

# Privadesa i seguretat en temps de megadades

Josep Domingo Ferrer



Chair in  
Data Privacy

Universitat Rovira i Virgili  
Valls, 27 de gener del 2016

1. Com s'obtenen les megadades (“big data”)?



# Activitat a la xarxa (1 minut)



# Metadades

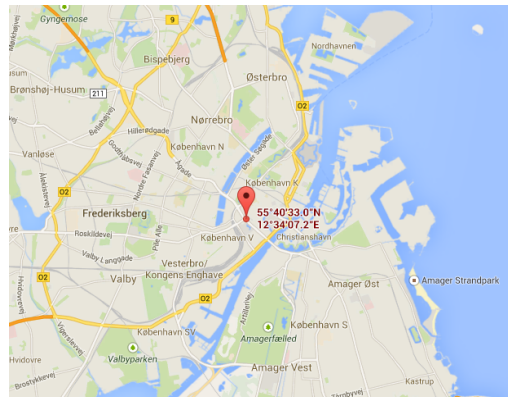


Dispositiu

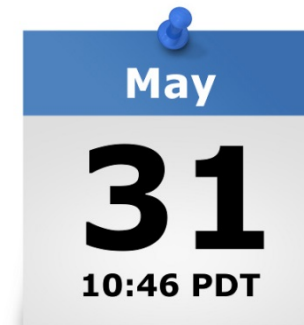


iPhone 5

Lloc

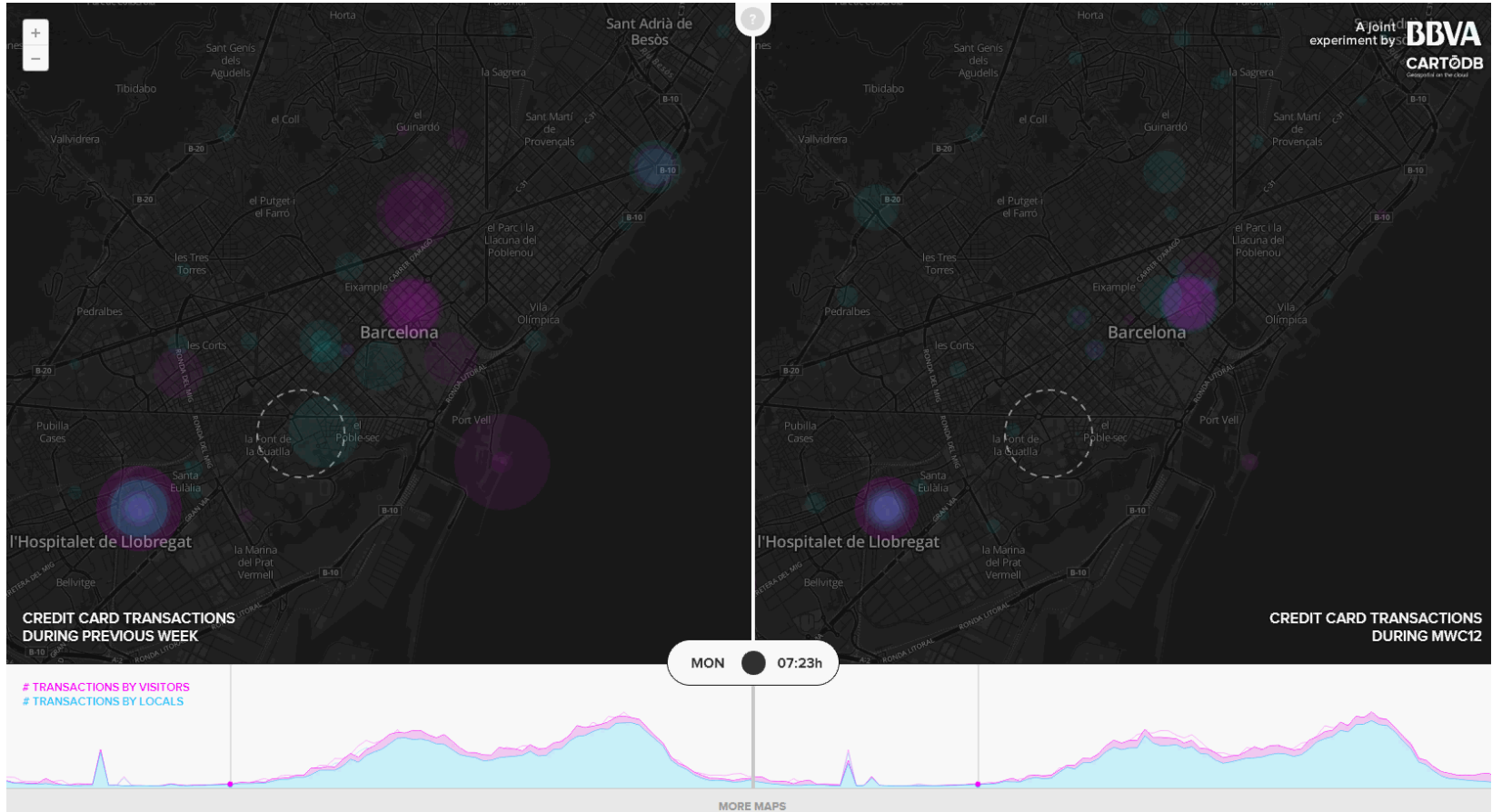


Dia i hora





# Targetes de crèdit



# <http://www.flightradar24.com/>

**flightradar24** LIVE AIR TRAFFIC

APPS ADD COVERAGE PRESS DATA / HISTORY SOCIAL ABOUT

09:05 UTC

Jump to area Map

**AF4376 / AFR4376**  
Air France

**TLS → RAK**  
Toulouse → Marrakesh

STD 10:10 CEST STA 10:45 WET  
ATD 10:20 CEST ETA 10:22 WET

3D

Aircraft: Airbus A319-111 (A319)  
Registration: F-GRHP (3944EF)

Altitude: 38,975 ft	Vertical Speed: 0 fpm
Speed: 449 kt	Track: 214°
Latitude: 39.79	Longitude: -2.69
Radar: T.LEGT1	Squawk: 2337

