Building geospatial capability in a statistical organisation

Eurogeographics General Assembly

3 October 2017

Building modern statistics

Target Operating Model

- Customer focussed
- Partnerships
- Curiosity
- Innovation

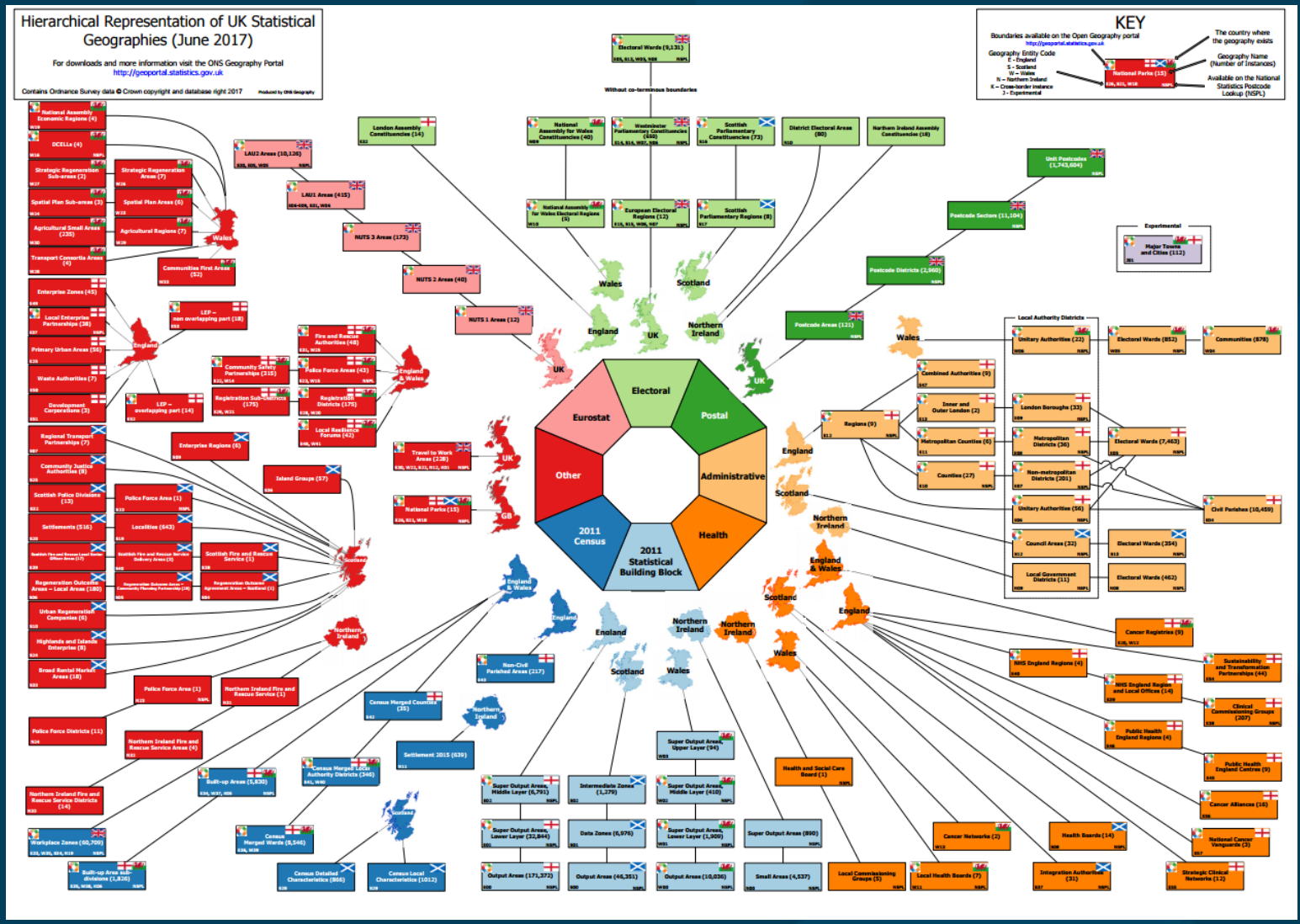
Stepping forward
Stepping up
Stepping on the gas

Drowning in data to swimming with information

“Everything happens somewhere”

- Place Matters: The Location Strategy for the United Kingdom, 2008

Complexity of UK geography



Geospatial data and statistics

GSS Geography Policy for National Statistics – a UK framework for consistent and comparable spatial statistics

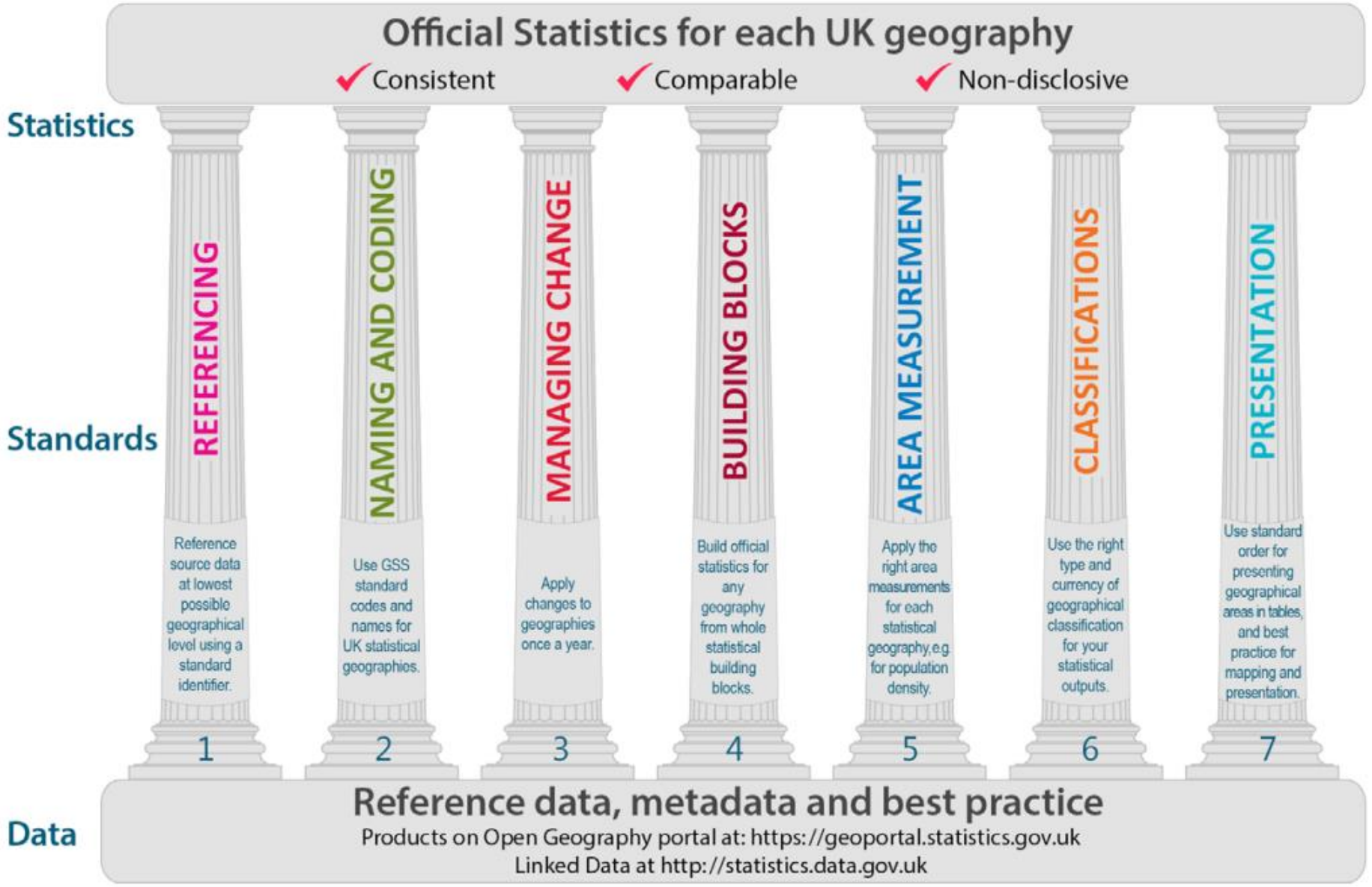
Enabler for spatial analysis

Main focus is address level geocoding, geographic aggregation of statistics and consistent identifiers

Doesn’t cover alternative data sources or the application of geospatial data to statistical analysis

Government Statistical Service

GSS Geography Policy Framework



The challenges of geospatial

Data management – only treated as a variable of statistics despite its complexity;

Authority – some data available at UK level, some at GB, some at England and Wales, some at devolved level;

Disclosure –having data on more geographies increases possibility of disclosure by differencing;

Interoperability –lack of standards between geospatial and statistical communities making it difficult to integrate;

Licensing – some progress has been made (NI open data) but still inconsistency in geospatial licensing (UPRN and Northern Irish postcodes not open)

Sustainable Development Goals

The 2030 Sustainable Development Agenda sets out 17 goals and 169 targets.

