crc•si *Spatially Enabling Australia and New Zealand*

Project 4.41 | Improving Usability of Online Health Geovisualisation Tools

Project Leader Dr S Moncrieff, Curtin University

Research Team Ms E-K Gulland*; Dr S Moncrieff*; Mr P Konings ANU; Prof G West*; Dr Z Mohamed-Ghouse GHD; Dr O Gudes* (*Curtin University)

Project Participants Curtin University, National Centre for Geographic & Resource Analysis in Primary Health Care at Australian National University, GHD

Objectives To improve discoverability of online spatial datasets in context for a user's problem domain, focussing on health applications

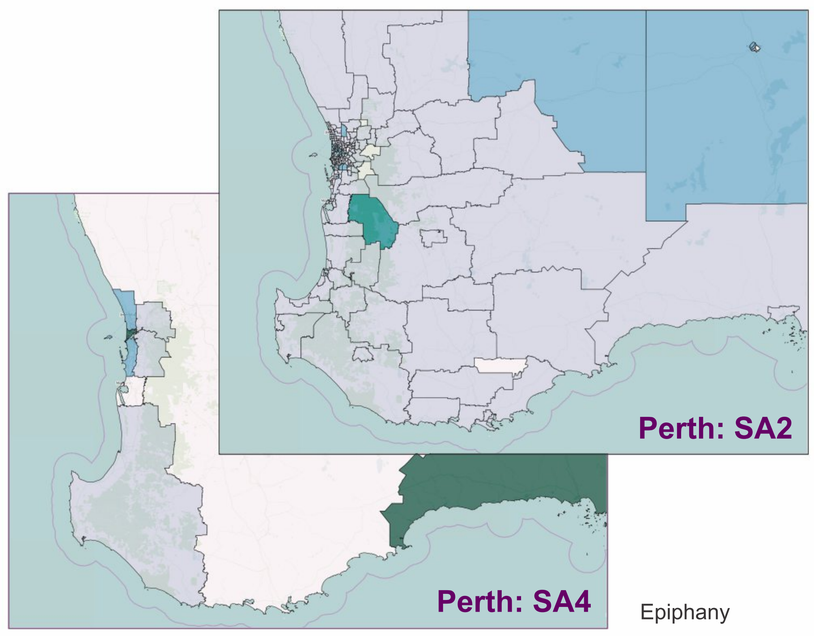
Outcomes This project aims to develop and test:

- Online distributed search tools to look for data from different data providers and in different themes
  - including spatial search for filtering by area and region type
  - returning more comprehensive dataset lists based on initial queries
- Display of results organised by topics to make it easier to choose between them
- Mechanisms to learn from users' choices, include expert users

Context: Thematic data,and where to start

- Thematic data is common in research fields including health and socioeconomics
- Data can be collected and reported for different region types
  - National, state, medicare locals, statistical areas (levels 1-4), … and many more!
  - Where regions are hierarchical, some data can be aggregated to a different region
- Multiple terminologies are used in different contexts
- Large volume of datasets, increasingly being made available online
  - Current online tools can "search in text" to match a query term to dataset names or other metadata

Finding the right data at the right time for the right purpose is difficult and time-consuming



Epiphany

Problems

Relevant datasets not returned

What if I use a term in a different way to the data custodian(s)?

"fire" – 5 datasets
"bushfire" – 1 dataset

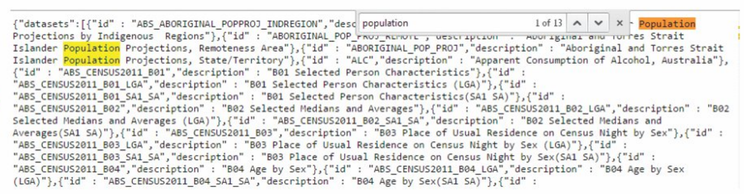
4 results are missing if you use the "wrong" term!

GeoNetwork (TERN AusCover)

Many irrelevant datasets returned

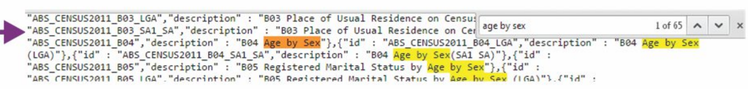
What if I want to find population counts from the last census?

Search for "population" →



13 matches (out of a possible 262); none are counts from the last census. The one you want is actually called "Age by Sex". Surprised?

Search for "age by sex" →



65 matches (133 for either of 'age' or 'sex')
How do I find the right one? What if I wanted specific counts for education, income,…?

ABS – Australian Bureau of Statistics (online REST interface)

Approach

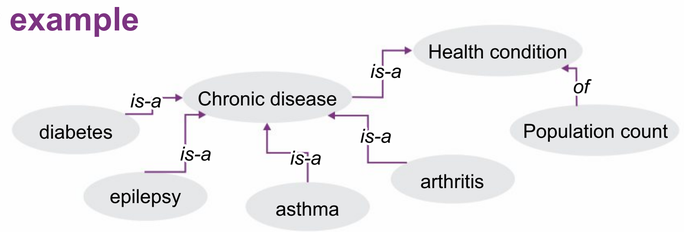
Original query: chronic diseases in my area

Missing link!: Wikipedia, Ontologies, Industry resources, chronic disease, *Related search terms*

Search datasets → **Datasets found**: No results → **Investigator's Choice**: X

Datasets found: Theme 1, Theme 2, Theme 3 → asthma

Ontology example



Industry resource example: Health

ICD-10 Disease classifier, with Chapter (e.g. respiratory), Major code (e.g. J45 – Asthma), Minor code (e.g. J45.0 – allergic asthma)

International Statistical Classification of Diseases & Related Health Problems 10th Revision

Summary

- More datasets available online means more robust search tools are needed
- Query terms do not always match indexes used to search data stores:
  - Different people organise data in different ways
  - Different terminologies are used in different ways depending on context
- Ontologies and crowd-sourced data are useful sources of related terms

PROB(A|B)

A: this is a dataset that is useful

B: this is the original query term or phrase