The dataset provided here is used in the paper: Web Service Discovery based on Goal-Oriented Query Expansion, which has been submitted to Journal of Systems and Software.

Dataset description
i) We crawled the descriptive data of 13,520 Web APIs crawled from ProgrammableWeb (http://www.programmableweb.com/, PW for short) on July 25 2016. For each Web API, we collected ten kinds of attributes, which are described below:
* name: the "API Name" of the Web API on PW
* description: the textual "Description" of the Web API on PW
* pw_link: the relative link of the Web API on PW
* primary_category: the "Primary Category" of the Web API on PW
* secondary_categories: the "Secondary Categories" of the Web API on PW
* provider: the "API Provider" of the Web API on PW
* submitted: the "Submitted" date of the Web API on PW
* architectural_style: the "Architectural Style" of the Web API on PW
* supported_request_formats: the "Supported Request Formats" of the Web API on PW
* supported_response_formats: the "Supported Response Formats" of the Web API on PW

   Original Web APIs: 13,520 Web APIs.zip
   Categorized Web APIs: Categorized Web APIs (466 categories in total).zip
NOTE: Each Web API is categorized into each of its primary category and secondary categories. For example, the subfolder "Mapping (747)" contains 747 Web APIs that belong to category "Mapping".
NOTE: Each Web API is categorized into each of its primary category and secondary categories.

ii) Service goals extracted from the textual descriptions of Web APIs in seven experimental categories/domains: Extracted Service Goals of Web APIs in 7 Categories.zip
The subfolders and files in each category (e.g., Mapping (747)) are described as follows:
1) **ig-res**: contains the initial goals and their extension results of each Web API in the category
2) **csg-res**: contains the candidate service goals of each Web API in the category
3) **sg-res**: contains the refined service goals of each Web API in the category
4) **sgsta.txt**: contains the statistics of service goals extracted for all Web APIs in the category. For example, "create map -> 15" indicates that service goal (create, map, null, null) occurs 15 times in the category.

NOTE: The structure of service goal is defined as (sgv, sgn, $vsgp$, $nsgp$), where $vsgp$ and $nsgp$ contain the sets of prepositional phrases associated with sgv and sgn, respectively. In other words, sgp defined in the paper consists of $vsgp$ and $nsgp$.