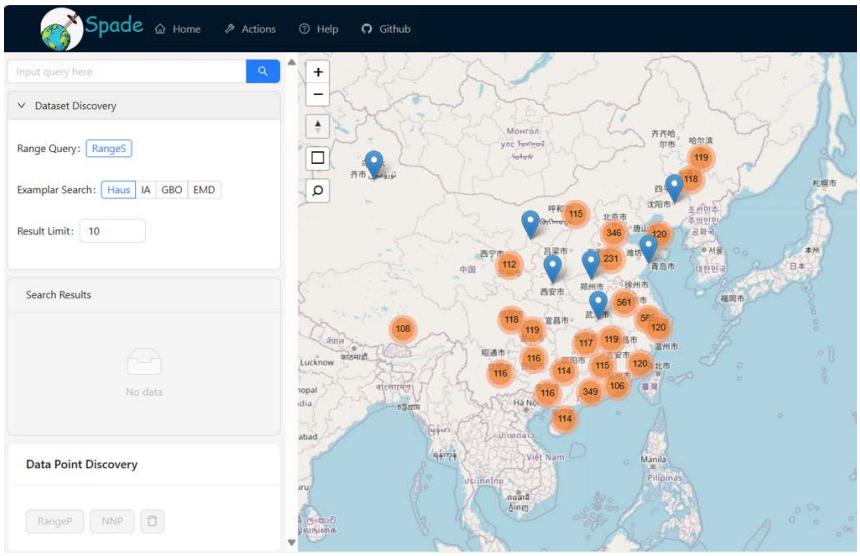
Spade Guideline

System Overview



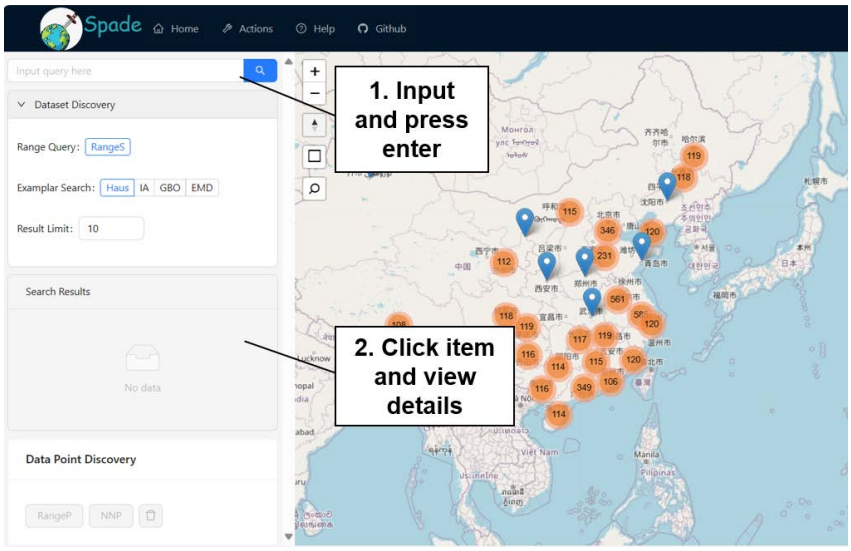
Instruction

- This is a demo system named Spade designed to support multi-granularity spatial data discovery, including coarse-grained dataset discovery based on multiple distance metrics and fine-grained data point discovery.
- In addition to basic keyword dataset search, the Spade system supports two types of dataset discovery queries (i.e., range-based dataset search and top-k exemplar dataset search) based on four distance measures (Haus, IA, GBO), and data point discovery operations (i.e., range-based data point search and nearest neighbor point search). Moreover, the Spade also implements the algorithm we proposed in Fast Dataset Search with Earth Mover's Distance. PVLDB 15(11) (2022).
- Users can easily explore the datasets they are interested in by operating the map component and the operation panel on the left. This system currently loaded the Argoverse 2 datasets in two cities, datasets from the New York City Open Data platform, a roadmap from an open street map organization, and POIs in China. Below is the user operational guide for the system functionalities.

