An (increasingly) visible college
Mapping and strengthening research and innovation networks with open data in Arloesiadur

Juan Mateos-Garcia
Semantic Analysis for Innovation Policy Workshop
Paris
13 March 2018
Why this matters

Scientific Research, Development and Innovation System

Data about research inputs, activities, outputs

Visual interfaces to analyse the data

National and local policymakers
- What are the emerging research trends?
- Who is collaborating with who and where are the gaps?
- What is the impact of research funding and other interventions?

Other innovation system agents
- What are the emerging research trends?
- What researchers and organisations are active in different disciplines?
- Where are they based?

Arloesiadur

Traditional ST&I indicators (# researchers, publications, patents, knowledge networks)

Open datasets of research collaboration: CORDA, STARMETRICS, GtR

15 month project, delivered exclusively with open source tools (Python + PostgreSQL + D3.js)
The data we worked with

In May 2017 we collected data about ~75K GtR projects, organisations and funders, and geo-coded organisation location.

- Analyse geographies
- Analyse networks
- Analyse topics
- Analyse trends
Analysis 1: From text to trends

We used topic modelling (Latent Dirichlet allocation) and community detection to identify 88 topics in 45k papers with research topics/abstracts.

'\textit{the use of rfid technology affords an opportunity for greater visibility in the supply chain and further supply chain automation}, making the processes more streamlined, providing accurate and timely automatic data capture, thereby improving shipment reliability...

Infrastructure
Data analytics

...
Visualising what’s going on

What are the research trends in Wales?

Wales has become more competitive in scientific and technological areas, with a particularly strong recent showing in Engineering and Technology, Medical Sciences and Mathematics and Computing, where projects led by Welsh organisations were awarded £71m in funding in 2015-16, compared with £55m in 2007-08.

Analysis shows growth in priority areas for Welsh Government, highlights research specialisms in different parts of Wales.
We mapped existing research collaborations between organisations in Wales and used their specialisms to recommend other potential collaborators.
To conclude:

Arloesiadur launched mid-October

We have shown:

• Significant opportunities to use open datasets like GtR to inform R&I policy at many different levels...
• ...but opening data up is only the beginning.
• We have to work with policymakers to understand their needs, use data science and visualisation to address them and consider what innovations in R&I policy (new processes and instruments) are enabled by new data

Watch this space!