REAL TIME 5 MIN DELAY ADS-B data FAA data

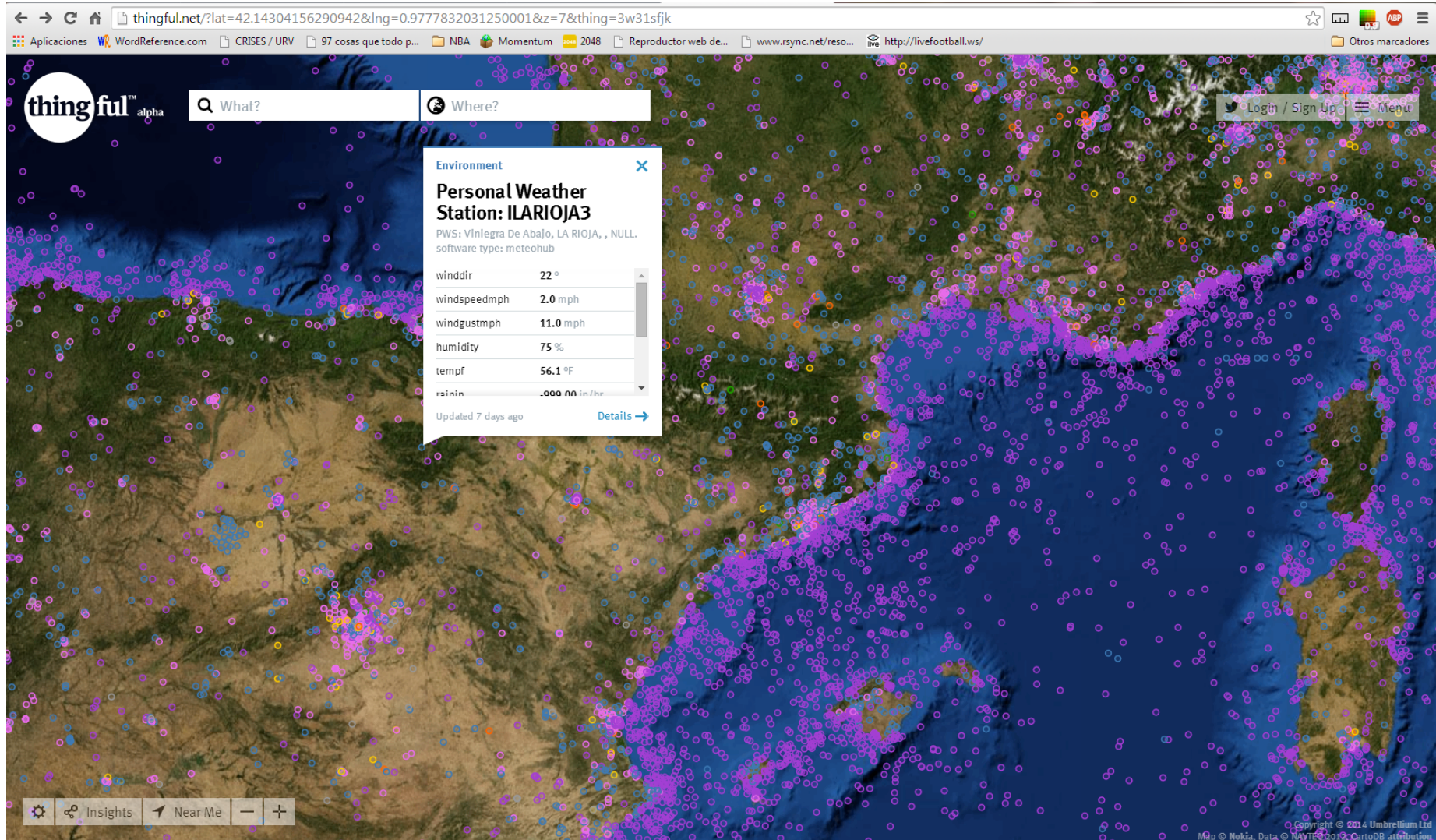
Want flightradar24 on your phone, tablet, Mac OS, or Windows 8?

Want flightradar24 without ads?

Map data ©2014 Basarsoft, GeoBasis-DE/BKG (©2009), Google, basado en BCN IGN España | 100 km | Terms of Use Report a map error

