

The European Commission's science and knowledge service

Joint Research Centre

Data4policy

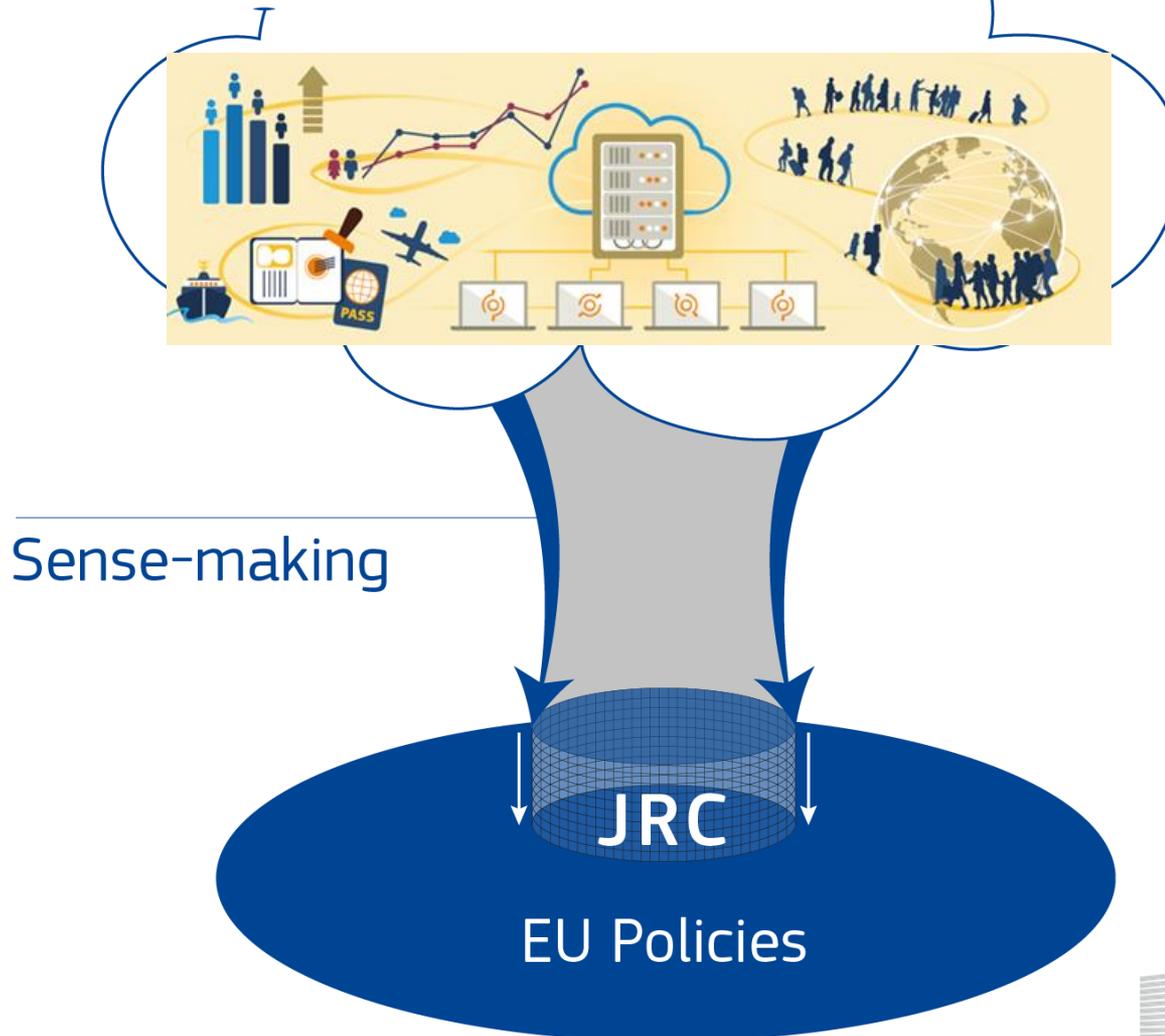
Bratislava 14 November 2016

Vladimír Šucha,
Director-General

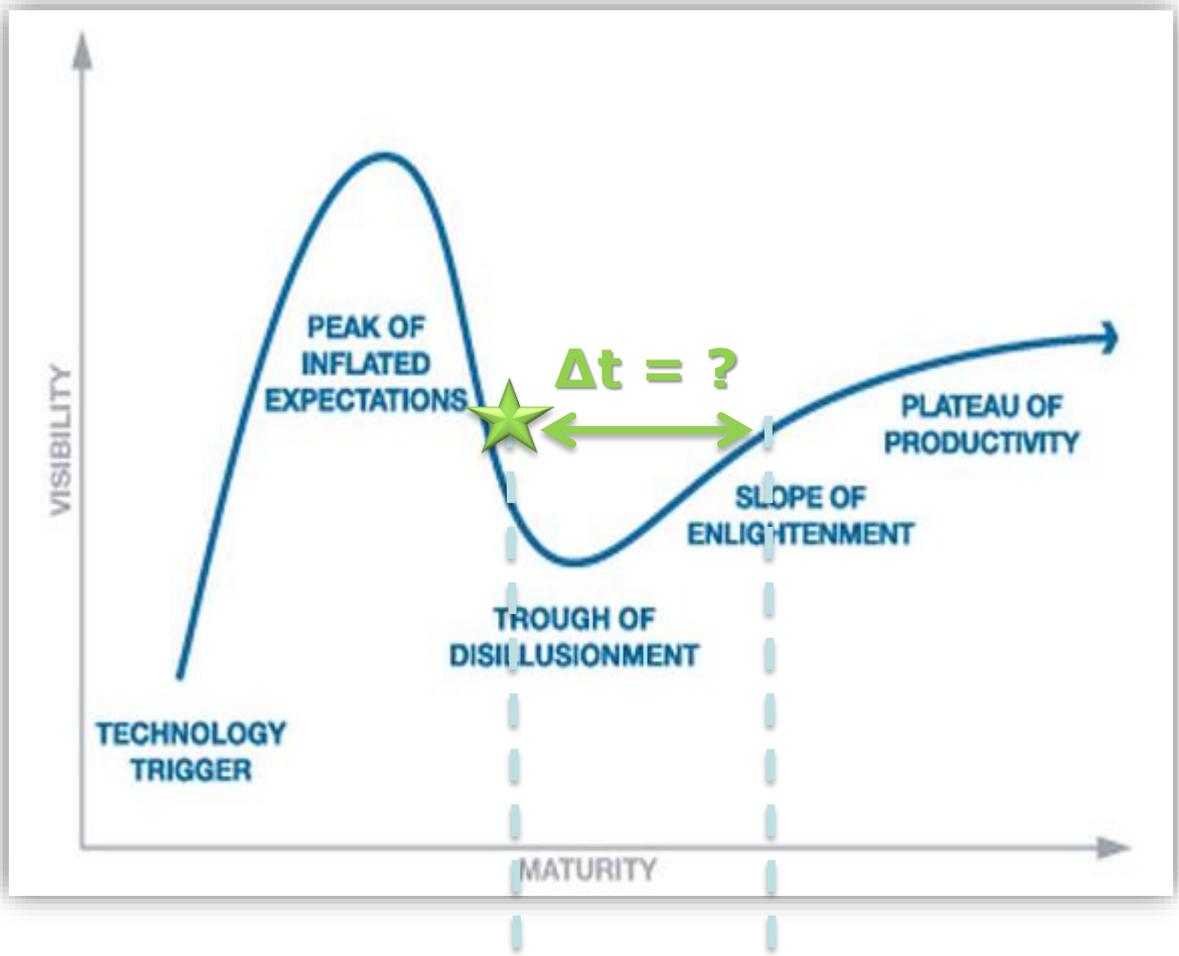


EC Science and knowledge service

Global knowledge production



Data4policy – where are we on the "hype cycle"?



Infrastructure

Hadoop On-Premise
cloudera Hortonworks MAPR Pivotal IBM InfoSphere bluedata jethro

Hadoop in the Cloud
amazon Microsoft Azure Google Cloud Platform IBM InfoSphere CAZENA TREASURE DATA altiscale Du bale

Spark
databricks GridGain TACHYON NEXUS

Cluster Services
amazon kubernetes amazon mesosphere docker HPCC SYSTEMS MESOSPHERE Core OS peppercuda StackIQ

Analytics

Analyst Platforms
Palantir AYASDI Quid enigma Digital Reasoning ORBITAL INSIGHT

Analytics Platforms
Microsoft Microsoft GUAVUS Datameer Bottlenose interana

Data Science Platforms
context relevant DataRobot CONTINUUM Alpine MODE plioty ARIMO datatiku tonian DOMINO sense yhat ALGORITHMIA

Visualization
tableau Google Cloud Platform Qlik looker Roambi SENSE SCHDATA dakaroma CHARTIO

Applications

Sales & Marketing
RADIUS Gainsight bloomreach Zeta EVERSTRING livefyre blueyonder Lattice kahuna infer SAILTHRU persado AVISO sence QUANTIFIND ACTIONIQ fusemachines JENGA GIO

Customer Service
MEDALLIA ATTENITY CLARABRIDGE CLICKFOX STELLA Service NGDATA Preact DigitalGenius appurri Wuseio

Human Capital
gild Connectifier entelo hiQ

Legal
RAVEL JUDICATA Everlaw Brevia PRESENTION

NoSQL Databases
amazon dynamodb Google Cloud Platform Microsoft Azure ORACLE mongoDB MarkLogic DATASTAX Couchbase KERO SPIKE SequoiaDB redislabs Influxdata

NewSQL Databases
SAP SAPANA Clustring Pivotal paradigm4 nuODB splice IMAGING MariaDB VOLTDB citusdata deopdb Trafodion Cockroach LABS

BI Platforms
Power BI amazon Domo Wave Analytics GoodData birst platform atscale

Statistical Computing
sas SPSS MATLAB

Log Analytics
splunk sumologic hibana Cloud PHYSICS loggly

Social Analytics
Hootsuite NETBASE DATASIFT track bitly synthiso simple reach

Ad Optimization
AppNexus critereal MediaMath OpenX rocket hub Intel Ad Science theTradeDesk Liventent TAPAD DataXu Cippier MOAT

Security
CYCLONE CounterTack cyberason ThreatMetrix AREA 1 SECURITY SentinelOne Recorded Future Guardian Analytics FORTSCALE sif science Keybase feedzai SICNIFYD

