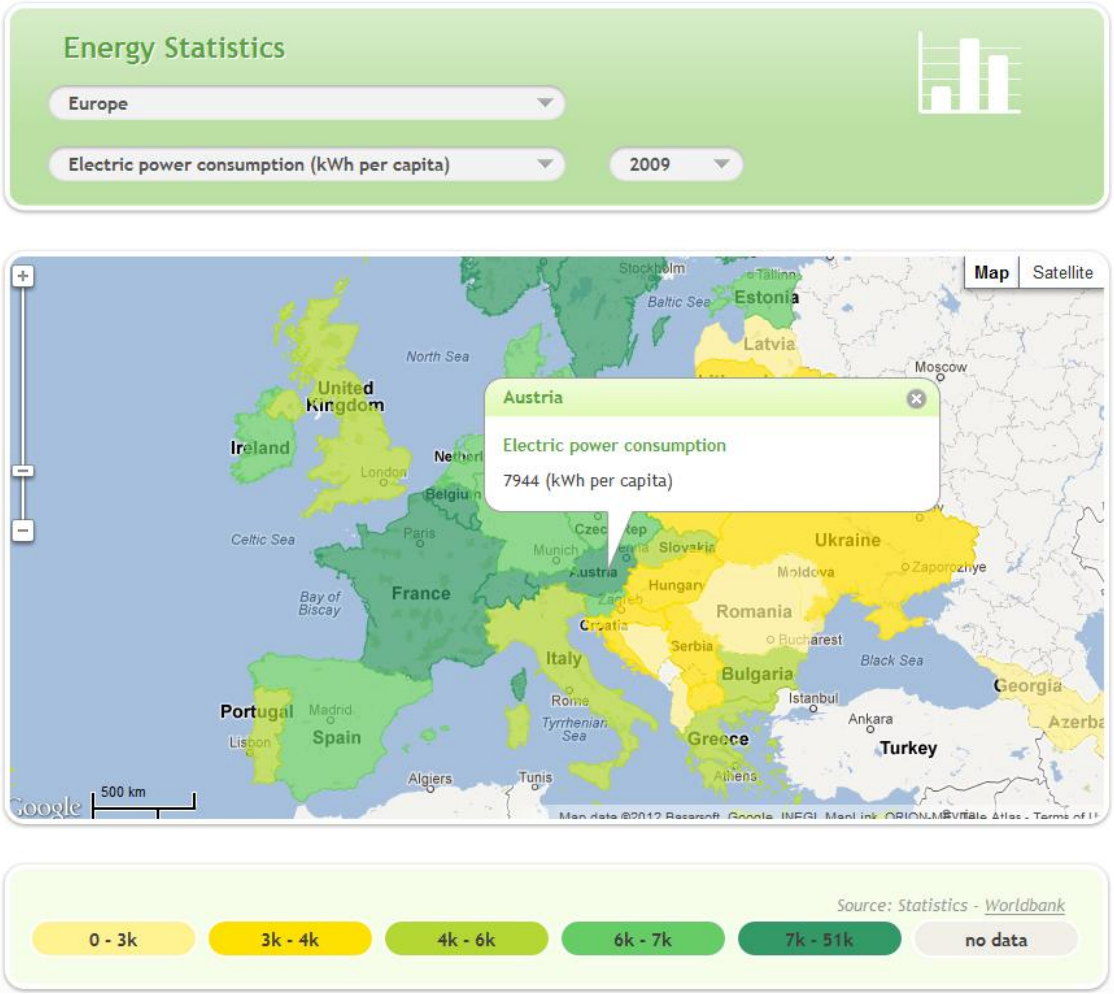
Florian Bauer, Operations and IT Director, REEEP
Jon Weers, Senior Web Applications Engineer, NREL

We use data for ...



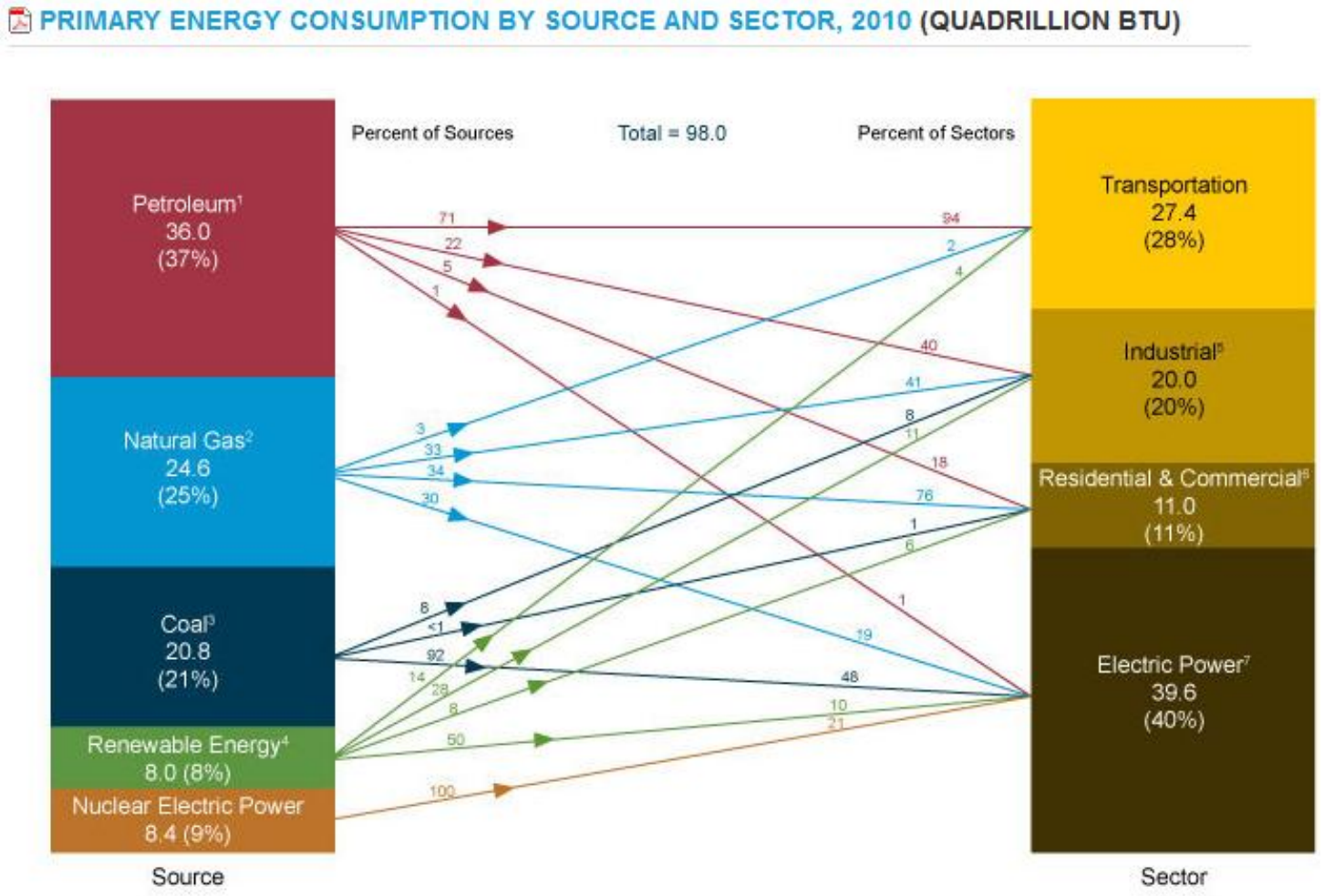
energy.publicdata.eu — Europe's energy

We use data for ...



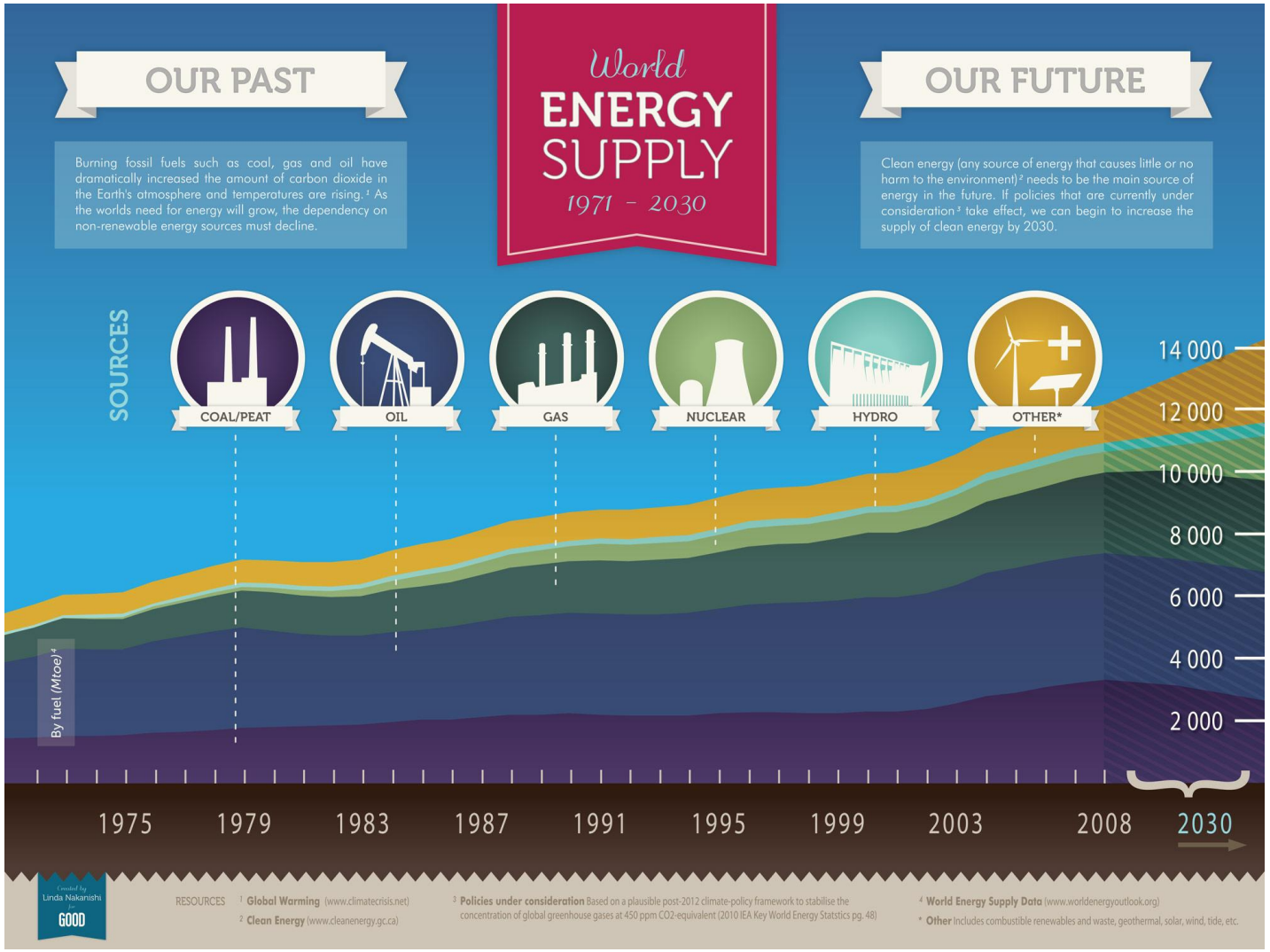
reegle.info – energy statics

We use data for ...