Monitored by 232 statistical indicators.

Geospatial data may be used to produce an indicator but can also be used to augment, validate, disaggregate or disseminate a statistic.

Intra-Agency Expert Group on SDGs has a working group looking specifically at geospatial datasets and disaggregation

The Global Approach

UN leading development of Global Geospatial Information Management through a Committee of Experts

↓

Setting agenda for development of global geospatial information and promoting its use to address key global challenges.

↓

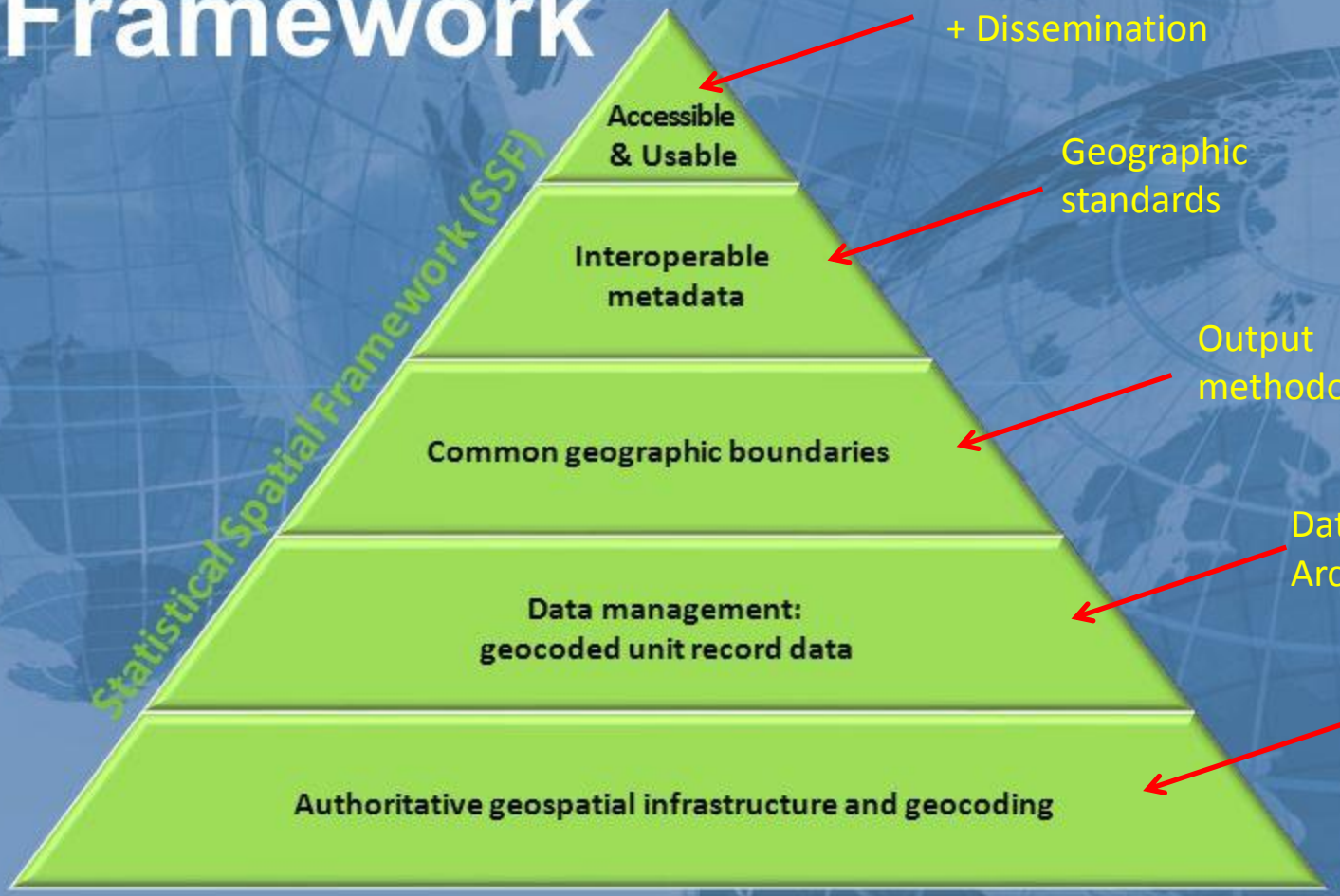
Integration of statistics and geospatial identified as one of these key global challenges

↓

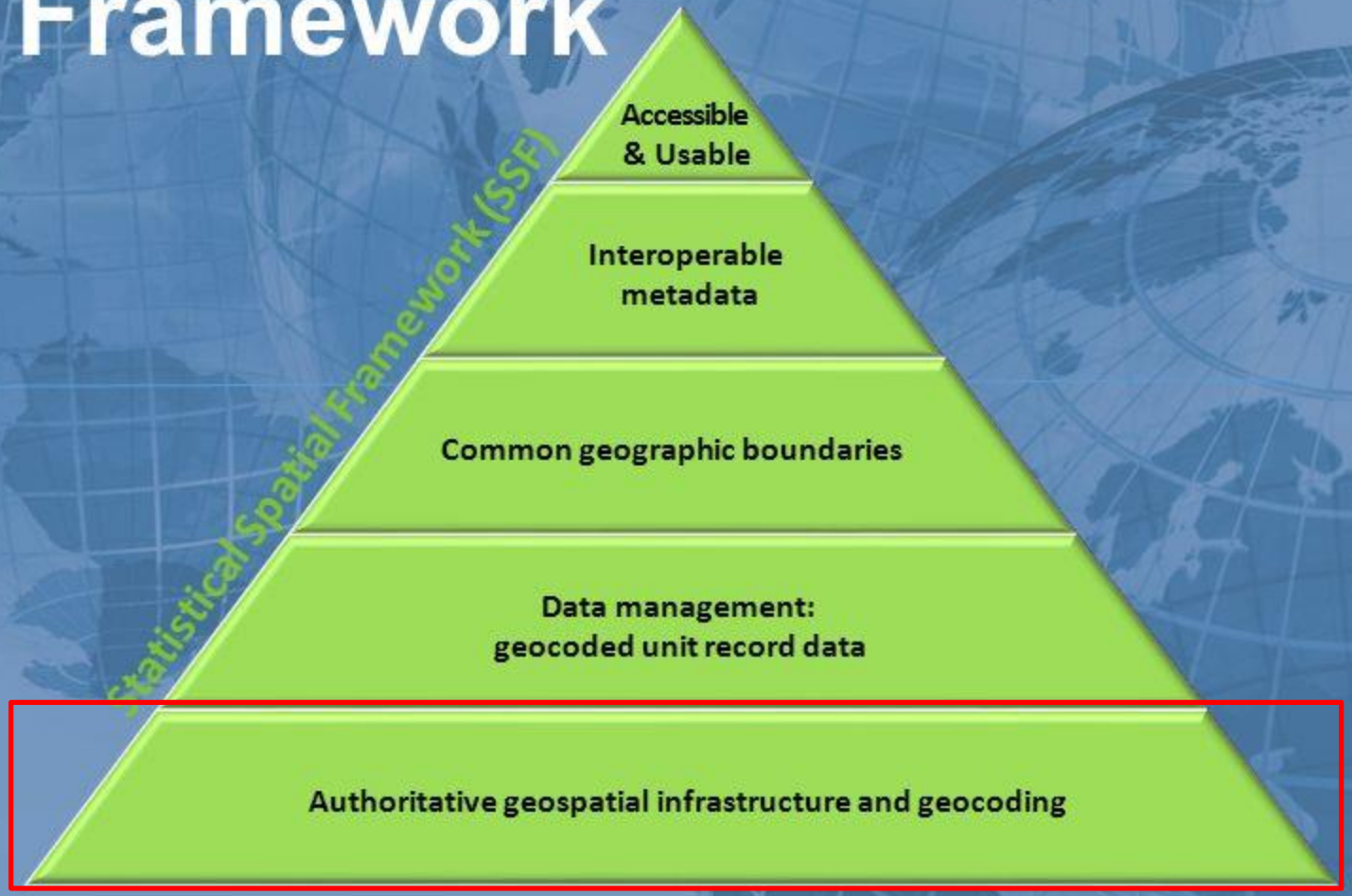
Expert Group on Integration of Statistics and Geospatial established to address this challenge – developing a statistical-spatial framework