Functions

- **Keyword Search**

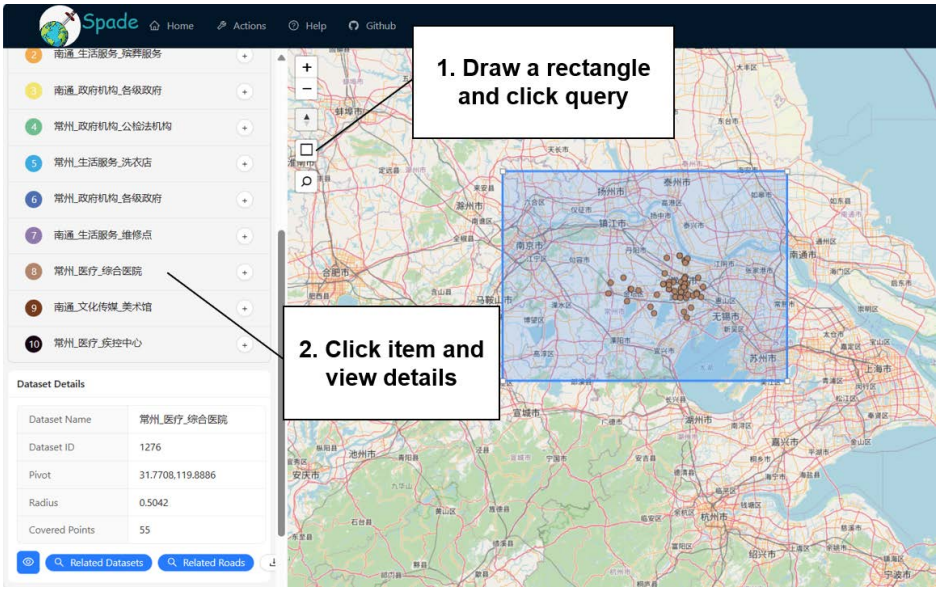
1. Input the keyword in the top query input box, click the magnifying glass button at the right, or press enter on the keyboard.
2. The top related datasets by dataset name are shown below. Click one result item to view the detailed information below and points in this dataset on the map.



● Dataset Discovery Query

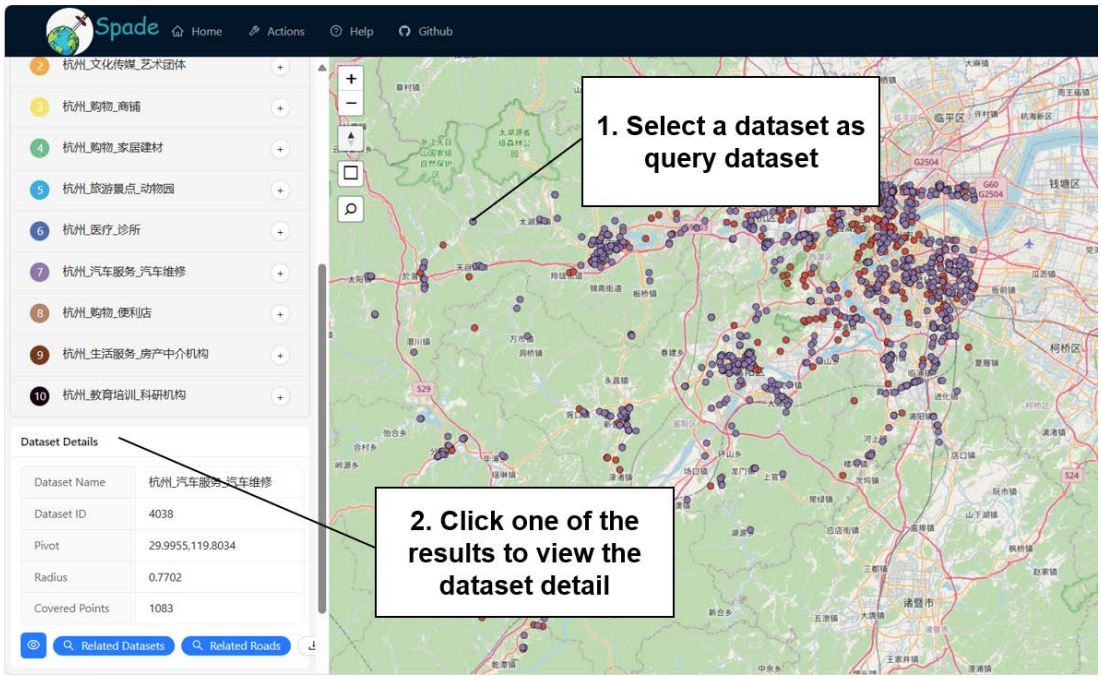
➢ Range-based Dataset Search

1. Click the rectangular control button in the top left corner of the map and drag the mouse to draw a rectangle on the map.
2. Click the magnifying glass control in the top left corner of the map to execute a range query. The query results are shown in the left control panel. Users can click the result item to view the dataset details.