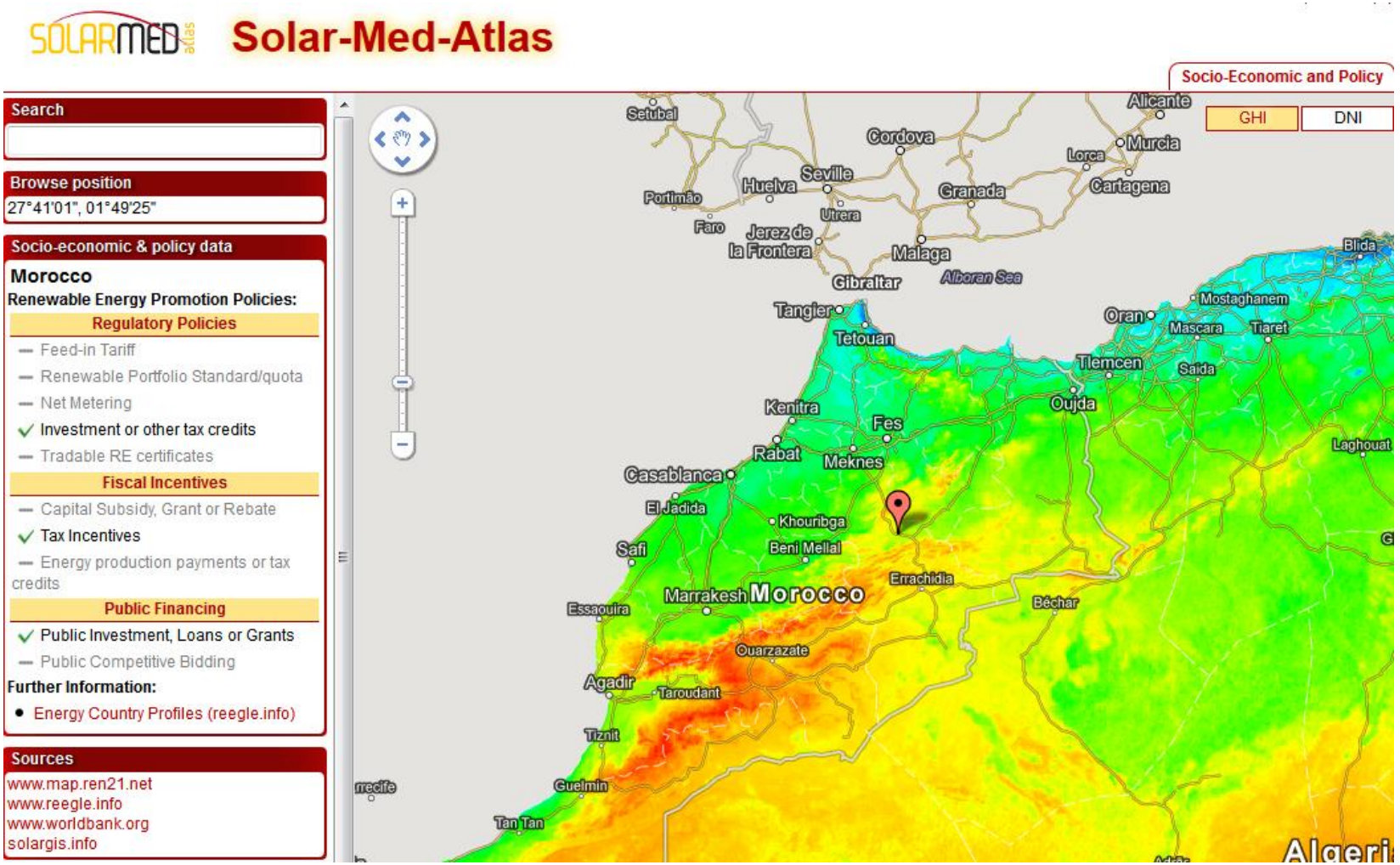
U.S. Energy Information Administration
Annual Energy Review 2010

We use data for ...



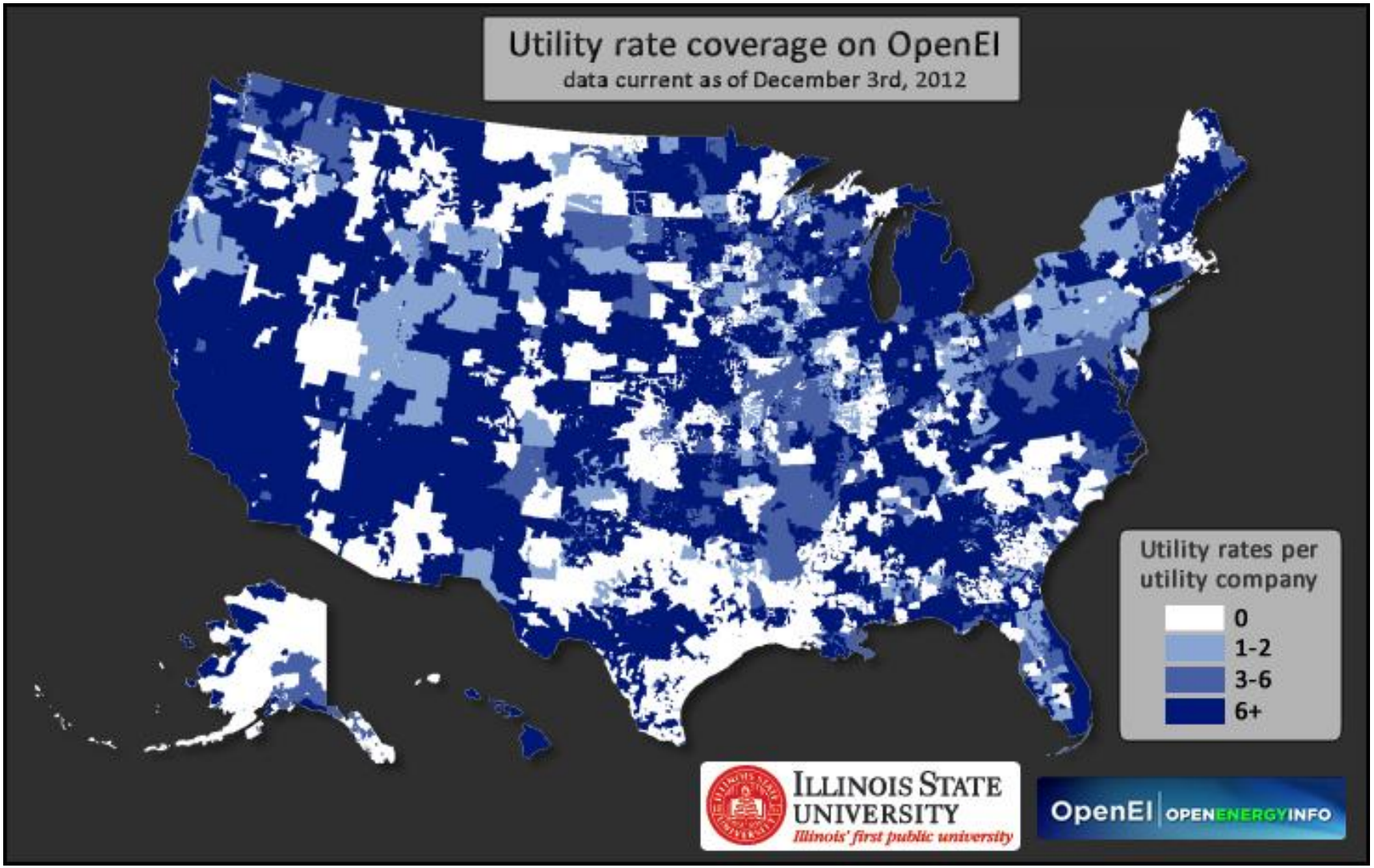
http://awesome.good.is/transparency/web/1012/energy-submissions/linda-nakanishi/flat.html

We use data for …



http://www.solar-med-atlas.org/solarmed-atla

We use data for ...



http://openei.org

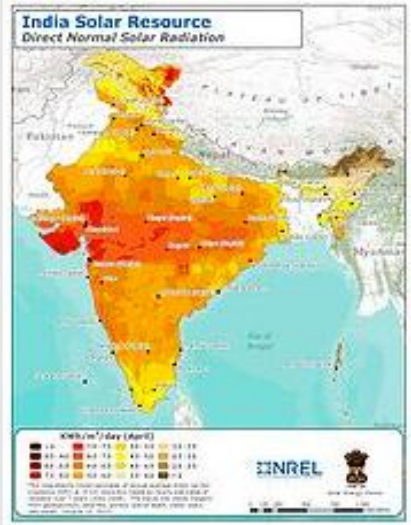
OpenEI | OPENENERGYINFO

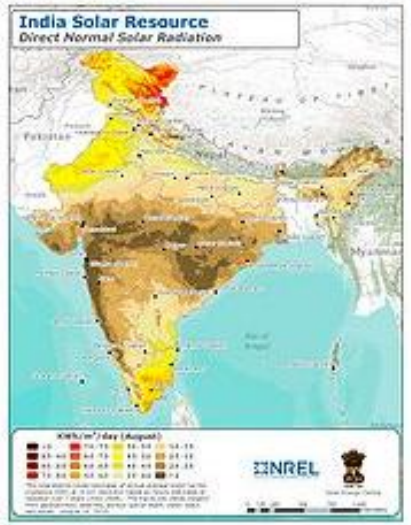
We use data for ...

Energy Resources

Resource	Value	Units	Rank	Period	Source
Wind Potential	0	Area(km^2) Class 3-7 Wind at 50m	120	1990	NREL
Solar Potential	9,877,095,200	MWh/year	8	2008	NREL
Coal Reserves	66,800.07	Million Short Tons	5	2008	EIA
Natural Gas Reserves	1,075,000,000,000	Cubic Meters (cu m)	26	2010	CIA World Factbook
Oil Reserves	5,800,000,000	Barrels (bbl)	23	2010	CIA World Factbook

Energy Maps featuring India





More Maps..

http://openei.org/wiki/Spain

Great, but how to get the data?

… and why do we need LOD?

"Every day, we create 2.5 quintillion bytes of data — so much that 90% of the data in the world today has been created in the last two years alone." - IBM, 2012 (http://www-01.ibm.com/software/data/bigdata)

Usual way to store and share data

- **Stores all information in its own database**

Source: Jon Weers, NREL

Usual way to store and share data

- **Stores all information in its own database**
- Other sites have similar design pattern

  => Duplication of effort and information

Source: Jon Weers, NREL

Using data from another site requires you to download a copy of it to install into your database.

