

# Developing a Science Keyword Taxonomy and SKOS/Semantic Relationships in NASA's Global Change Master Directory (GCMD)



Member

T. Stevens<sup>1,2</sup>, S. Ritz<sup>1,2</sup>, S. Areu<sup>1,2</sup>, T. Cherry<sup>1,2</sup> (1 NASA/GSFC, 2 WYLE)

## Introduction

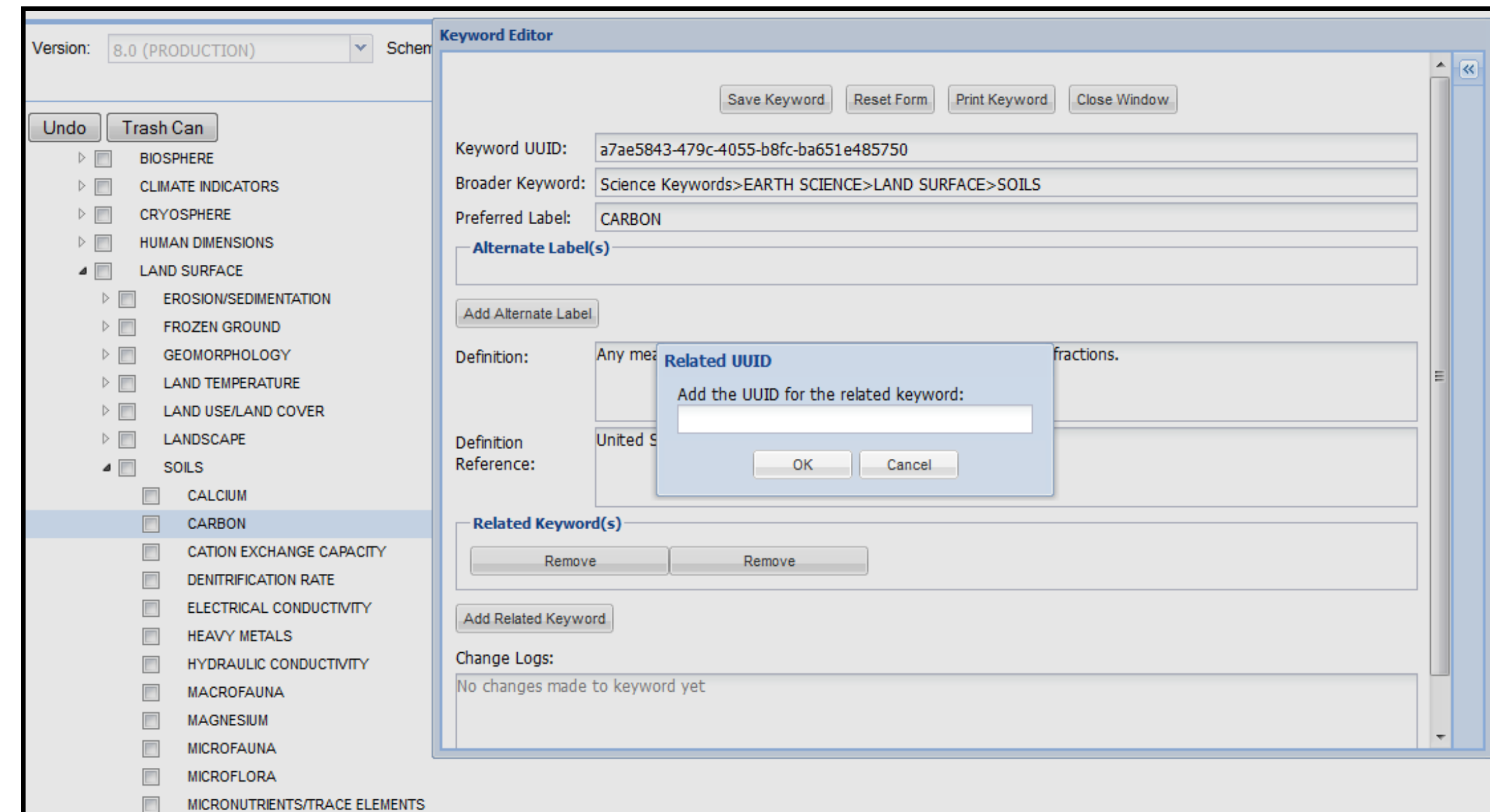
Controlled keyword vocabularies ensure that metadata is described in a consistent manner and allow for the precise searching of metadata records and subsequent retrieval of data and services. There are seven sets of controlled keywords in the GCMD: (1) Earth Science, (2) Data Services, (3) Data Centers, (4) Locations, (5) Instrument/Sensors, (6) Platforms/Sources, and (7) Projects. In addition, the Climate Diagnostics descriptions include two unique keyword sets: visualization type and analysis type.

In order to provide smarter and more enhanced searches on metadata records, the GCMD is working on extending the Keyword Management System (KMS) to allow GCMD Science Keywords to be related to any other existing keyword within the broader keyword hierarchy. This functionality is needed to allow soft linkages (hard linkages being SKOS style parent-child relationships) to be made between science keywords and other keyword hierarchies such as Projects, Platforms, or Instruments.

