

Summer Internship Report

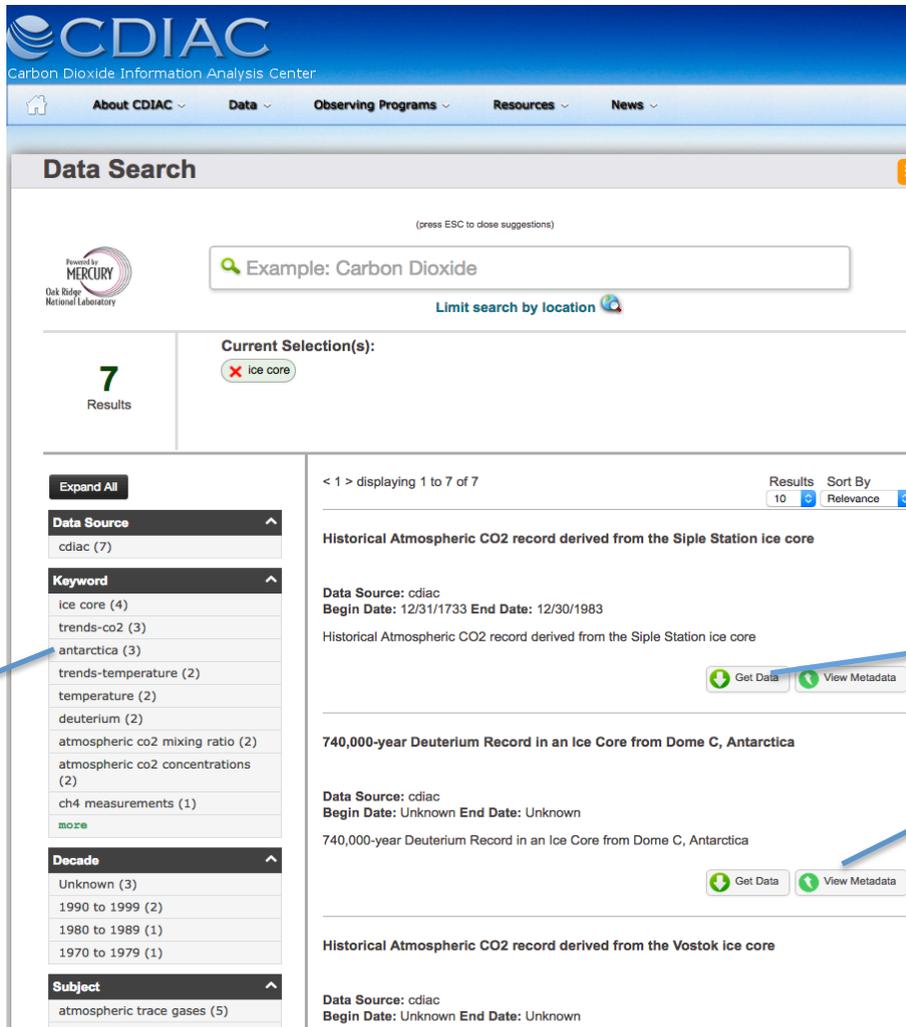
Supriya Ramireddy

Department of Computer Science, University of Georgia

Summary:

During my Appointment at the Oak Ridge National Laboratory (ORNL), I was part of Climate Change Science Institute (CCSI), in Environmental Sciences Division (ESD). As part of my summer internship my task was to develop a data search tool for Carbon dioxide Information and Analysis Data Center (CDIAC). CDIAC is a US Department of Energy's (DOE) data archive for climate change-related data and serves the information needs of users worldwide. CDIAC's data holdings include estimates of carbon dioxide emissions from fossil-fuel consumption and land-use changes; records of atmospheric concentrations of carbon dioxide and other radiatively active trace gases etc. Sharing data with colleagues and the broader scientific community is highly desirable. Long-term preservation of this data generally requires some form of institutional archive. Effective standards for data and metadata exchange are important factors for users to be able to find and retrieve data. The CDIAC's search interface based on Mercury [1] provides a single portal to very quickly search for data and information contained in disparate data management systems, each of which may use different metadata formats.

I am part of the Software Development Team in CDIAC. As mentioned, my task was to develop CDIAC's Data Search interface. The task was split into three main parts, Harvesting Metadata, Building the search index, and developing graphical user interface (GUI) to query the index. The technology used for indexing is Apache Solr 4.3, which is an open source enterprise search platform, written in Java. The indexes generated by the Solr are made available to the search GUI. The GUI then allows users to perform a variety of fielded, spatial, and temporal searches across the metadata sources which are some of the built-in features provided by Solr. The technologies I used in this regard are Java, Spring MVC, Hibernate, JavaScript, HTML/JSP Pages, Bash Shell Scripting, PERL and Apache Solr along with the application server Apache Tomcat 6. It was a challenge for me as I was working on several of these technologies for the first time but the CDIAC's development team helped me to get through issues and challenges. Fig. 1 below represents the CDIAC's Search interface.



Faceted navigation breaks up search results into multiple categories

Button to download Data

Button to view Metadata

Fig.1 CDIAC's Search Interface

Overall, I have had an incredible experience at ORNL. I got a chance to meet several scientists and engineers from different areas and learn about various domains. Everyone was friendly and ready to help irrespective of his or her busy schedule. It was a great opportunity to meet other interns from various top schools in US and around the world. Team was extremely helpful and provided me with enough guidance so that I can work on my own and learn. Definitely, this experience has improved me both personally and professionally. I would certainly consider ORNL as my future employer.

Acknowledgements:

ORNL Staff:

Mentor: Ranjeet Devarakonda

Team Member: Michael Crow

Group Leader: Tom A. Boden

References:

[1] Devarakonda, Ranjeet, et al. "Mercury: reusable metadata management, data discovery and access system." *Earth Science Informatics* 3.1-2 (2010): 87-94.