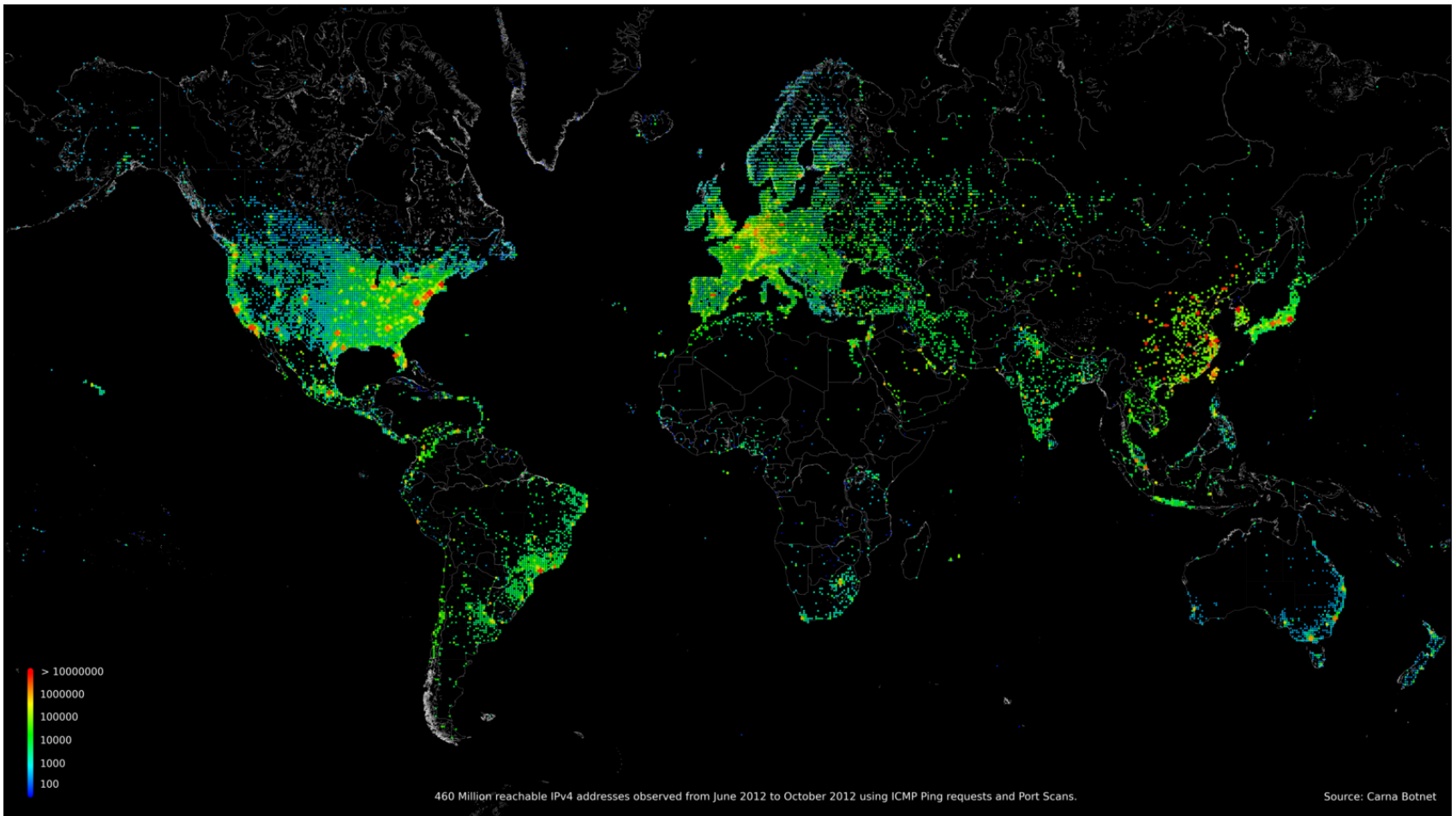


# Thingful.net (Internet de les coses, IoT)





# Densitat d'Internet



# Ecosistema Google

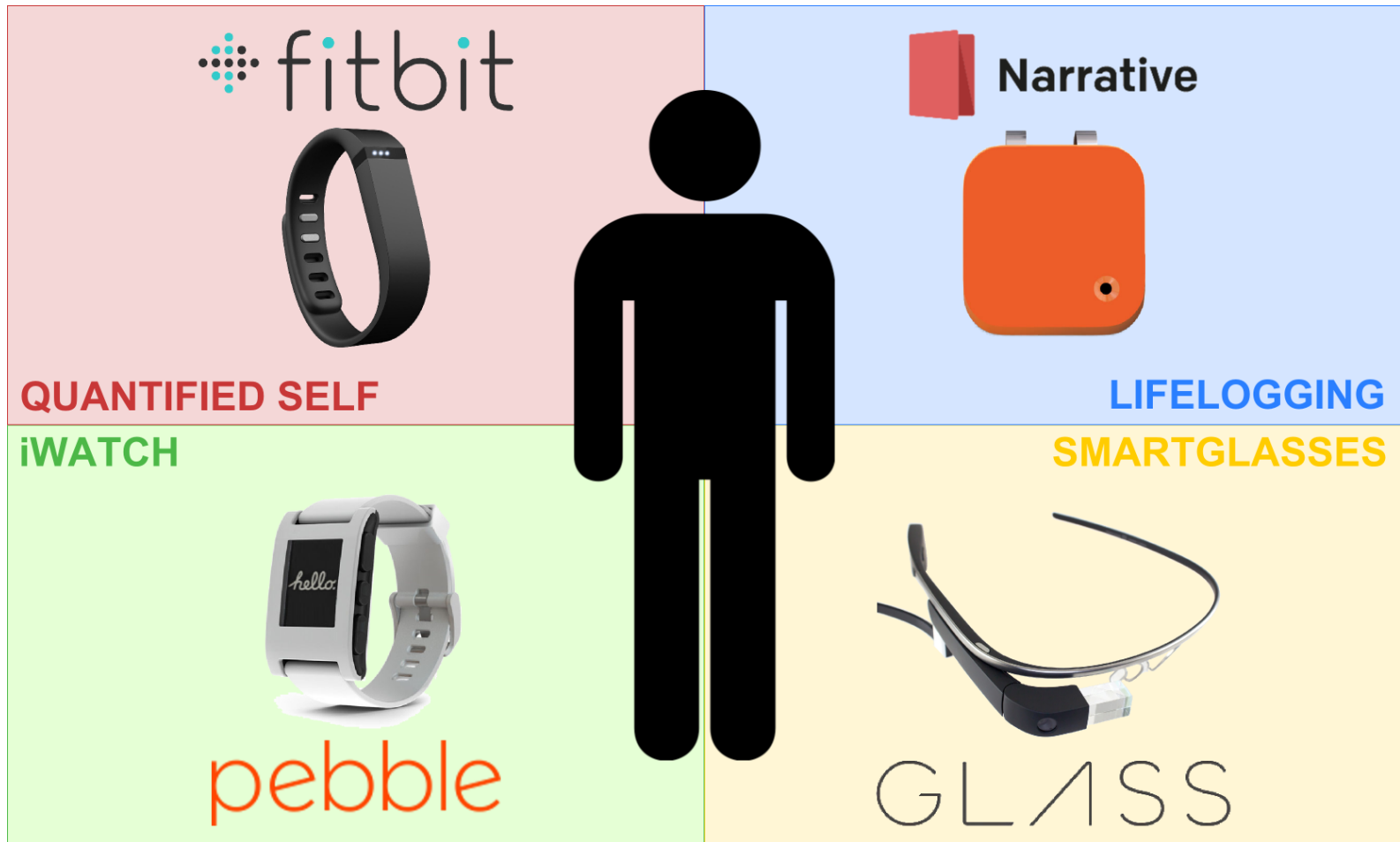


# Sensors en un téléphone intel·ligent





# Dispositius portables



# Dades genètiques



**Find out what  
your DNA says  
about you  
and your family.**

- Learn what percent of your DNA is from populations around the world
- Contact your DNA relatives across continents or across the street
- Build your family tree and enhance your experience with relatives

**order now**

**\$99**

# Dades sanitàries

## New Health Sciences Data Sources



Drug Research



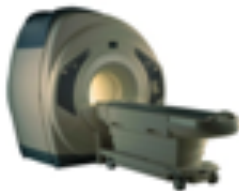
Social Media



Patient Records



Gene Sequencing



Test Results



Claims



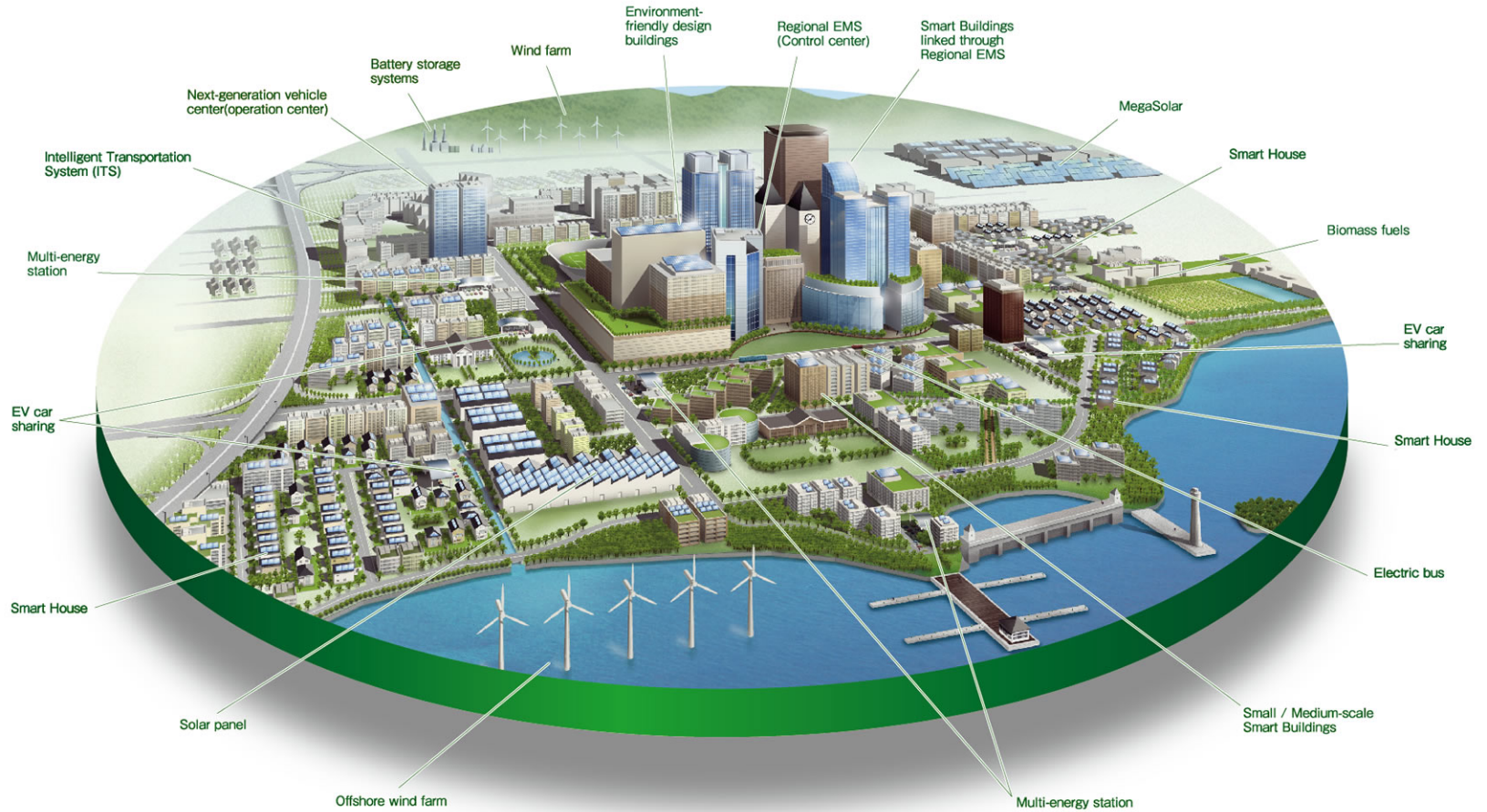
Home Monitoring



Mobile Apps



# Ciutats intel·ligents



# Ciutats intel·ligents + IoT

## Libelium Smart World

### Air Pollution

Control of CO<sub>2</sub> emissions of factories, pollution emitted by cars and toxic gases generated in farms.