Usual way to store and share data

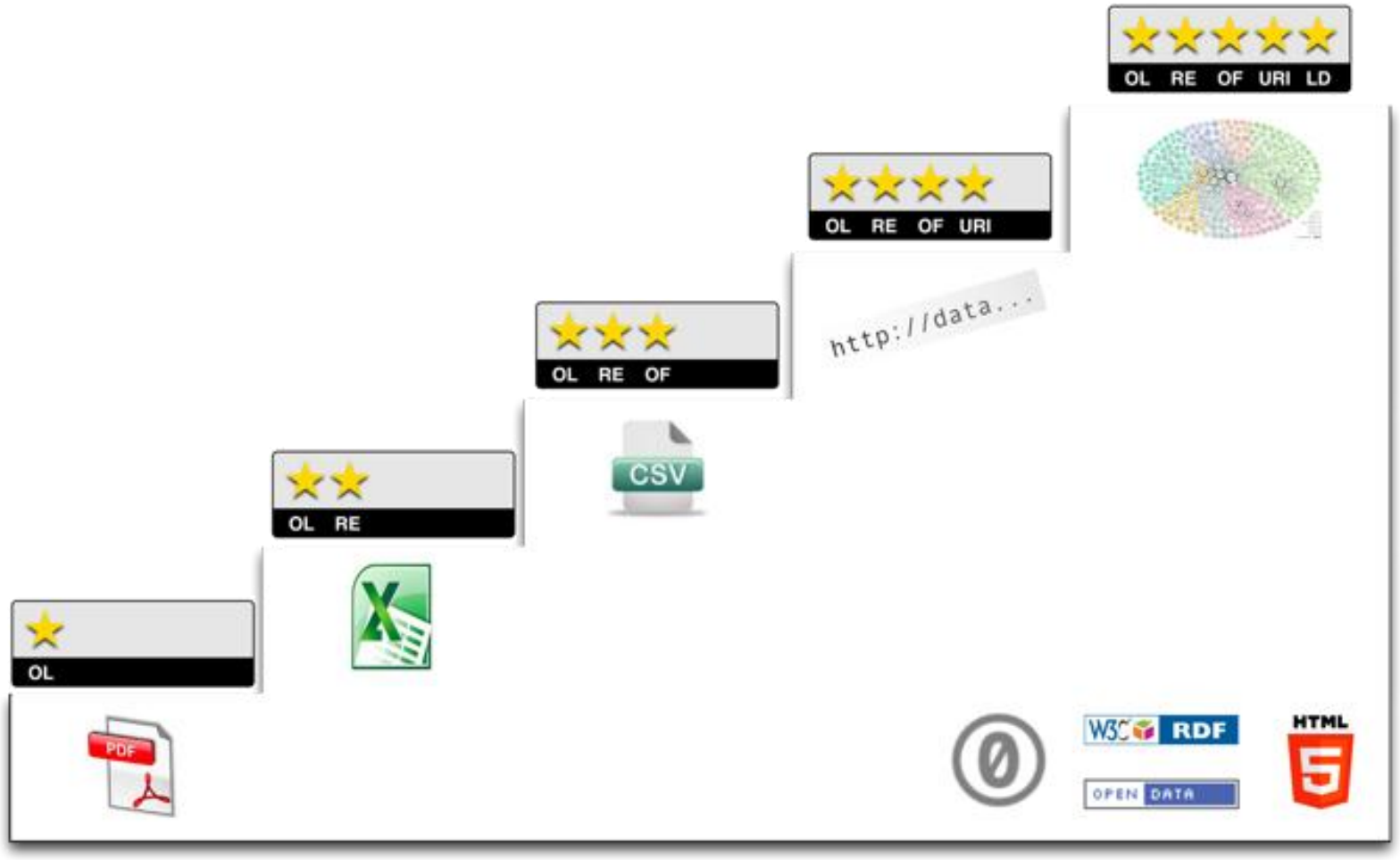
- **Stores all information in its own database**
- Other sites have similar design pattern
  => Duplication of effort and information
- Both sites responsible for updating information

  => Potential for online community to be presented with conflicting information

Source: Jon Weers, NREL

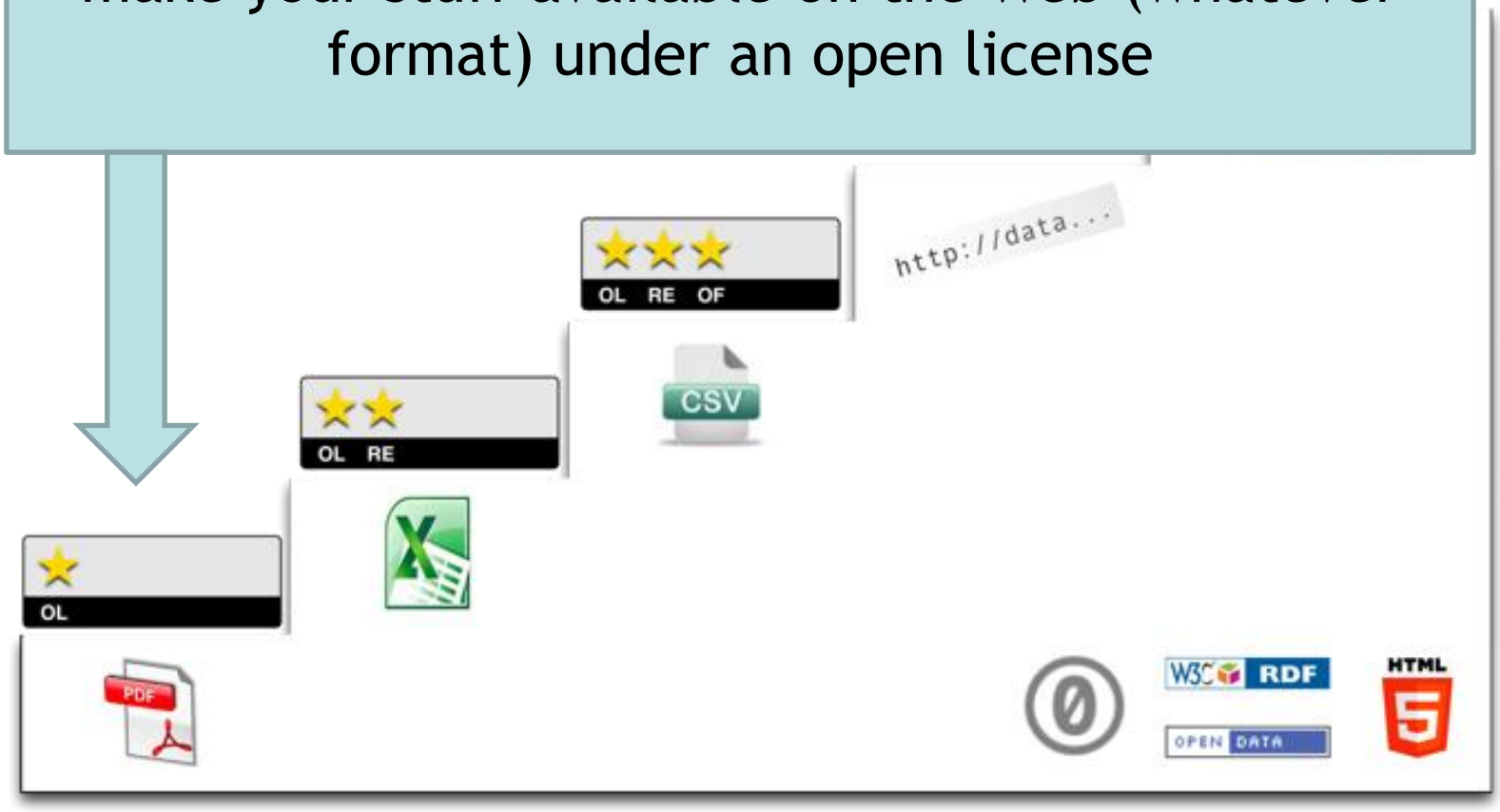
If the original site updates its data, the two sites become out of sync. How does the online community know which site is more accurate?