UN-GGIM
UNITED NATIONS INITIATIVE ON GLOBAL GEOSPATIAL INFORMATION MANAGEMENT

Statistical Spatial Framework

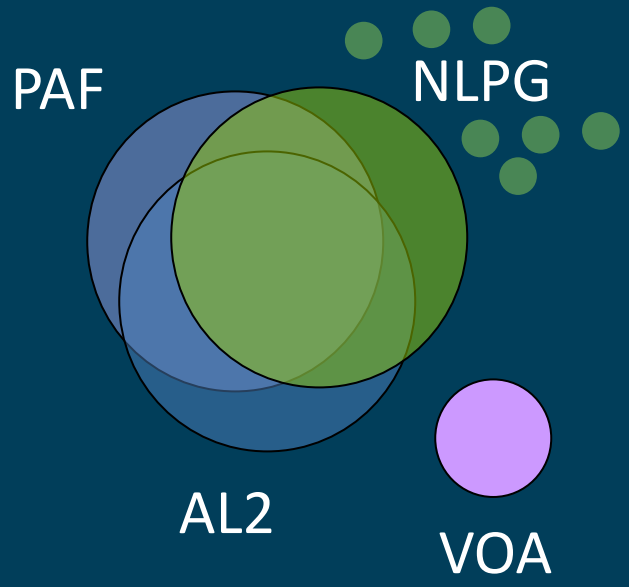


Statistical Spatial Framework



Additional CE sources

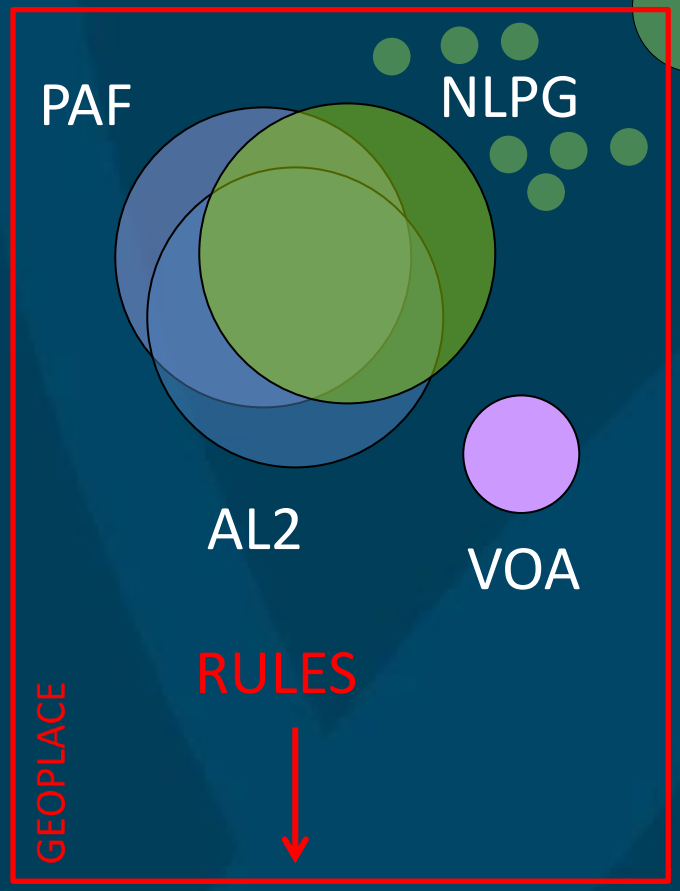
2011 Census AR



RULES

Residential address list

AddressBase



RULES

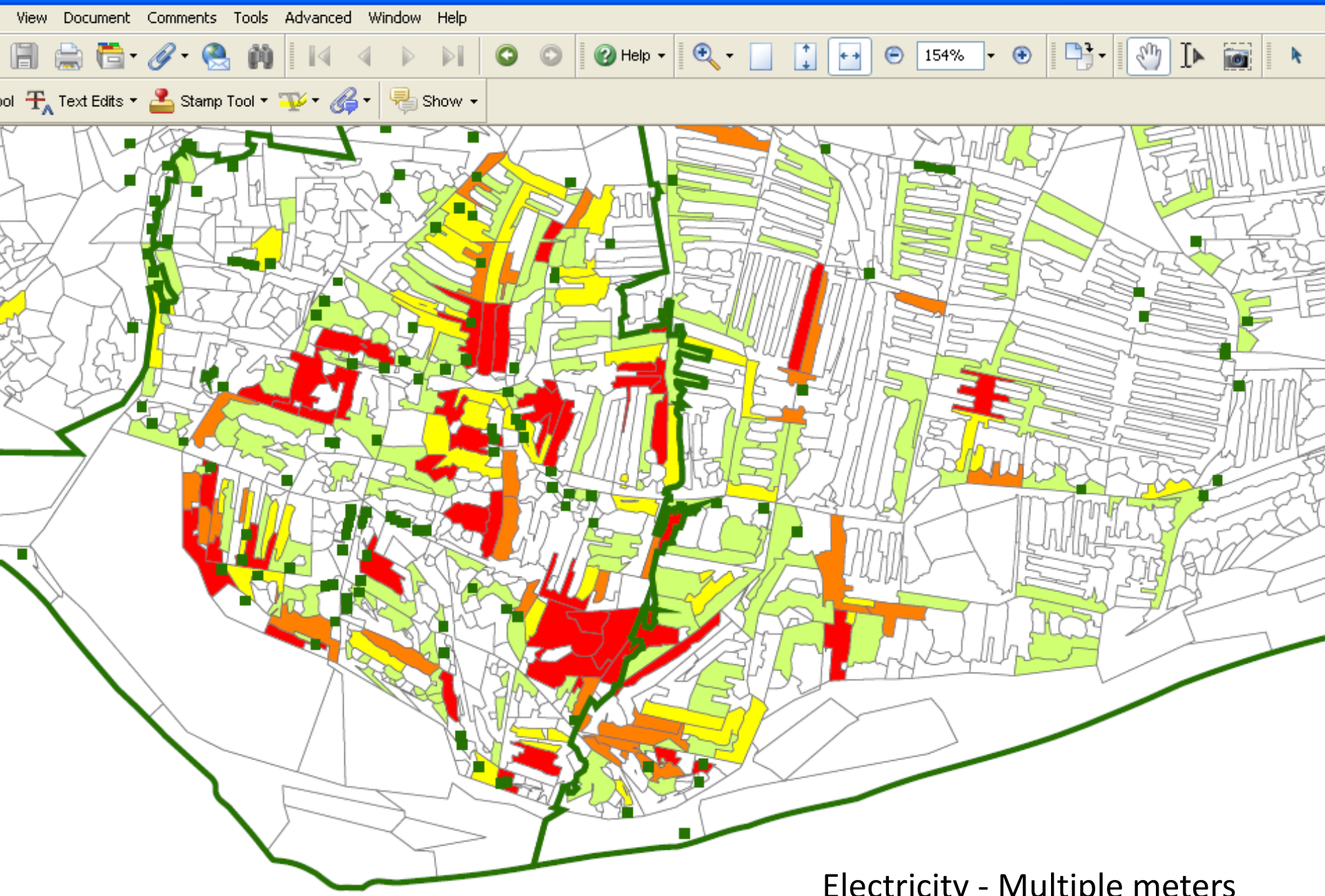
Residential address list

Communal address list

The challenge ….. Why it's hard this time

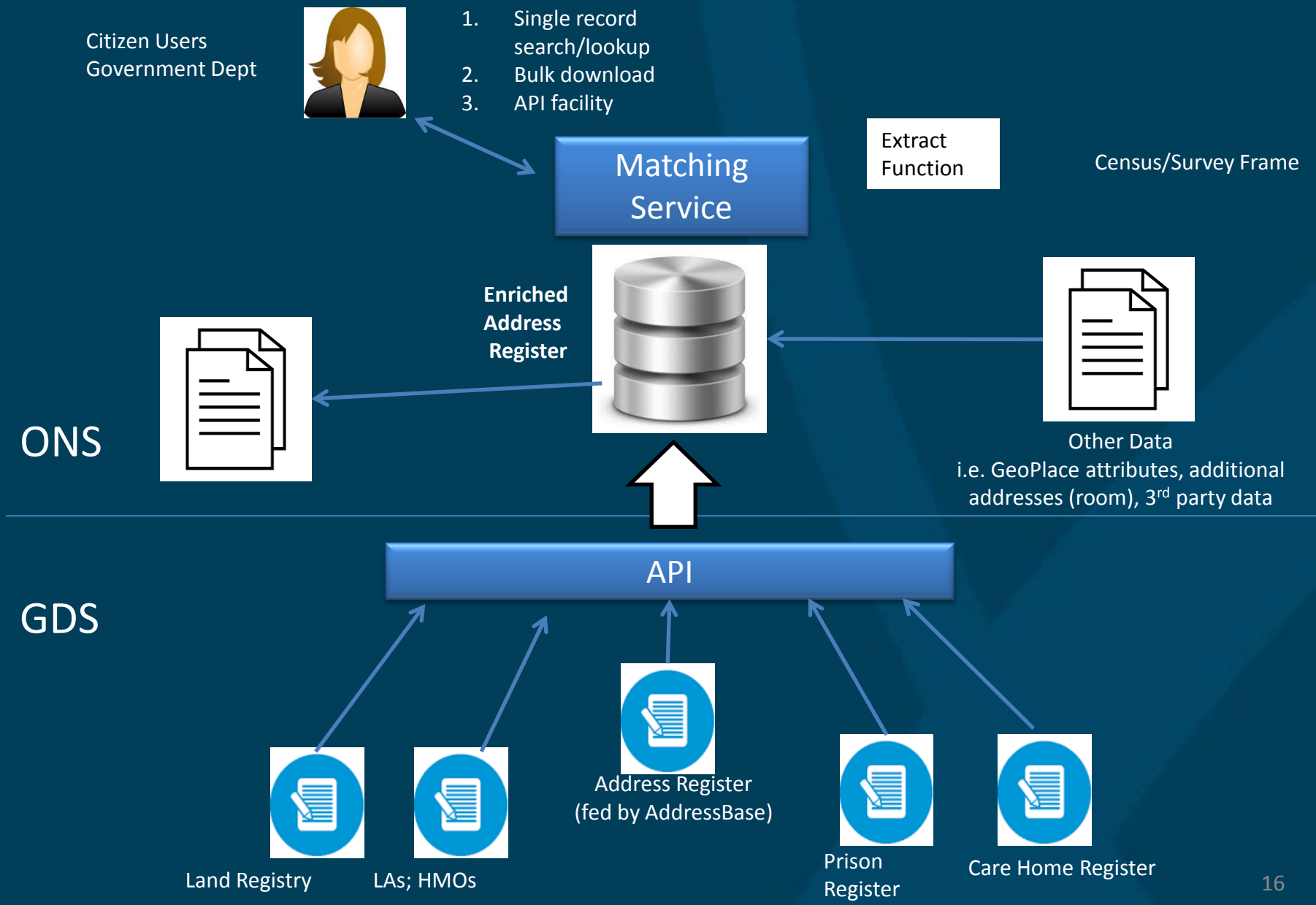
- We have an excellent starting point but addresses are complicated and change a lot. There will be error & error clusters itself in the areas we care about the most – **Very difficult to check quality**
- Extracting the right ones is difficult. Small errors can be significant – and cause trauma
- Communals are important and particularly challenging
- Addresses are complex so matching is really hard
- We plan to do MUCH more with the register than post-out – huge opportunity but attribute thinking is new