Vertical AI Applications
facebook Clara KASIST lumiatia

Graph Databases
neo4j OrientDB InfretraGraph

MPP Databases
TERADATA VERTICA NETEZZA COTION koginito SQL dremio

Cloud EDW
amazon Microsoft Azure Pivotal snowflake WATERLINE DATA Infoworks

Data Transformation
alteryx talend tamr StreamSets Alation

Data Integration
informatica MuleSoft snapLogic Bedrock Data xplenty

Real-Time
amazon METAMARKETS Streamlit confluent dataArtisans

Machine Learning
Azure Machine Learning H2O amazon SKYTREE rapiminer DATASIRI deepsense YISENZE Prediction IO flowflask

Speech & NLP
NarrativeScience NUANCE semantic technologies ARRIA cortico.ai Mindfield IDIBON YSECOPI

Horizontal AI
IBM Watson Cortana. sentient n.ryana viv nara Numenta clarifai MetaMind

Publisher Tools
Outbrain Taboola quantcast Chartbeat yieldbot Yieldmo

Govt / Regulation
Socrata OPENGOV FN FiscalNote enigma mark43 PREDPOL OpenDataSoft

Finance
affirm OnDeck LendingClub Kreditech LendUp Kabbage tidemark INSIKT UORA Dataminr Lenddo KENSHC AIDYIA ISENTIUM Quantopian sentient

Management / Monitoring
New Relic APPDYNAMICS amazon octifio Numerify splunk DATADOG TROICANO DRIVEN Anodot

Security
TANUM illumio CODE42 DataGravity CipherCloud VECTRA sqrrl BlueTalon

Storage
amazon Microsoft Azure panasas nimblestorage COHO Qumulo

App Dev
apigee CRASK Typesafe DRIVEN

Crowd-sourcing
amazon mechanicalturk CrowdFlower WorkFusion

Search
hp ORACLE ENDECA EXPLERD Lucidworks elastic ThoughtSpot MAANA swifttype Algolia SINEOUI

Data Services
UC OPERA Mu Sigma EXL DATA SCIENCE kaggle dallascopa DataKind

For Business Analysts
OrigamiLogic ClearStory CIRRO import IO

Web / Mobile / Commerce
Google Analytics mixpanel RJMetric BLUECORE AMPLITUDE granify sumall Airtable retention custora

Education / Learning
KNEWTON Clever eclarra PANORAMA knowre

Life Sciences
23andMe COUNSEL Recombine KYRUS FLATIRON zymergen HealthTop METABIOTA ZEPHYR HEALTH ovia SwiftKey Gingerio transcriptic Glow enlitic AICure Atomwise

Industries
OPOWER eHarmony RetailNext STITCH FIX WorkFusion BLUE RIVER TACHYUS Seeq FarmLogs SwiftKey HowGood collect SHIRT MACHINE statmuse B@XEVER

Cross-Infrastructure/Analytics

amazon Google Microsoft IBM SAP SAS data hp Autonomy VERTICA vmware TIBCO TERADATA ORACLE NetApp

Open Source

Framework
hadoop HADOOP YARN Spark MESOS TEZ Flink CDAP

Query / Data Flow
SLAMDATA ARAPACHE DRILL Google Cloud Dataflow HIVE

Data Access
cassandra HBASE mongoDB CouchDB riak OPEN TSCOE kafka nifi

Coordination
talend Apache Zookeeper Apache Ambari

Real-Time
STORM Spark APEX Flink TACHYON druid

Stat Tools
ScalaLab Nuanfy SciPy

Machine Learning
mlilb Apache SINGA Aerosolve Caffe WEKA FeatureFu DIMSUM jupyter DL4J

Search
elasticsearch Solr lucene

Security
Apache Ranger

Visualization
Zeppelin

Data Sources & APIs

Health
JAWBONE GARMIN practicefusion fitbit Withings VALIDIC netatmo kinsa Human API

IOT
UPTAKE ThingWorx helium samsara

Financial & Economic Data
Bloomberg THOMSON REUTERS DOW JONES YODLEE PREMISE S&P CAPITAL IQ quandl xignite CBINSIGHTS mattermark Stocktwits @estimote PLAID

Air / Space / Sea
PLANET LABS WINDWARD SPIRE CRUISE SETCATCH Airware DroneDeploy

Location / People / Entities
axioma Experian EPSILON InsideView GARMIN foursquare STREETLINE esri Crimson Hexagon CARTODB factual. PlaceIQ CIRCULATE placemeter BASIS Sense

Other
qualtrics panjiva DATA.GOV

Incubators & Schools
GA PLURAL SIGHT INSIGHT DataCamp DataElite The Data Incubator METIS

Last Updated 3/23/2016

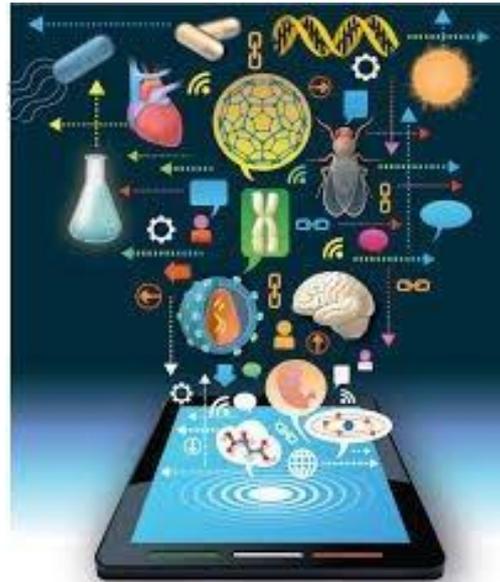
© Matt Turck (@mattturck), Jim Hao (@jimhao), & FirstMark Capital (@firstmarkcap)

FIRSTMARK



The need for evidence based policies

**Robust science
& evidence /
policy options**

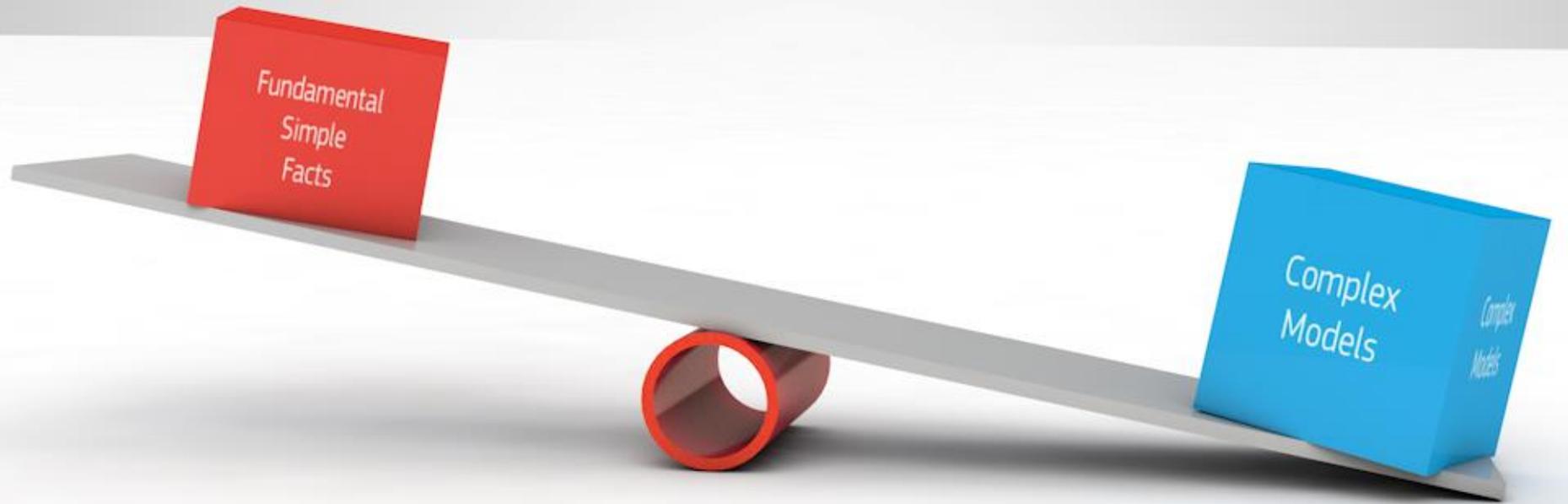


**Data /
feedback**



Policies





Public data



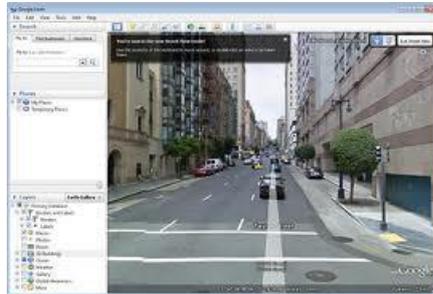
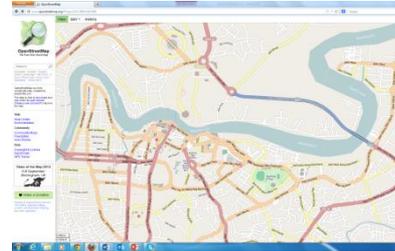
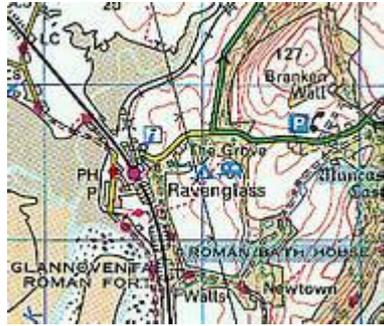
Geo-data landscape evolution

<2000-2016
Government

2005-2016
Private Sector

2008-2016
VGI

2010-2016
Social media



New data sources – opportunities and paradoxes

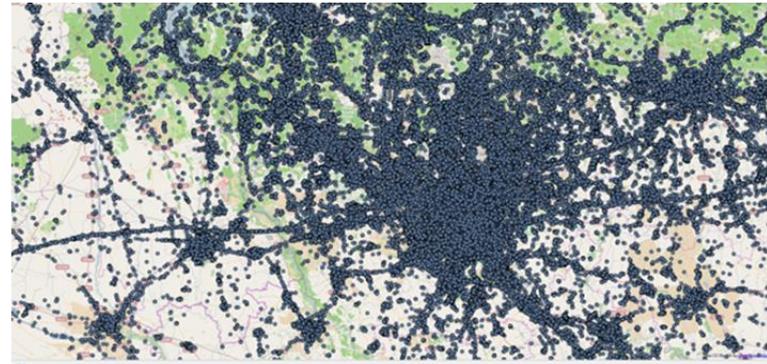
Multiple sources:

- same entity (city of Milan)
- different views

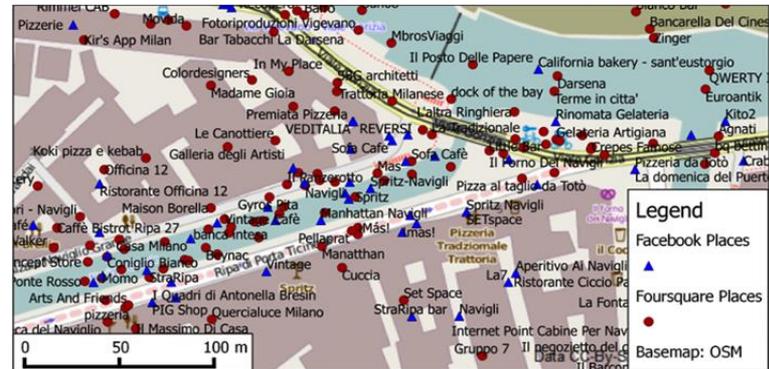


Milan by Copernicus

Source: [ESA](#)



Milan by Twitter



Milan by FourSquare & Facebook

Use of mobile phone data: Up-to-date population

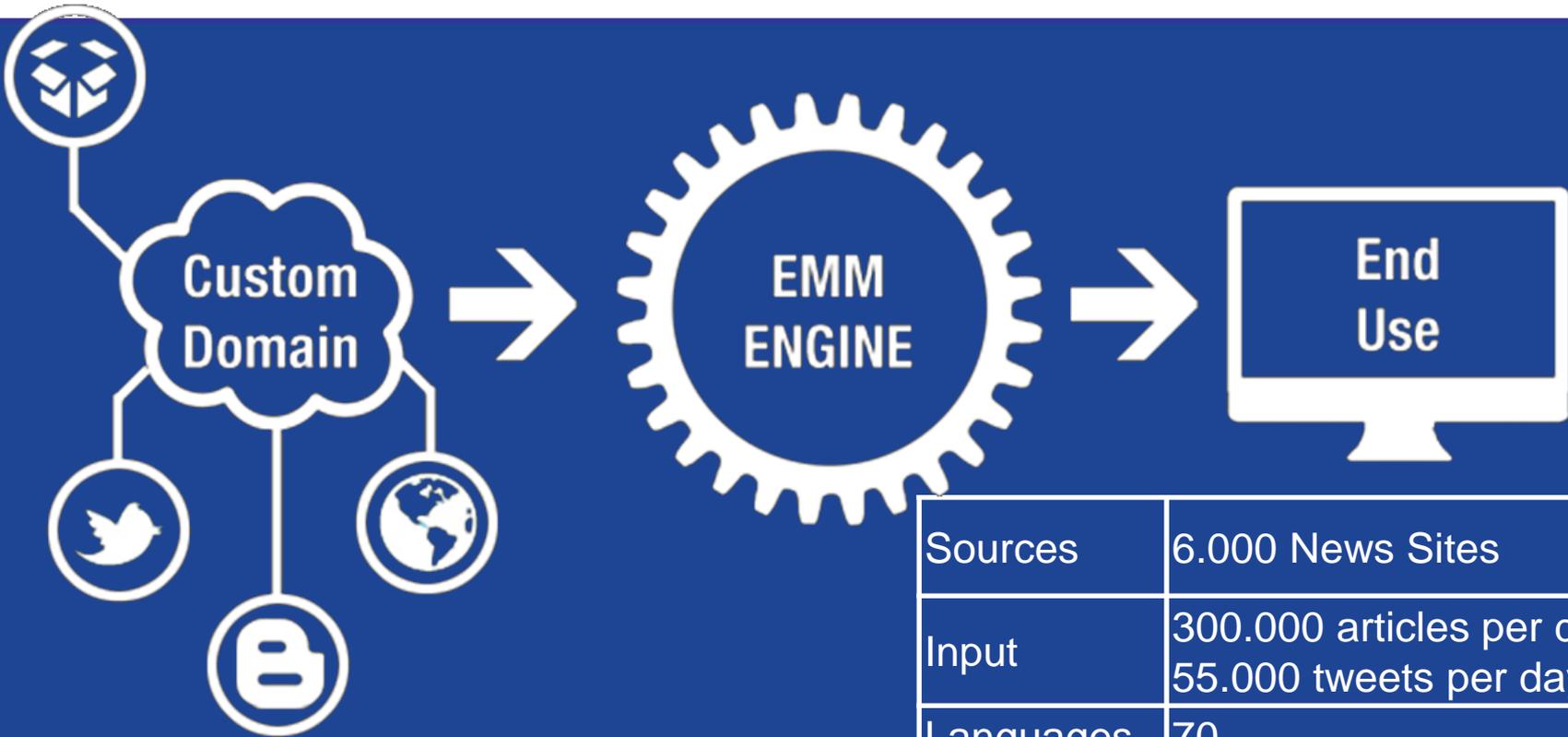


Development of new indicators on mobility from car sharing platforms



European
Commission

News wires



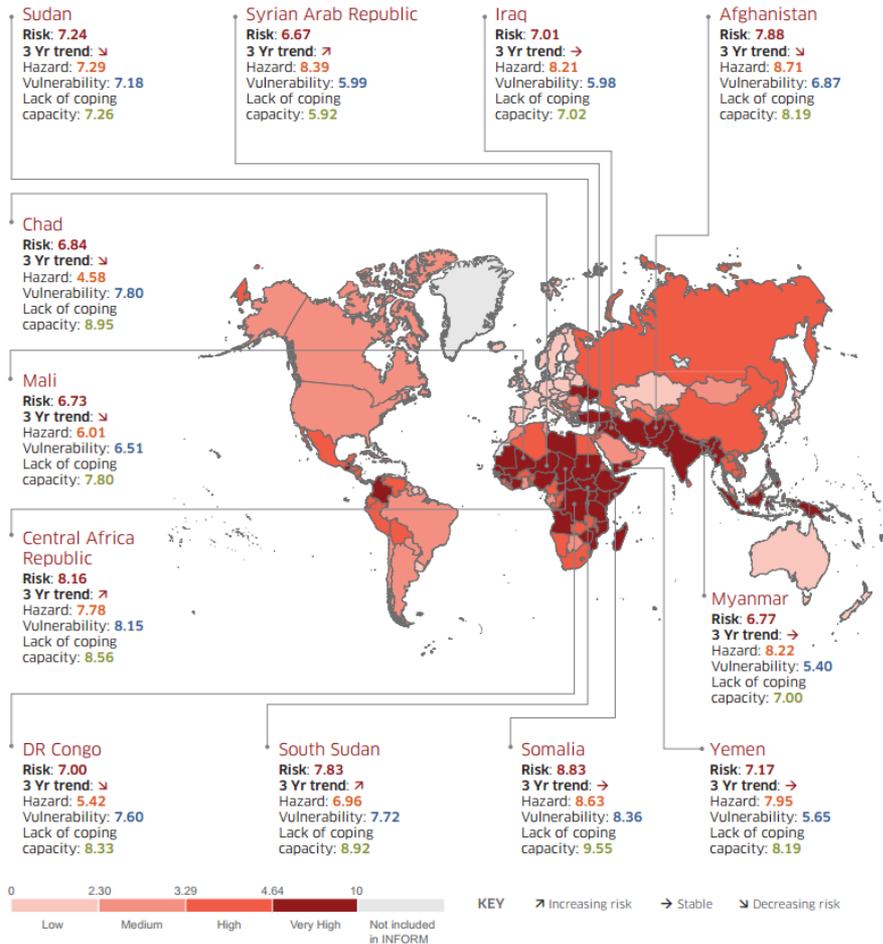
Sources	6.000 News Sites
Input	300.000 articles per day 55.000 tweets per day
Languages	70
Categories	1.000 classes
Classes	30.000 keywords
Runs	24/7
Visitors/day	25.000



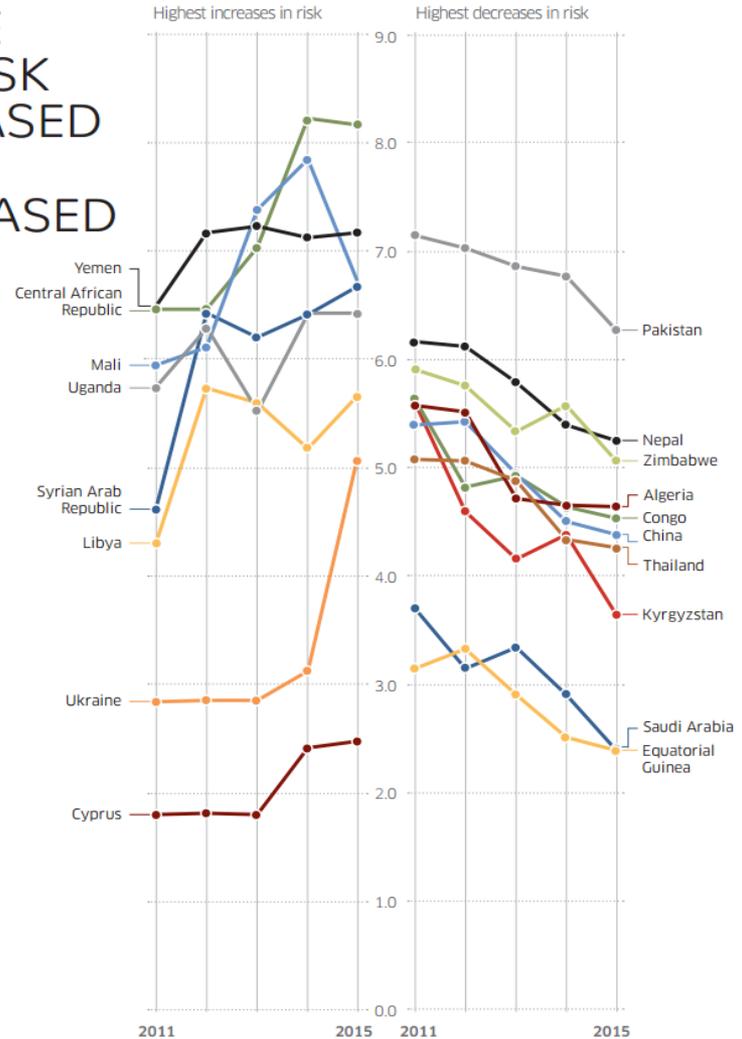
RISK OF HUMANITARIAN CRISES AND DISASTERS

The overall INFORM risk index identifies countries at risk from humanitarian crises and disasters that could overwhelm national response capacity. It is made up of three dimensions - hazards and exposure, vulnerability and lack of coping capacity. This map shows details for the 12 countries with the highest overall risk.