➢ Top-k Examplar Dataset Search

1. For ease of usage for users without prior knowledge, Spade allows users to select a dataset on the map as a query dataset.
2. Select the distance measure on the left side of the system and set the k value.
3. Click the "Related Datasets" button below the "Dataset Details" to search for the datasets most related to the current dataset.
4. Click one of the results to view the dataset detail in the "Dataset Details" table and on the map.



● Data Point Discovery

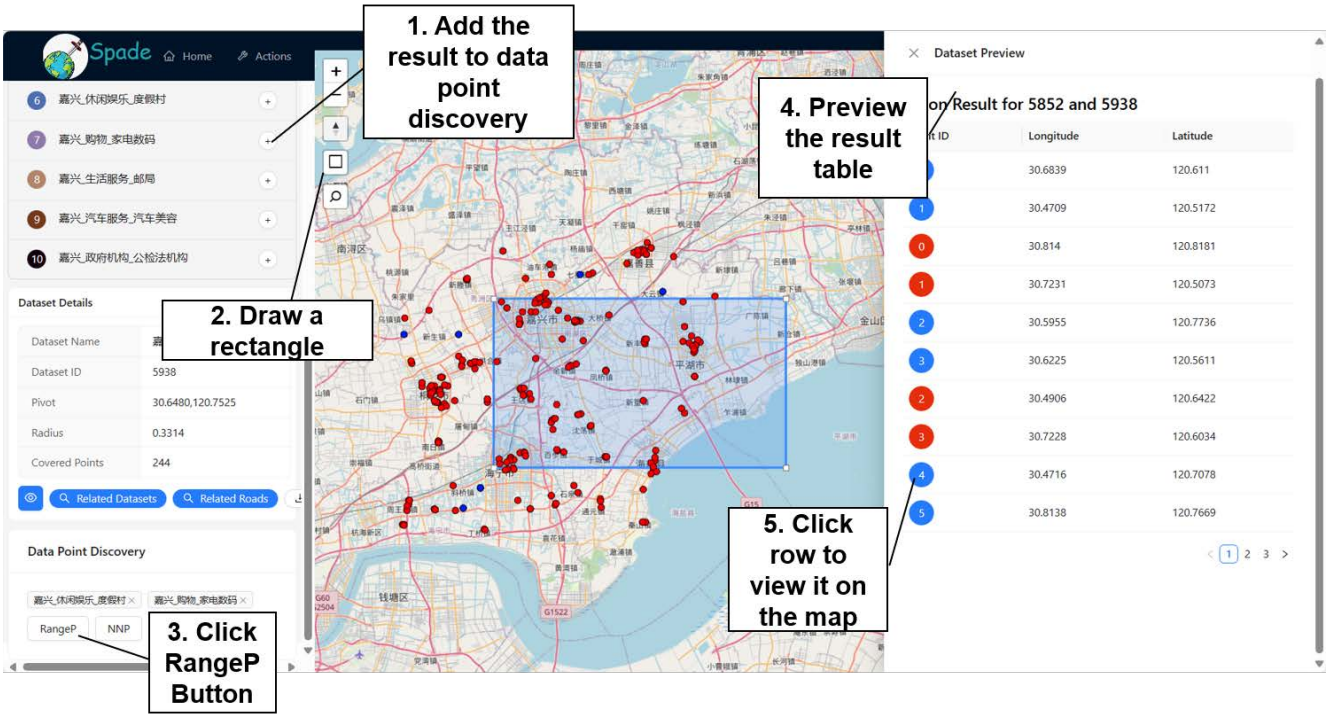
1. After finding the datasets the user is interested in from the above steps, click the "+" button next to the dataset name to add it to "Data Point Discovery".
2. Spade supports both range-based data point search and nearest neighbor point search on two datasets. Click the "RangeP" or "NNP" button to perform the data point discovery.
3. The preview table will appear from the right border. The points in the query dataset are displayed in blue, and the points in the related dataset are displayed in red.
4. Click each item from the preview table to view the point-wise result highlighted on the map.

