

# Open Data in Agrifood: A Tutorial

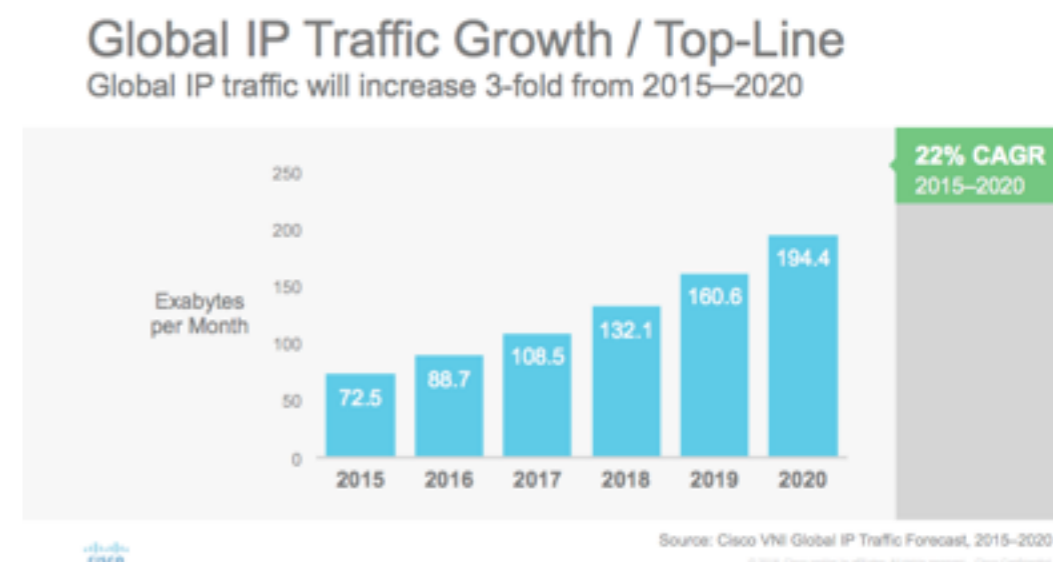
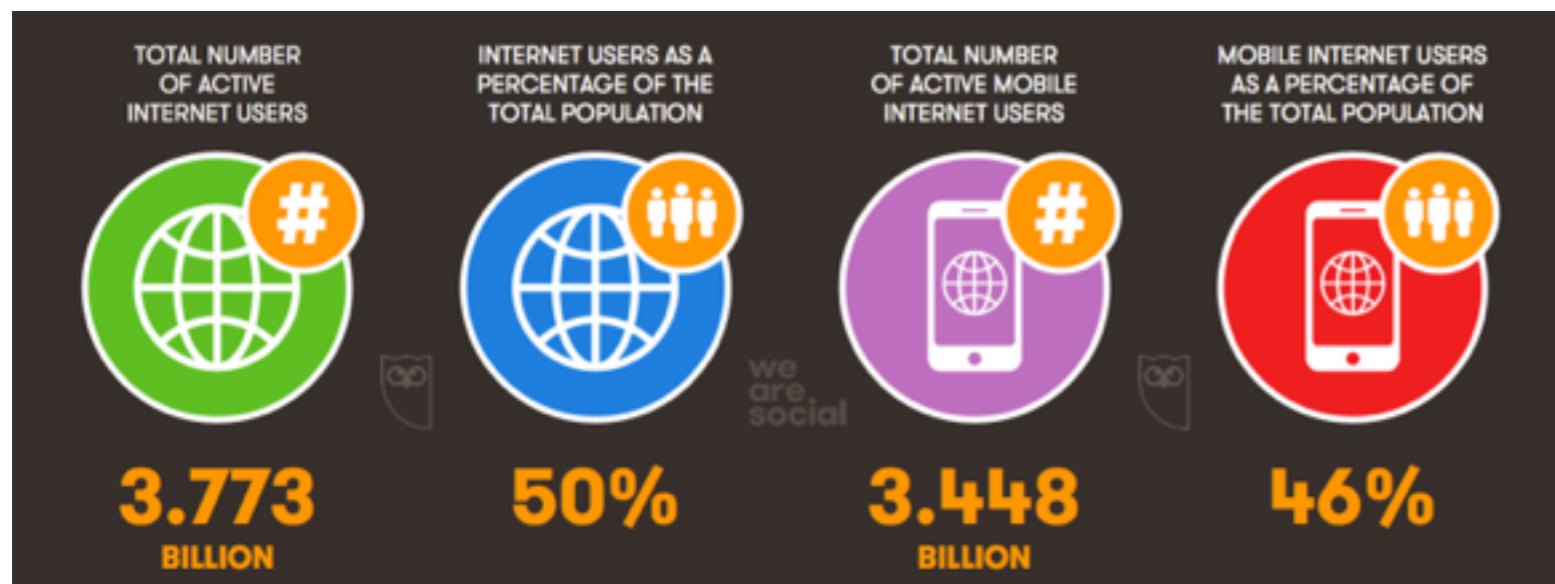
Christopher Brewster  
TNO, The Netherlands

# OUTLINE

- ▶ A Digital world
- ▶ Origins of Open Data and the OD movements
- ▶ Initiatives for Open Data in agrifood
- ▶ Types and Sources of Open Data in agrifood.
- ▶ Uses of Open Data and Examples
- ▶ Issues and challenges

# A Digital World

- ▶ Huge quantities of digital data produced every hour, minute, second — and uploaded to platforms
- ▶ Global IP traffic is 1.2 Zb (Zetabytes) per year (2016), predicted to be 3.2 Zb by 2021
- ▶ 2.3 networked devices per capita in 2016 (17.1Bn), reaching 3.5 in 2021 (27Bn)
- ▶ 73% of this data is video.
- ▶ All these devices generate data. Some of that data is/will be **OPEN DATA**



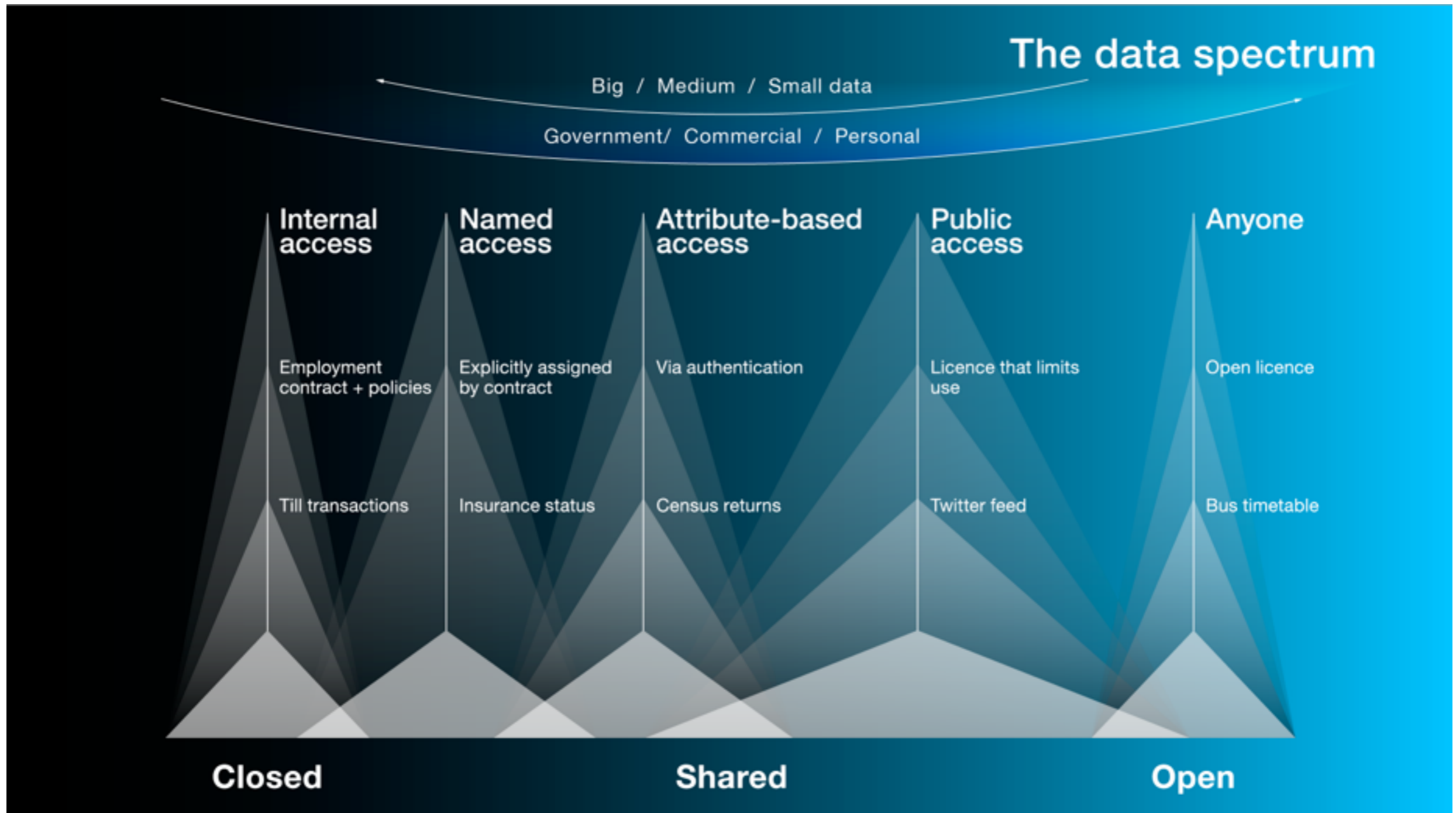
# Closed Data

- ▶ Most data is closed, even most government data.
- ▶ Businesses afraid of competition, loss of value, loss of control (what happens in the future?).
- ▶ Government afraid of use/abuse by third parties. Also a matter of culture.
- ▶ For example, GPS data - originally for US military only (1974-85), eventually made available to all (airlines 1983, everyone 1998) - slow change of culture.