The 5 steps to Linked Open Data



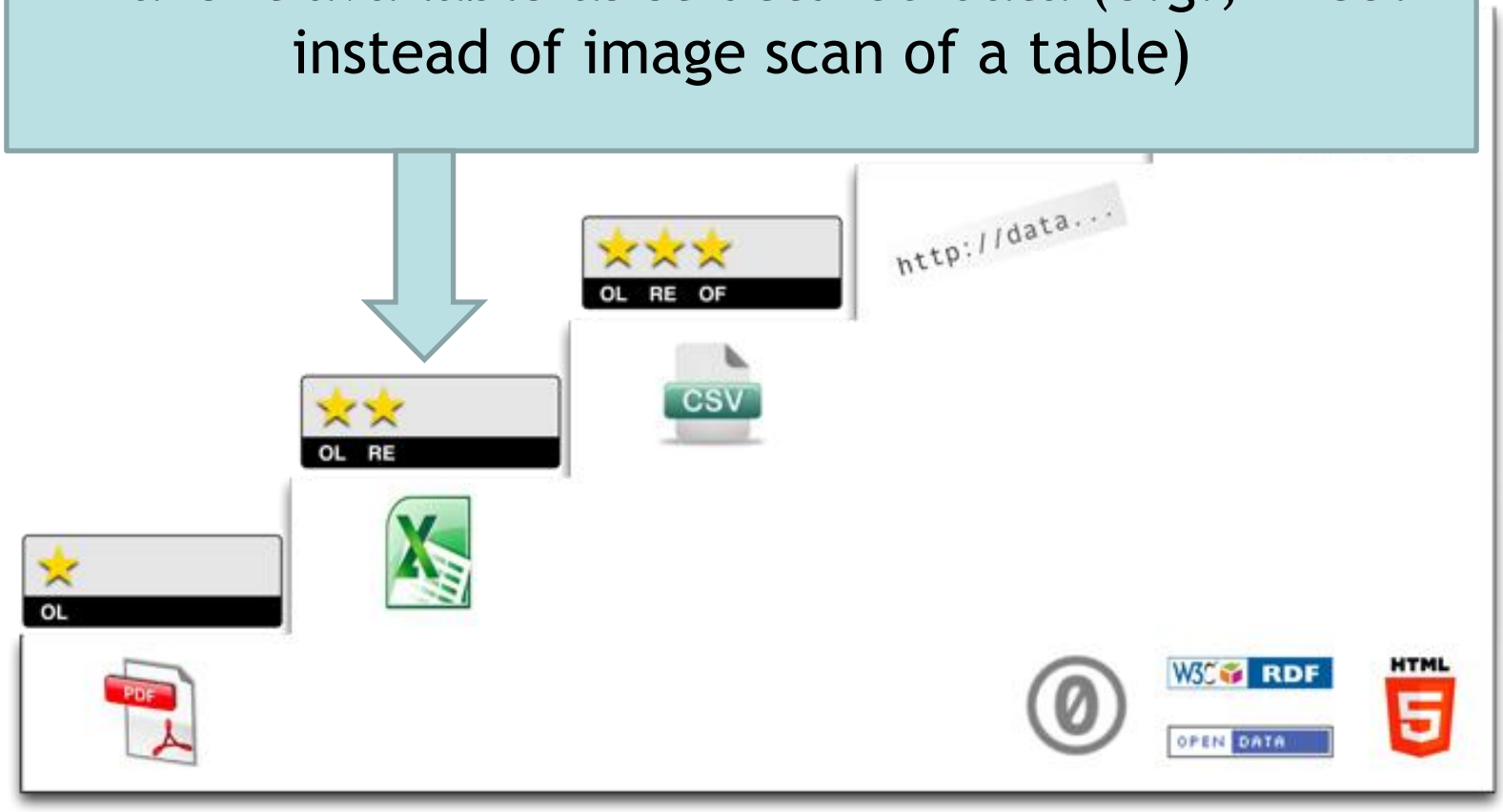
The 5 steps to Linked Open Data

make your stuff available on the Web (whatever format) under an open license



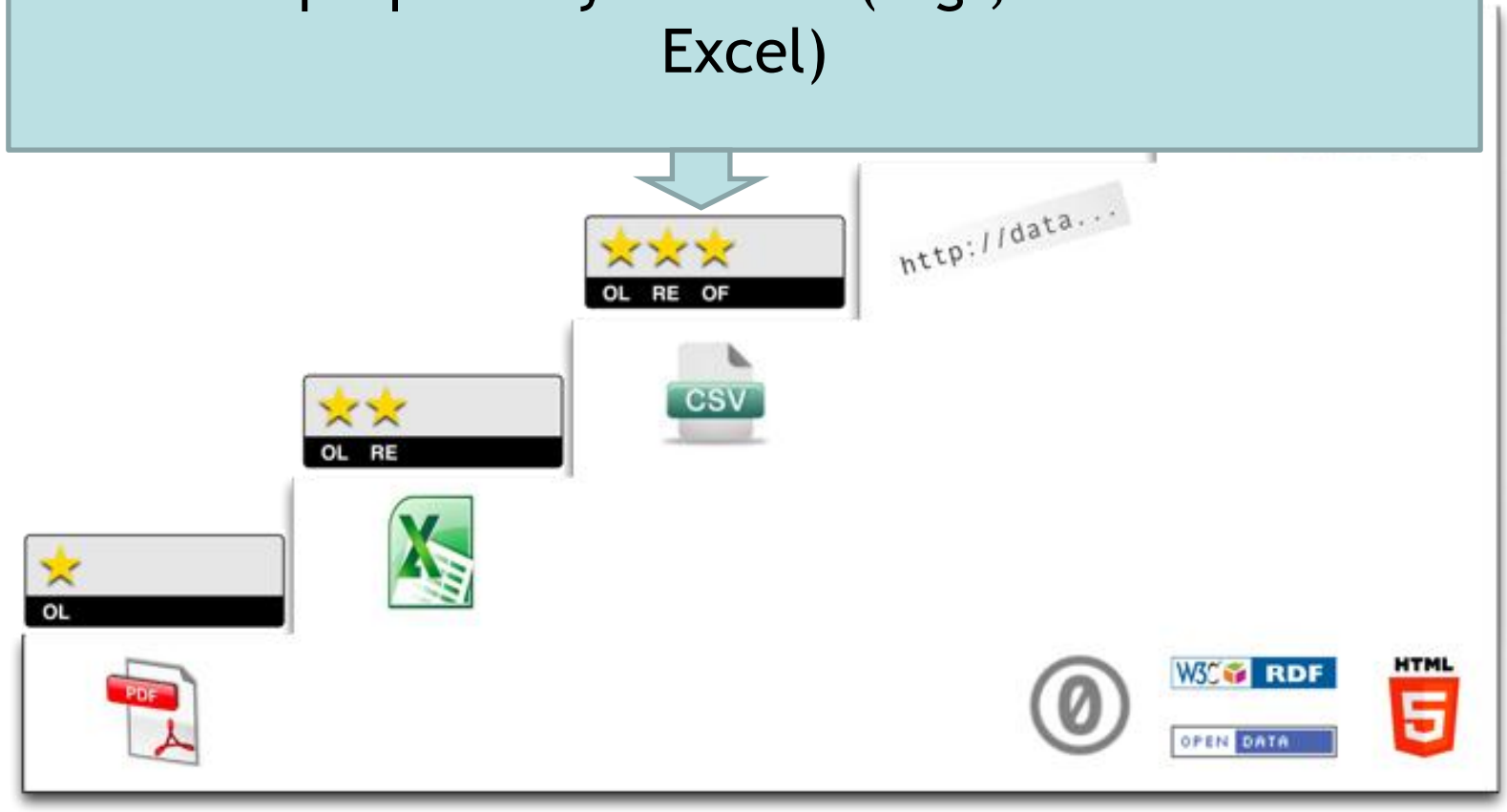
The 5 steps to Linked Open Data

make it available as structured data (e.g., Excel instead of image scan of a table)

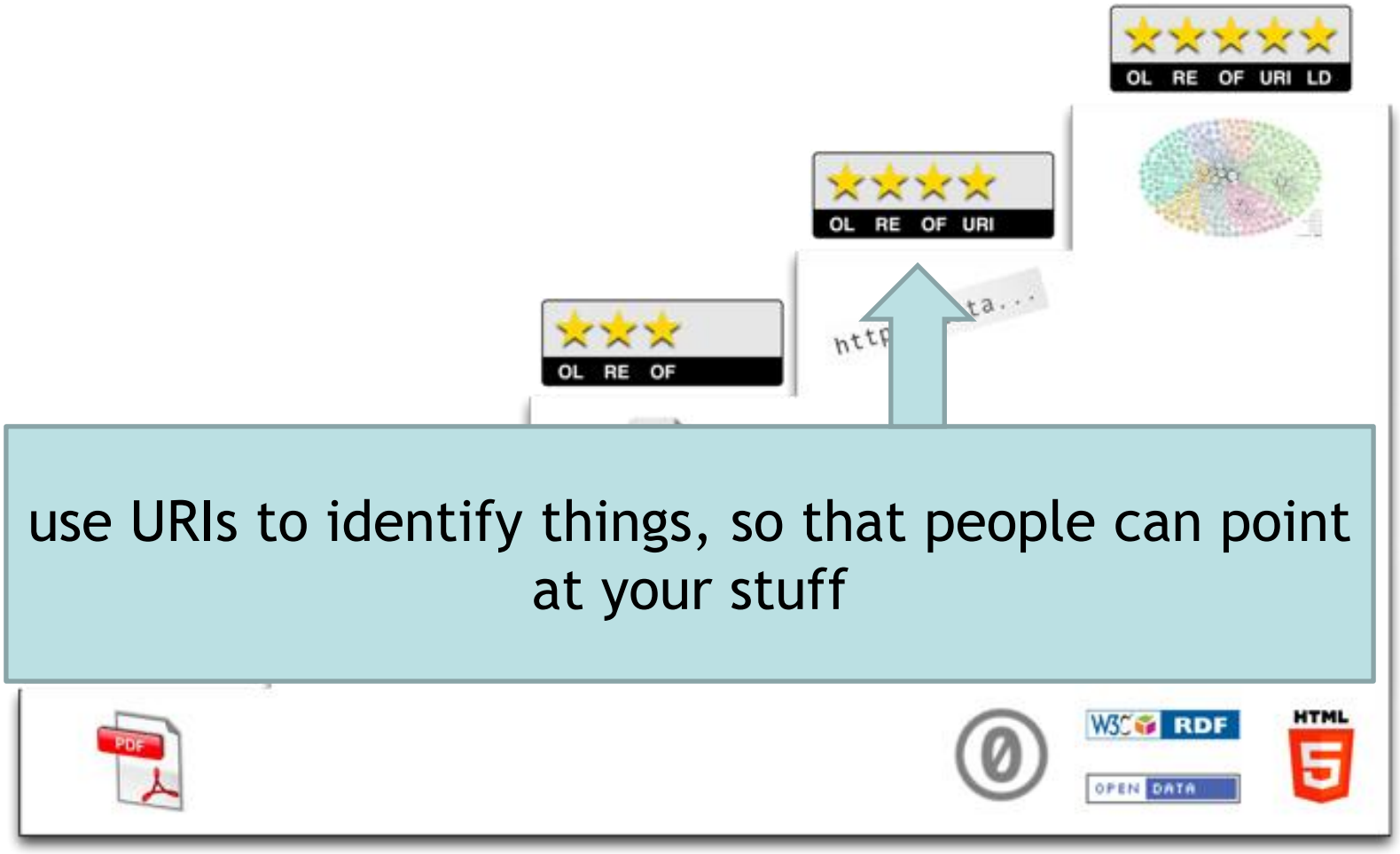


The 5 steps to Linked Open Data

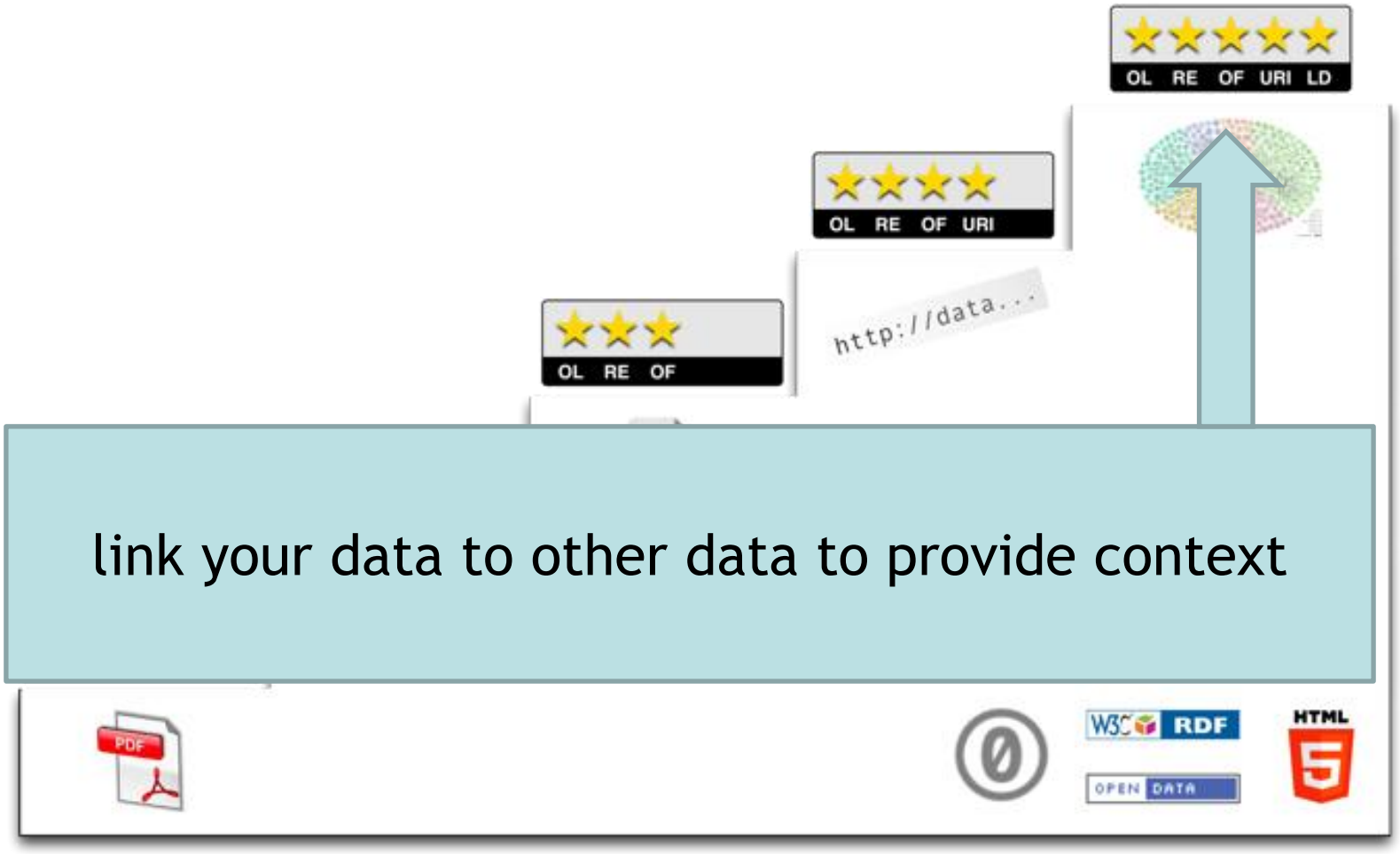
use non-proprietary formats (e.g., CSV instead of Excel)



The 5 steps to Linked Open Data



The 5 steps to Linked Open Data



What changes with LOD?

With Linked Open Data

- **Datasets are shared behind the scenes**
  => Each site can focus on key data and import supplemental data

Source: Jon Weers, NREL

What changes with LOD?