### Forest Fire Detection

Monitoring of combustion gases and preemptive fire conditions to define alert zones.

### Wine Quality Enhancing

Monitoring soil moisture and trunk diameter in vineyards to control the amount of sugar in grapes and grapevine health.

### Offspring Care

Control of growing conditions of the offspring in animal farms to ensure its survival and health.

### Sportsmen Care

Vital signs monitoring in high performance centers and fields.

### Structural Health

Monitoring of vibrations and material conditions in buildings, bridges and historical monuments.

### Smartphones Detection

Detect iPhone and Android devices and in general any device which works with Wifi or Bluetooth interfaces.

### Perimeter Access Control

Access control to restricted areas and detection of people in non-authorized areas.

### Radiation Levels

Distributed measurement of radiation levels in nuclear power stations surroundings to generate leakage alerts.

### Electromagnetic Levels

Measurement of the energy radiated by cell stations and and WIFI routers.

### Traffic Congestion

Monitoring of vehicles and pedestrian affluence to optimize driving and walking routes.

### Smart Roads

Warning messages and diversions according to climate conditions and unexpected events like accidents or traffic jams.

### Smart Lighting

Intelligent and weather adaptive lighting in street lights.

### Intelligent Shopping

Getting advices in the point of sale according to customer habits, preferences, presence of allergic components for them or expiring dates.

### Noise Urban Maps

Sound monitoring in bar areas and centric zones in real time.

### Water Leakages

Detection of liquid presence outside tanks and pressure variations along pipes.

### Vehicle Auto-diagnosis

Information collection from CanBus to send real time alarms to emergencies or provide advice to drivers.

### Item Location

Search of individual items in big surfaces like warehouses or harbours.

### Waste Management

Detection of rubbish levels in containers to optimize the trash collection routes.

### Smart Parking

Monitoring of parking spaces availability in the city.

### Golf Courses

Selective irrigation in dry zones to reduce the water resources required in the green.

### Water Quality

Study of water suitability in rivers and the sea for fauna and eligibility for drinkable use.

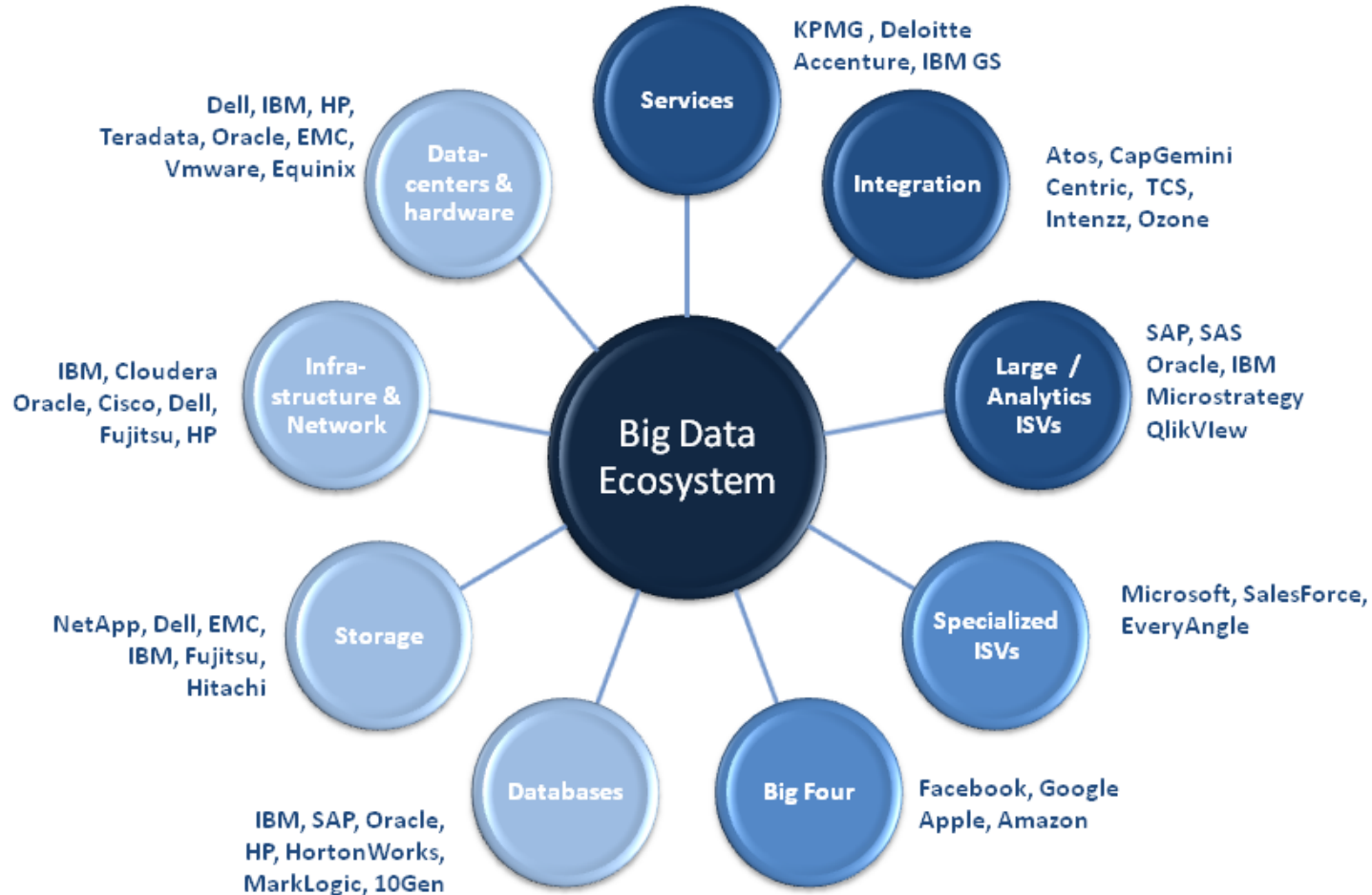
### Quality of Shipment Conditions

Monitoring of vibrations, strokes, container openings or cold chain maintenance for insurance purposes.