# Why is data closed?

- ▶ “Closed” means inaccessible
- ▶ Forgotten, lost, impractical formats (e.g. binary or proprietary)
- ▶ Seen as too sensitive, given only on a “need to know” basis
- ▶ Seen as too strategic, maybe of competitive advantage
- ▶ Seen as too costly - to adapt, reformat, provide an API, follow a data standard
- ▶ Culture/time and place: data coming off farm machinery (dairy robots vs. tractors + field machinery)

# The Data Spectrum



# Origins of Open Data

# Open Data

- ▶ “Open data and content can be freely used, modified, and shared by anyone for any purpose” — The Open Definition
- ▶ Open data also needs a **License** specifying that is open data, who to credit, and what happens when remixed.
- ▶ Typically (traditionally) collected and made available by government agencies (e.g. statistics office)
- ▶ Classic example is UK school league tables (1994), or meteorological data (even older).



# Why is data important?

- ▶ Data is key to decision making. Governments collect all kinds of data, corporations collect all kinds of data.
  - ▶ Historically censuses and offices of statistics
  - ▶ Now every website, every app, every digital interaction
  - ▶ Used to make decisions from school building to advertisement delivery
- ▶ Open data is a public good!
- ▶ Open data is infrastructure!

# Open Data as a public good

- ▶ A “public good” is non-excludable (everyone can use it) and non-rivalrous (use by one does not reduce use by another) cf. [https://en.wikipedia.org/wiki/Public\\_good](https://en.wikipedia.org/wiki/Public_good)
- ▶ Like clean air or public parks
- ▶ Public goods cost money - usual paid for by government, collaborations, cross-subsidy, volunteering/crowd sourcing, social norms etc.
- ▶ However we usually pay for infrastructure .....

# Open Data as Infrastructure

- ▶ Data is like a road system. You choose to build them.
- ▶ A “data infrastructure” provides a foundation for social and business activity.
- ▶ Some data sets are “trunk roads” - important, frequently used, some data sets are “side roads” less important — but they all must be interconnected ...
- ▶ Core data sets are references by other data sets e.g. address, postcodes, etc. **Importance of identifiers!**
  - ▶ Important entities = companies, schools, hospitals, stations and buss stops, and (perhaps) **land parcels**
  - ▶ Most common example is **post codes** (open data in most countries)

# Origins of Open Data

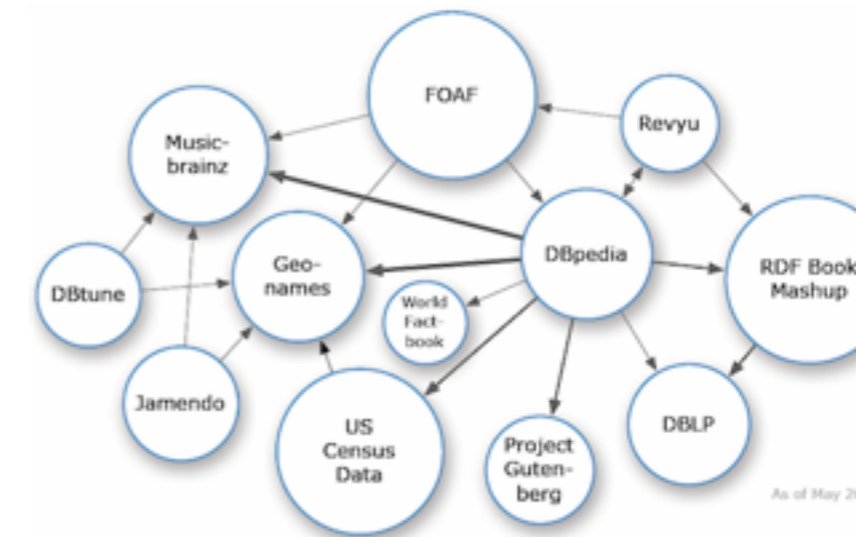
- ▶ Five sources:
  - ▶ Semantic Web vision of Berners-Lee to go from “web of documents” to “web of data” (2001) - proposal for Linked Data (2006)
  - ▶ Transparency movement - political significance of transparency of expenditure and other activity (from 90s but much more mid 2000s especially UK and US). Part of “Open Government” movement. (Tauberer 2014)
  - ▶ Fight in UK for “Free our Data” from 2006 was gradually very successful (<http://www.freeourdata.org.uk/>)
  - ▶ Scientific Open Data - move towards open access, open data for reuse
  - ▶ Open Data for innovation - enable new business services
  - ▶ Data as new gold - in view of Google, Facebook etc. allow others new business opportunities (globally late 2000s)

# Three movements: Open Data

- ▶ Essentially about releasing government data - not only but mainly
- ▶ Establishing standards for format and licenses,
- ▶ Key players: Nigel Shadbolt and Tim Berners-Lee
- ▶ Key institutions: Open Data Institute, Open Knowledge Foundation
- ▶ Web sites: data.gov.uk, data.gov, https://www.europeandataportal.eu/, http://openindex.gr/ , http://geodata.gov.gr

# Three movements: Linked Open Data

- ▶ Result of need to built a “web of data” (part of Semantic Web initiative) cf. Berners-lee 2006)
- ▶ Core rapidly became DBpedia (<http://wiki.dbpedia.org/> )
- ▶ Grew quite rapidly until today - dominated by government data (e.g. CIA fact book), life sciences, and lingusitics. Not much agrifood ....



2007

# Publishing Linked Data

- ▶ Berners-Lee (2006) wrote “The Semantic Web isn't just about putting data on the web. It is about making links, so that a person or machine can explore the web of data. With linked data, when you have some of it, you can find other, related, data.”
  - ▶ Use URIs as names for things
  - ▶ Use HTTP URIs so that people can look up those names.
  - ▶ When someone looks up a URI, provide useful information, using the standards (RDF\*, SPARQL)
  - ▶ Include links to other URIs. so that they can discover more things.

# Publishing Five Star Data

▶ He also proposed five levels of openness and linked-ness:

1. Available on the web (whatever format) but with an open licence, to be Open Data



2. Available as machine-readable structured data (e.g. excel instead of image scan of a table)



3. as (2) plus non-proprietary format (e.g. CSV instead of excel)