With Linked Open Data

- **Datasets are shared behind the scenes**
  => Each site can focus on key data and import supplemental data
- **Imported data updates automatically**
  => Provides users with consistent information across multiple sites

Source: Jon Weers, NREL

Data is shared at the database level. Updates to a linked database appear instantly on partner sites.

What changes with LOD?

With Linked Open Data

- **Datasets are shared behind the scenes**
  => Each site can focus on key data and import supplemental data
- **Imported data updates automatically**
  => Provides users with consistent information across multiple sites
- **Other Websites can consume** LOD resources to present new content in exciting and unanticipated ways

Source: Jon Weers, NREL

Third party websites can combine (or "mashup") linked open data to form innovative content, or new data.

Summary: Why LOD in Clean Energy

There is a need to focus efforts

- We want to display all relevant information about a topic but need to focus on providing only the information we are subject matter experts

We need to avoid replication

- Re-using existing datasets avoids replication of work already done and saves costs

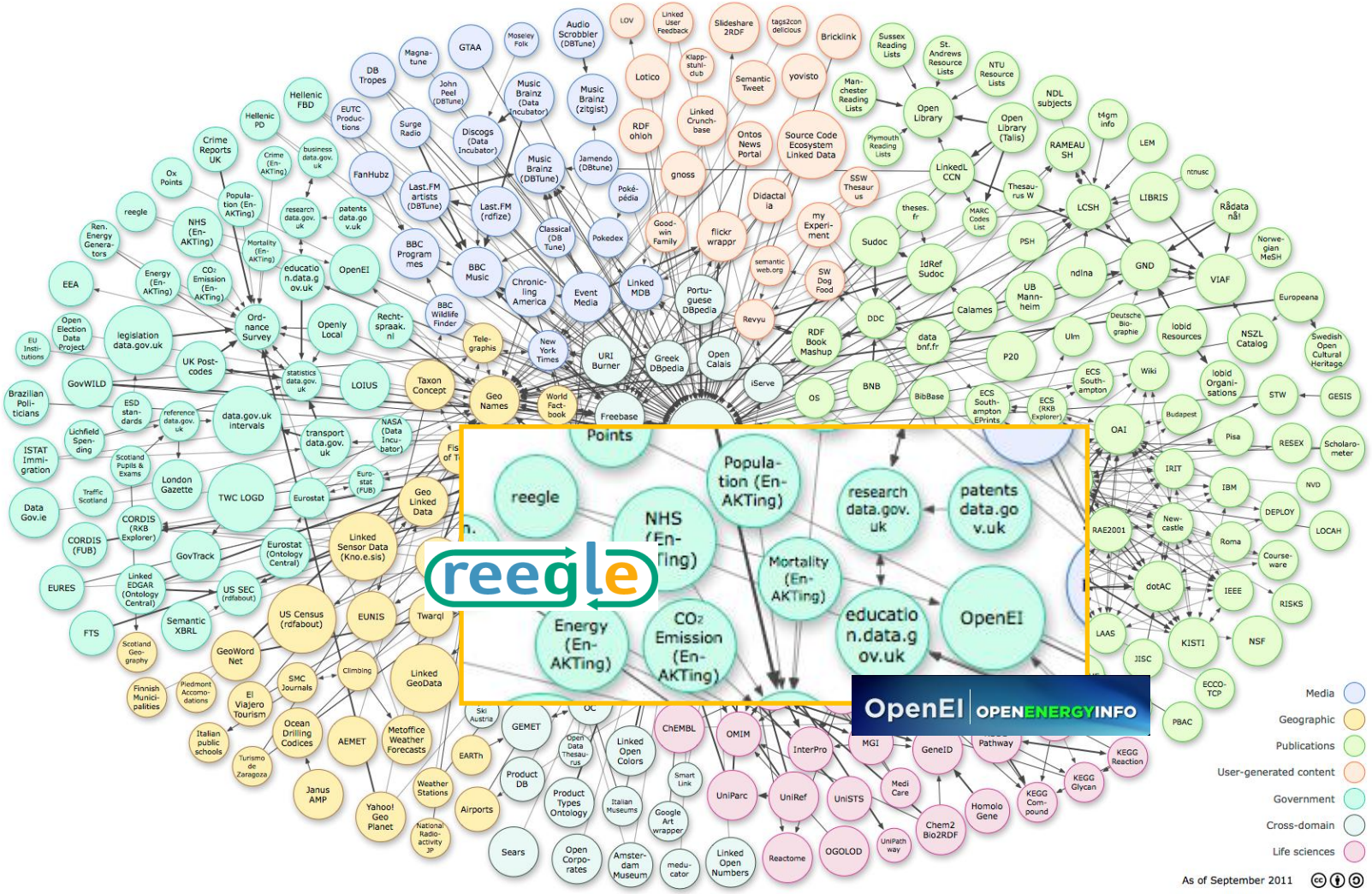
We want to reduced maintenance and effort

- Updates to linked open data are propagated instantly

Our aim is to move towards semantic linkages and interoperability

- Concepts become part of the semantic web
  - Data mash-ups and utilizations never before imaged
  - SPARQL queries can span multiple data sources

LOD cooperation – an example



Linking Open Data cloud diagram, by Richard Cyganiak and Anja Jentzsch. http://lod-cloud.net/

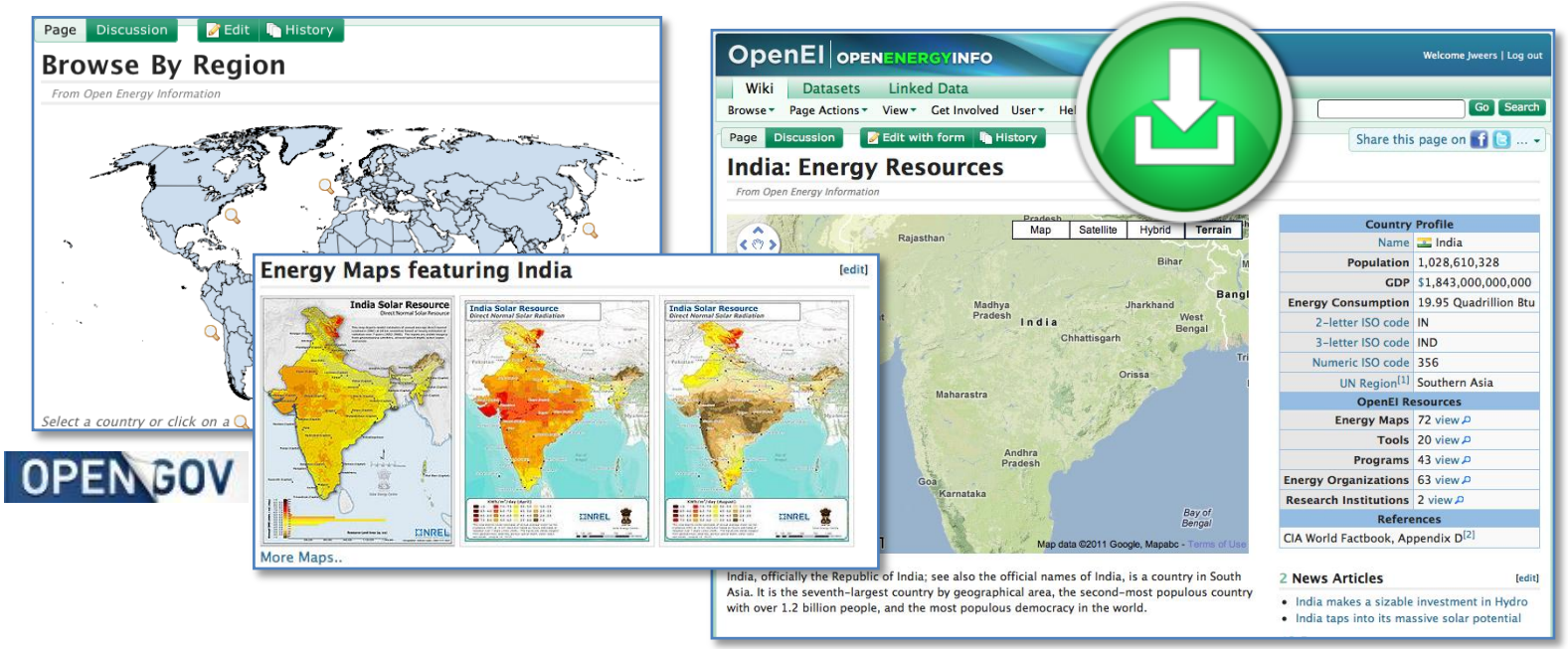
Linked Open Data on OpenEI

http://en.openei.org

OpenEI is a collaborative knowledge-sharing platform with free and open access to energy-related data, models, tools, and information.

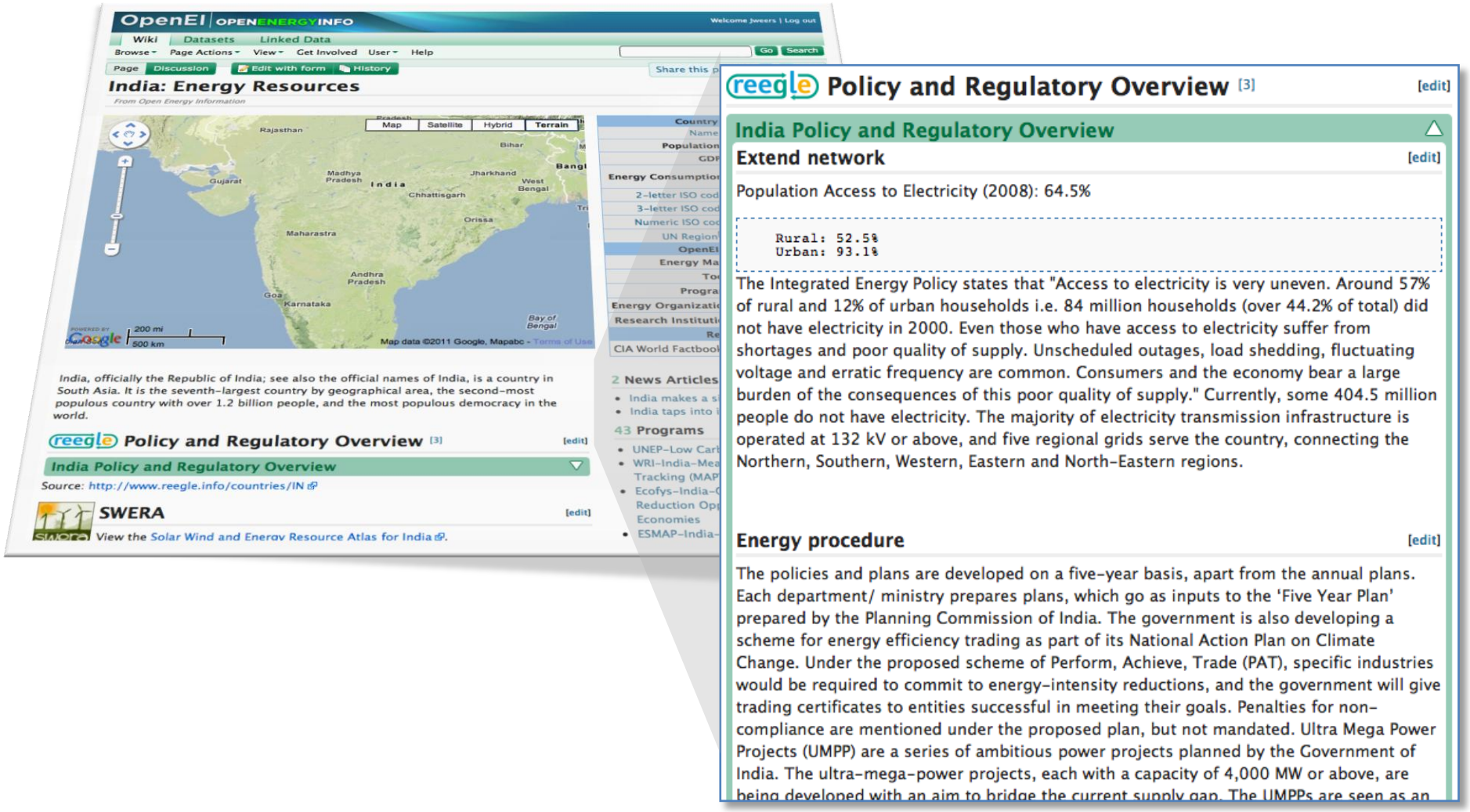
OpenEI features:

- more than 56,000 content pages
- more than 840 downloadable datasets
- regional gateways on a variety of energy-related topics
- Over 160 registered apps, including 64 Green Button apps.



