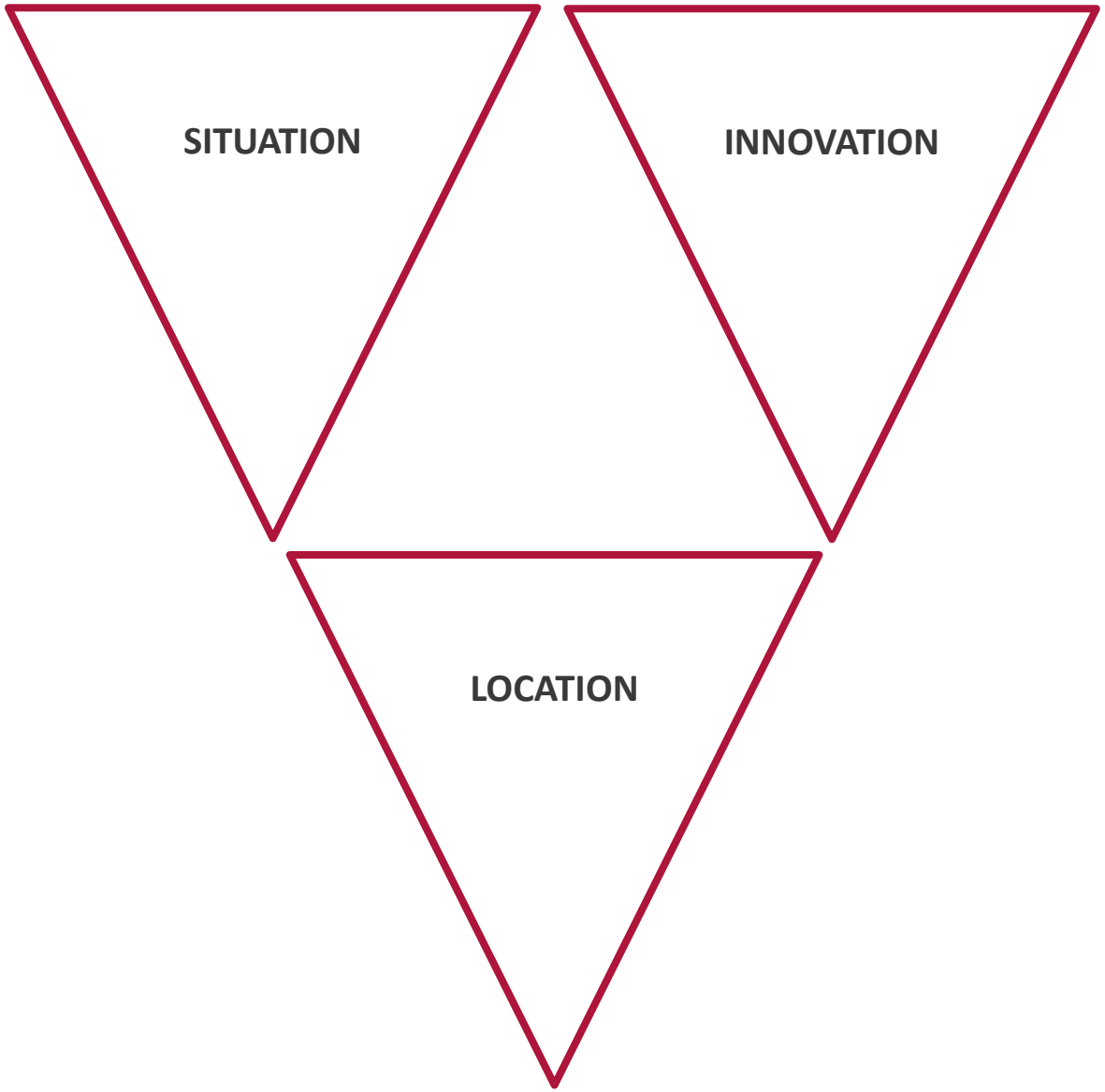
Addressing the problem: using geospatial science to unlock infrastructure intelligence

Andrew Button - andrew.button@ofcom.org.uk

Ofcom Technology Group, Data Services

Data science for bridging the digital divide and beyond, 16-17 July 2019

- **Situation**; what characteristics am I interested in?
- **Innovation**; what techniques do I have available?
- **Location**; what is the appropriate scale to investigate?