4. All the above plus, Use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff



5. All the above, plus: Link your data to other people's data to provide context

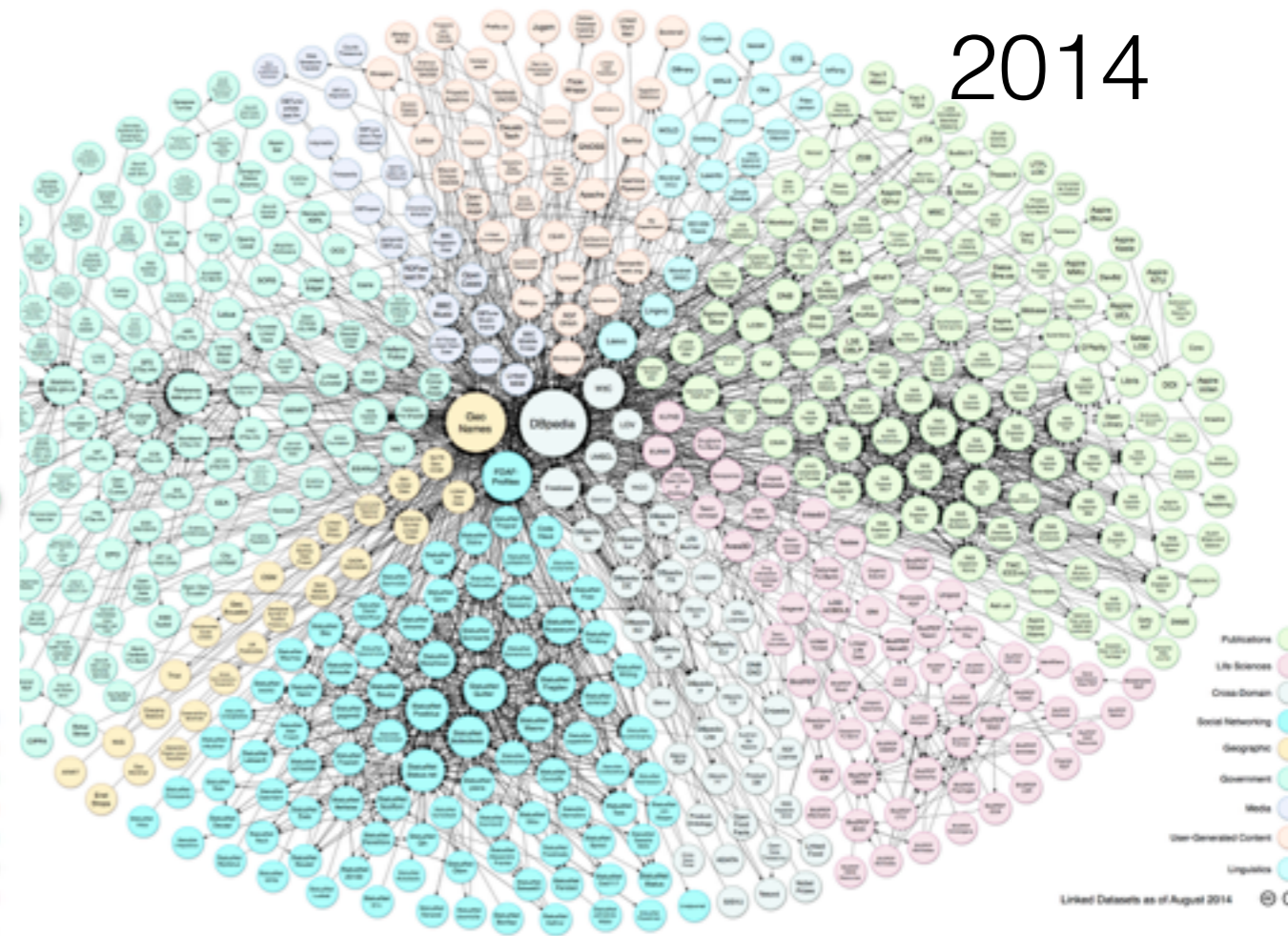
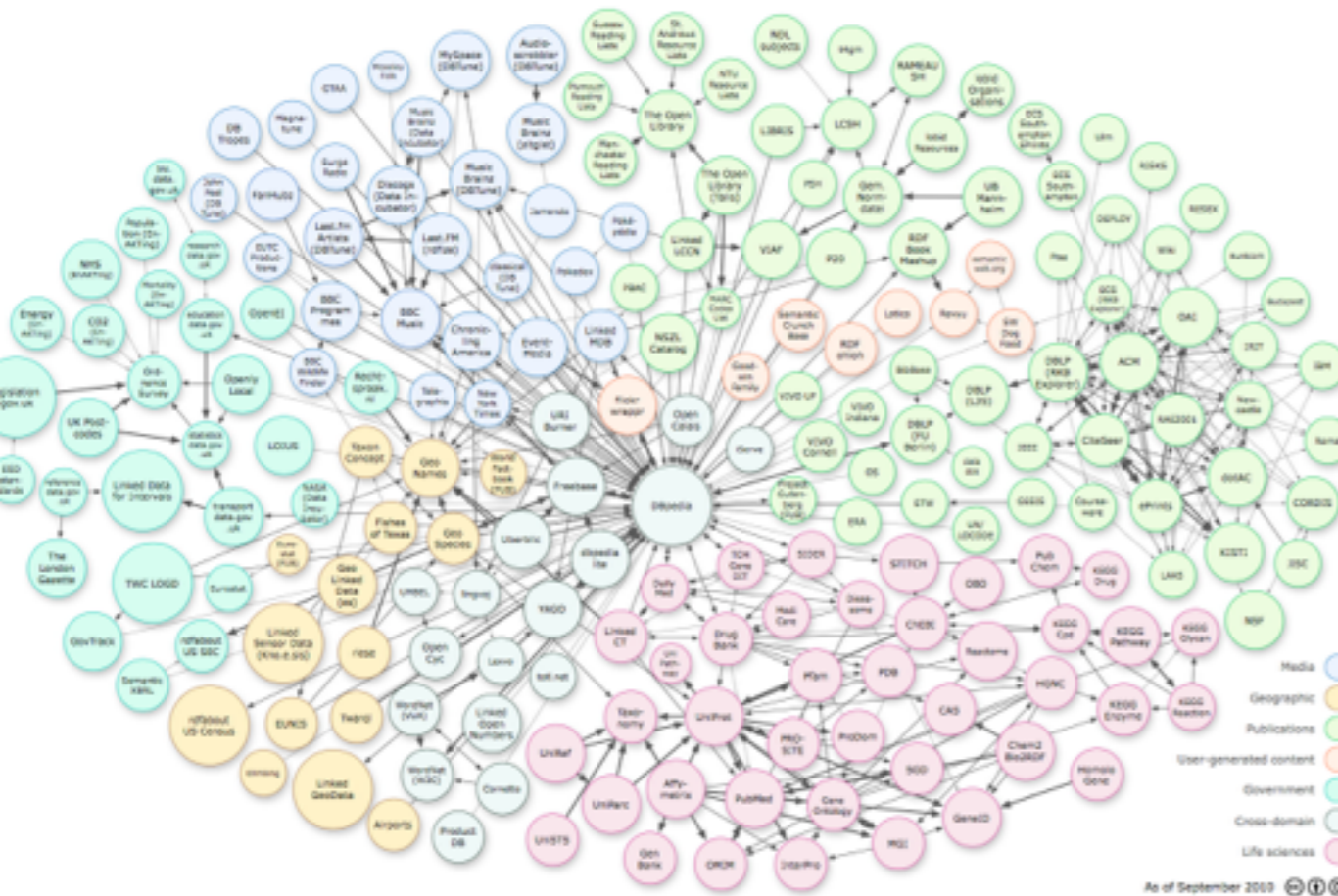




# Growth of the LOD cloud

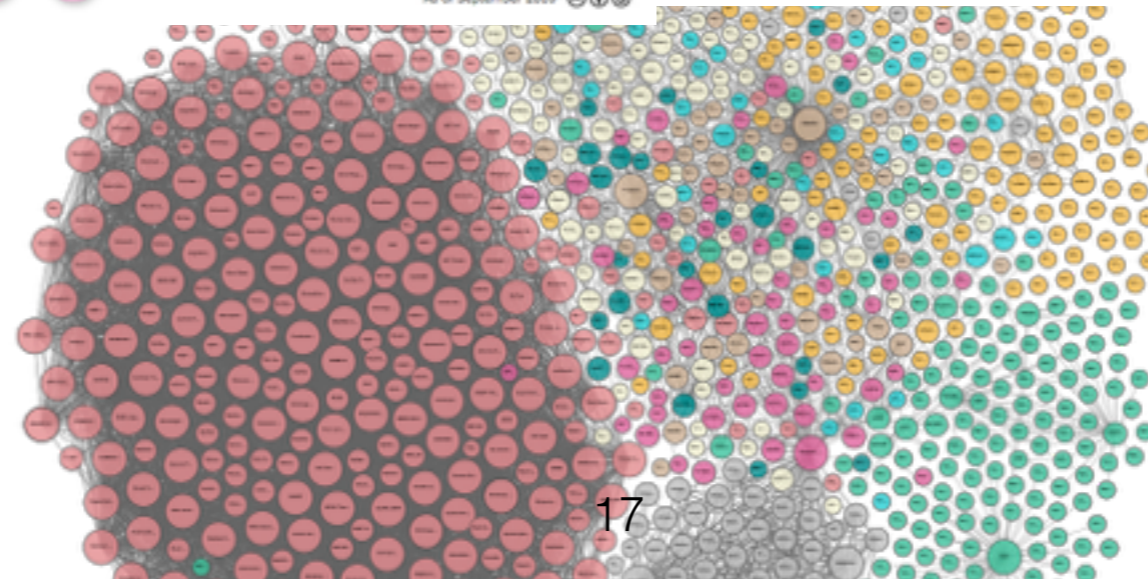
2010

2014



<http://lod-cloud.net/>

2017



# Three movements: FAIR Principles

- ▶ Initiated by the FORCE11 group “a community of scholars, librarians, archivists, publishers and research funders” in 2011. Very influenced by Berner-Lee’s proposals
- ▶ Created a set of principles for “data intensive science”. Data should be:

## TO BE FINDABLE:

- F1. (meta)data are assigned a globally unique and eternally persistent identifier.
- F2. data are described with rich metadata.
- F3. (meta)data are registered or indexed in a searchable resource.
- F4. metadata specify the data identifier.

## TO BE ACCESSIBLE:

- A1 (meta)data are retrievable by their identifier using a standardized communication protocol.
- A1.1 the protocol is open, free, and universally implementable.
- A1.2 the protocol allows for an authentication and authorization procedure, where needed.
- A2 metadata are accessible, even when the data are no longer available.

## TO BE INTEROPERABLE:

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles.
- I3. (meta)data include qualified references to other (meta)data.

## TO BE RE-USABLE:

- R1. meta(data) have a plurality of accurate and relevant attributes.
- R1.1. (meta)data are released with a clear and accessible data usage license.
- R1.2. (meta)data are associated with their provenance.
- R1.3. (meta)data meet domain-relevant community standards.

# FAIR principles

- ▶ Resulted in a key paper: Wilkinson et al 2016 “The FAIR Guiding Principles for scientific data management and stewardship”
- ▶ Major impact for Life Sciences, and therefore agricultural sciences.
- ▶ FAIR principles adopted by EC in general, e.g. for H2020 projects.
- ▶ Primary focus is **research data** but not only ...