POST
2
4
5
?
5
B
10

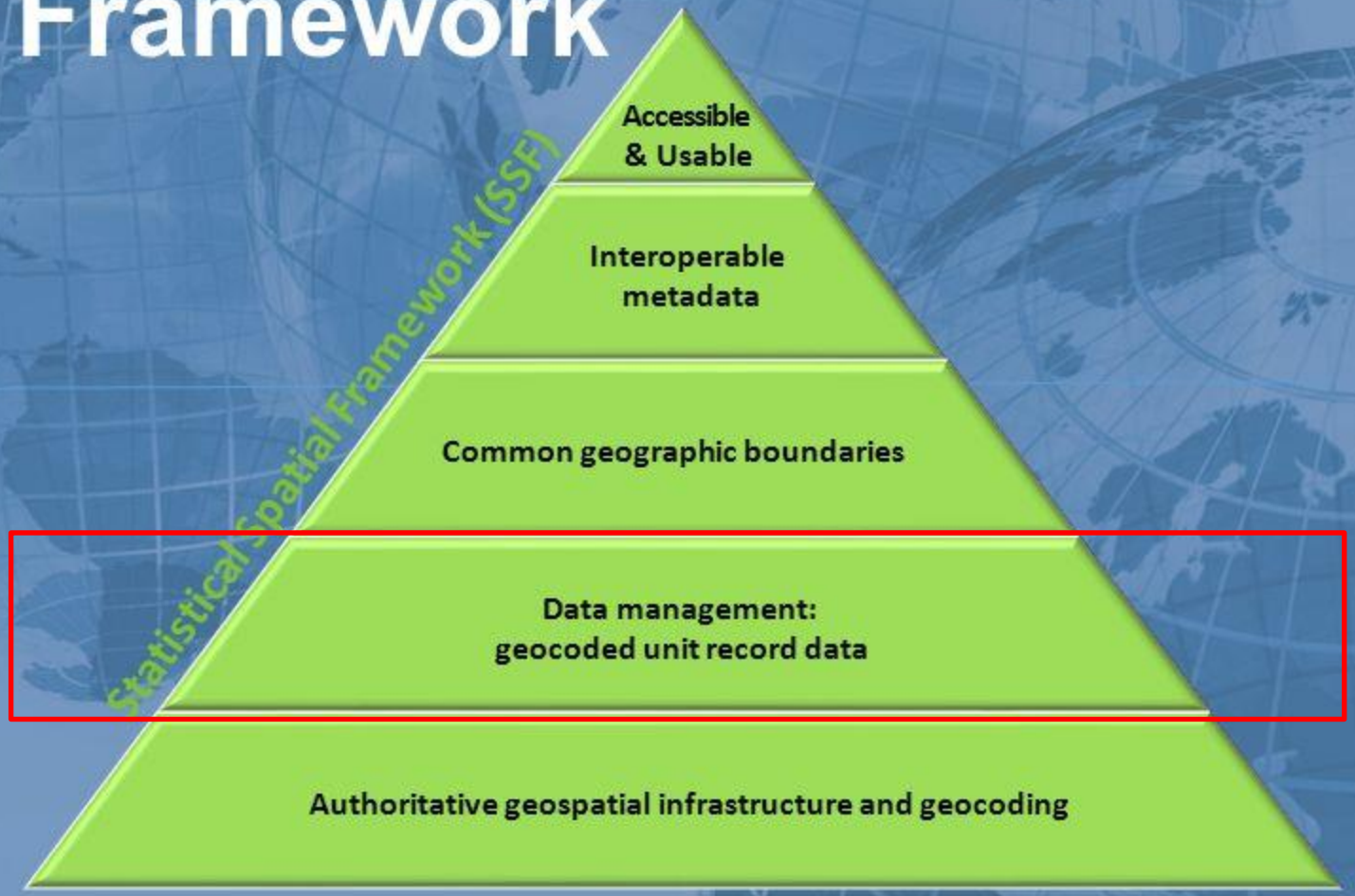


Electricity - Multiple meters

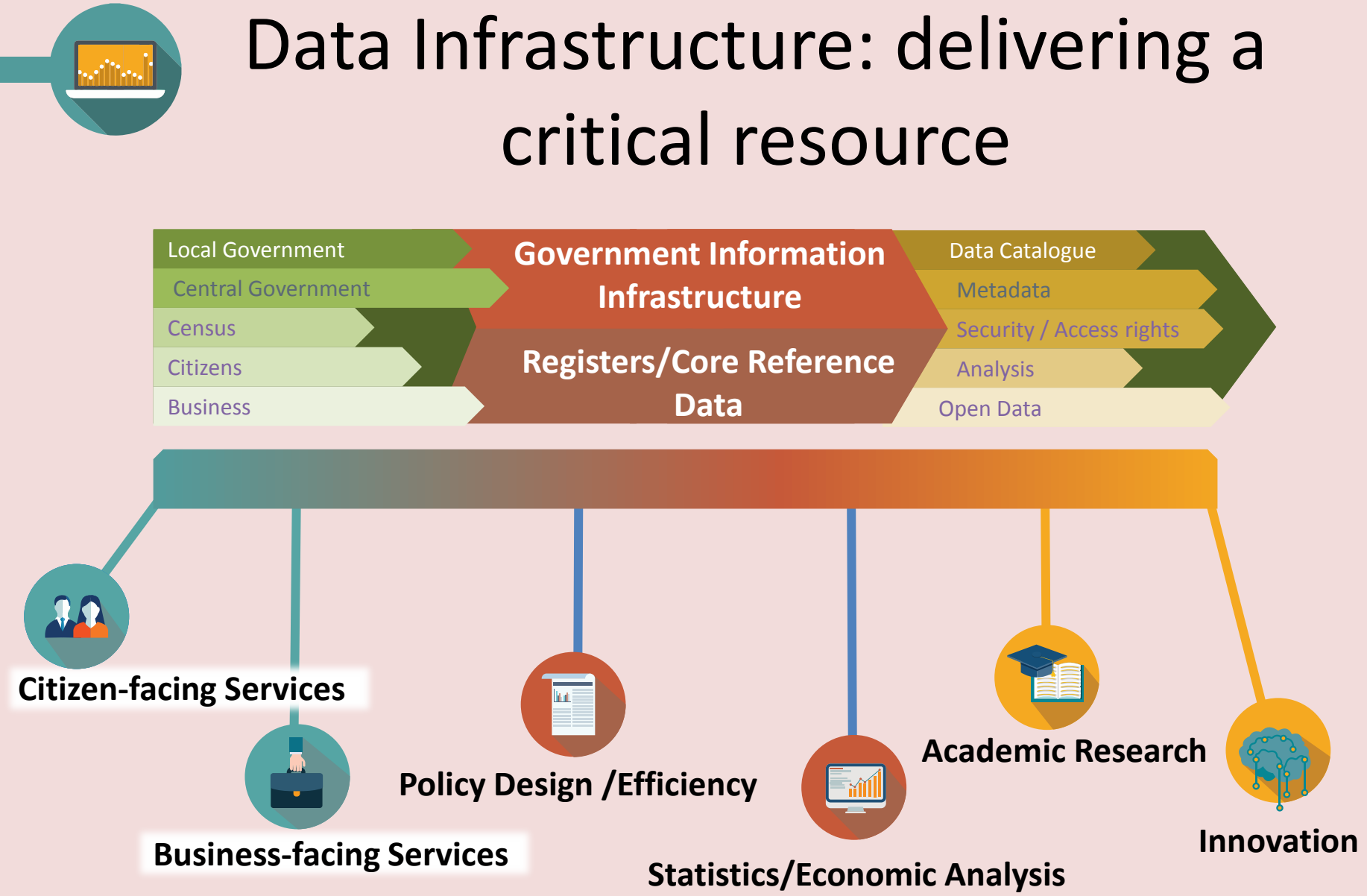
GDS & ONS; Registers, Platforms & Services



Statistical Spatial Framework



Data Infrastructure: delivering a critical resource



