Enabling Flexible Search for ArcGIS Online Using Semantic Web Technologies

SATHYA PRASAD
Applied Research Architect
Applications Prototype Lab
Esri
Email: sprasad@esri.com

ArcGIS Online (http://www.arcgis.com/) is a Web platform developed by Environment System Research Institute (ESRI). As a worldwide geoportal, it enables GIS users from different countries to create, edit, and upload geo-data, Web maps, services, web applications and GIS tools, as well as to share these GIS resources with colleagues, friends, and the general public. ArcGIS Online resources (including all the maps, data, and resources) are called items, and they are accompanied by a rich set of metadata, including information about title, snippet, description, user, user groups, thumbnail image, creation date, modification date, comments, ratings, and so forth. Based on these metadata, users can browse through the collection of items, or sort them by features such as the popularity or themes.

In order to efficiently find the right items in the ocean of ArcGIS Online data, a RESTful Web API has been developed which enables search and retrieval of the items along with the users and groups. However, these existing APIs also restrict the way that the ArcGIS Online resources can be searched. For example, if the users would like to understand the information about both an item and the user who have created this item, then two queries have to be sent out separately to get the corresponding information. In some other cases, a user may want to know the information about the layers contained within a map, but such information cannot be easily searched simply because no corresponding API functions are there. These issues emerge because the developed APIs restricts the search capabilities, and therefore we cannot search the items when no corresponding APIs have been built even if there exist items for the users’ queries.

While it is always possible to expand the API functionalities by adding more code, users in different organizations and domains may easily come up with specific queries, which are beyond the expectation of the developers. Consequently, our goal in this position paper is to enable flexible search based on the existing ArcGIS Online search architecture. By “flexible search,” we mean that our system should allow users to input any queries without being restricted by the capabilities of the pre-designed APIs. While this work has been experimental with ArcGIS Online, the proposed methods can also be applied to other geoportals (e.g., Data.gov).

The solution we propose in this position paper is based on the Semantic Web technologies and the principles of Linked Data. Semantic Web is the third generation of the Web proposed by Tim Berners-Lee, the father of the World Wide Web. Semantic Web employs a set of standards for organization, publishing, and retrieving data from the Web. Resource Description Framework (RDF) is a standard data model used on Semantic Web to store data, and it organizes
data into subjects, predicates, and objects. The term “Linked Data” can be used to represent two meanings. On one side, it refers to the four standard principles to organize data, and on the other side, it can refer to the data that gets organized by following the four principles.

The idea we have in this paper is to organize ArcGIS Online metadata into Linked Data that can be represented as a graph of nodes and edges. Then, users can query items by entering this graph from any node and in any direction. Thus, instead of adding a huge number of API functionalities, we organize the data in a “smart” way, that can answer any query from the users as long as the corresponding items exist.

To test this idea, we experimented with a sample of ArcGIS Online data. Ontologies have been defined for items on ArcGIS Online, which can be primarily divided into classes of items, item types, users, and groups. The ontologies are not designed from scratch but based on the schema ArcGIS Online already has in the JSON response from the REST API. Therefore, developers who are familiar with the REST API can learn how to query the RDF data more quickly.

Using Java and Jena API, we have designed an application called “RDFConverter” to convert data from its original style to RDF. We then published these data following the Linked Data principles onto a triple store called Parliament. We choose Parliament instead of other triple stores as the SPARQL endpoint because it has implemented GeoSPARQL, which we are very interested to investigate. After publishing the data, we built a front-end web application which directly talks to the SPARQL endpoint and consumes RDF data. The implemented prototype portal supports the flexible queries without having to design corresponding API functions. For example, a quick SPARQL search can find out the owners of the maps that use a particular basemap. We can also find out the top 10 most popular basemaps based on their usage. In addition, we can search into the map and find information about layers (e.g., a Web map with layers about “population” and “disaster”). These are some of the examples, and all these queries can be complete within a single request.

To sum up, this work employs Semantic Web technologies for ArcGIS Online to enable a more flexible search. Developing “smart data” instead of only “smart applications” (i.e., adding many API functions) could be an important direction for enhancing spatial search.