# Open Data in Agrifood

# Open Data in Agrifood

- ▶ This has “happened” in last 10-15 years
- ▶ Motivation:
  - ▶ Growing awareness of government data sets, and potential applications
  - ▶ growing availability of public earth observation data from satellites (especially ESA and EC’s Copernicus Programme 1998-2020, but mostly 2014-2020)
  - ▶ Desire to “help” farmers with digital technologies and data
  - ▶ Sudden fashion for precision agriculture/smart farming (since 2010 mostly)

## Important Initiatives: Open Data Institute

- ▶ ODI create in UK in 2012 by Tim Berners-Lee and Nigel Shadbolt
- ▶ Mission to educate, train and encourage adoption of Open Data
- ▶ **Major effort in agriculture** - collation of open data sets, hackathons, support for UK start-ups
- ▶ It has produced various reports, and worked closely with GODAN
- ▶ <https://theodi.org/>

# Important Initiative: GODAN

- ▶ “Global Open Data for Agriculture and Nutrition”
- ▶ Followed a G8 commitment in 2012, announced in 2013, funded by US, UK and NL governments + support from FAO, GFAR, CGIAR etc.
- ▶ Major emphasis on developing countries, helping to open up data, and enable new applications and services.
- ▶ Several research reports: <http://www.godan.info/resources/research>
- ▶ Several “success stories”: <http://www.godan.info/resources/success-stories>

# Movement: Agrifood hackathons

- ▶ With growth of Open Data, so growth of hackathons focussed on agrifood open data
- ▶ (Past) examples:
  - ▶ Hackathon on "The power of Linked Data in Agriculture and Food Safety" (2015)
  - ▶ Climate change, Agriculture and Food Security Hackathon (2014)  
Lima, Peru
  - ▶ FarmHack NL (repeated hackathons) <http://www.farmhack.nl/en/>
  - ▶ Hack4Farming hackathon for the future of agricultural development in India (2016)
  - ▶ Hackathon for Open Agricultural Data, China (2014)



# Types and Sources of Agrifood Open Data

# Government Statistics

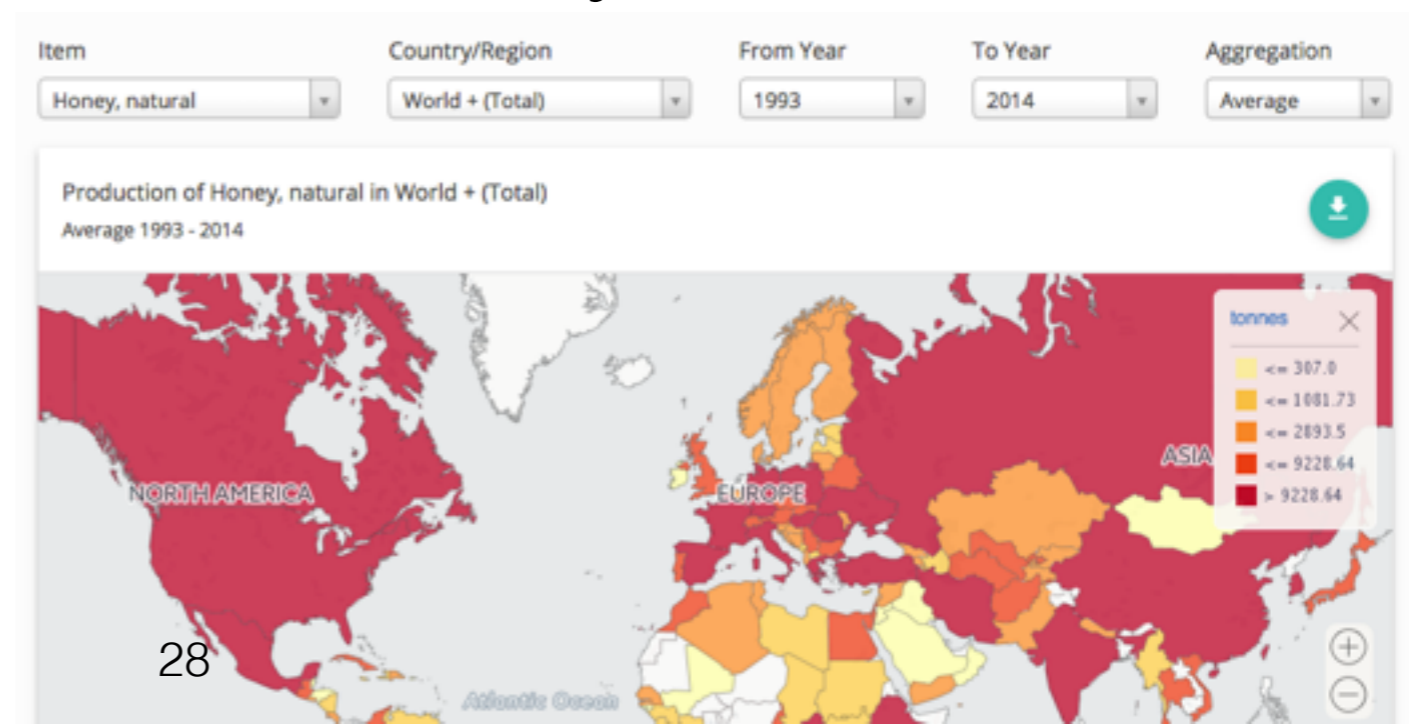
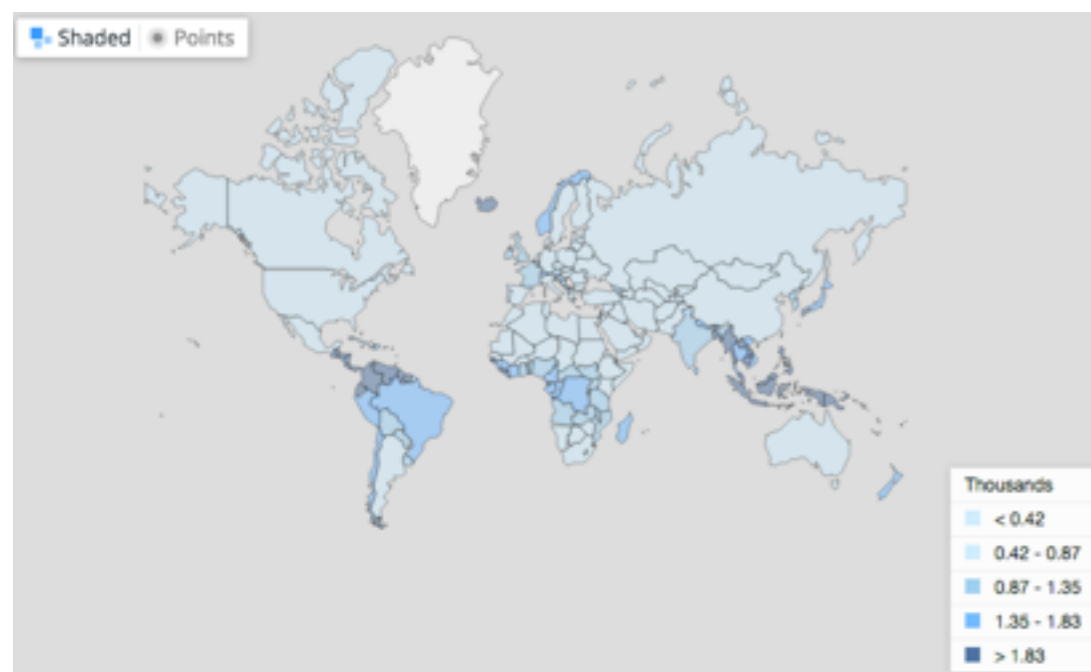
- ▶ Governments collect a lot of data about food and agriculture, mostly economic and social
- ▶ For example:
  - ▶ information about the structure of farms, orchards & vineyards, including parcel data
  - ▶ agricultural production
  - ▶ economic accounts for agriculture
  - ▶ agriculture and environment
  - ▶ agricultural prices
  - ▶ animal populations
  - ▶ milk production, animal slaughter
  - ▶ etc.

# Government data (non statistical)

- ▶ Plenty of other government data, not really statistics:
  - ▶ Meteorological data ← **probably the most important**
  - ▶ Crop disease outbreaks
  - ▶ Food recalls
  - ▶ Health and safety inspections
  - ▶ Maximum residue levels (for pesticides)