The challenge: multiple data user lenses across estate

Statistical production:
Secure, in-house access
ONS staff

Statistical services:
Match, link, annonymise data: internal and 3rd party use

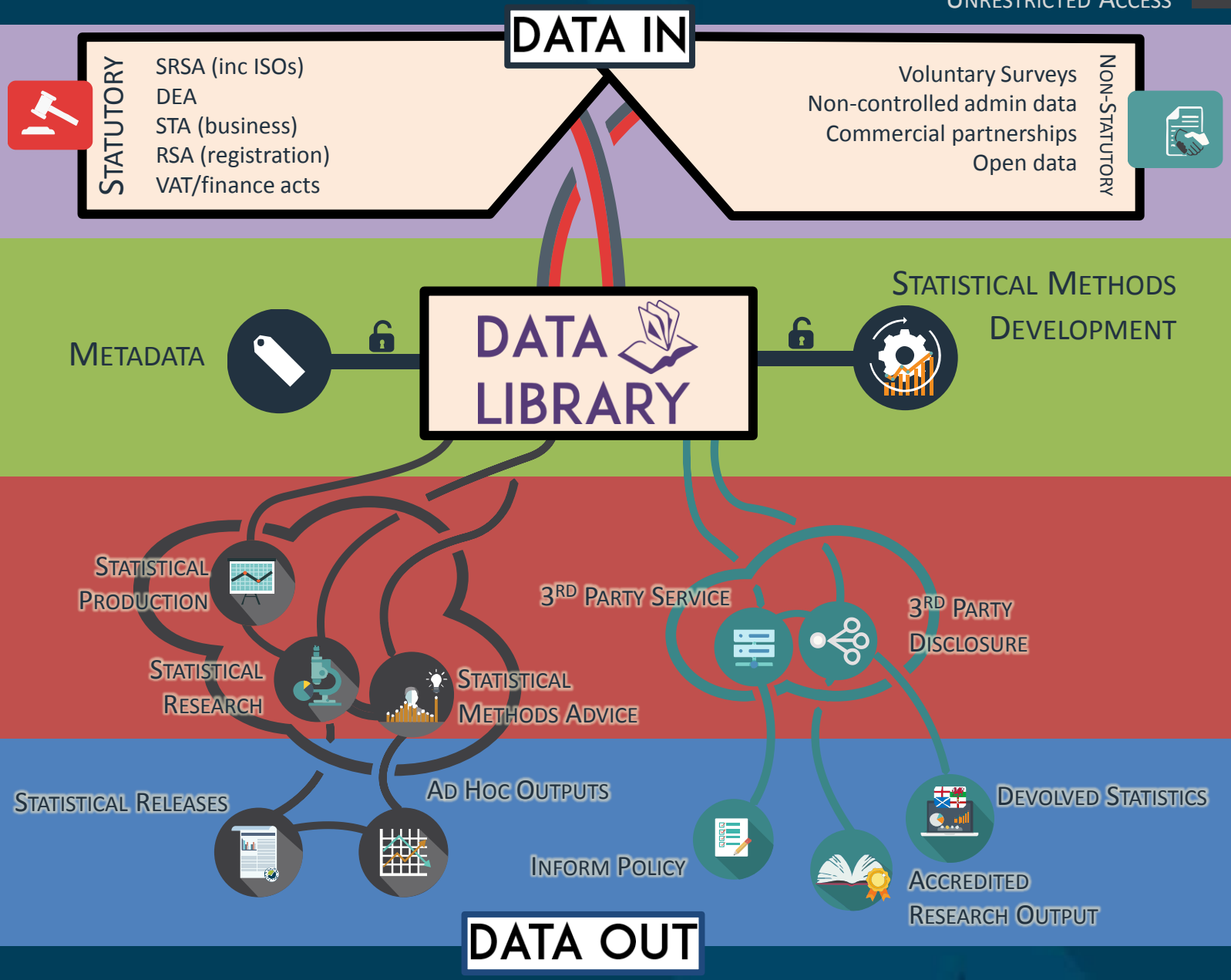
Statistical research:
Working in partnerships, project-based access