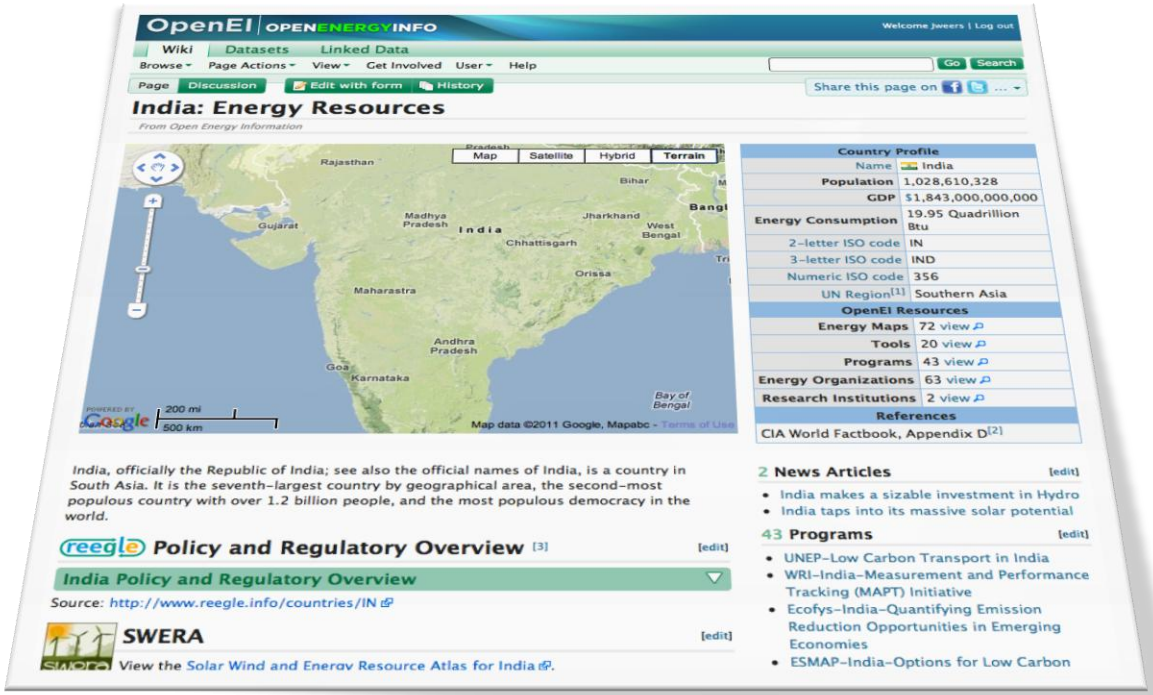
Linked Open Data on OpenEI

One of the featured pieces of content on OpenEI's country pages is the reegle Policy and Regulatory Overview:



Linked Open Data on OpenEI

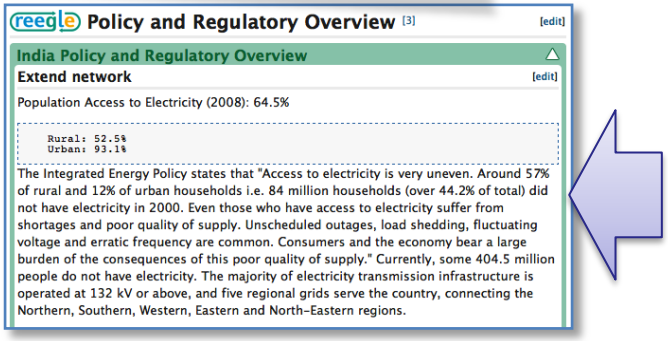
Behind The Scenes, Linked Open Data at Work:



reegle profiles are consumed in real time using SPARQL

use of semantic concepts allows the correct profile to be pulled

SPARQL



reegle DATA

SPARQL Endpoint

Print page

Contact reegle

… a LOD producer and consumer

http://www.reegle.info

- Well established information gateway for high quality information on renewable energy, efficiency and climate compatible development
- More than 220,000 users per month
- Data portal data.reegle.info launched in 2011

Available as Linked Open Data:

- key datasets including energy statistics
- over 1,700 stakeholders worldwide
- extensive glossary enriched with DBpedia linked data
- country energy profiles including policy & regulation data

reegle.info – clean energy info portal

DBpedia

UNdata
A world of information

THE WORLD BANK
Working for a World Free of Poverty

eurostat

Energy Profile Germany

Germany, officially the Federal Republic of Germany, is a federal parliamentary republic in Europe. The country consists of sixteen states while the capital and largest city is Berlin. It covers an area of 357,021 km and has a largely temperate seasonal climate. With 81.8 million inhabitants, it is the most populous member state and the largest economy in the European Union. It is one of the major political powers of the European continent and a technological leader in many fields. A region named Germania, inhabited by several Germanic peoples, was documented before AD 100. During the Migration Age, the Germanic tribes expanded southward, and established successor kingdoms throughout much of Europe. Beginning in the 10th century, German territories formed a central part of the Holy Roman Empire of the German Nation. During the 16th century, northern German regions became the centre of the Protestant Reformation while southern and western parts remained dominated by Roman Catholic denominations, with the two factions clashing in the Thirty Years' War. Occupied during the Napoleonic ... read more

Source: dbpedia

Capital	Berlin
ISO Code	DE, DEU
Area	357.021 km²
Population	81.799.600 (2010)
GDP, PPP	3.044.241.583.883,- current international $ (2010)

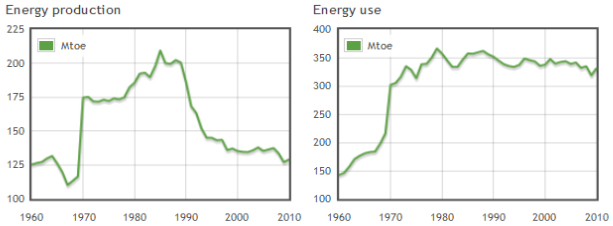
Events

23.04.2012
Hannover Messe Industrial Greentec

15.07.2012
Sustainable Energy Finance Summer Academy

Key Statistic Charts (26)

Energy production and use



Legal sources on support schemes and grid issues

Project Outputs (28)

Stakeholders (104)

Source: reegle Actors

- 500 PPM GmbH (500 PPM GmbH)
- Abo Wind
- AeroCratf Energitec
- Aircon International
- Aufwind
- BioKraftstoff
- BioStrom Energie System

reeep renewable energy & energy efficiency partnership

reegle

http://reegle.info/countries

Tagcloud – City of Edmonton (Canada) Workshop

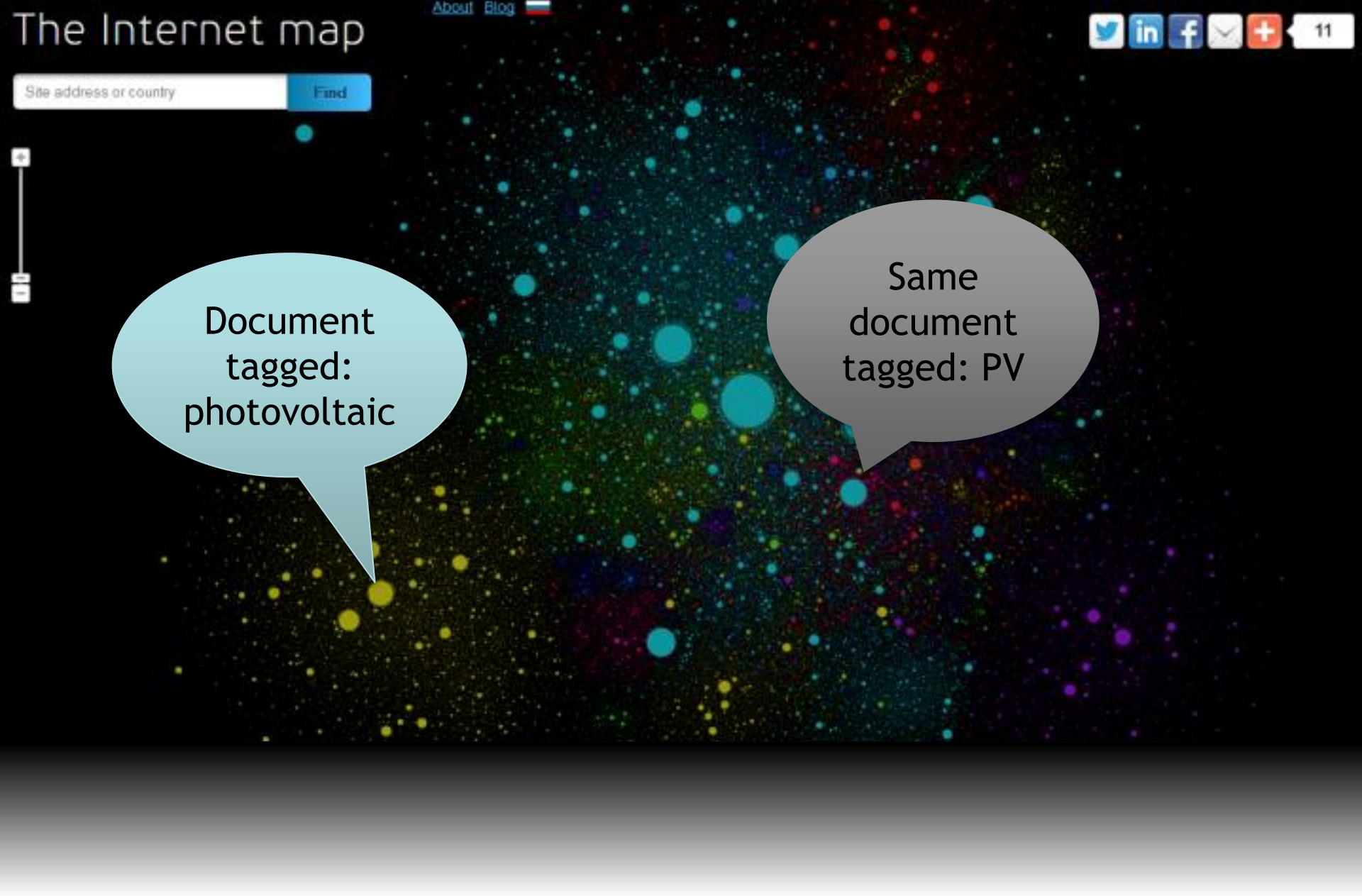
Providing standards and putting data in context is important