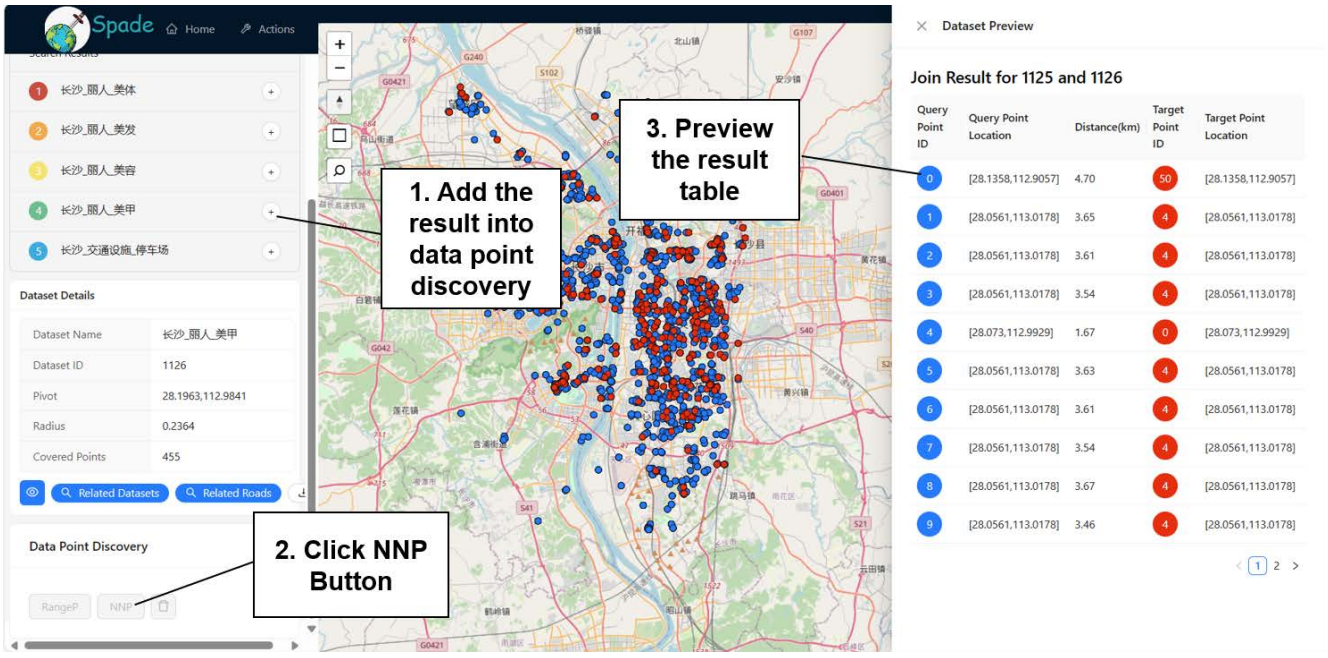
➢ Range-based Data Point Search

1. Select two datasets and click on the rectangular control in the top left corner of the map to draw a rectangular box on the map.
2. Click the "RangeP" button. The preview table will appear from the right border.

➢ Nearest Neighbor Point Search

1. Find datasets from the above steps and click the interested dataset to view the dataset details below and points on the map.
2. Click the "NNP" button below to perform the NNP operation. The preview table will appear from the right border.


1. Add the result to data point discovery
2. Draw a rectangle
3. Click RangeP Button
4. Preview the result table
5. Click row to view it on the map
Spade
Home
Actions
Dataset Details
Dataset Name
Dataset ID
5938
Pivot
30.6480,120.7525
Radius
0.3314
Covered Points
244
Related Datasets
Related Roads
Data Point Discovery
RangeP
NNP
Dataset Preview
Longitude
Latitude
30.6839
120.611
30.4709
120.5172
30.814
120.8181
30.7231
120.5073
30.5955
120.7736
30.6225
120.5611
30.4906
120.6422
30.7228
120.6034
30.4716
120.7078
30.8138
120.7669


1. Add the result into data point discovery
2. Click NNP Button
3. Preview the result table
Spade
Home
Actions
Dataset Details
Dataset Name
Dataset ID
1126
Pivot
28.1963,112.9841
Radius
0.2364
Covered Points
455
Related Datasets
Related Roads
Data Point Discovery
RangeP
NNP
Dataset Preview
Join Result for 1125 and 1126
Query Point ID
Query Point Location
Distance(km)
Target Point ID
Target Point Location
0 [28.1358,112.9057] 4.70 50 [28.1358,112.9057]
1 [28.0561,113.0178] 3.65 4 [28.0561,113.0178]
2 [28.0561,113.0178] 3.61 4 [28.0561,113.0178]
3 [28.0561,113.0178] 3.54 4 [28.0561,113.0178]
4 [28.073,112.9929] 1.67 0 [28.073,112.9929]
5 [28.0561,113.0178] 3.63 4 [28.0561,113.0178]
6 [28.0561,113.0178] 3.61 4 [28.0561,113.0178]
7 [28.0561,113.0178] 3.54 4 [28.0561,113.0178]
8 [28.0561,113.0178] 3.67 4 [28.0561,113.0178]
9 [28.0561,113.0178] 3.46 4 [28.0561,113.0178]