3rd party disclosure:
Accredited research and statistics

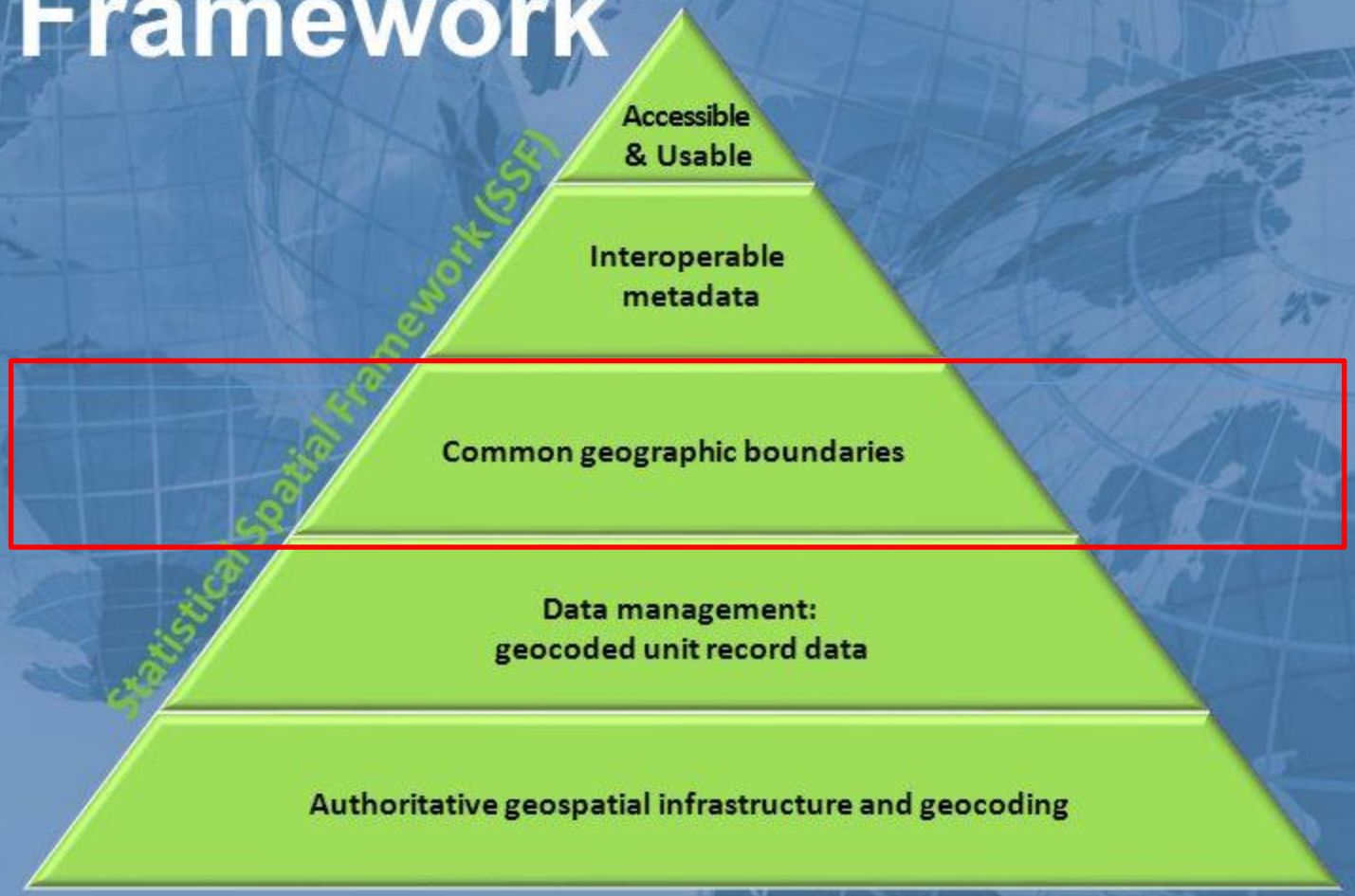
Office for National Statistics

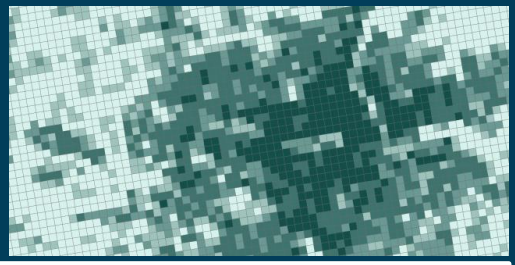
Data use and access

IDENTIFIED
SAFE
CONTROLLED ACCESS
UNRESTRICTED ACCESS

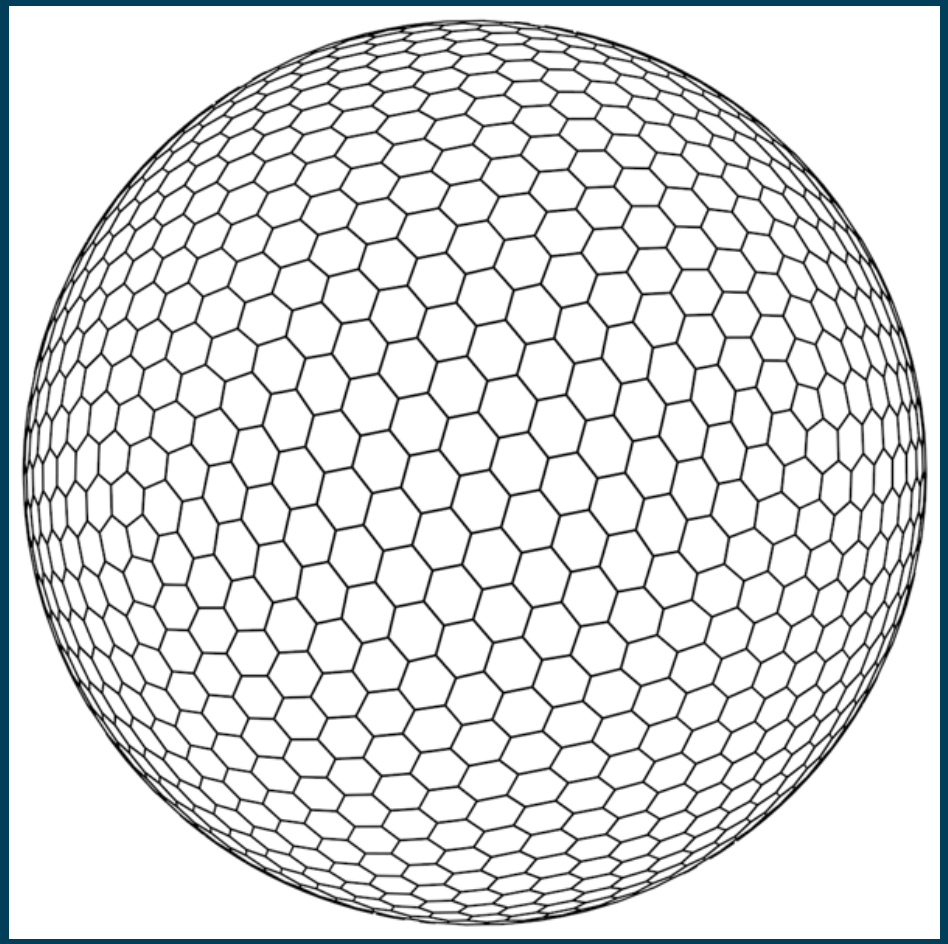


Statistical Spatial Framework



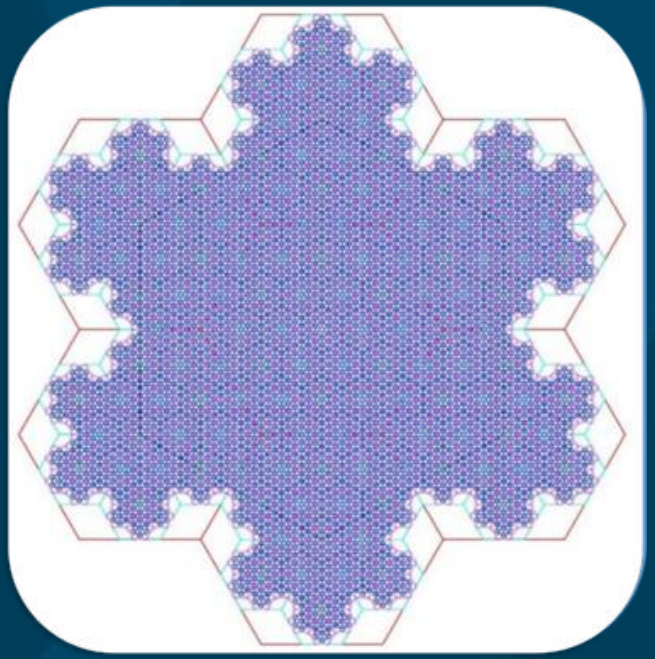
	administrative	statistical	grid
			
BENEFITS	• Relatable boundaries • Authoritative sources • Contain nesting relationships • Relationships with cadastral boundaries	• Allows statistics for small areas • Non disclosive • Comparability • Built around social homogeneity • Stability	• Global scope • Cross border studies • Comparability • Can locate people in space with more precision • Ready to use with GIS analysis • Able to see clusters
DIS-BENEFITS	• Comparability • Boundary changes • Number of different geographic areas • Irregular size • Irregular shape	• Geography size in rural areas • Don't exist at European or Global levels • Disguises spatial patterns • Boundaries are abstract	• Disclosure control/cell size • Grid cell sizes in rural areas • Different grid projection systems needed globally • Different grids exist