INFORM 2015 RISK INDEX



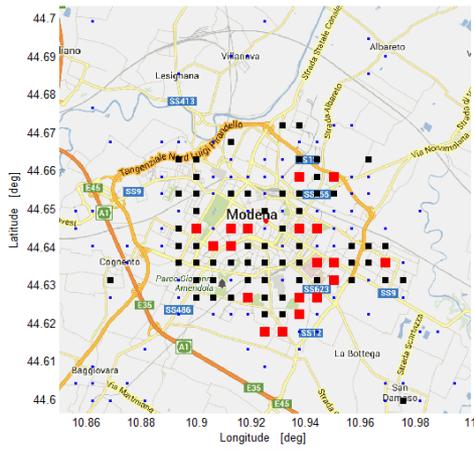
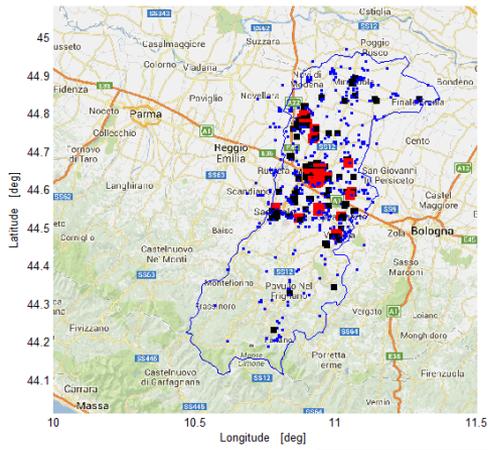
WHERE HAS RISK INCREASED AND DECREASED MOST?



Data Mining: real world travel behaviour

Medium size vehicle – Smart-AC recharges

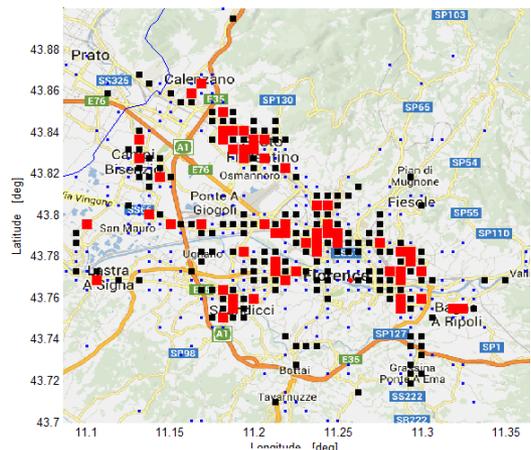
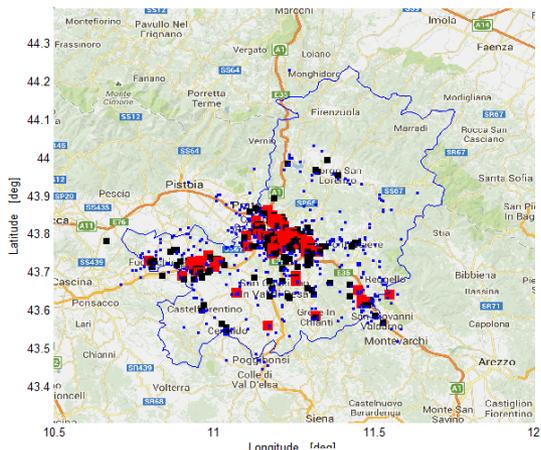
Modena



100km x 100km
500m edge mesh

• 20 < kWh/day < 100 ■ 100 < kWh/day < 250 ■ kWh/day > 250

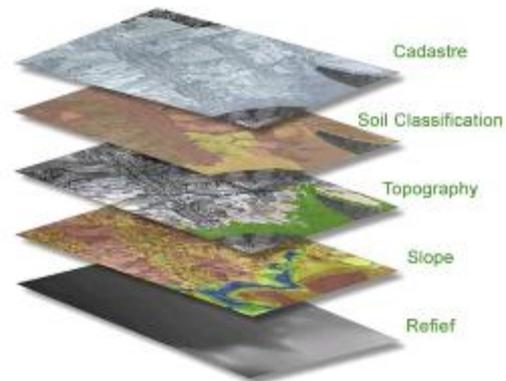
Firenze



• 20 < kWh/day < 100 ■ 100 < kWh/day < 250 ■ kWh/day > 250

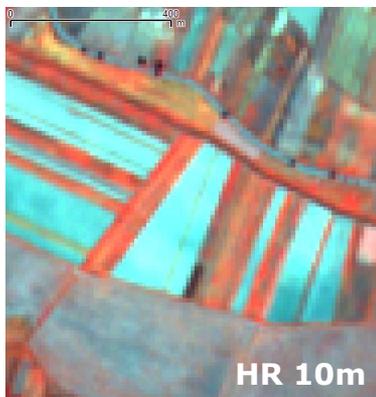


INSPIRE



Positioning and measurements
Anytime, anywhere in Europe

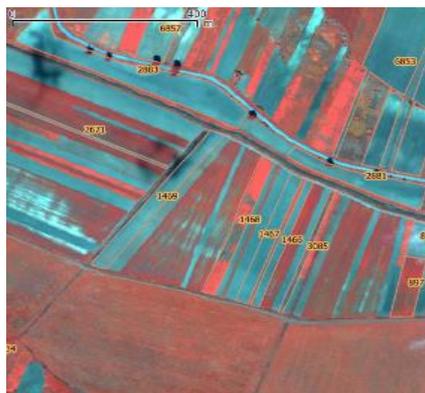
Land use \ Land cover



Today 126 million parcels in EU

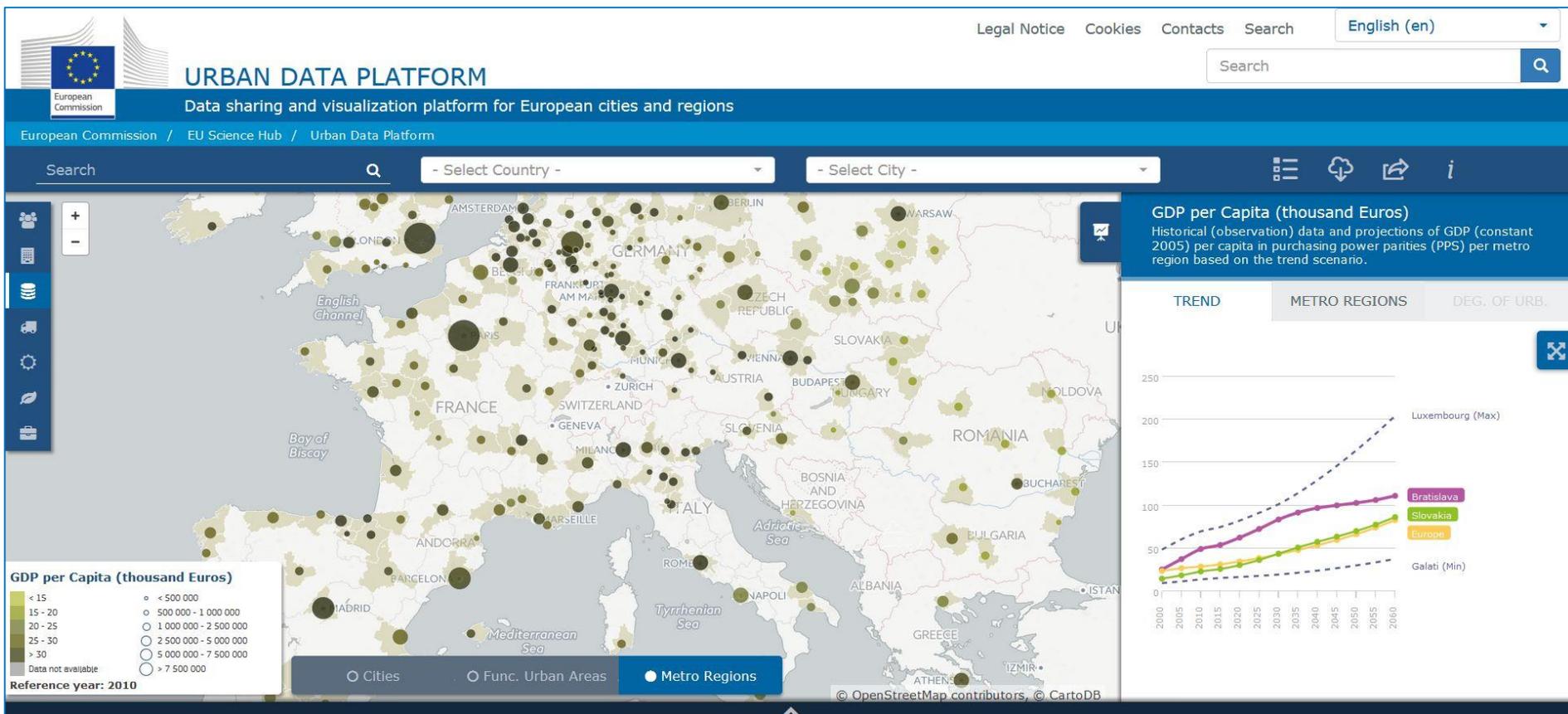
1:5.000 scale

Land Parcel identification System (LPIS)



How is your city performing?

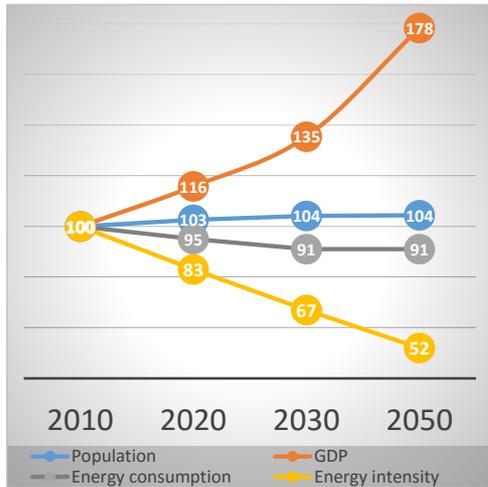
- Explore the performance of your city in various fields such as **demography, urban development, economic development and resource efficiency.**
- Compare more than 800 urban areas using **60 indicators.**



- <http://urban.jrc.ec.europa.eu/>

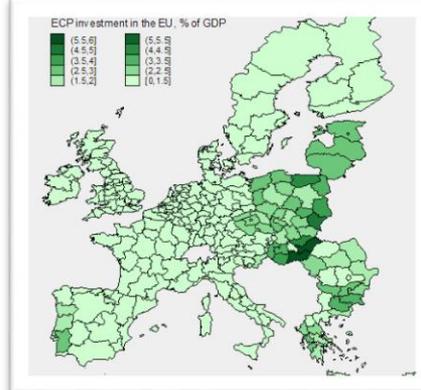
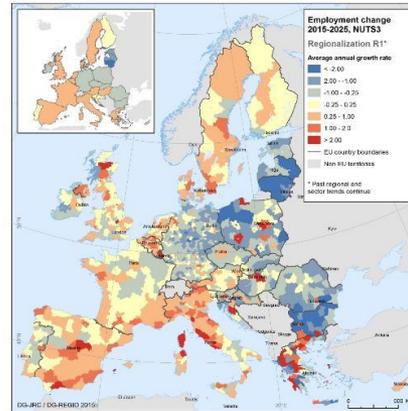
From macro-trends to place-based integrated assessments

EU Reference projections and scenarios (ESTAT, ECFIN, ENER, AGRI..)