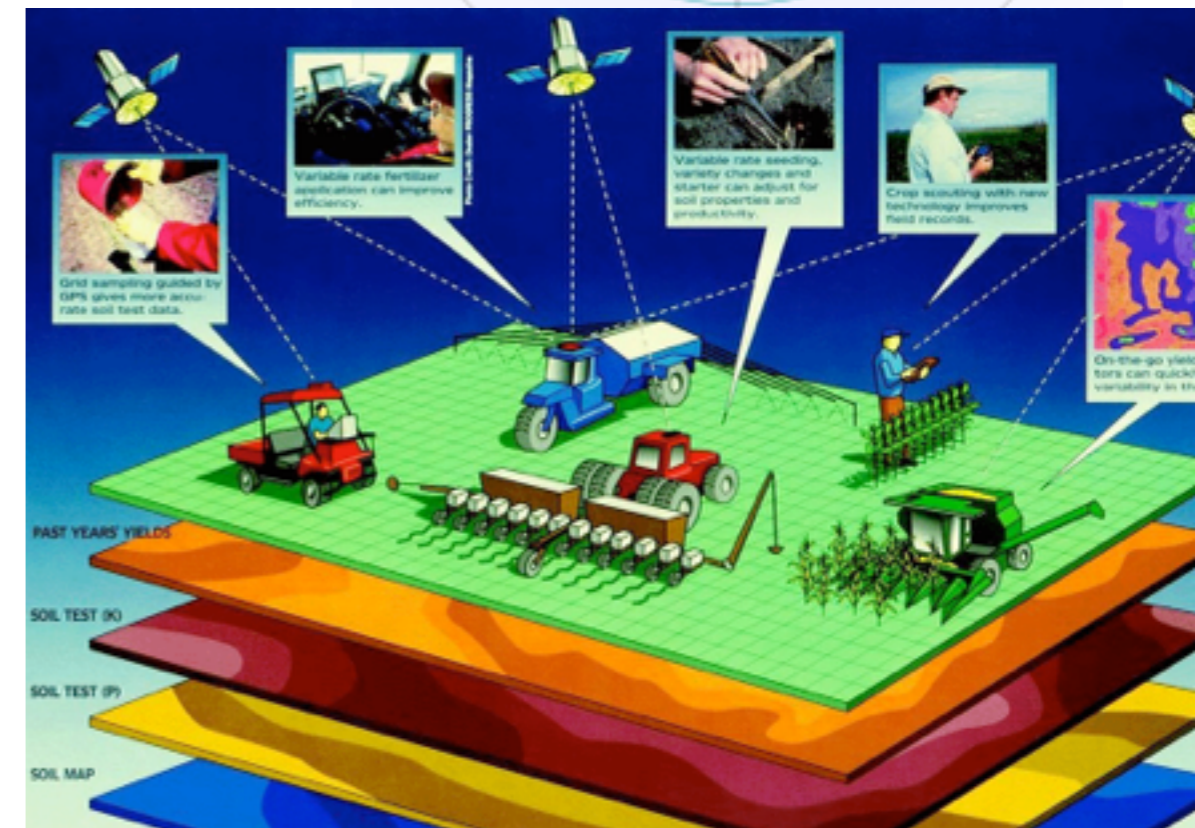
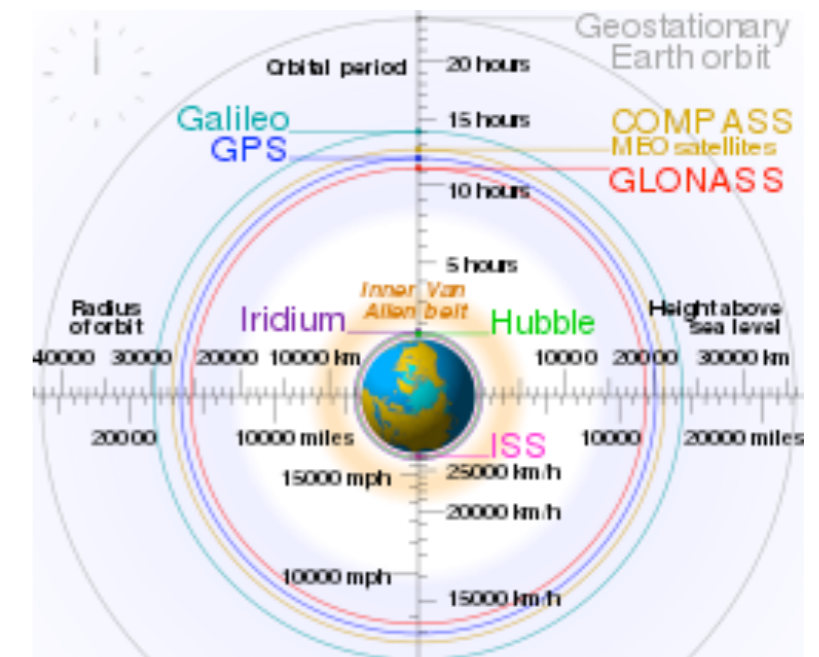
# International Government Data

- ▶ UN - specifically Food and Agriculture Organisation (FAO) - has lost of global statistics
- ▶ World Bank - usually derived from national statistical offices
- ▶ Mostly comparative across countries e.g. wettest country in the world, or global production of honey



# Other International Data - GPS

- ▶ Global Positioning System (GPS) - now in every smart phone, already extensively used in Precision Agriculture
  - ▶ US system, now complemented by Galileo (EC), GLONASS (Russia), Beidou (China)
- ▶ Much expectation that GPS will facilitate more environmentally sustainable agriculture.



# Other International Data - Copernicus

- ▶ European system for monitoring the Earth ([www.copernicus.eu](http://www.copernicus.eu)). Collects data from multiple sources: earth observation satellites and in situ sensors such as ground stations, air-borne and sea-borne sensors.
- ▶ Data collected includes:
  - ▶ land changes (it can map different classes of cover such as forest, crops, grassland, water surfaces and artificial covers like roads and buildings;
  - ▶ water health and pollution;
  - ▶ agricultural practices, to estimate crop acreage, to provide soil moisture information and to forecast yield;
  - ▶ distinguishing between different crop types as well as data on numerous plant indices, such as leaf area index, leaf chlorophyll content and leaf water content;
  - ▶ disaster mapping.

## COPERNICUS AND ITS SENTINELS

European Earth Observation Programme Copernicus: observing our planet for a safer world

**SENTINEL-1**

- All-weather, day and night radar imaging satellite for land and ocean services
- Able to "see" through clouds and rain
- Data delivery within 1 hour of acquisition
- Airbus Defence and Space developed Chirp radar instrument

**SENTINEL-2**

- Medium Res. Multippectral optical satellite for observation of land, vegetation and water
- 12 spectral bands with 10, 20 or 60m resolution and 290km swath width
- Global coverage of the Earth's land surface every 5 days
- Airbus Defence and Space prime contractor for satellites and instruments

**SENTINEL-3**

- Measures sea surface topography with a resolution of 300m, sea and land surface temperature and colour with a resolution of 10m
- Measures water vapour, cloud water content and thermal radiation emitted by the Earth
- Estimates global sea surface temperatures with an accuracy greater than 0.2 K
- Airbus Defence and Space supplies Microvelocimetry Radiometer

**SENTINEL-5P**

- Global observation of key atmospheric constituents, including ozone, nitrogen dioxide, sulphur dioxide and other environmental pollutants
- Improves climate models and weather forecasts
- Provides data continuously during Europe gap between the retirement of Envisat and the launch of Sentinel-5
- Airbus Defence and Space prime contractor for satellite and TROPOMI instrument

**SENTINEL-4**

- Provides hourly updates on air quality with data on atmospheric aerosol and ozone gas concentrations
- Spatial sampling is 5km and spectral resolution between 0.12um and 0.5um
- Airbus Defence and Space prime contractor for spectrometer
- Contract awarded EUMETSAT's MetOp Second Generation (MSG) satellite

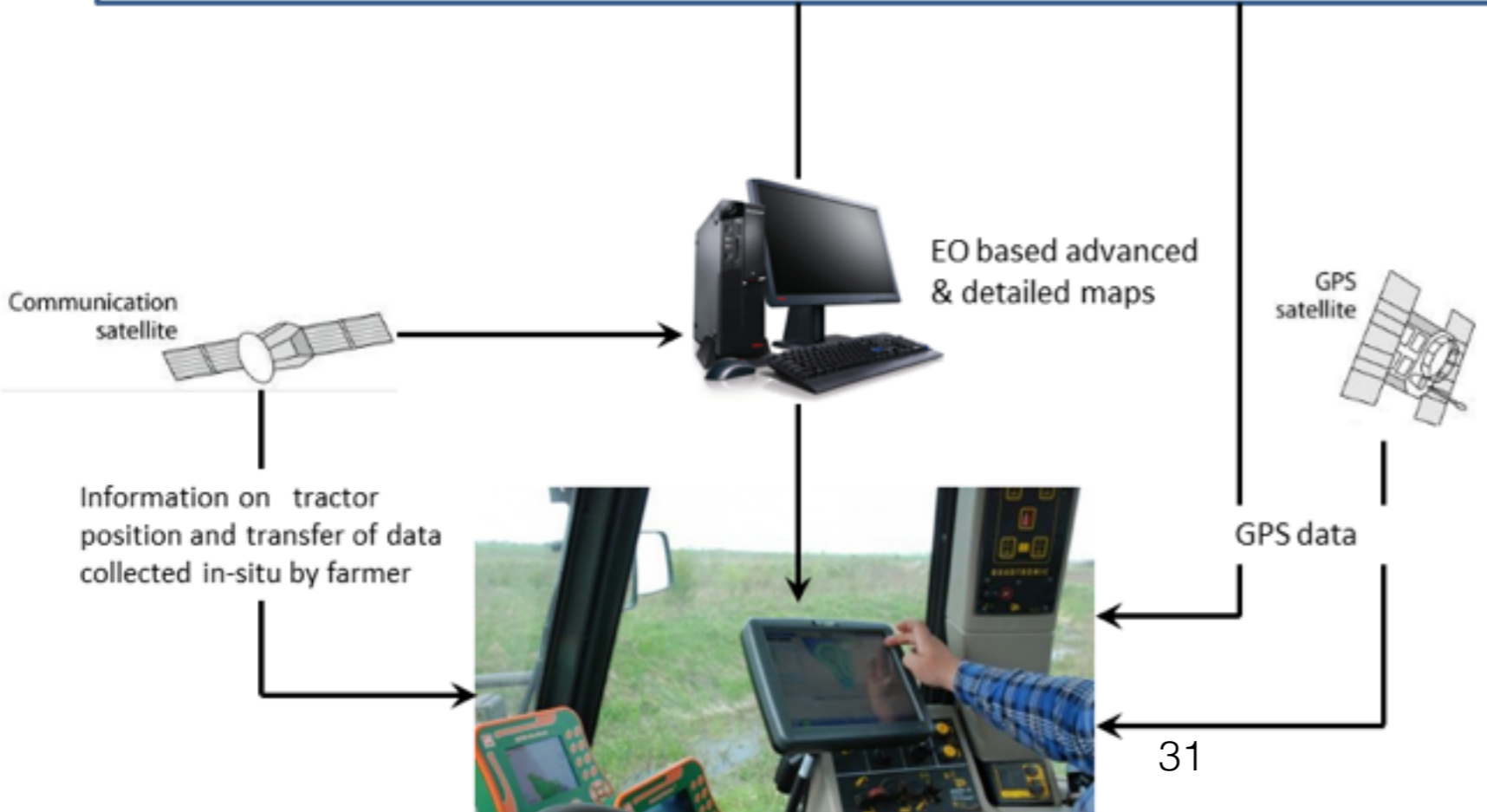
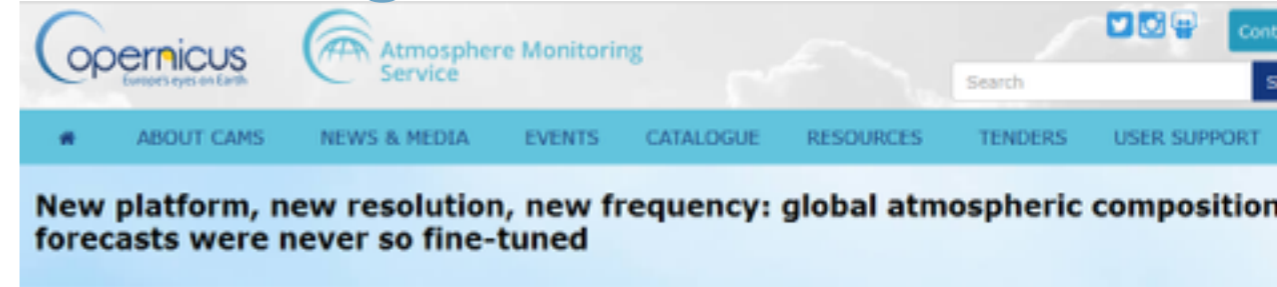
**SENTINEL-5**