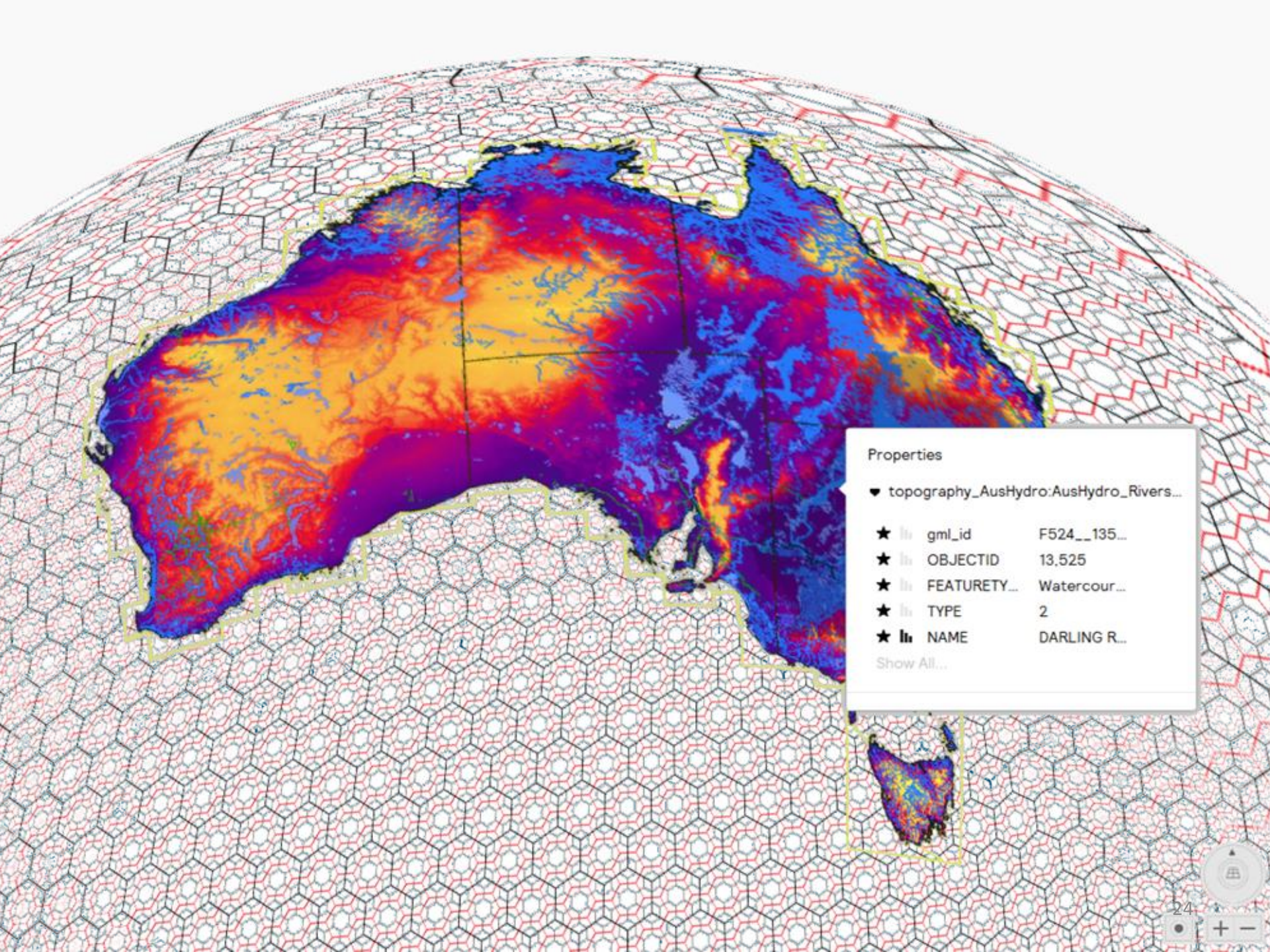
Discrete Global Grids



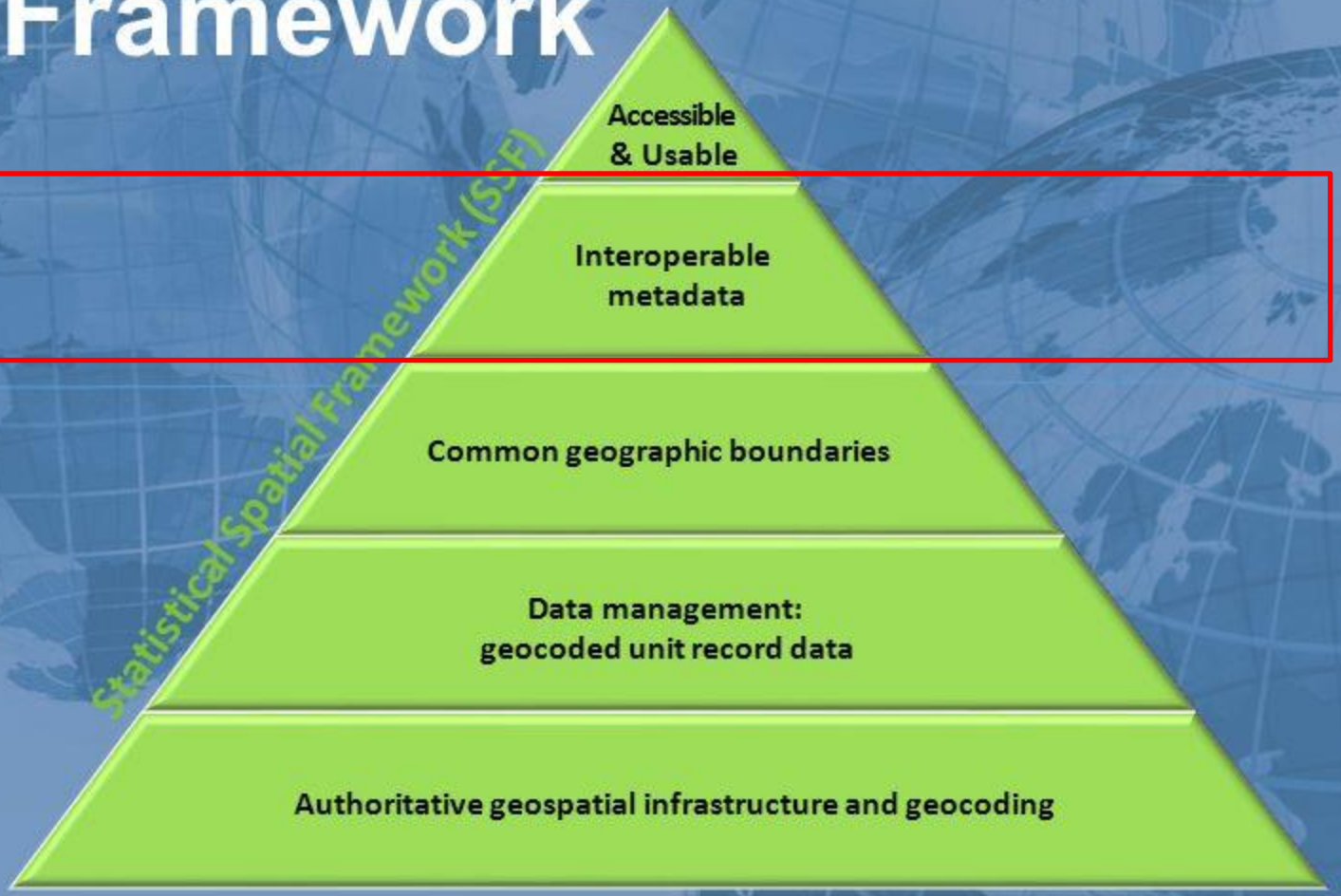
HEXAGONS ARE BETTER FOR A SPHERE THAN SQUARES

NOT AS SCALABLE AS SQUARES BUT STILL POSSIBLE




Properties
topography_AusHydro:AusHydro_Rivers...
gml_id F524__135...
OBJECTID 13,525
FEATURETY... Watercour...
TYPE 2
NAME DARLING R...
Show All...