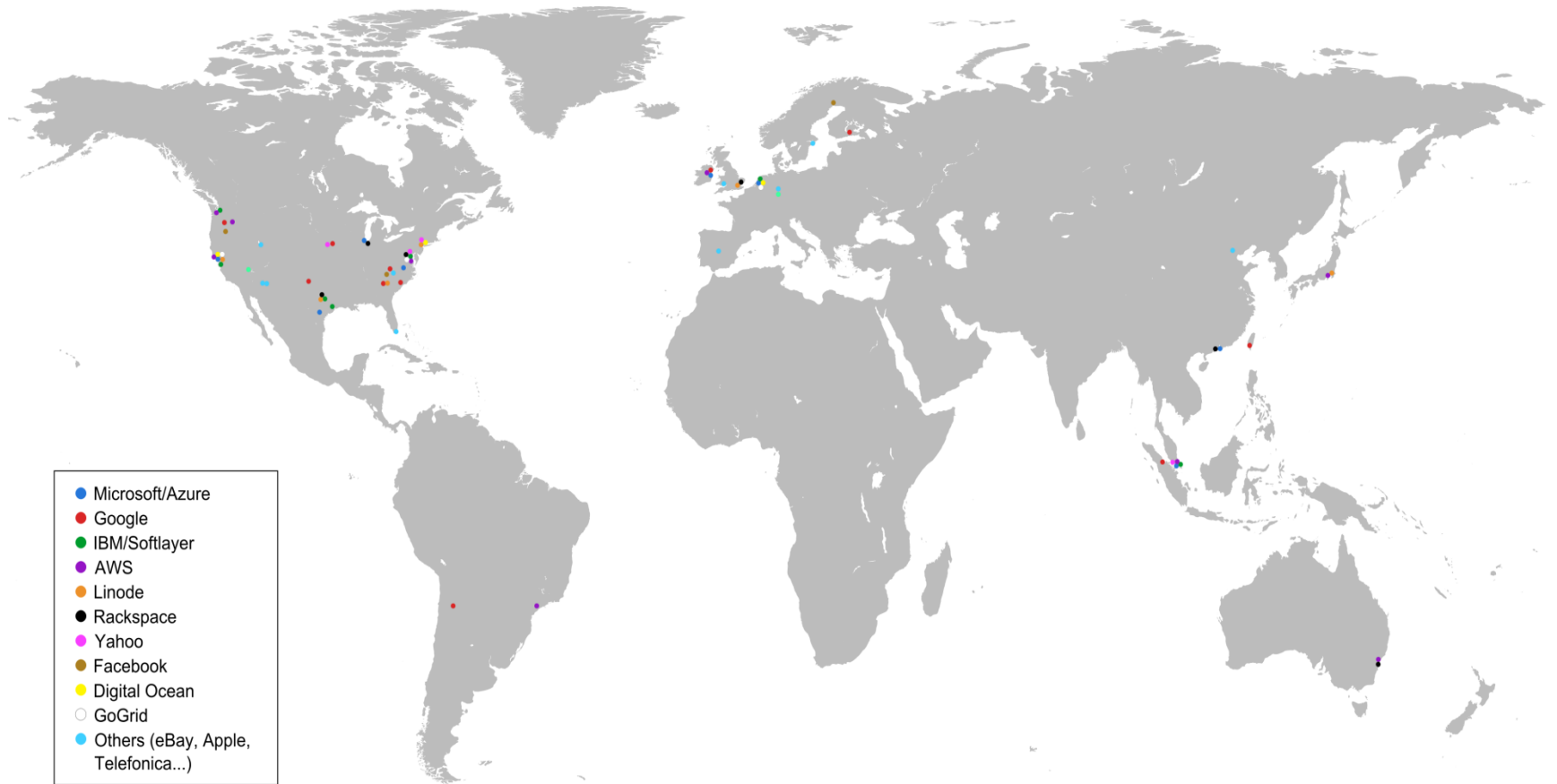
2. Qui i on té les megadades?