- Measures air quality and solar radiation, monitors atmospheric ozone and the climate
- Global coverage of Earth's atmosphere with an unprecedented spatial resolution
- Airbus Defence and Space prime contractor for instrument
- Contract awarded EUMETSAT's MetOp Second Generation satellite

**SENTINEL-6**

- Observes changes in sea surface height with an accuracy of a few centimetres
- Global mapping of the sea surface topography every 10 days
- Enables precise observation of ocean currents and ocean heat storage, vital for predicting rises in sea levels
- Airbus Defence and Space prime contractor for satellite

# Earth Observation data for agriculture



» New platform, new resolution, new frequency: so fine-tuned

**SERVICE THEMES**

- AIR QUALITY & ATMOSPHERIC COMPOSITION
- CLIMATE FORCING
- OZONE LAYER & UV
- SOLAR RADIATION
- EMISSIONS AND SURFACE FLUXES

**ANALYSES**

- European Air Quality
- Fire Monitoring
- Reactive Gases
- Aerosols

**FORECASTS**

- Reactive Gases

in by the Copernicus Atmosphere Monitoring m-Range Weather Forecasts (ECMWF) is now rovements for its global products: moving to a y two atmospheric composition forecasts per lays after the launch of the service's new

# Academic Data

- ▶ More and more academic data available
- ▶ Both research publications (AGRIS) and research data sets (look ODJAR journal, MACSUR project)

The screenshot shows the AGRIS search interface. At the top, it features the FAO logo and the text 'Food and Agriculture Organization of the United Nations'. Below this is a search bar with the text 'Find resources...'. To the left of the search bar are options to 'Refine your search' and 'Sort by: Relevance' (selected) or 'Submission Year'. Below the search bar, there are options to 'Order: Ascending' or 'Descending' (selected), and a 'Change order' button. To the right of the search bar, the search results are displayed, showing 'Query: greek agriculture' and 'Results 11 - 20 of 618'. Below the search bar, there are navigation buttons: '< first', '< prev', 'next >', and '> last'. Below the search bar, there is a 'Multilingual search' section with options to 'Enable' or 'Disable' (selected 'Disable') and a 'Refine search' button. Below the 'Refine search' button, there is a list of search results with the following items: 'greece (219)', 'trichosporium br', 'kingdom fascicul', 'biomass productio', 'per cent colonizat', 'agriculture (38)', 'cyprus (32)', 'rain (16)', 'air temperature (5)', and 'european union (1)'. Below the search results, there is a section for 'ODJAR.org' with a navigation menu: 'HOME', 'ABOUT', 'LOGIN', 'REGISTER', 'SEARCH', 'CURRENT', 'ARCHIVES', 'ANNOUNCEMENTS', 'MANUAL'. Below the navigation menu, there is a section for 'Open Data Journal for Agricultural Research' with logos for 'AgMIP', 'cimsans', and 'MACSUR'. Below the logos, there is a paragraph of text: 'Agricultural research uses and produces many relevant data sets in studying agricultural systems across the globe, through its efforts (in)security at different spatial scales (from regional to national to continental). These data sets have a value to the specific research results and conclusions, that are published in peer-reviewed scientific journals or presented at scientific conferences. These data have more value than the specific research in which they are collected. Other researchers or experts can use these data in new analysis, meta-analysis tools, leading to new insights for the future. The Open Data Journal for Agriculture Research (ODJAR) acts as a central hub for storing and sharing research data for the future where publications and their authors get appropriate credit through citations and digital object identifiers for future research. Many different data sets exist, that are of value and deserve accreditation: experimental data, surveys, model inputs, model outputs, and mark-ups, maps, measured data points. Unlike journal articles describing the main new insights and the most important lessons learned, these data sets are often not available after the funding period ends or the research is published, leading to a situation where these are difficult to reuse for other purposes, or difficult to access. With the advance of Open Access, Linked Open Data and Open data portals of governments, there is increasing awareness of the value of research data, leading to increased innovation, creation of jobs and better services. Also, governments and science funders are increasing their support with tax-payer financial resources, and should thus have a public benefit. These recent developments of Open Access to data and the acknowledge value of data archiving lead to four global networks in agriculture to support the Open Data Journal for Agricultural Research: + Agricultural Model Improvement and Intercomparison Project (AgMIP) + Global Yield Gap Atlas (GYGA) + Center for Integrated Modeling of Sustainable Agriculture and Nutrition Security (CIMSANS) + Modelling European Agriculture with Climate Change for Food Security (MACSUR)'



# Crowd Sourced: Open Street Maps

- ▶ <https://www.openstreetmap.org/>
- ▶ Started in 2004 in reaction to absence of open mapping data in UK
- ▶ Over 1M registered users, 30% have contributed data
- ▶ Many governments and companies contributed data sets.
- ▶ Widely used in websites, services and applications.

The screenshot displays the OpenStreetMap website interface. At the top, there is a navigation bar with the 'OpenStreetMap' logo and buttons for 'Edit', 'History', and 'Export'. Below this is a search bar with the text 'Search' and 'Where am I?' followed by a 'Go' button and a share icon. A large map of Greece is visible on the right side, with labels for cities like Αθήνα (Athens), Χαλκίδα (Chalkida), and Καλαμάτα (Kalamata). A welcome message is overlaid on the map, stating: 'Welcome to OpenStreetMap! OpenStreetMap is a map of the world, created by people like you and free to use under an open licence. Hosting is supported by UCL, Imperial College London and Bytemark Hosting, and other partners.' Below the message are two buttons: 'Learn More' and 'Start Mapping'. At the bottom, there is a dark green banner for 'STATE OF THE MAP ASIA 2017' with the text 'Mapping. Community. Resiliency.' and the dates 'September 23-24' and location 'Kathmandu, Nepal'.

# Crowd-sourced: Open Food Facts

▶ <https://world.openfoodfacts.org/>



**Open Food Facts**  
The free food products database.

Tell me what you eat and I will tell you what you are.  
— Jean Anthelme Brillat-Savarin - 1825

▶ Started 2012 - now has over 350,000 branded products

▶ Data exists in GS1 data bases but it is not “open”. This has many applications.

Open your food and know what you eat

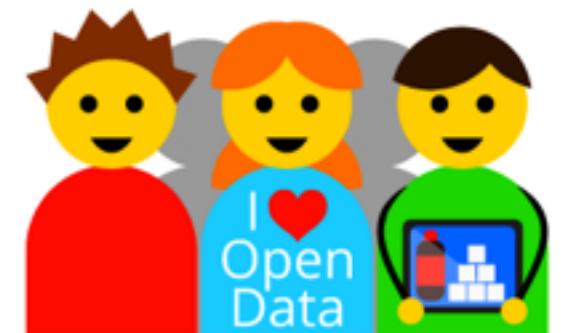
Be part of our collaborative, free and open database of food products from around the world!



A food products database



Made by everyone



For everyone



# Data Sources

- ▶ There is a profusion of data sources around the world
- ▶ Most national governments now have a (sort of) open data portal
- ▶ Several international ones too (EC, UN, World Bank)
- ▶ Data comes in many different formats
- ▶ Some example sources:

# Data Source: European Open Data Portal

- ▶ <http://data.europa.eu/euodp/en/home/>
- ▶ “open data produced by EU institutions and bodies”
- ▶ 730 data sets in agriculture, forestry and fishing

EU **Open Data** Portal  
Access to European Union open data

EUROPA > EU Open Data Portal > Home

Home Data Applications Linked data Developers' corner About

The **European Union Open Data Portal (EU ODP)** is your single point of access to open data produced by EU institutions and bodies. All the data you can find via this catalogue are free to use and reuse for commercial or non-commercial purposes.