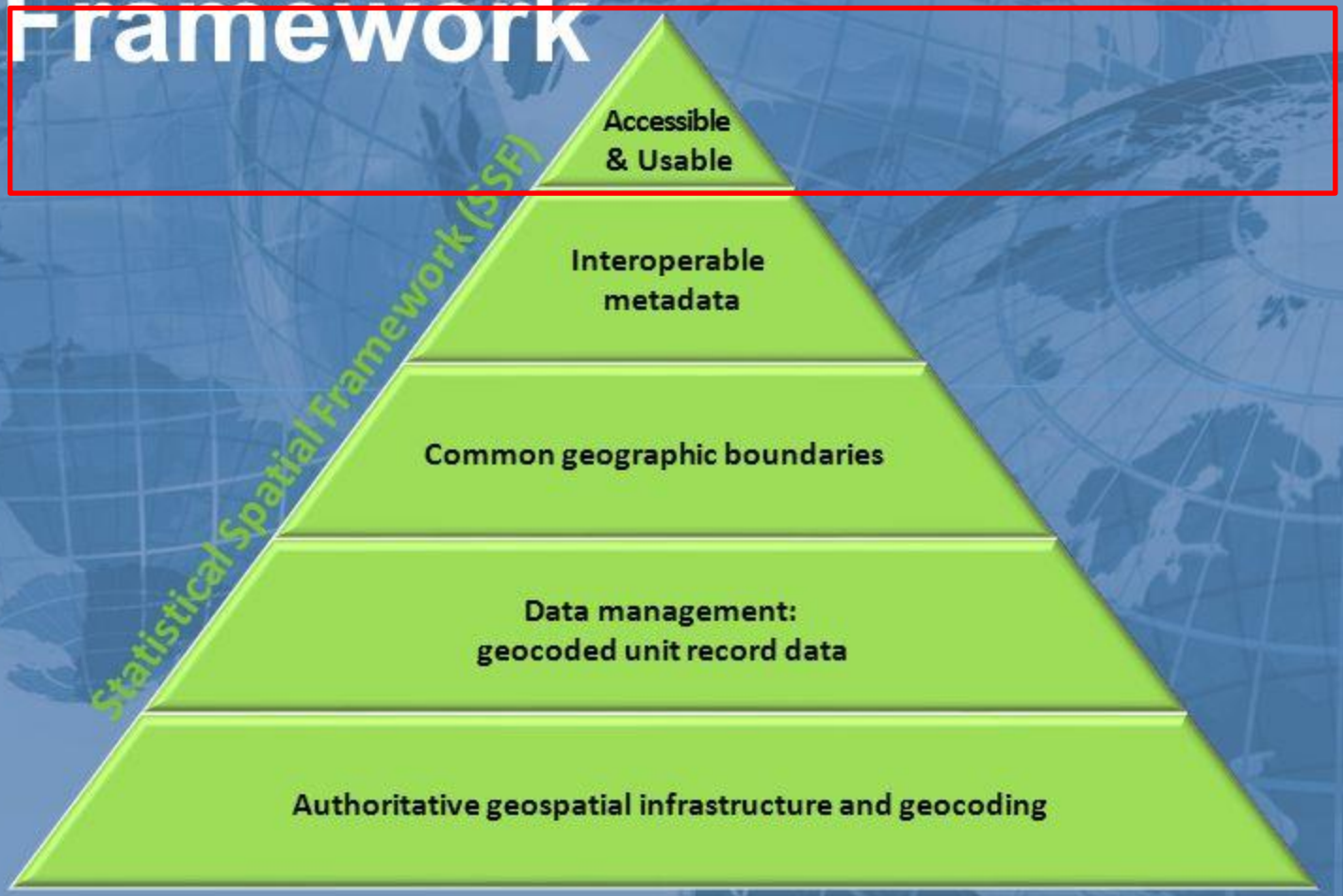
Statistical Spatial Framework



Standards

- Most geographic standards focussed on dissemination and driven by the EC INSPIRE Directive
  - GEMINI – standard for harmonised metadata
  - RDF – data standard for publishing data using unique identifiers and linking those identifiers to allow queries across distributed datasets
  - Table Joining Service – a standard for joining together statistical tables and geographic boundaries and offering them up as a web service

Statistical Spatial Framework



Explore the GeoSpatial Hub without a sign-in. To create content and use the Hub to its full potential, please contact ONS Geography for a login.

The ONS GeoSpatial Hub is an internal platform for geographic information, smart mapping and spatial analysis. It provides free and open access to the definitive source of geographic products, web applications, story maps, services and APIs. All content is available under the Open Government Licence v3.0, except where otherwise stated.

Get started! Learn how to...

Map data from the Hub

Make a web application with a

Add authoritative content to a

"Although better use of [data] has the potential to transform the provision of economic statistics, ONS will need to build up its capability to handle such data.

This will take some time and will require not only recruitment of a cadre of data scientists but also active learning and experimentation.

That can be facilitated through collaboration with relevant partners – in academia, the private and public sectors, and internationally."

Independent Review of Economic Statistics
Professor Sir Charles Bean, March 2016, p.11

London Transport workers manually examine over 4 million tickets to identify most and least popular routes, March 1939

Gerry Cranham/Fox Photos/Hulton Archive/Getty Images

Building the Data Science Campus

- Funding approved in Mar 2016
- Start-up team in place in Jul 2016
- Temporary Campus open Aug 2016
- 1st Data Scientist in Aug 2016
- 1st Academic Manager in Sep 2016
- Research commenced in Sep 2016
- 1st Apprentices in Nov 2016
- 1st research output in Dec 2016
- Managing Director joined in Jan 2017
- Headcount reached 26 in Feb 2017
- Formal launch 27 March 2017
- Move to bespoke Campus in May 2017
- Headcount to reach 50-60 by March 2018

Mapping the urban forest

Urban forest refers to the trees and vegetation present in the streets, parks, gardens, balconies and even green roofs within our towns and cities.

Trees are known to provide a wide variety of **social**, **environmental** and **economic** benefits to the inhabitants of a city. For example, the presence of trees can improve social cohesion, encourage exercise, filter air, reduce flooding, provide habitat for wildlife and even stimulate the local economy.

This project aims to examine ways to **automatically count and classify trees in cities**, using computer vision and alternative data sources such as Google Street View.

Video available at - https://www.youtube.com/watch?v=xinSdVMf5ug

Spatial Disaggregation

The 2030 Sustainable Development Agenda includes a principle that “leaves no-one behind”. This means that all of the Sustainable Development Indicators need to be provided at sub-national geographic levels, even where that data does not currently exist.

To support this, methods for disaggregation national level datasets need to be developed and these need to be supported by covariate datasets that can support accurate apportionment of national level data.

This project aims to identify geographic datasets that can be used as covariates and use this to disaggregate data at higher administrative areas onto 100m grids.

Using land cover for natural capital

Joint ONS-DEFRA project to identify the value of natural capital to the UK economy

Previous definitions have used Centre for Ecology and Hydrology Land Cover Map based on LandSat-8 data

Sentinel-2 data may offer better resolution and more frequent outputs.

Data Science Campus will evaluate the differences between LandSat-8 and Sentinel-2 to quantify the value that a move to the latter would offer to the Natural Capital Accounts outputs

Summary and Discussion

- To modernise official statistics, geospatial data is needed as core reference data within the statistical process
- Key to this will be building spatial awareness and literacy
- Geospatial Frameworks are important in delivering a greater integration of statistics and geospatial
- Interoperability between the statistical and geospatial domains will be important but not clear how we will manage the velocity of this work.
- Increased use of geospatial data will create privacy challenges – how do we manage the ethical use of geospatial data?

"Cyberpoaching" Feared as New Threat to Rare Wildlife

Email hacking incident in India raises concerns, conservationists say.

By **Sasha Ingber**, for National Geographic

PUBLISHED OCTOBER 11, 2013

A tiger is fitted with a radio tracking collar by researchers in Thailand.

PHOTOGRAPH BY STEVE WINTER, NATIONAL GEOGRAPHIC

RELATED CONTENT

Now We Know Why Poison Frogs Don't Poison Themselves

This Spider Isn't From Mars—So Why Is It Named for David Bowie?

ANY QUESTIONS?

Ian Coady
Head of Geospatial Policy, Research and International
Ian.coady@ons.gov.uk