Situation

People's ability to participate in communications markets and society is affected by many factors that can lead to the temporary or permanent exclusion or reliance on certain services

- Regulators will have specific legal frameworks to ensure that there is regard to the needs and interests of citizens and consumers
- Ofcom has to take account of these groups and has intervened to promote participation and address vulnerability including
  - Ensuring that telephony and broadband communication providers offer services for disabled consumers
  - Improving the rules to require all UK communications providers to have clear, effective policies and procedures for identifying vulnerable customers
  - Requiring (under the Universal Service Obligation) consumers to be given connection to landline telephone services
  - Requiring social tariffs to be available for people on low incomes
- Enables non-spatial solution to the challenge and regulation of industry

Location

The UK Digital Economy Act 2010 required a report every three years on the state of the UK's communications infrastructure and granted formal powers to Ofcom to collect relevant data

- We use generic, but formal, information requests to ensure consistency in reporting and clarification of purpose:
  - Annual, triannual and monthly collections of up to 150m records (2bn data points) at property resolution
  - Collected from national, regional and local providers
  - Mobile operators provide signal predictions – we then apply category thresholds to derive coverage reports
  - Fixed operators provide address level predictions, and measured performance
  - Fixed wireless access operators provide sufficient information that enables us to apply a line-of-sight modelling method based on ITU-R P.530 recommendations
- The information collected tells us much about telecommunications, but not about the consumer situation
  - Opportunities to link information to other data sources to derive further insights
  - The spatial granularity of this information, whilst valuable for our analysis, is commercially sensitive, and collected for a specific purpose

Location value and challenge of address data

In 2016 we started to collect address level data from Communication providers (previously at postcode level)

We use the Ordnance Survey AddressBase® product to describe the characteristics of a premise

- CP data is linked to the Ordnance Survey data
  - Matched via Unique Property Reference Number
  - Or through advanced matching algorithms
- Address level data enables us to aggregate information for reporting at higher, and disclosable, geographies
- UPRN does provide an opportunity for linking information between organisations
  - Providing just the necessary information, at limited volume, such as via an API solution
- However, the management of address information can be a burden for any organisation

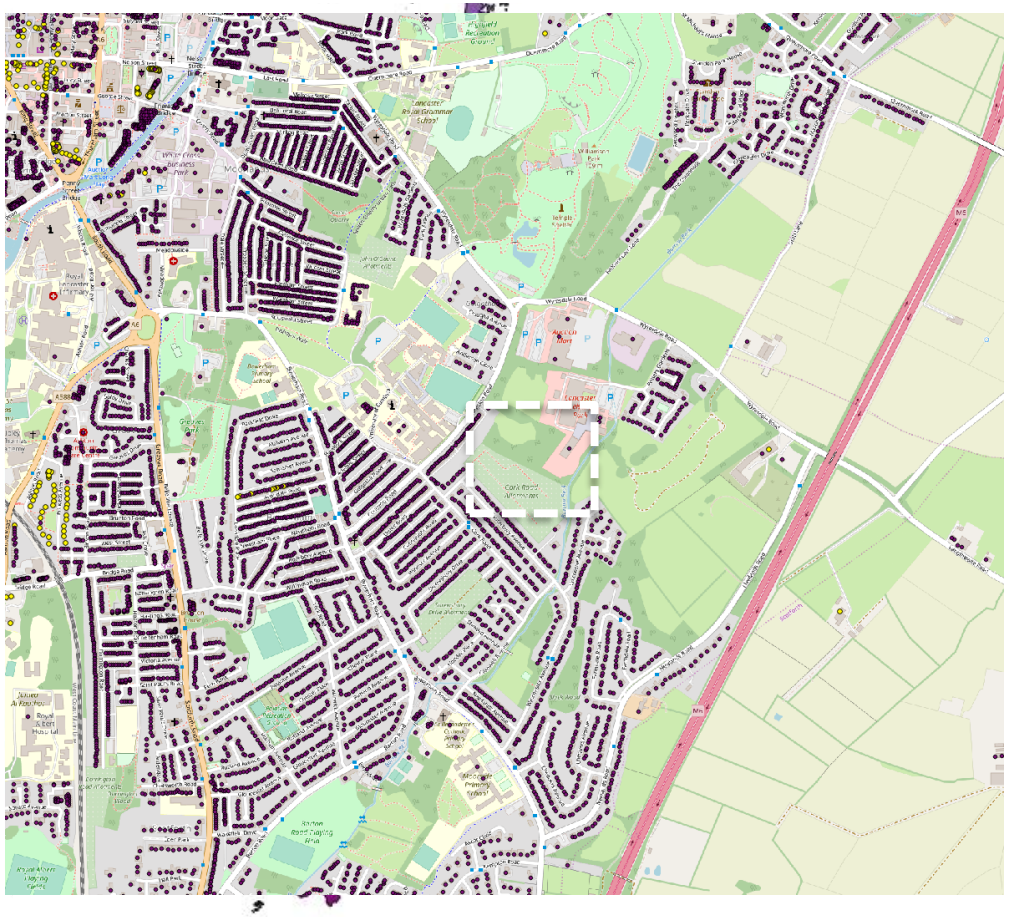
Location analysis

With access to such granular data we can apply a range of methods to unlock its value

- Address matching and linking records is now utilising Natural Language Processing methods
- Neural networks are enabling us to explore coverage availability based on infrastructure inputs
- Classification algorithms are providing new insights into the complex target groups that inform our work
  - Identification of “where consumers are”
  - Characteristics of ‘digital’ exclusion
- Statistical analysis provides insights to principal decision drivers
  - Economic geography
  - Physical infrastructure and business connectivity