Show results with:  
 all of these words |  any of these words |  the exact phrase

Search for metadata using our [SPARQL endpoint query editor](#) or access the application programming interface.

Discover our datasets View datasets by subject View all datasets View all

**Focus on**

Public procurement notices  
> Publications Office

**Ted**  
Tenders electronic daily

Twitter

Europeana @Europeanaeu  
 .@franky\_abbott from @dpla  
 what #OpenCulture means he  
 @oeconsortium #yearofopen  
 bit.ly/2j8Wg1Y  
 pic.twitter.com/aneEzWLCrE  
 Retweeted by EU ODP

# Data Source: European Data Portal

- ▶ <https://www.europeandataportal.eu/data/en/group/agriculture-fisheries-forestry-and-food>
- ▶ “The European Data Portal harvests the metadata of Public Sector Information available on public data portals across European countries”
- ▶ 13700 data sets in Agriculture, fisheries, forestry and food

European Data Portal > Categories > Agriculture, fisheries, forestry ...

What we do - Data - Providing Data - Using Data -

**Agriculture, fisheries, forestry and food**

Followers: 4 Datasets: 13.7k

[Follow](#)

**13,700 datasets found** Order by:

**Kreisstrassenflächen Landkreis Nordwestmecklenburg**

The dataset includes the Kreisstrassen strassenbegleitflächen strassenflächen as w district of Northwest Mecklenburg. Information on the nature of the land...

Germany (7537)

Italy (2159)

Spain (1231)

Poland (837)

**Development plan No 177 – analogue and digital**

Naufstellungsbeschluss schleswiger road/rude\ : 1986-11-13\ nsatzungsbeschluß: 2

GDI-DE | July 22, 2016

GDI-DE | July 21, 2016

# Data source: [data.gov.uk](https://data.gov.uk)

- ▶ <https://data.gov.uk/data/search>
- ▶ UK Government Open Data portal - pioneer in this field
- ▶ 5700 data sets for agriculture, 6000 for food

The screenshot shows the DATA.GOV.UK website interface. At the top, the logo reads "DATA.GOV.UK Beta" with the tagline "Opening up Government". Navigation links include "Home", "Data", "Datasets", "Map Search", "Data Requests", "Publishers", "Data API", "Organograms", "Site Analytics", and "Reports". The main content area displays "42,333 results" for a search query. On the left, there are filters for "PUBLISHED STATUS" (38792 published, 3541 unpublished), "COLLECTION" (318 National Information Infrastructure, 249 Organogram, 14 Registers, 2 Intelligent Transport Systems), and "API" (42314 datasets with APIs). On the right, three dataset results are visible: "LIDAR Composite DSM - 1m" by Environment Agency, "Planned road works on the HE road network" by Highways England, and "Road Safety Data" by Department for Transport.

# Data Source: data.gov

- ▶ data.gov of US government.
- ▶ “Under the terms of the 2013 Federal Open Data Policy, newly-generated government data is required to be made available in open, machine-readable formats, while continuing to ensure privacy and security”
- ▶ 565 data sets
- ▶ It has lost a bit of momentum ...

The screenshot shows the data.gov website. At the top left is the logo with an American flag and the text 'DATA.GOV'. To the right are navigation links: 'DATA', 'TOPICS', 'IMPACT', 'APPLICATIONS', and 'DEVELOP'. Below the navigation is a blue banner with the text 'The home of the U.S. Government's open data' and a sub-header 'Here you will find data, tools, and resources to conduct research, develop web and mobile applications, design data visualizations, and more.' Below the banner is a 'GET STARTED' button with the text 'SEARCH OVER 196,904 DATASETS'. Below the button is a search bar containing the text 'Health Care Provider Charge Data'. At the bottom is a 'BROWSE TOPICS' section with icons and labels for Agriculture, Climate, Consumer, Ecosystems, Education, and Energy.

# Data Source: World Bank

- ▶ <https://data.worldbank.org/>
- ▶ “At the World Bank, the Development Data Group coordinates statistical and data work and maintains a number of macro, financial and sector databases.”
- ▶ Open Data is subset of all data it possesses
- ▶ 236 datasets for agriculture and rural development

The screenshot shows the 'World Bank Open Data' website. At the top, it says 'Free and open access to global development data'. Below this is a search bar with the placeholder text 'Search data e.g. GDP, population, Indonesia'. There are also links to 'Browse by Country or Indicator'. The main content area is divided into several sections:

- MOST RECENT:** A list of recent news and blog posts, including 'Chart: What Are the World's Wettest Countries?' by T. Khokhar (Sep 15, 2017), 'Achieving good budgetary governance: What have we learned from PEFA in the past decade?' by L. Hawke (Sep 13, 2017), and 'Introducing Data360R — data to the power of R' by R. Onglao (Sep 12, 2017). There are buttons for 'View all news' and 'View all blogs'.
- WHAT YOU CAN LEARN WITH OPEN DATA:** A section featuring a chart titled 'Poverty headcount ratio at \$1.90 a day (2011 PPP) (% of population)'. The chart shows a downward trend in poverty headcount ratio over time, with a 'WORLD' button. Below the chart is the heading 'Extreme Poverty' and the text 'The proportion of the world's population living in extreme poverty has dropped significantly'.
- Atlas of Sustainable 2017:** A section titled 'Atlas of Sustainable 2017 From W' with a grid of world maps showing various indicators. Below the maps is the text 'A Visual Guide to' and the date 'Apr 17, 2017'.
- RECENTLY UPDATED DATASETS:** A section titled 'World Development Indicators'.

At the bottom of the page, there is a SoundCloud player and a logo for a tiger.



# Data source: [data.gov.gr](http://data.gov.gr)

- ▶ <http://www.data.gov.gr/>
- ▶ “Το data.gov.gr είναι ο κεντρικός κατάλογος των δημόσιων δεδομένων που παρέχει πρόσβαση σε βάσεις δεδομένων των φορέων της ελληνικής κυβέρνησης.”
- ▶ Γεωχωρικά/  
Γεωσκόπηση και  
Περιβάλλον - 49 data sets

**Αναζήτηση δεδομένων**

π.χ. Περιβάλλον

Ετικέτες Γενικά Αρχεία του Κ... Αρχεία Εργαλεία έρευνας

Το data.gov.gr είναι ο κεντρικός κατάλογος των δημόσιων δεδομένων που παρέχει πρόσβαση σε βάσεις δεδομένων των φορέων της ελληνικής κυβέρνησης.

**DATA**

Welcome to Our New Website!

**data.gov.gr στατιστικά**

<b>5,7k</b>	<b>257</b>	<b>9</b>
σύνολα δεδομένων	Φορείς	ομάδες

**data.gov.gr links**

- Αποφάσεις διάθεσης δεδομένων
- Οδηγός Διαχειριστή Φορέα
- Ετήσια Έκθεση

# Data Source: Eurostat

- ▶ <http://ec.europa.eu/eurostat/web/main/home>
- ▶ Main statistical office for EC.
- ▶ Much data can be downloaded (but is not strictly “open”)

The screenshot shows the Eurostat website interface. At the top, there is the Eurostat logo and the tagline 'Your key to European statistics'. A search bar is located on the right. Below the logo, there are navigation tabs for 'News', 'Data', 'Publications', and 'About Eurostat'. The main content area is titled 'European Commission > Eurostat > Agriculture > Overview'. On the left, there is a sidebar menu with options like 'Data', 'Methodology', and 'Legislation'. The main content area is titled 'Agricultural statistics' and contains text about the purpose of agricultural statistics and a list of topics covered.

Legal notice | RSS | Cookies

Type a keyword, a code, a title...

News Data Publications About Eurostat

European Commission > Eurostat > Agriculture > Overview

AGRICULTURE OVERVIEW

OVERVIEW

– Data

- Main tables
- Database
- Methodology
- Legislation
- ESS Agreement
- Publications
- Statistics illustrated

**Agricultural statistics**

Initially, the purpose of Agricultural statistics was to **monitor the Common Agricultural Policy's main objectives**, such as the production and supply of agricultural products and income in the farming sector.

In recent decades, new indicators and statistics have emerged that reflect changing EU policy. Some, for example, such as the use and impact of **pesticides**, have an environmental focus. Another development in agricultural statistics has been to establish statistics related to the **organic farming**.

In 2015, a **strategy for agricultural statistics for 2020 and beyond** was agreed, designed to modernise this area of statistics.

Today, agricultural statistics cover topics as diverse as:

- ① information about the structure of farms, orchards & vineyards;
- ② agricultural production;
- ③ economic accounts for agriculture;
- ④ agriculture and environment.

Comprehensive information is available at both a national and regional level.