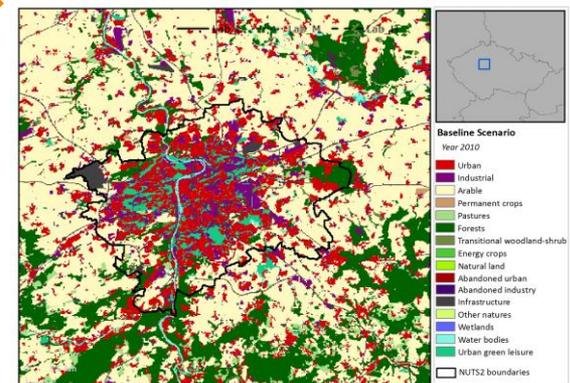
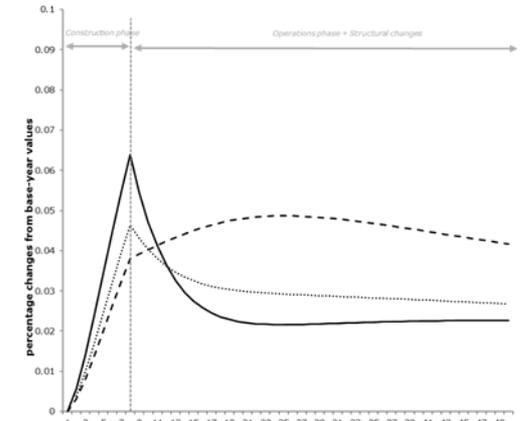
28 Member States

Regionalisation
Shocks & Investments



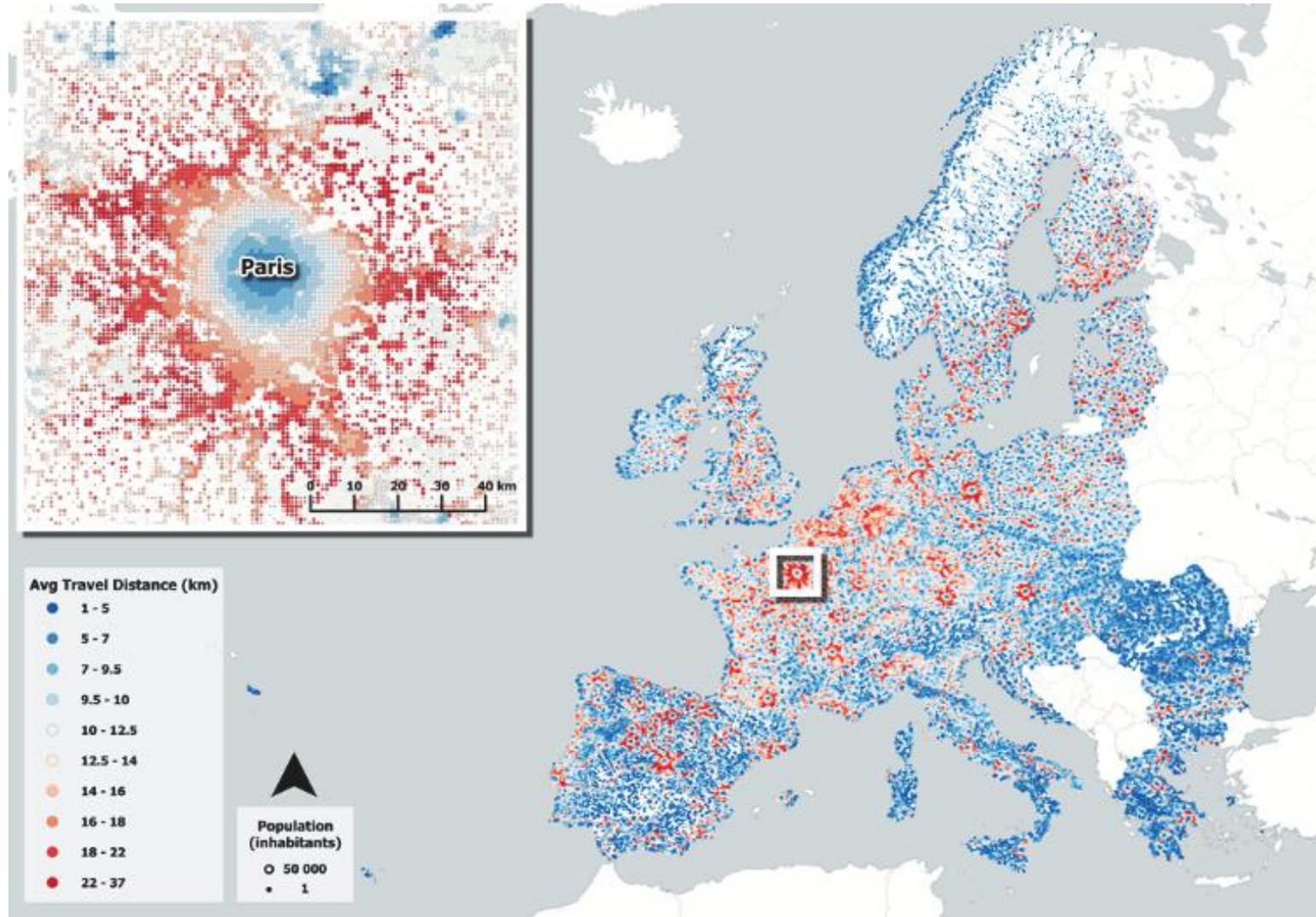
272 NUTS-2
1300 NUTS-3
~130 000 LAU-2

Regional Performances Indicators
Territorial Impact Assessment



> 4 10⁶ 1x1 km cell
> 4 10⁸ 100x100 m cell

Is transport infrastructure efficient?



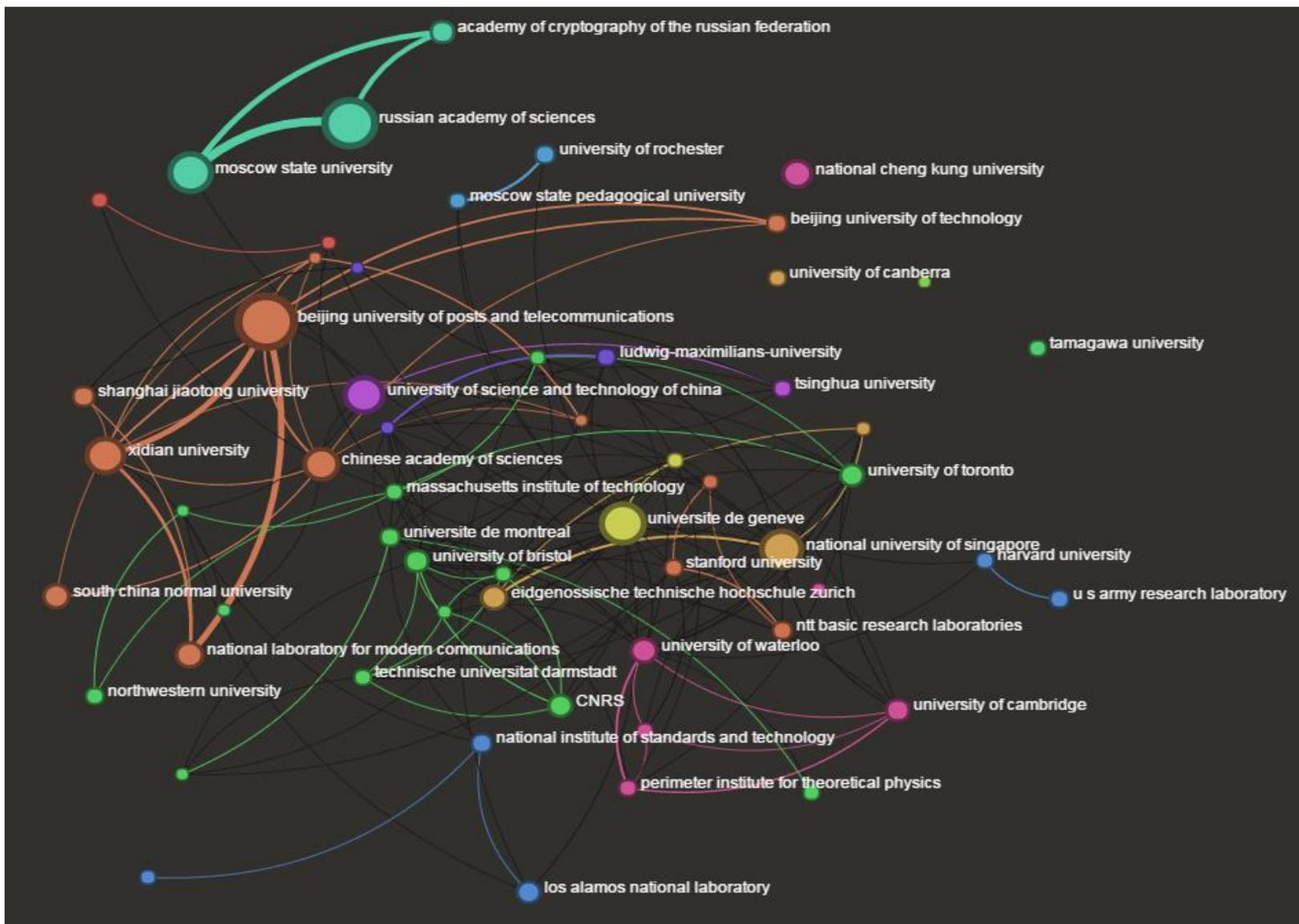
Estimated average travel distances of EU citizens.

The way in which cities develop often affects the efficiency of their transportation systems. City peripheries can have high average travel distances, hence need specific transport-related investments.