Standardisation and consitency is key

Based on our experience in establishing knowledge broker portals we know:

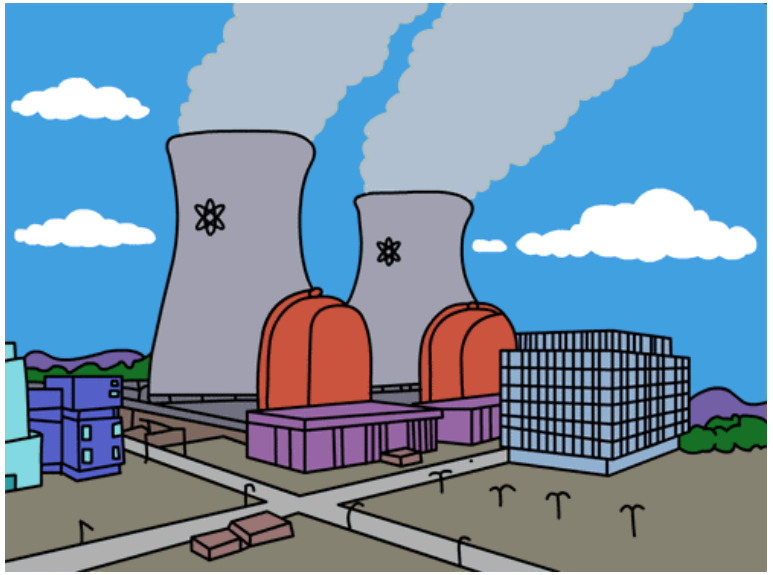
- There is a strong need to **increase consistency when tagging** climate and energy resources
- We need to ensure the **consistency of message** being delivered to the public to avoid confusion using terms in different ways
- This needs **standardization** of the used categories and tags

Can we support that with an automated system ?



Understanding synonyms & relations?

Dealing with disambiguation ...

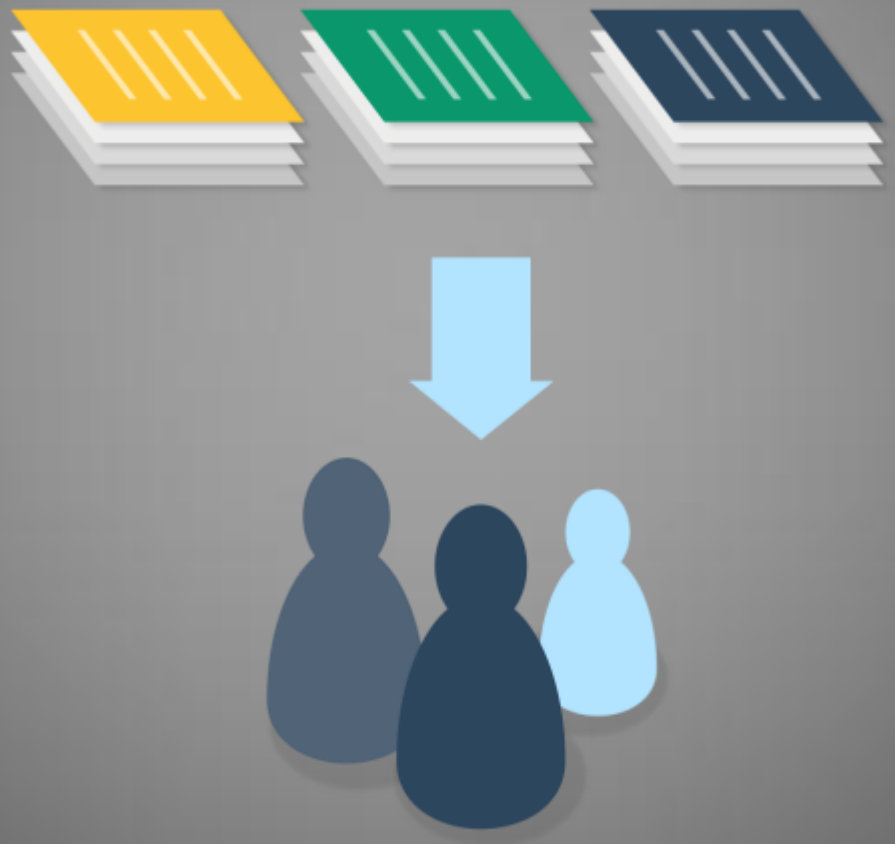


A new tool to help
with conistent tagging:

The reegle tagging API

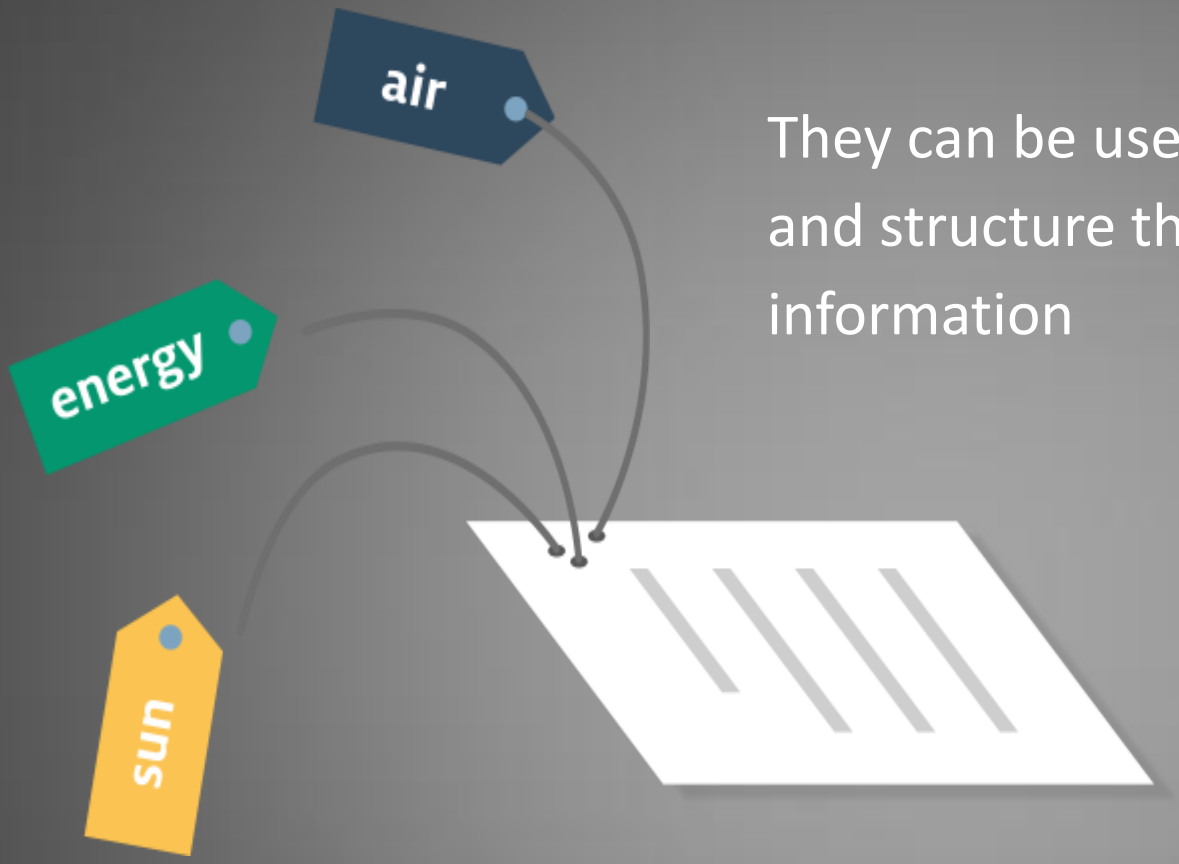
Unstructured Data

reegle tagging API turns data into knowledge



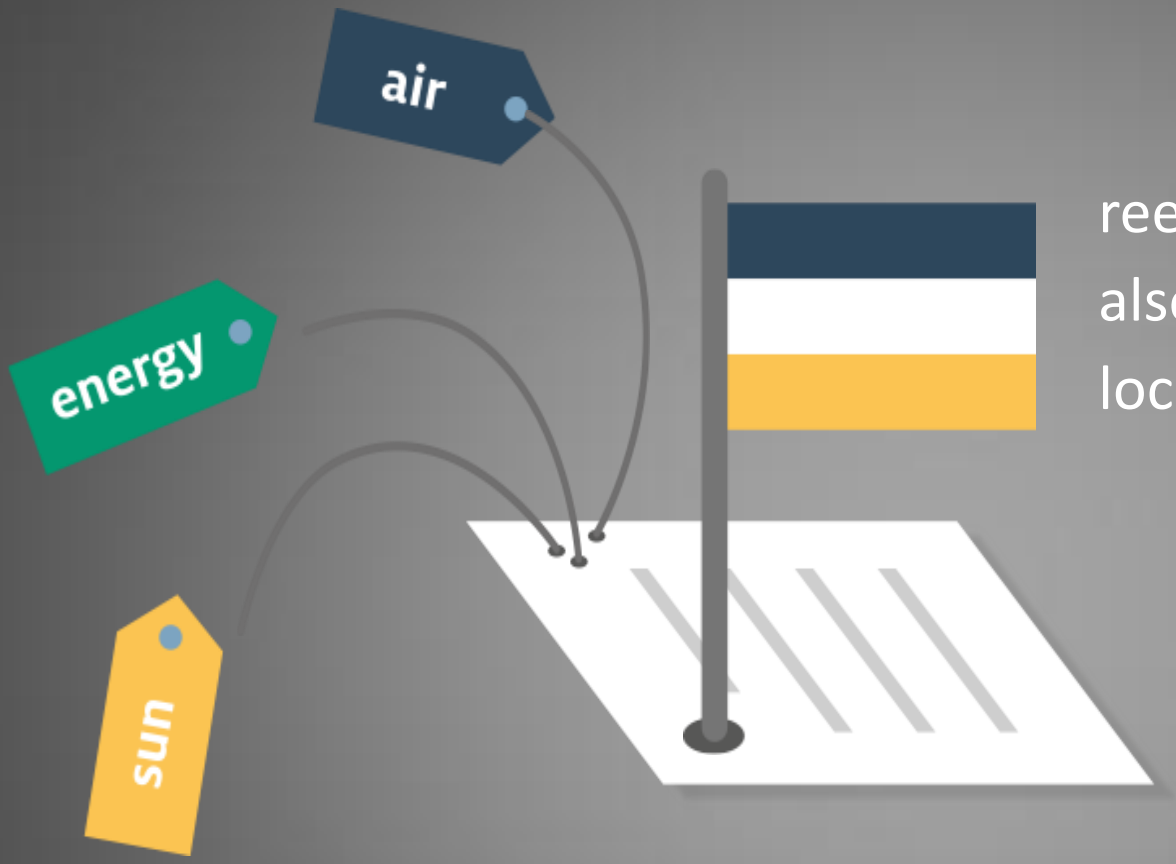
A file is sent …

… to the reegle tagging API,
where its content is analyzed.