Premises passed, Jan 2019

Premise

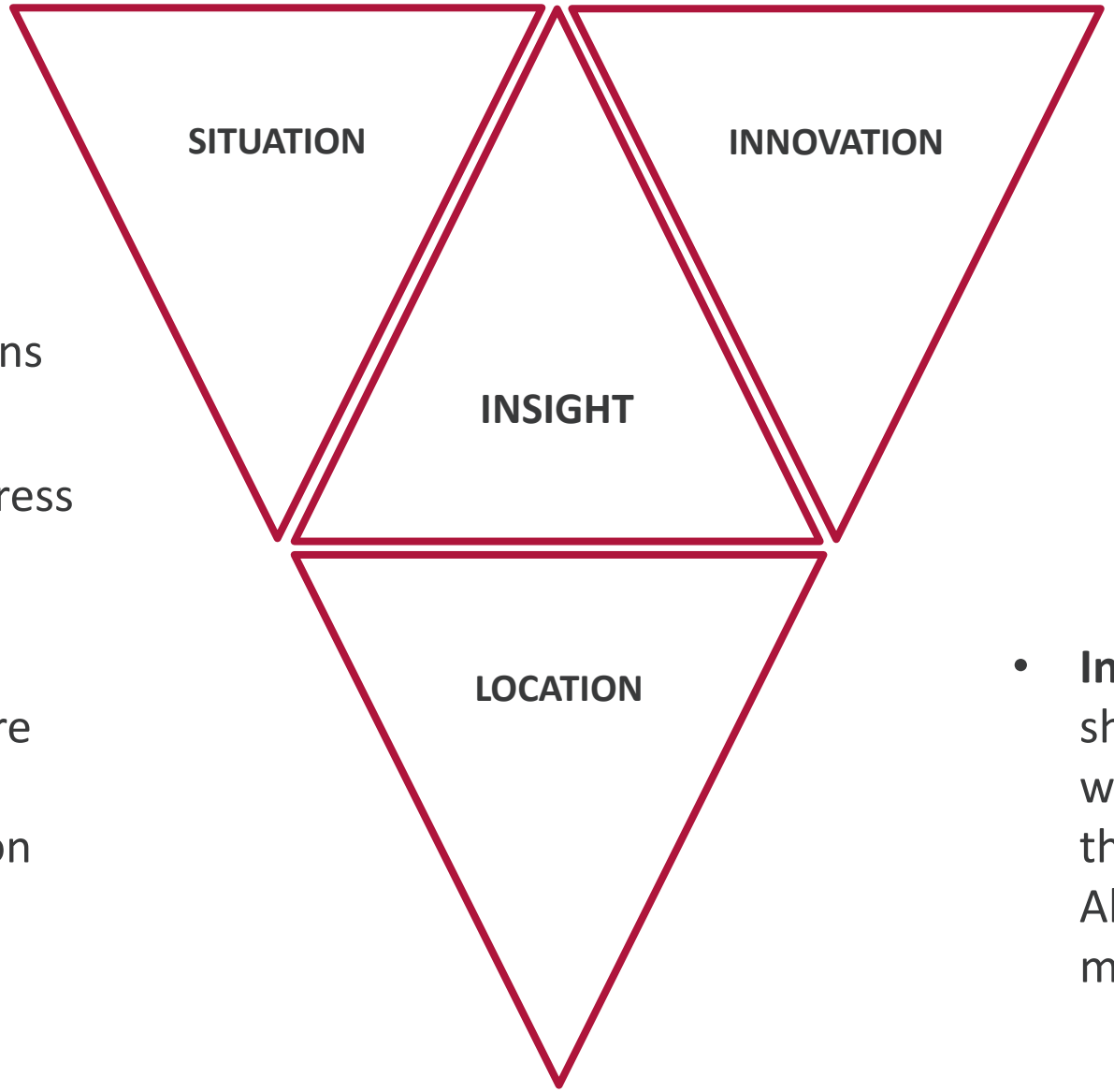


Innovation

Data analytics is an important part of Ofcom's work, and effective data pipelines present both opportunities and challenges for us as we aim to make communications work for everyone

- There is more data, but it is presented in containerised ways, and we need to adapt our digital skills to comprehend how to access and analyse it
- As a consequence we need to govern, manage and engineer our data appropriately
- Data science methodologies have returned:
  - To give us new approaches to address regulatory matters
  - A need to understand how data analytics and algorithms can, and are being used, especially with regard to enabling consumer services
- Delivery, both through reporting, and providing data for use by interested parties

- **Situation**; spans an increasing range of personal and technology dimensions
- **Innovation**; project opportunities to address the data pipeline
- **Location**; unlocking and sharing data more effectively at the appropriate resolution



- **Insight**; “we can only see a short distance ahead, but we can see plenty there that needs to be done” Alan Turing, Computing machinery and intelligence