Blue Hub



TIM



vodafone IT 13:48



Invasive Alien Species in Europe

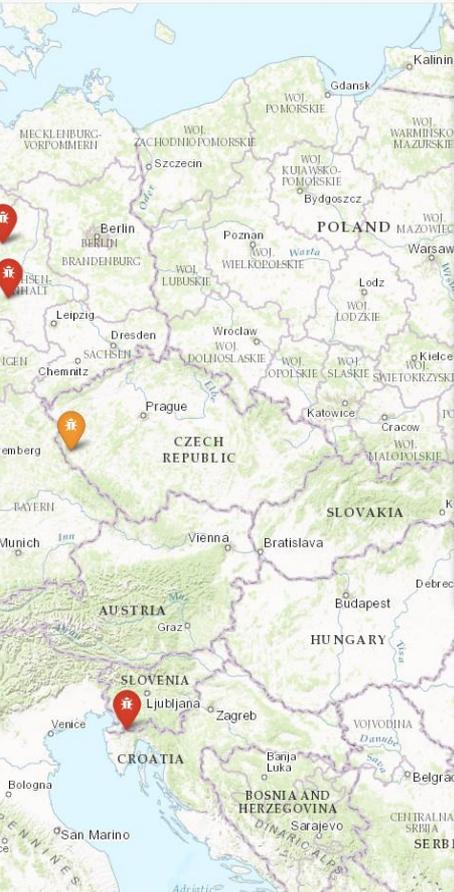




European Commission > MyGeoss

About | Contact | Legal Notice | Cookies

Admin New Report Exit





Status: Submitted

Coordinates
45.512 / 14.032

Abundance
1 number of individuals

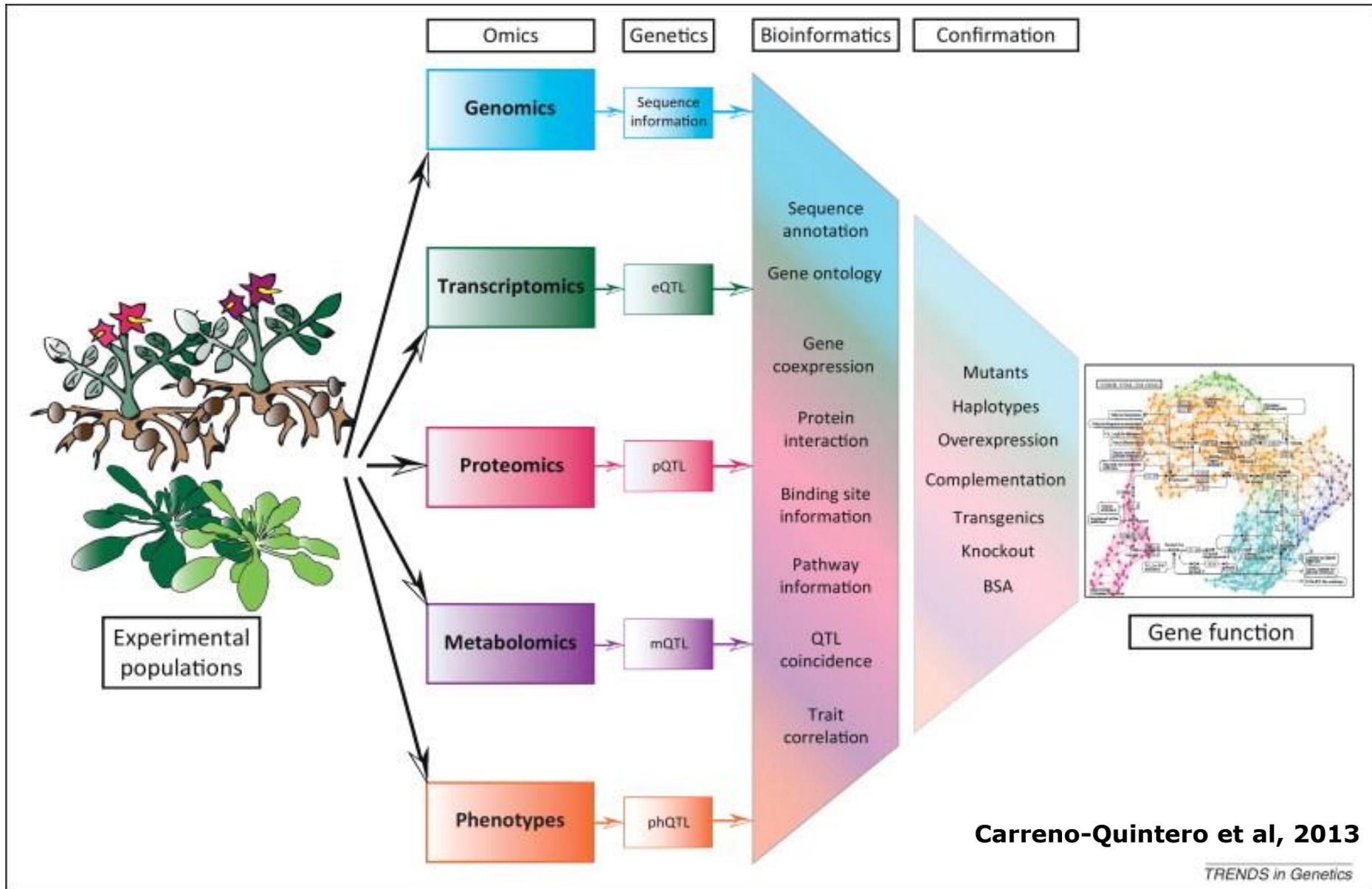
Precision
Measured

Comment
Habitat : terrestrial. Comment :

ICCID
629dd5f87c98d892

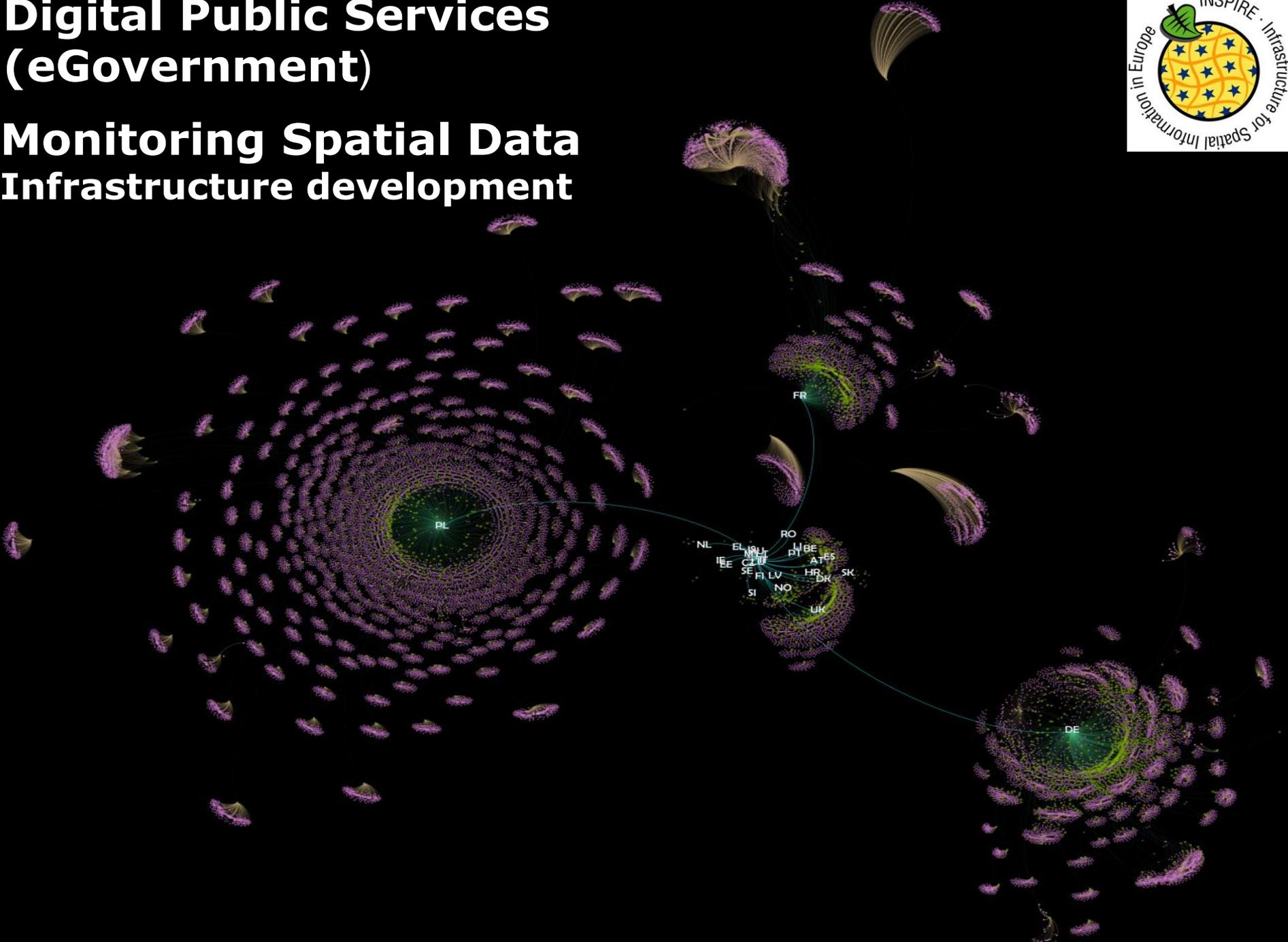
Observed by (OAUTH-ID):
opendayjrc@gmail.com