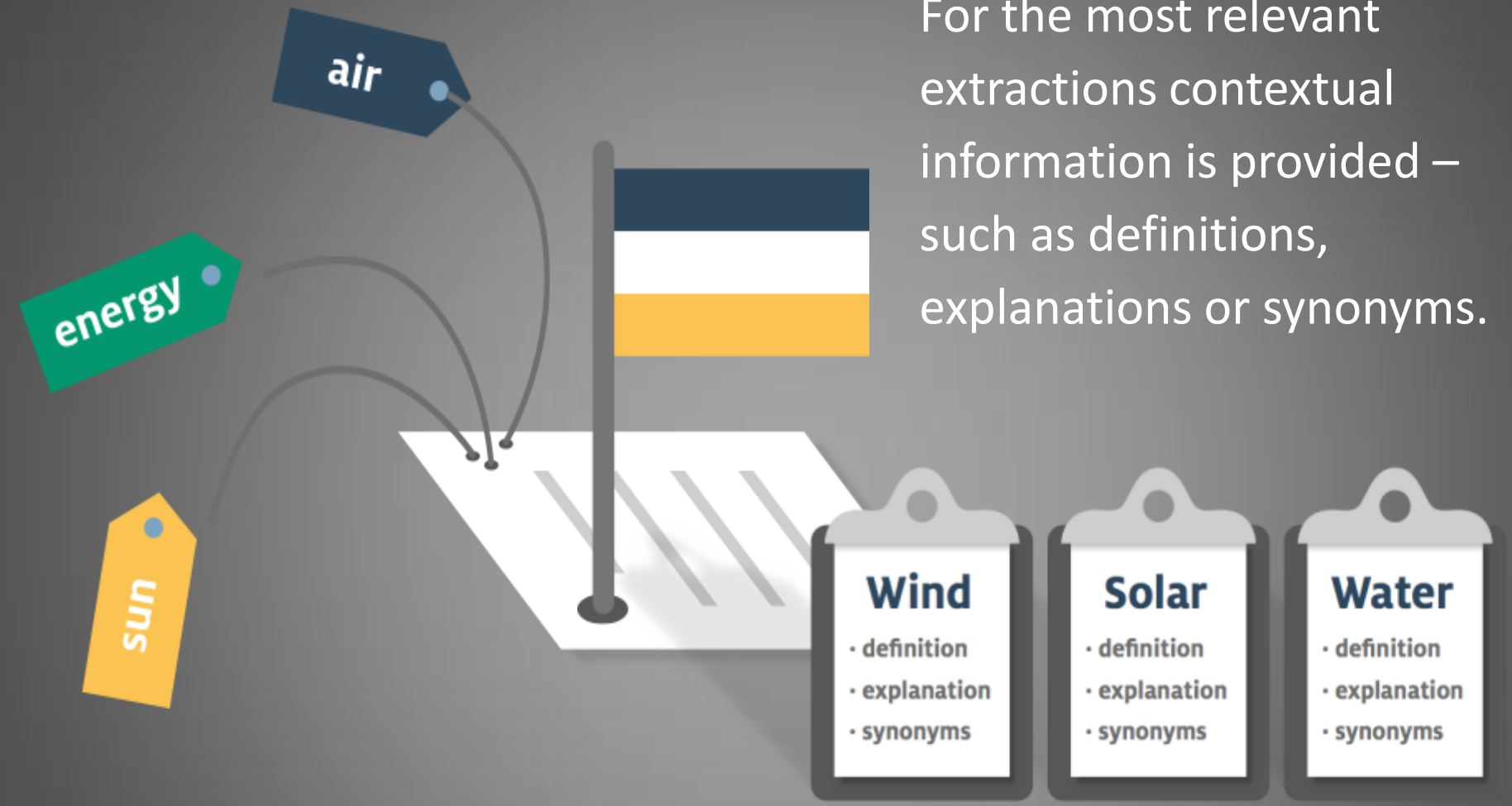


They can be used to tag and structure the information

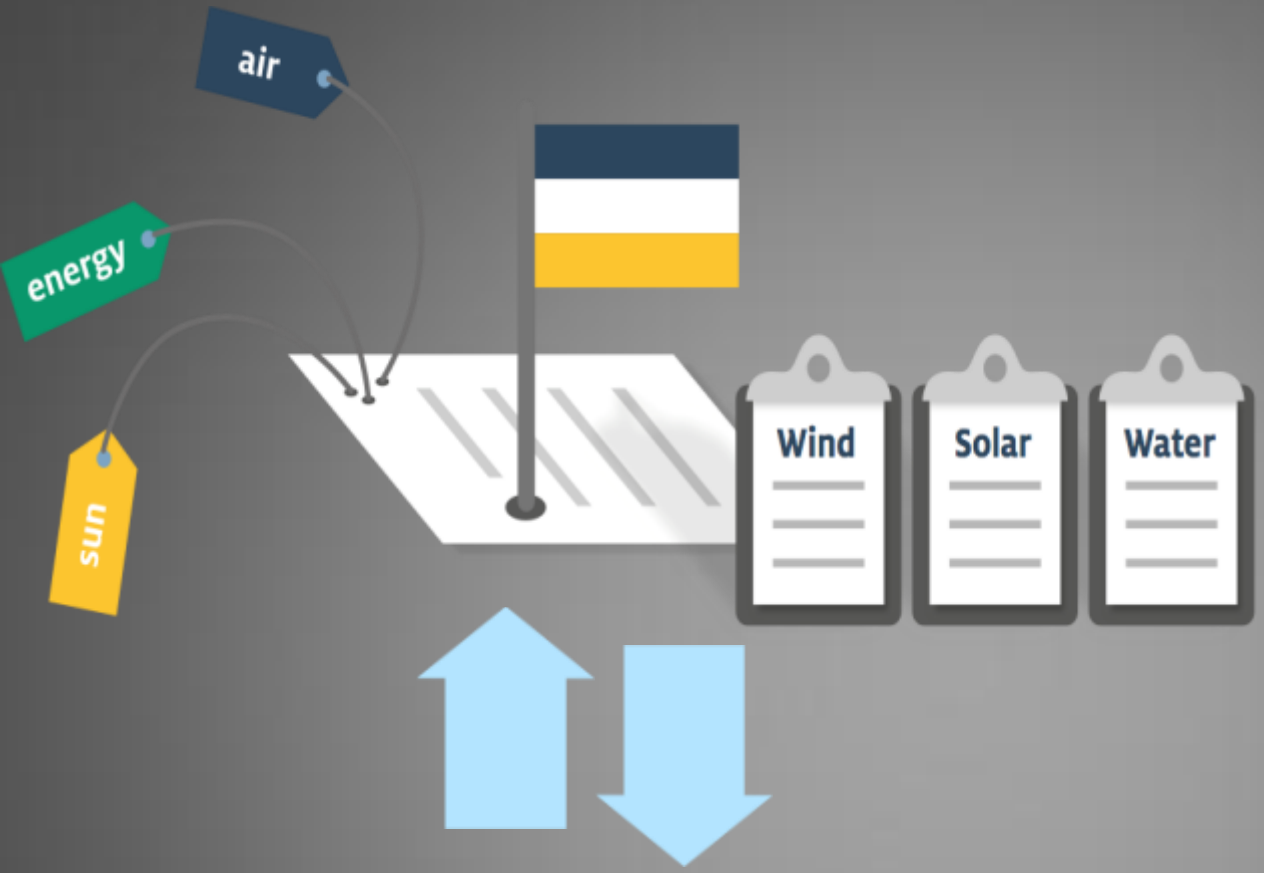
The most relevant terms are extracted



reegle tagging API
also extracts geographic
locations



For the most relevant extractions contextual information is provided – such as definitions, explanations or synonyms.



Enrich your documents with related articles from our content pool.

You can also add documents to the pool, making them available to others.

Content Pool

Use the tagging API to ...

... ensure consistency in tagging your information

... make your documents better searchable and accesible

... share your documents with others

... receive suggestions for similar other existing documents

Have a look on http://api.reegle.info !

reegle tagging API funded by:

Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit

Climate & Development Knowledge Network

Department of Energy & Climate Change

Use the tagging API to …

… enrich content by matching terms to definitions

Have a look at http://openei.org/wiki/Gateway:Wind!

reegle tagging API funded by:

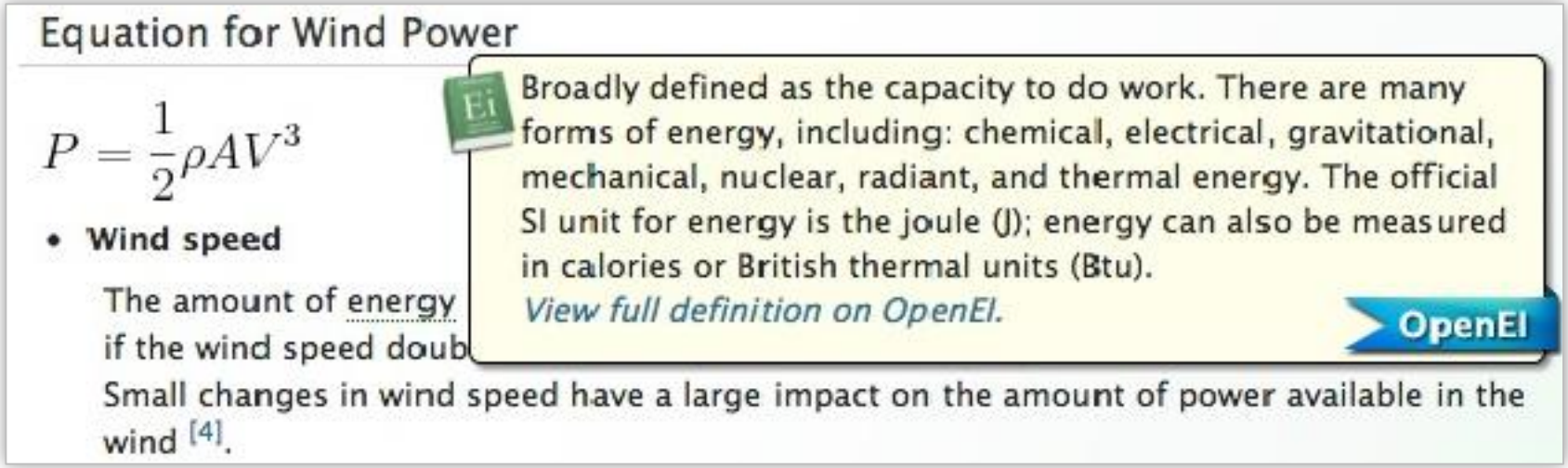
Bundesministerium
für Umwelt, Naturschutz
und Reaktorsicherheit

CDKN Climate & Development Knowledge Network

DEPARTMENT OF ENERGY & CLIMATE CHANGE

Use the tagging API to …

… providing users with real-time defintions of key terms



… recommending related articles by searching similar terms

Have a look at http://openei.org/wiki/Gateway:Wind!

reegle tagging API funded by:

Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit

Climate & Development Knowledge Network

Department of Energy & Climate Change

Questions? Please contact us …

Florian Bauer
Operations & IT Director, REEEP

Wagramerstrasse 5,
A-1400 Wien, Austria

Web	http://www.reeep.org , http://www.reegle.info
Blog	http://blog.reegle.info
Twitter	http://twitter.com/reegle
LinkedIN	http://at.linkedin.com/in/florianbauer
Mail	florian.bauer@reeep.org
Skype	**reeep_f.bauer**
Phone	+43-1-26026 3714

Jon Weers
Senior Web Applications Engineer, NREL

15013 Denver West Pkwy
Golden, CO 80401-3305

Web	http://www.nrel.gov, http://en.openei.org
Blog	http://blog.openei.org
LinkedIN	http://www.linkedin.com/in/jonweers
Mail	Jon.Weers@nrel.gov
Phone	+1-303-275-3701