To learn more about the science keywords and the KMS, see <http://gcmd.nasa.gov/learn/keywords>.

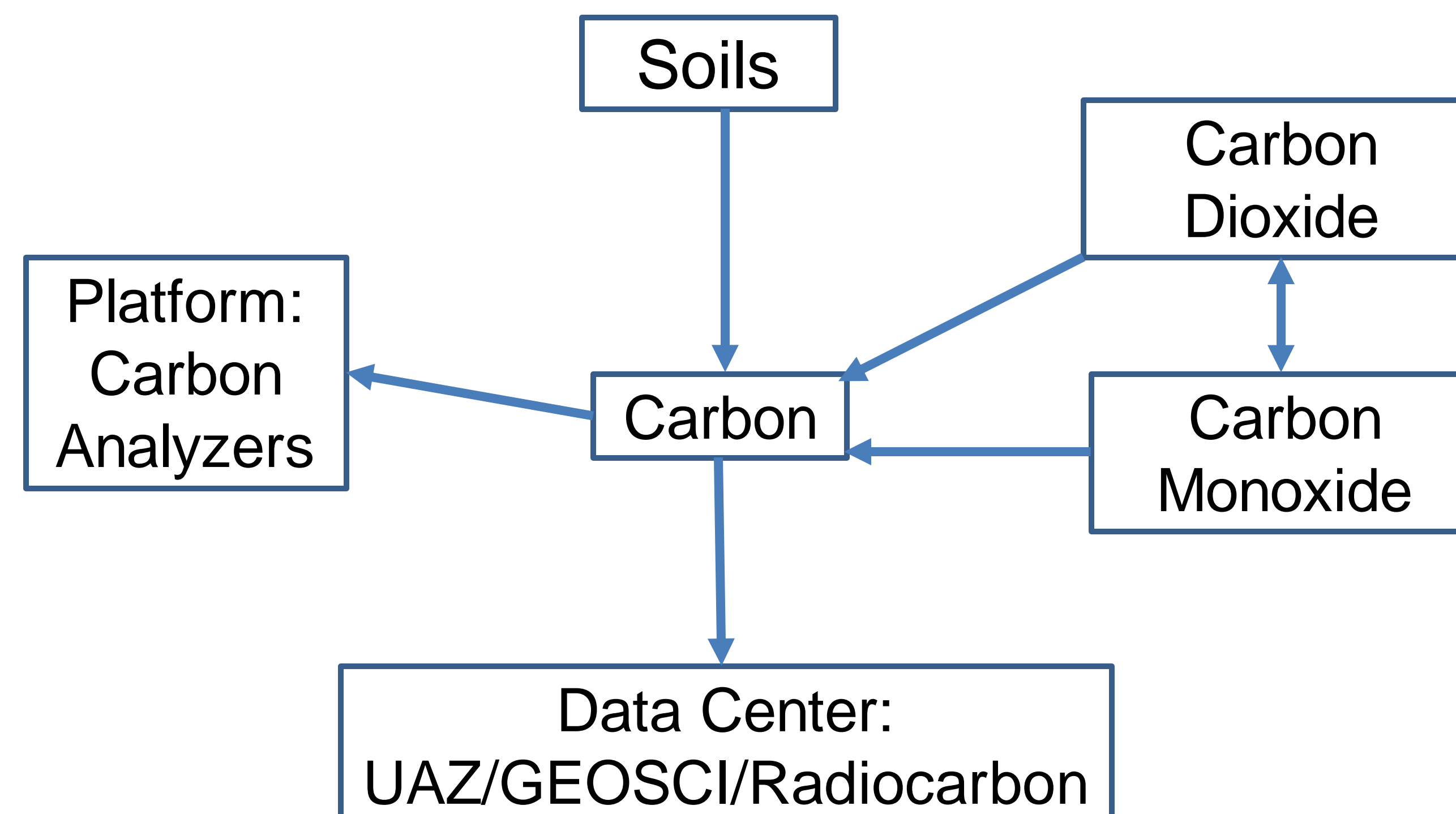
## Adding the Semantic Relationship

The GCMD keyword manager allows you to edit the keyword and add the related keyword (via UUID). All of the related keywords will show up in the keyword editor panel.