# Ecosistema de megadades



# On? Centres de processament de dades

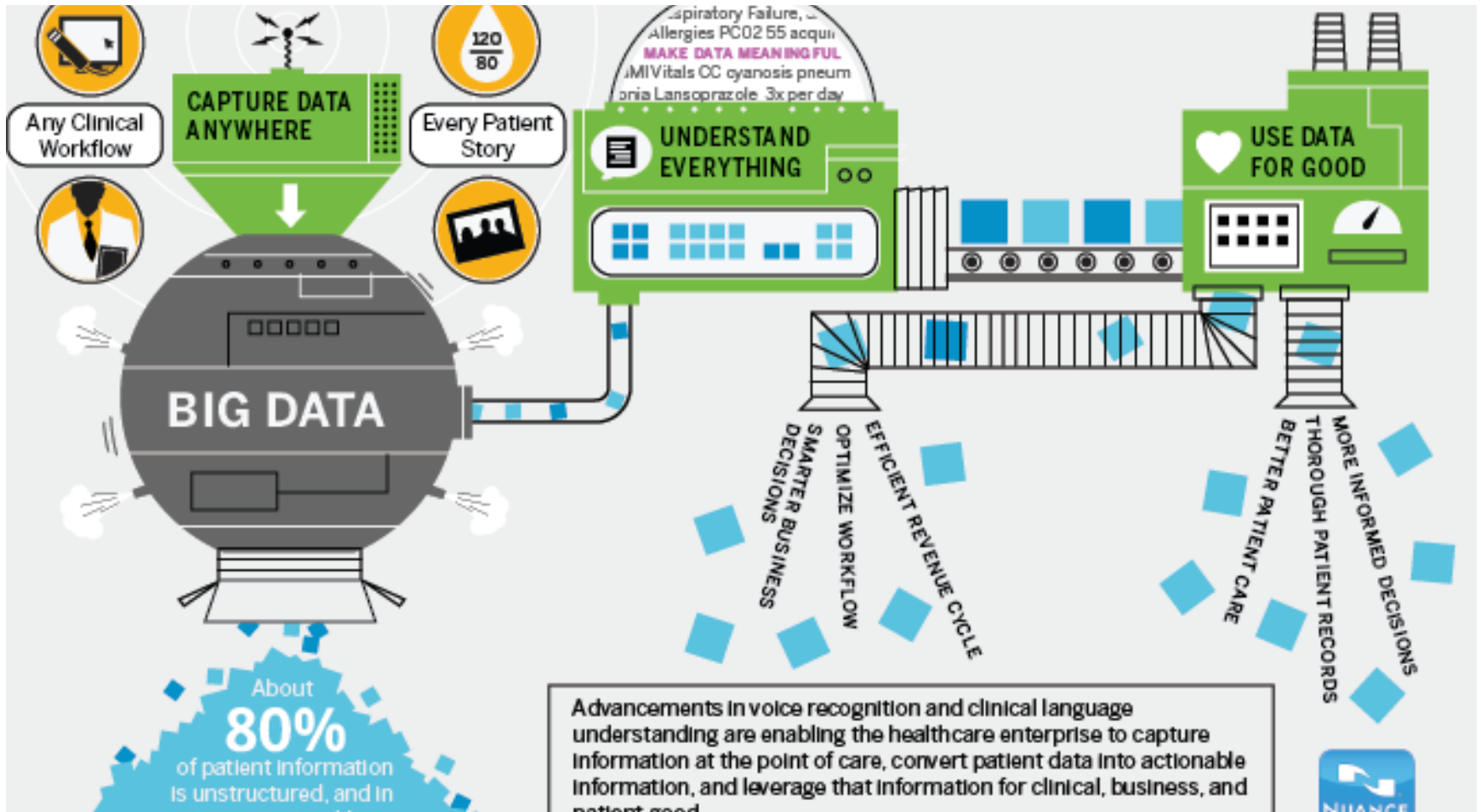


### 3. Doble ús de les megadades

# FuturICT

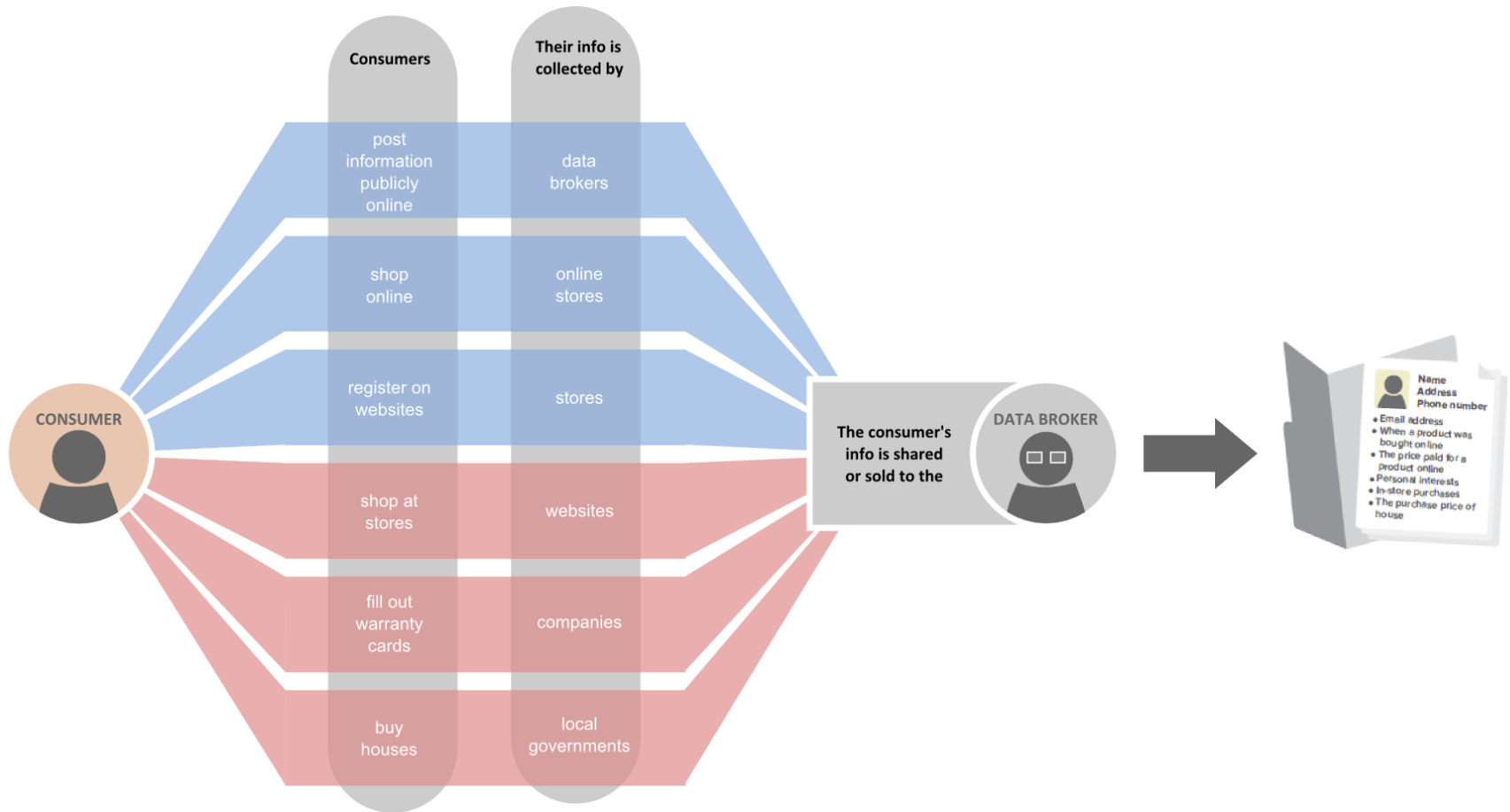
- Projecte ex FET Flagship (7PM UE)  
<http://www.futurict.eu>
- Objectius:
  - Entendre i gestionar sistemes complexos, globals i socialment interactius
  - Construir una “Living Earth Platform”, per simular el funcionament del món i preveure crisis econòmiques, sanitàries, meteorològiques, etc.

# Doble ús de megadades en sanitat





# Els desconeguts: “data brokers” (marxants de dades)



## 4. Privadesa i doble ús

# Hi ha qui ven les seves dades...

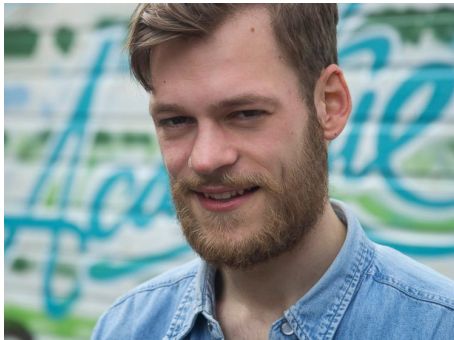


Federico Zannier

<http://myprivacy.info>

Juny 2013

2.090,70 €



Shawn Buckles

<http://shawnbuckles.nl/dataforsale/>

Abril 2014

350,00 €



Un perfil vostre

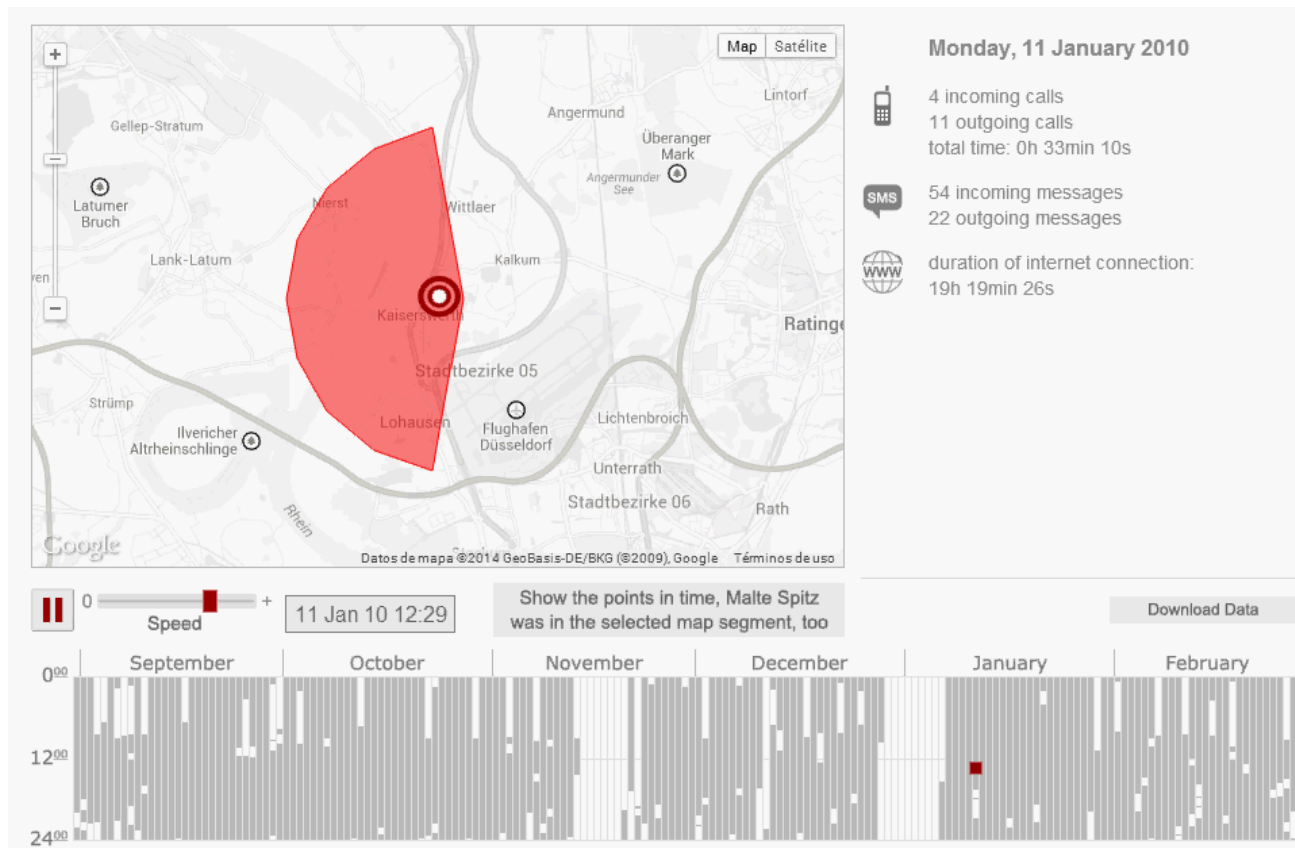
<http://on.ft.com/1tmi7fO>

Juny 2013

Menys d'1 €

# Hi ha qui regala les seves dades...

<http://www.zeit.de/datenschutz/malte-spitz-data-retention/>



# Però, en general, a Europa, cal...

- Que qui processi les dades tingui un **mandat legal exprés** para fer-ho en exercici de les seves competències