Big Data and Life Sciences

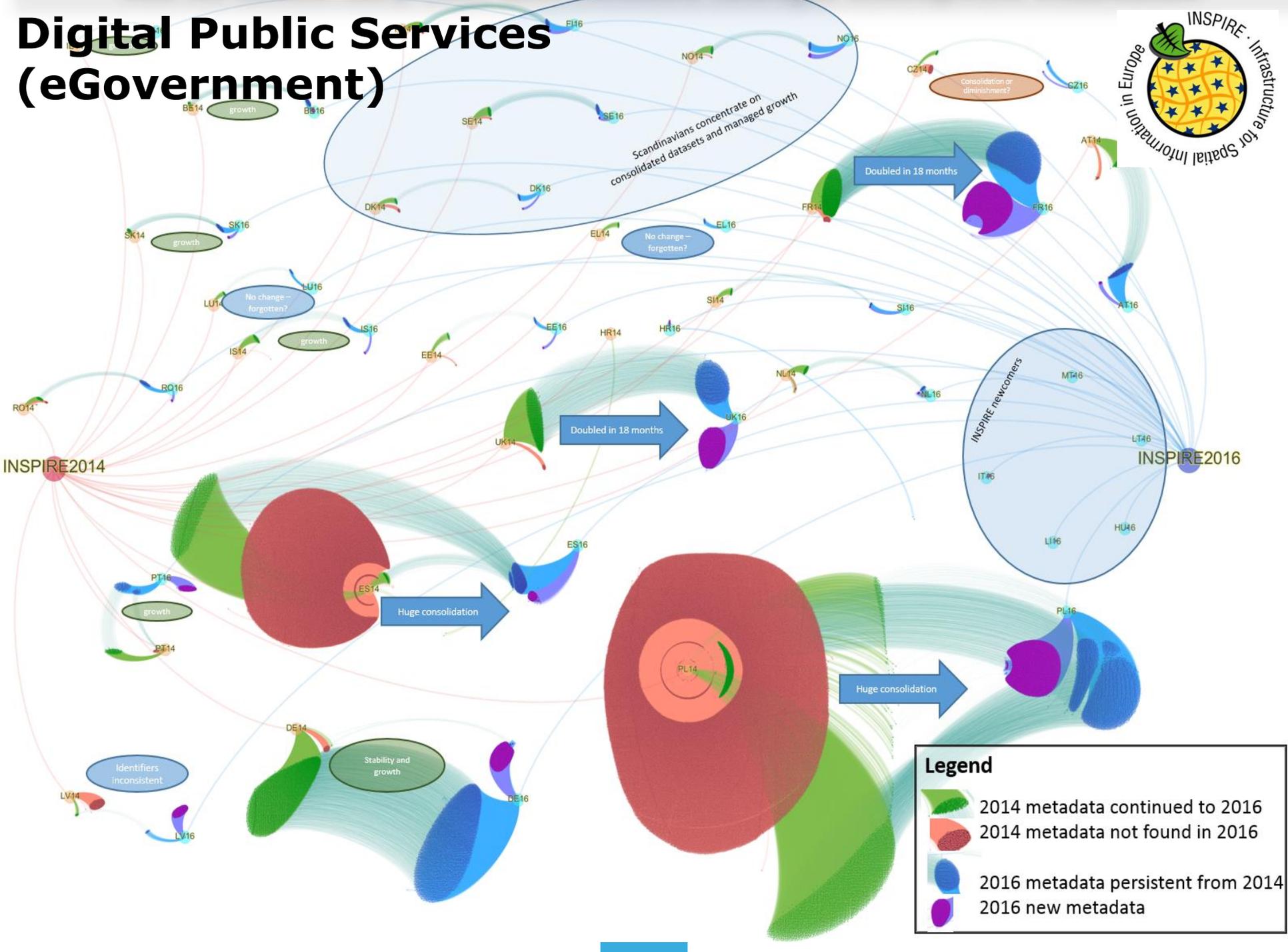


Digital Public Services (eGovernment)

Monitoring Spatial Data Infrastructure development

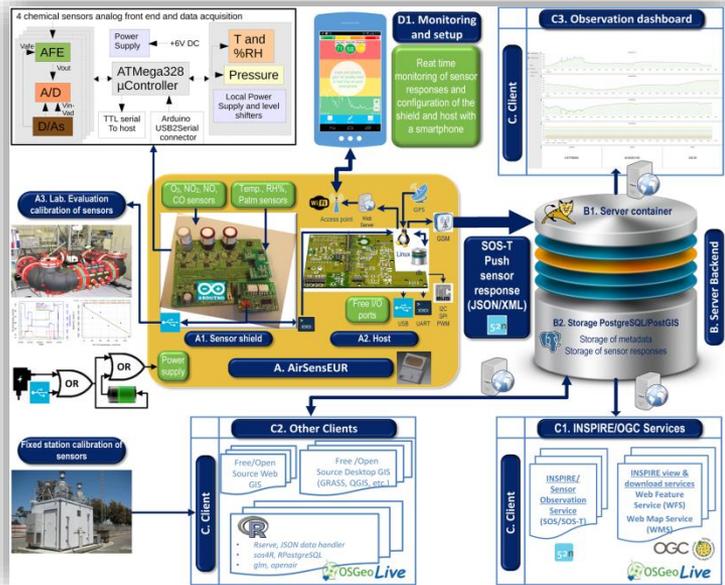


Digital Public Services (eGovernment)

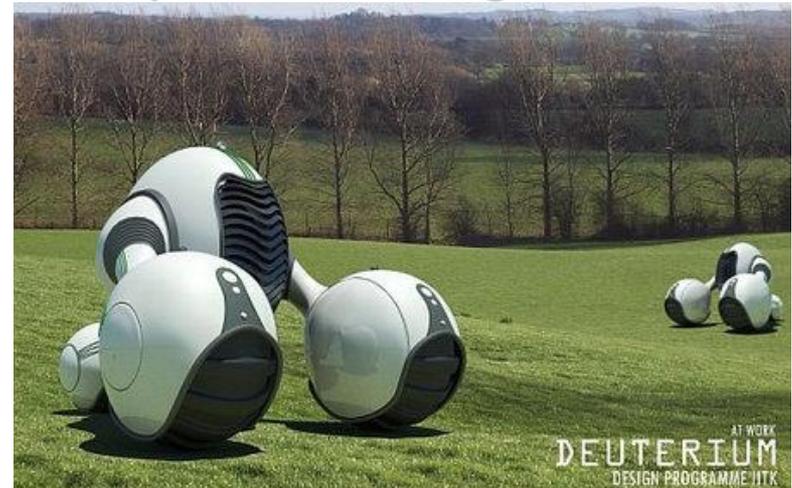
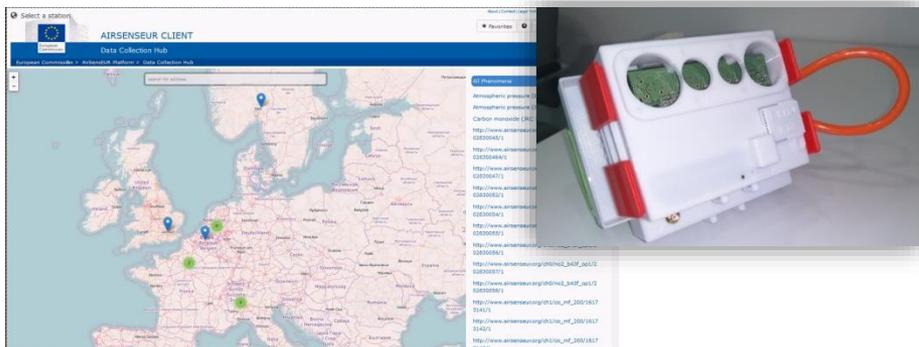


AGRI

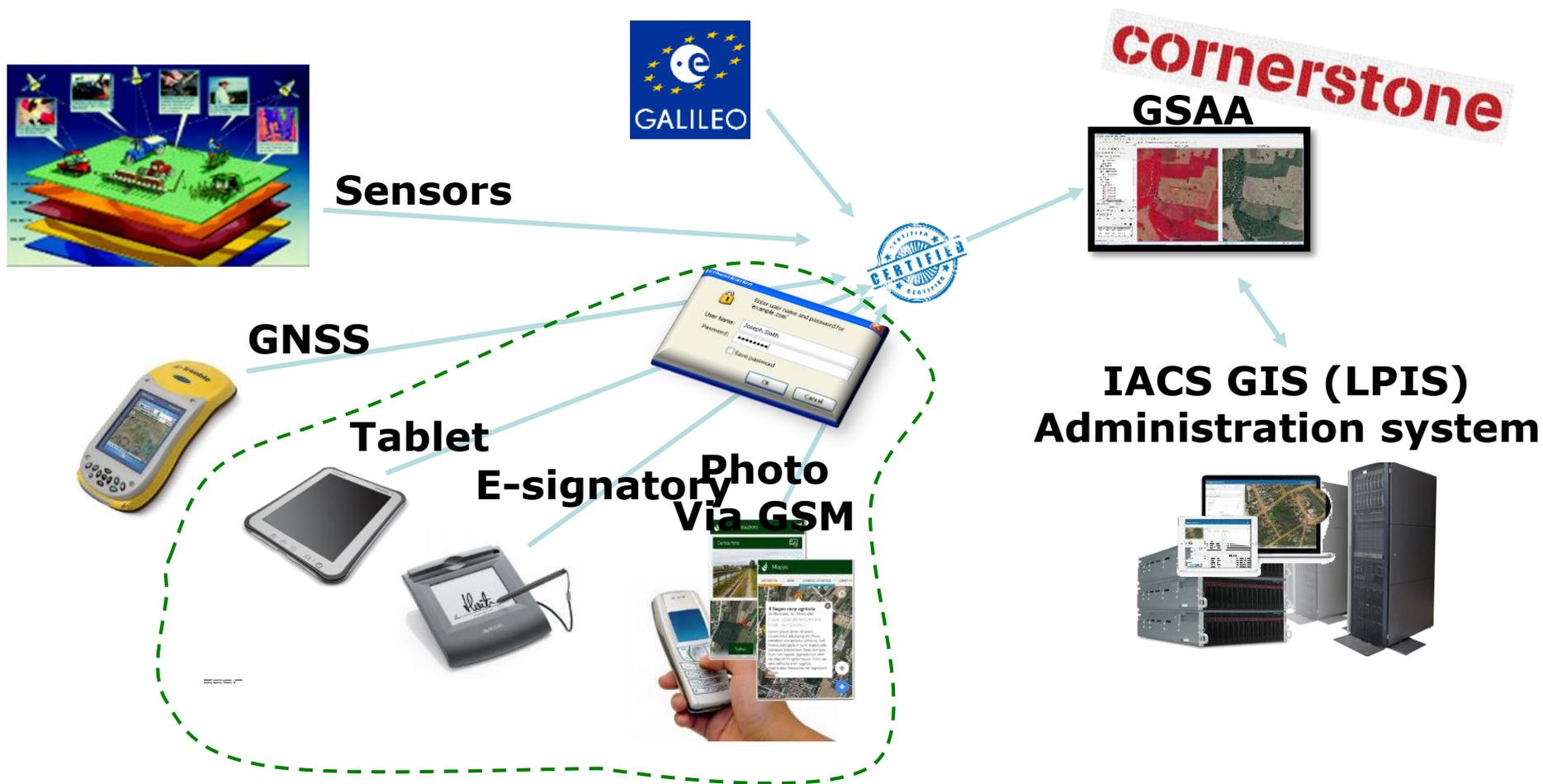
Internet of Things Sensorizing the world...



European Farming tomorrow ?