# Data Formats

- ▶ CSV - single common denominator
- ▶ EXCEL - most statistical data in this format
- ▶ XML - much better, but can be confusing
- ▶ JSON - favourite format for developers, not so frequently offered
- ▶ RDF - essential for **Linked Data**
- ▶ ARCGIS - (not a format, a tool) some open data in this form

# Insufficient use of standards

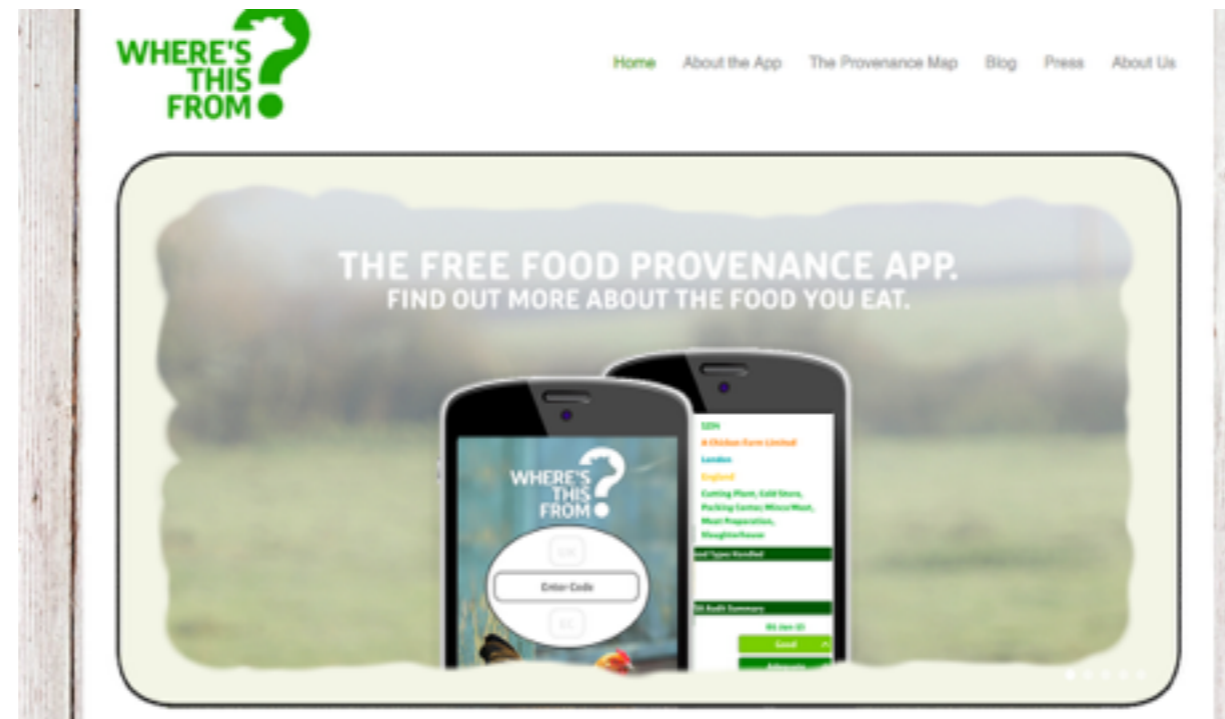
- ▶ Many different kinds of standards across three communities (research, precision farming, supply chain)
- ▶ Great failure to align and make interoperable.
- ▶ Most data is not compliant with a standard, especially not a semantic standard.
- ▶ Failure to agree on unique identifiers
- ▶ People revert to lowest common denominator (csv files).
- ▶ Creates barriers to wider uptake of data in general.

# The Uses of Open Data in Agrifood

# Benefits of Open Data

- ▶ Early accurate detection and prediction of problems (pest outbreaks, water shortages, floods)
- ▶ Planning what to grow, what treatments to apply, when to plant or harvest
- ▶ Risk management (insurance, hedging), and damage control (from drought and pests)
- ▶ Managing agricultural subsidies
- ▶ Informing consumers and supply chain participants

## Classic Example: The “Where’s this from” App



- ▶ Used UK’s FSA meat audit reports linked packaging codes
- ▶ Reported on: animal welfare, slaughtering hygiene, animal hygiene, HACCP (bacteria)
- ▶ Weaknesses: audits are carried out by local authorities, budget cuts, and often no inspection has occurred. **NOW DEFUNCT**

# Example: GroenMonitor

- ▶ Protecting crops from pest outbreaks with vegetation maps in Netherlands
- ▶ Shows a current vegetation map of the Netherlands, based on satellite images and maps made publicly available through the European Space Agency (ESA)
- ▶ In 2014, identified 12,000 ha affected by mice
- ▶ <http://www.groenmonitor.nl/>





# Example: Plantwise

- ▶ Boosting crop yields with a best practice knowledge bank in developing countries. Developed by CABI and based on the Open Access Plantwise Knowledge Bank.
- ▶ Plantwise helps smallholder farmers in developing countries deal with plant health issues.
- ▶ Combines global and local open access data from sources such as CABI's databases, research publications and governmental data.
- ▶ <https://www.plantwise.org/>



# Example: aclimatecolombia

- ▶ Led by CIAT in Columbia, in collaboration with Colombian ministry, farmer and private sector, helping farmers take precautions in rice harvests.
- ▶ Combined open and private data, to analyse issues behind the decreasing rice crop yields.
- ▶ Resulted in climate-smart agriculture decision-making tool, open to everyone.
- ▶ <http://www.aclimatecolombia.org/>

## Clima y Sector Agropecuario Colombiano Adaptación para la Sostenibilidad Productiva



[Bases de datos](#) [Glosario](#) [Acerca De](#)

### Predicción Climática

del clima, para la ciudad de Yopal, para el período comprendido entre Septiembre, 2017 y ... para hacer estas predicciones fue Aeropuerto Yopal - 35215010, provista por el IDEAM. Ten ... mes central del trimestre pronosticado. Por ejemplo: si el mes es Marzo, el trimestre que s ... dicción climática es dada en porcentaje de probabilidad con respecto al rango normal de precipi ... abajo, usted podrá encontrar cual es el rango normal de precipitación para un municipio s

Septiembre-2017  
probabilidad de precipitación (%)



### Escenarios de la predic

Variable	Minimo	Promedio
Precipitación	150 mm	188 mm
T. Máxima	31 °C	31 °C
T. Mínima	22 °C	22 °C
Predicción color	207 col/cm2d	492 col/cm2d

# Example: Good Growth Plan

- ▶ Syngenta undertook to open up a large data set from its experimental farms in 2013
- ▶ Six commitments to improve crop productivity, protect soil and biodiversity, train smallholders and ensure labour standards, with targets to be achieved by 2020.
- ▶ Data tracking this is open data.
- ▶ Objective is to build an open, collaborative platform to co-create solutions that minimise the use of resources required to feed a growing population and preserve habitats for biodiversity.
- ▶ <https://www4.syngenta.com/what-we-do/the-good-growth-plan/progress/progress-open-data>

The screenshot shows a webpage with a navigation bar at the top containing 'Who we are', 'What we do', and 'How we do it'. Below the navigation bar is a breadcrumb trail: 'Who we are > About our business > Our stories > Improving yields and livelihoods'. The main content area features a large photograph of a lush green field with rows of crops supported by wooden stakes. Overlaid on the bottom of the photo is the text 'Improving yields and livelihoods'. Below the photo is a 'Join the conversation' button and a 'Share' button with a Facebook icon. The text below the photo reads: 'How the people behind The Good Growth Plan are helping empower smallholders'. A short paragraph follows, mentioning Priscar Mwangangi, a smallholder in Kenya who grows tomatoes, peppers, and beans, and discusses challenges like heavy rains and crop disease.

# Example: How Much Sugar?

- ▶ Created by Open Food Facts, as an educational game. Uses their huge data set.
- ▶ Targets end consumer.
- ▶ Tests your knowledge of sugar content of foods.
- ▶ <http://howmuchsugar.in/>



# Greek example: GAIA Sense

- ▶ Neuropublic + Gaia Epichirein building network of sensors combined with Copernicus data for agricultural advice
- ▶ Gradual roll out 2016-2020



🛡️ Crop protection
🌱 Automate Irrigation
📄 Traceability
📡 Atmo Station
🌱 Soil Station
📡 Satellites
📡 RPAs
**Pilots 2016**

📡 @Services GAIA InFarm
📡 @Infrastructure GAIA Sense
📡 @EO GAIA Remote Sense
📡 Running
📡 Schedule

# Issues and Challenges: Data and Control

# Big Data

- ▶ Recent growth of interest in Big Data in farming
  - ▶ Poster child is “Climate Corp” bought by Monsanto. This uses “**open**” meteorological data as a basis.
  - ▶ Many other actors seeing agriculture as “big data” domain.
- ▶ BUT - big data approaches need access to a lot of data, and if farmers provide data will they benefit?
- ▶ The political problem here is the danger of crowding out small farmers

# Precision Agriculture

- ▶ Precision agriculture depends on combining open data (GPS, Meteo, remote sensing) with local/private data (on field sensors).
- ▶ It has both a “big data” aspect and a “micro data” aspect.
- ▶ Depends on farmers making data available, but also **ongoing availability of open data** (GPS, satellite observation, crop models etc.)
- ▶ The major push we are seeing for PA **must (?)** also mean that agrifood open data is supported even more.



# Privacy

- ▶ According to one big agricultural machinery manufacturer “GDPR has destroyed our business model”
- ▶ GDPR = General Data Protection Regulation of the EC (<http://www.eugdpr.org/>)
- ▶ Farm data is also private data in some regards.
- ▶ Eating/food purchasing data may considered private too.
- ▶ Lots of scope for complex demarcations between private vs. open data

# Thank You

# QUESTIONS

# Resources

- ▶ All references and copies of slides available from:
  - ▶ [www.cbrewster.com/haicta2017](http://www.cbrewster.com/haicta2017)