O bé

- Que tingui el **consentiment informat i explícit** del subjecte en el moment de recollir les dades per fer-ne doble ús per a un propòsit concret

O bé

- Que es limiti a processar dades **anonimitzades** (no tan sols desidentificades!).

Directiva europea 95/46/EC de protecció de dades i projecte de General Data Protection Regulation



5. Com cedir legalment dades sensibles per a doble ús?

# Fitxers de microdades d'us públic

- Poden descarregar-se lliurement d'Internet
- Han estat *desidentificats* (se n'han suprimit els identificadors, com ara nom, DNI, etc.)
- Han estat *fortament anonimitzats*:
  - Mostra d'1% o menys
  - Supressió o recodificació de quasi-identificadors (p.e. interval d'edat en comptes d'edat, província en comptes de ciutat, etc.)
  - Supressió o modificació d'atributs sensibles

# Exemple de fitxer d'ús públic

- Enquesta de salut i d'hàbits sexuals a Espanya (INE, 2003)
- 13600 ciutadans i cap atribut que permeti d'identificar ciutadans concrets ni tan sols sabent que són a la mostra



# Fitxers de microdades d'ús científic

- Han estat *desidentificats*
- Han estat *anonimitzats*, però menys que els fitxers d'ús públic
- Se cedeixen a investigadors amb nom i cognoms que signen un contracte
  - limitant-ne l'ús a un estudi concret
  - restringint la publicació de resultats a dades agregades i/o gràfics (però no dades individuals)

# Exemples de fitxers d'ús científic

- EU Structure of Earnings Survey (SES, Eurostat)
- EU Road Carriage microdata (RC, Eurostat)



(Per a tots dos jo he dissenyat el procediment d'anonimització)

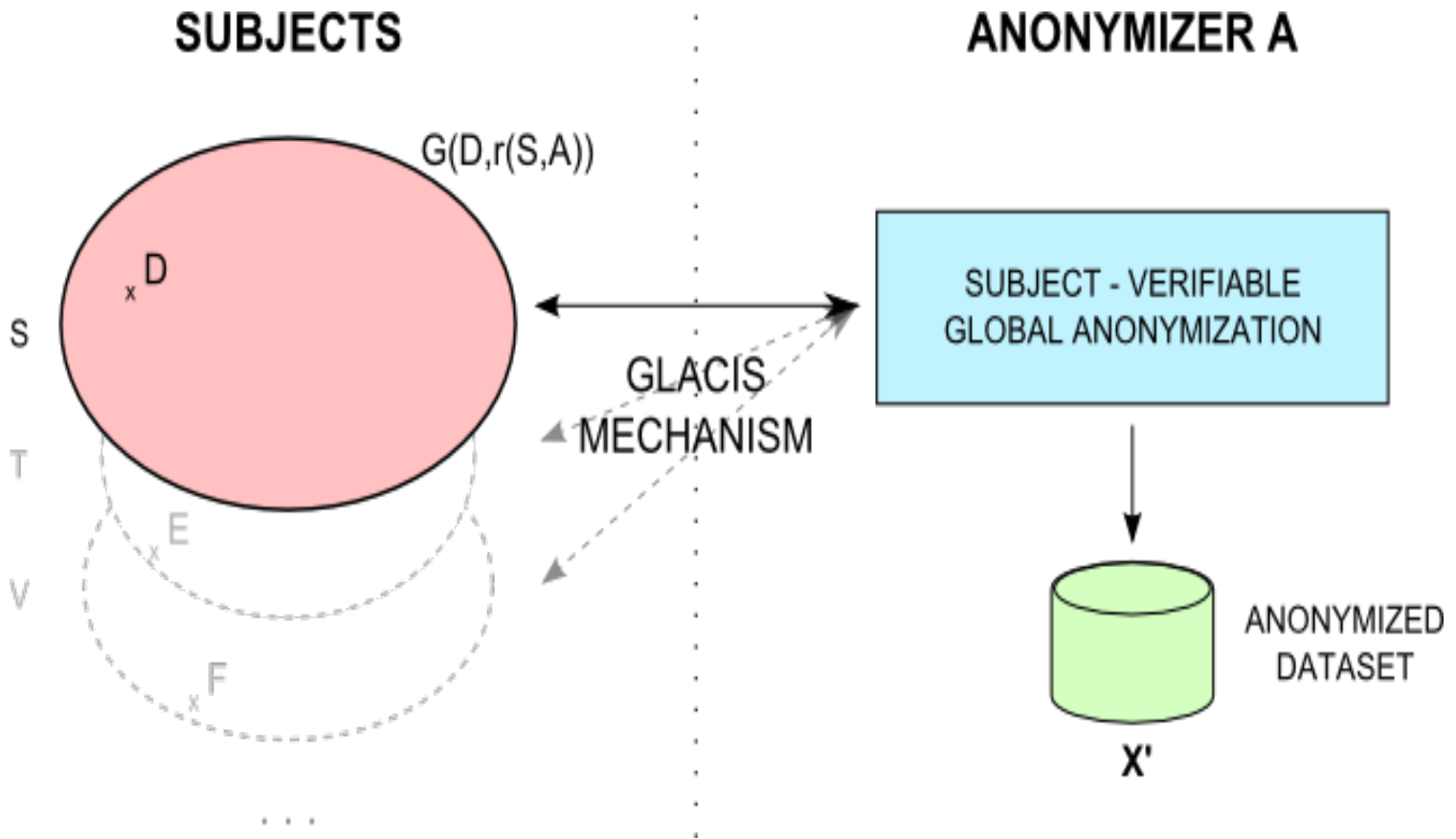


# Centres d'accés segur

- Permeten d'accedir a dades *desidentificades*, però no *anonimitzades*
- L'investigador ha d'acreditar el seu projecte per accedir al centre
- Les dades no es mouen dels ordinadors del centre, en els quals l'investigador fa els tractaments estadístics que li calen
- Abans de retornar els resultats a l'investigador, el centre comprova que no comportin revelació de dades individuals
- Normalment, l'investigador ha d'anar al centre, tot i que en algun cas l'accés pot ser a distància

