

# P3.01 | Improving Search and Discovery of Geospatial Data

**Project Leader** Dr David McMeekin, Dept of Spatial Sciences, Curtin University and Kylie Armstrong, CRCSI

**Research Team** Tristan Reed, Curtin University; Dr Simon Moncrieff, Curtin University; Em Prof Geoff West, Curtin University

**Project Participants** Curtin University, University of Canterbury  
AAM, Amristar, OmniLink, Semantic Sciences  
Landgate WA, DELWP VIC, DNRM QLD

**Objectives** To allow easier discovery of geospatial data with minimal effort from data custodians.

**Outcomes**

- Creation of new algorithms to better rank search results
- Integration of algorithms with existing catalogue tools (eg. CKAN)
- Creation of interoperable Resource Description Framework schema to integrate domain-specific information

## Outcomes

DiscovGO is a tool that provides improved search capabilities for metadata catalogues. By using modern semantic and spatial algorithms, the system aims to identify data sets that are more relevant to the users query compared to simple 'query matching' techniques where the software attempts to find what the user typed in to the software inside its records.

To achieve this, semantic technologies are used to predict similar queries to what the user typed in and to then weight them based on the difference from the original query. Spatial semantic filtering is also applied to further increase the relevancy of results by extracting locations from the search term and looking at similar locations.

## Application

Much productivity is wasted looking for the correct data set, as often the metadata of a data set is described very differently to how a user would describe it. This can extend to business opportunities being missed as often it is easier to give up than to find the right data.

Through the use of DiscovGO, time can be saved by better identifying more relevant data sets and presenting them to the user. It was estimated by ACIL Tasman that in 2008, over \$500m AUD was wasted annually due to the inefficient access of spatial data. This number has surely grown over the years with the increasing use of spatial data in non-traditional industries.

## Adoption

DiscovGO is designed to be used by data custodians to allow their end users to better discover data suited to their purposes. The system can also be integrated in other tools used for a variety of other cases, such as enterprise data management systems and the like.

Investigations are currently ongoing with Landgate of WA with regards to using the system over their catalogue as a demonstration of the research.

Further investigations are ongoing using tools provided by 43pl partners to identify the best way to build domain thesauri from institutional data stores.

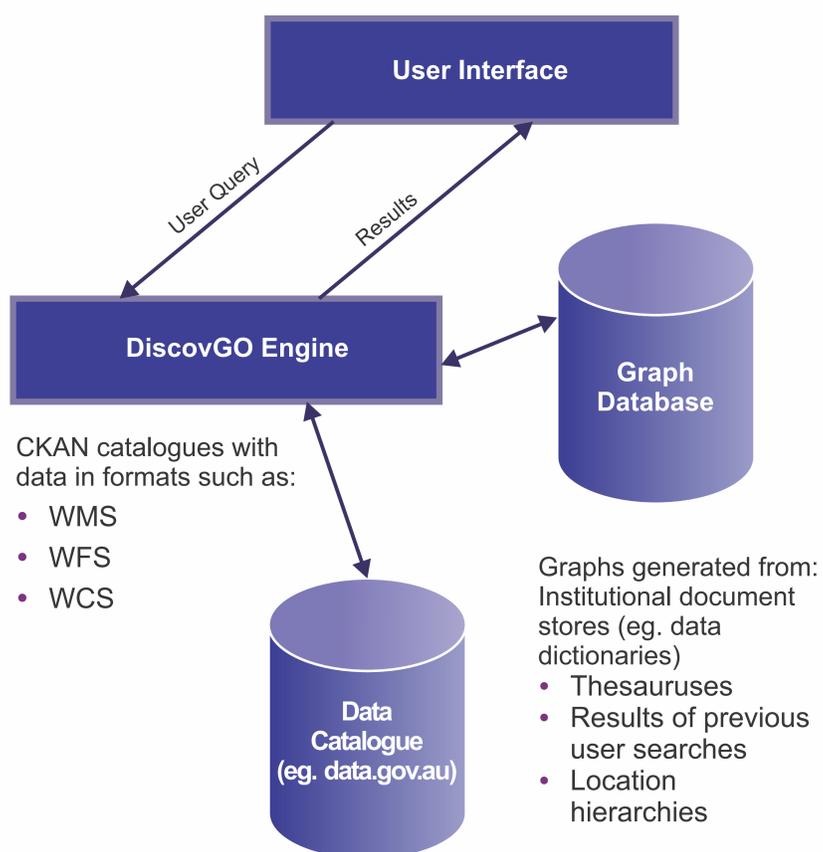


Fig 1 System Design and Architecture.

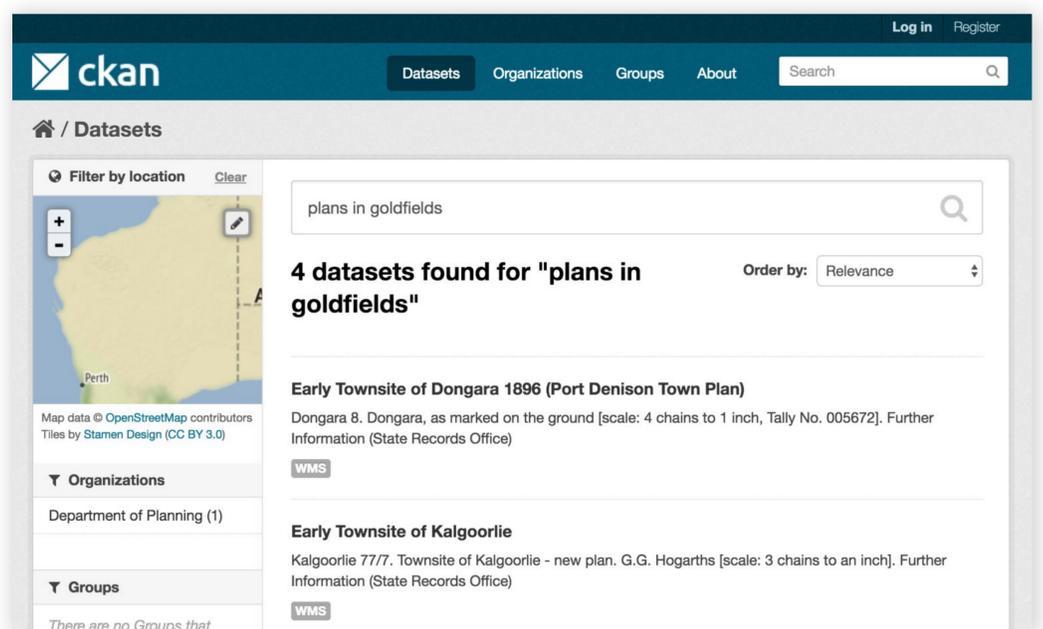


Fig 2 Screenshot of Demonstration System.