AGRI



AGRI

<https://agricoltura.servizirl.it/>



Users log in with an individual ID :

- RSC o NSC
- PIN
- One or more user profiles

RSC – NSC card using :

- System security increasing
- Digital signature
- By 2019 all users will be able to access using the personal ID card for the public health system

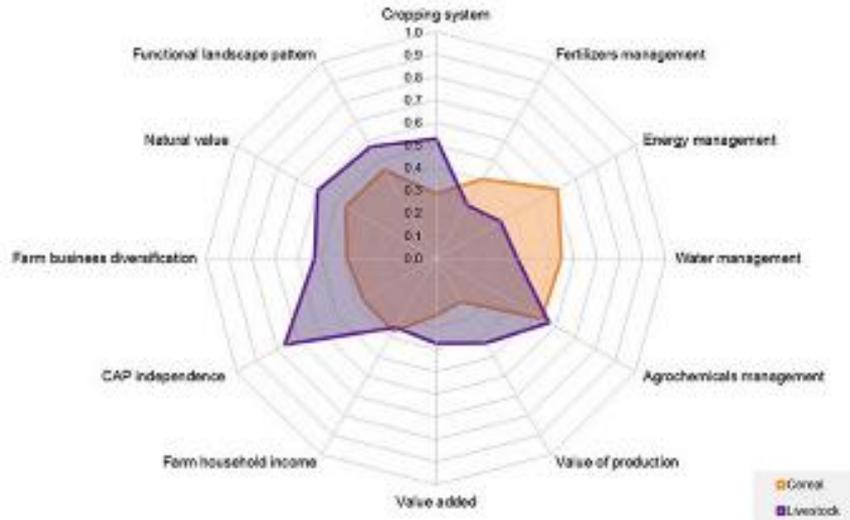
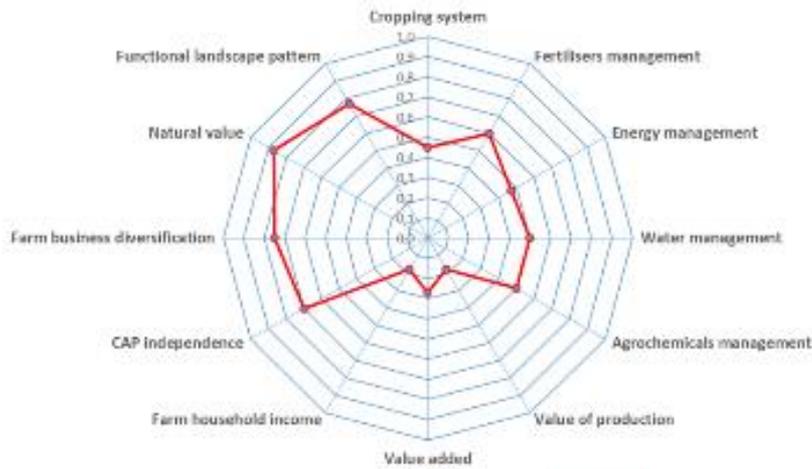


European
Commission

AGRI

Data assimilation within Farm Level Tools

SOSTARE model Regione Lombardia



Farm performance



↕
Equivalence

Certification systems

Low carbon
Organic
Greening compliant



Comparison of farm type performance

Make it easy to use for the farmer



General considerations

4th industrial/societal revolution

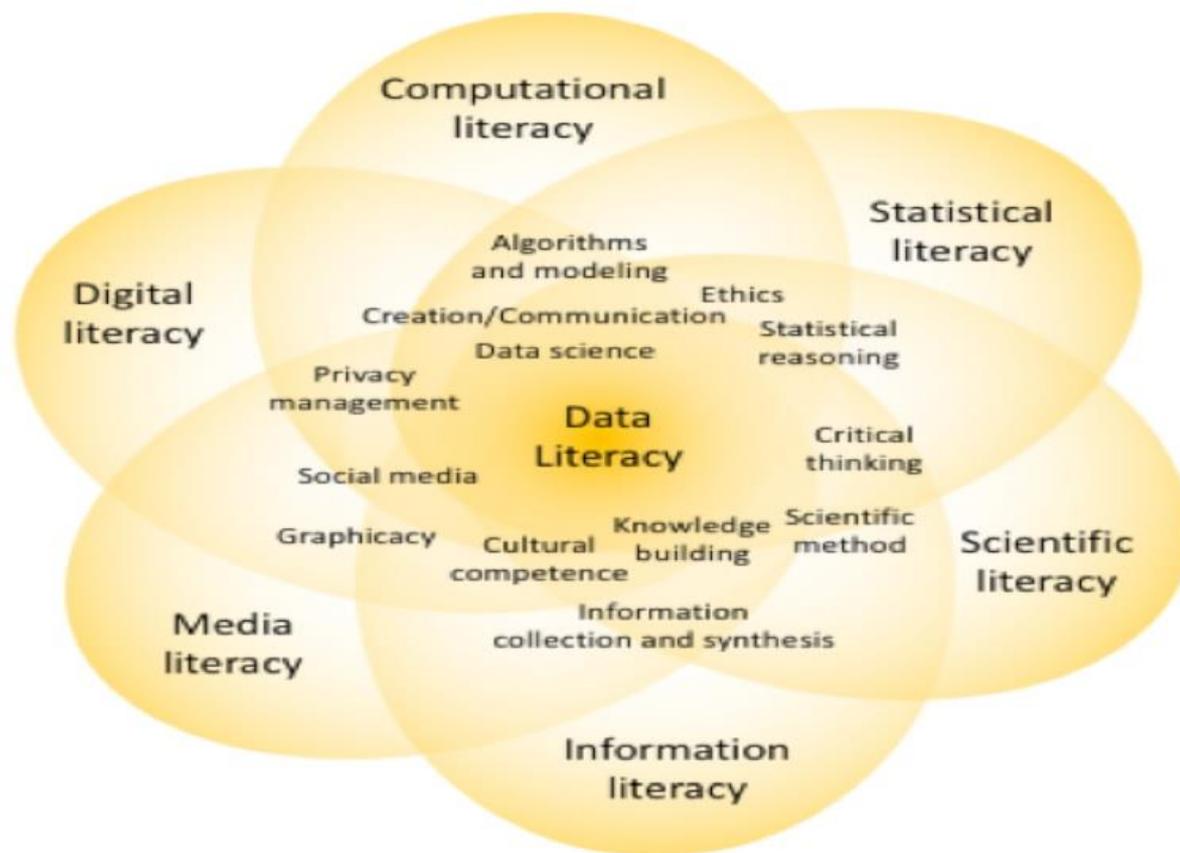
"Platformisation" of the economy

Challenge 1 – use of new data source to measure and forecast overall trends in Digital Economy

Challenge 2 – access to critical data

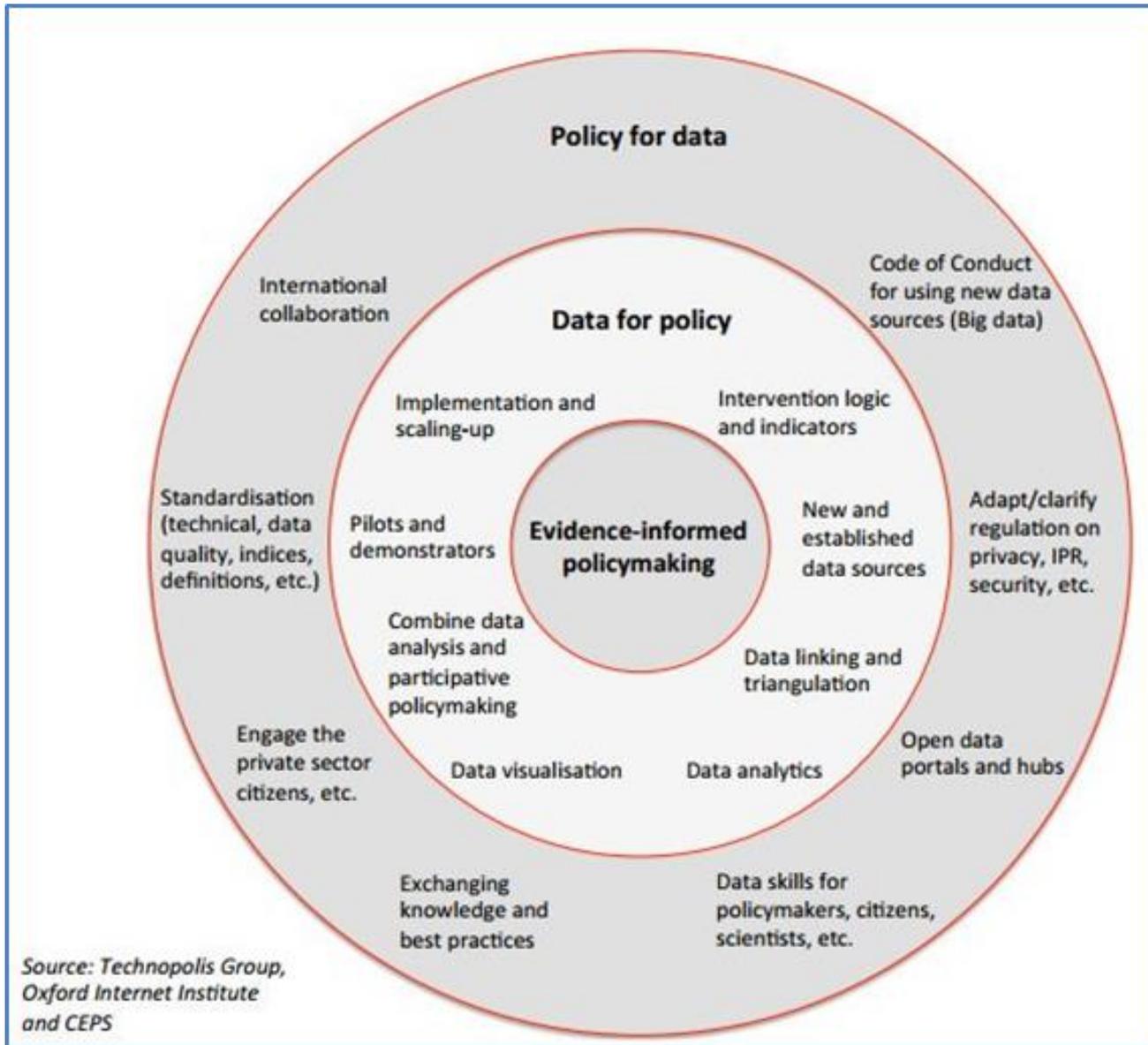
The paradox of the digital economy is that never before have so much data been collected and never before has it been so difficult to access these data.

Data Literacy



Source:

<https://datatherapy.files.wordpress.com/2015/10/beyond-data-literacy-2015.pdf>



Stay in touch



EU Science Hub: ec.europa.eu/jrc

KC-MD Web Portal: ec.europa.eu/jrc/en/migration-and-demography

KC-TP Web Portal: ec.europa.eu/jrc/territorial-policies

Urban Data: urban.jrc.ec.europa.eu



Twitter: @EU_ScienceHub



Facebook: EU Science Hub - Joint Research Centre



LinkedIn: Joint Research Centre



YouTube: EU Science Hub