6. I si el ciutadà no se'n refia, què pot fer-hi?

# Anonimització verificable



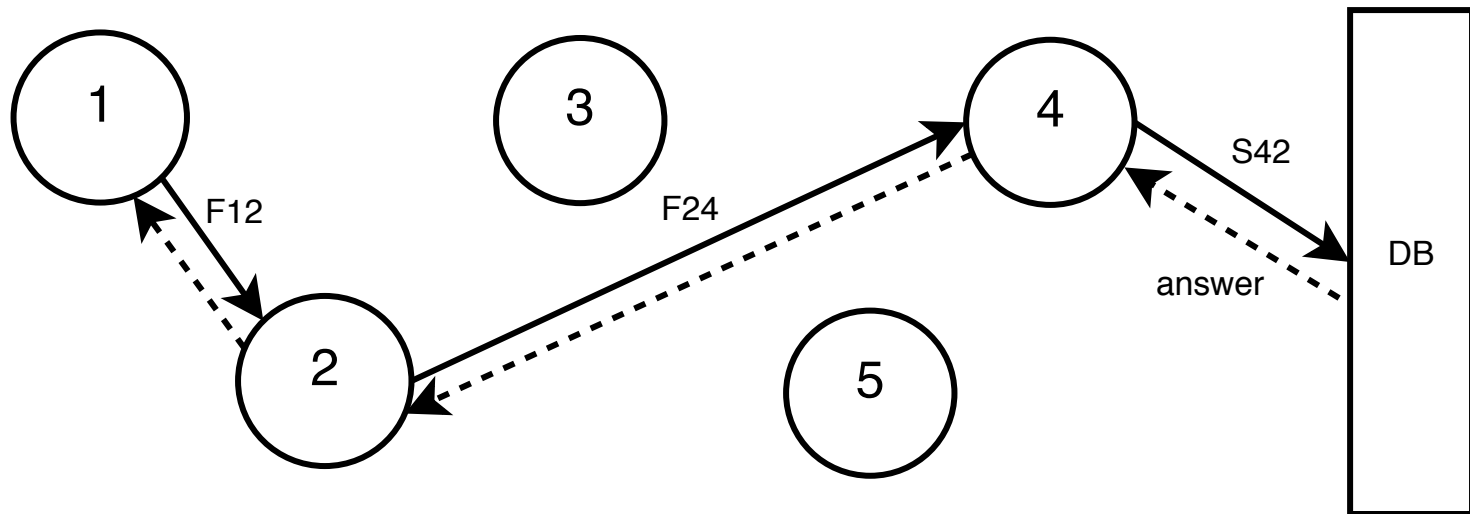


# Anti-perfilatge

- Autoprotecció: GooPIR, TrackMeNot, etc.



- Protecció P2P



# Coutilitat/coprivadesa

- Ajudar els altres a mantenir llur privadesa pot beneficiar la meva pròpia privadesa (coprivadesa)
- Es poden dissenyar protocols en els quals l'ajuda mútua sigui l'opció més racional per a qualsevol usuari, per egoista que sigui
- Projecte "CO-UTILITY" (URV)  
<http://crises-deim.urv.cat/co-utility/>



**CO-UTILITY**

Conciliating individual freedom and  
common good in the information society



# 7. Conclusions

- Qualsevol cessió de dades sensibles sense consentiment informat diferent de les anteriors no compliria els estàndards en vigor
- La General Data Protection Regulation en procés d'aprovació per la UE extrema les garanties en aquest sentit
- Si el ciutadà malfia, es resistirà a donar informació fiable

# Per saber-ne més



G. D'Acquisto,  
J. Domingo-Ferrer  
et al. (2015),  
European  
Network and Information  
Security Agency (ENISA)



Privacy by design in big data  
An overview of privacy enhancing technologies in  
the era of big data analytics

FINAL

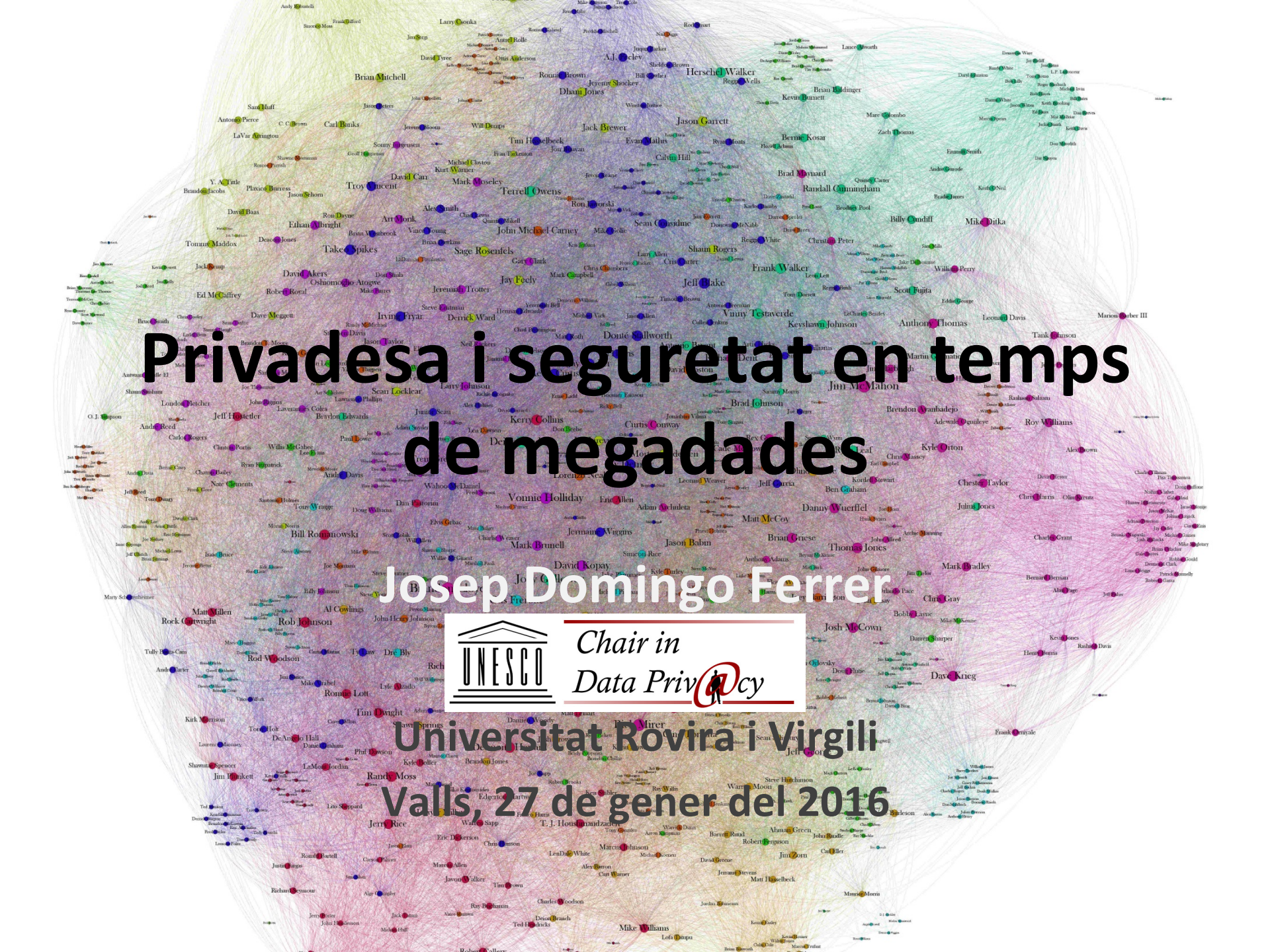
1.0

PUBLIC

DECEMBER 2015







# Privadesa i seguretat en temps de megadades

Josep Domingo Ferrer



Chair in  
Data Privacy

Universitat Rovira i Virgili  
Valls, 27 de gener del 2016