## Visualizing the Relationship

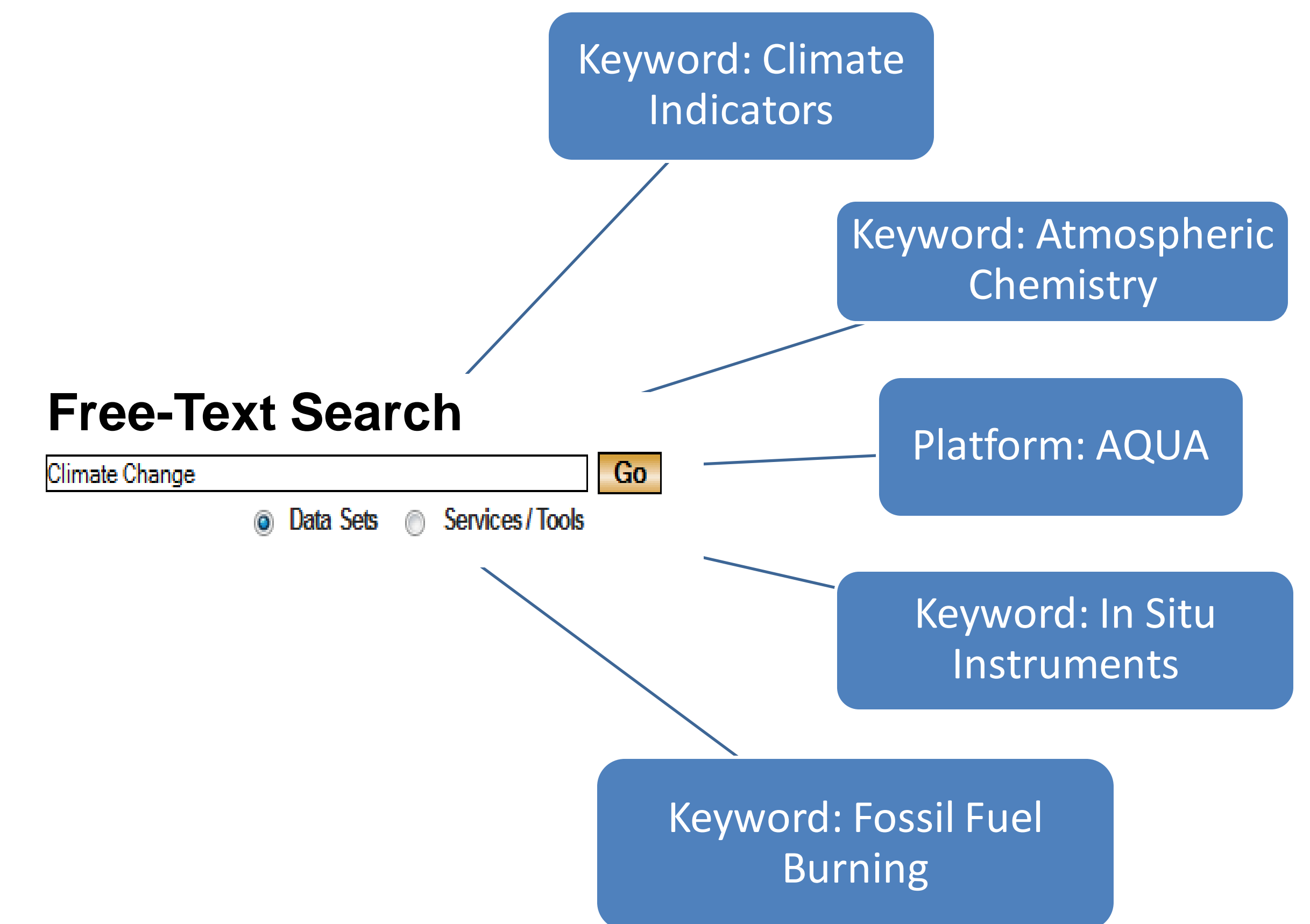
Relate the Science Keyword 'Soils > Carbon' to the Instrument 'Carbon Analyzers' that analyzes carbon and a Data Center 'UAZ/GEOSCI/RADIOCARBON' that does carbon research. There can also be other interrelationships within the seven sets of controlled keywords.



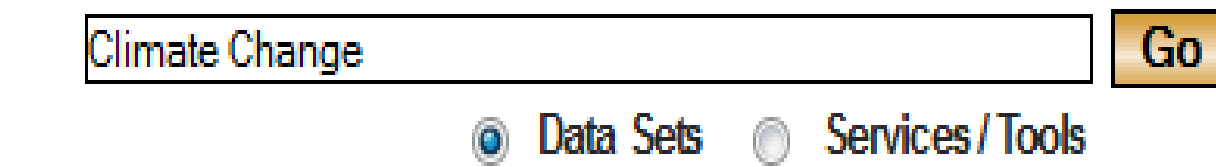
## Benefit to Search

The new keyword relationships can then be used in the future to provide enhanced search capabilities and search refinements with optional expanded lists of metadata records not directly searched for within the GCMD search results. This functionality would be similar to how web sites provide a "you may also be interested in..." section in the search results.

### Related Searches



### Free-Text Search



## Keywords Relationship

Any of the seven sets of controlled keywords can be related to each other. The related keywords will have three attributes to denote information about their semantic relationship:

- Type (Current)** • Text field allowing the relationship to be defined in a systematic way.
- Reason (Future)** • Free text describing why or how the concepts are related.
- Strength (Future)** • Float between 0-1 denoting the magnitude of the relationship.

## Future Work

- Add the keyword relationships into the search engine (via Keyword Manager) based on provider and science input.
- Develop an intuitive user interface to show metadata results from related and suggested searches.
- Implement weighted relationships (type and strength of relationship).

For questions or comments, please contact the GCMD staff at [GSFC-GCMDUSO@mail.nasa.gov](mailto:GSFC-GCMDUSO@mail.nasa.gov)