**Geospatial Analysis with Social & Web data**

# Basic Idea

The aim of this project will be to answer the following questions which are separated in 2 categories:

**Land**

* **What distinguishes a land?**
  + The most prominent aspect of the land.
  + What category does the land belong to (entertainment, rest, educational, business, …).
* **Who visited/mentioned a land?**
  + Location of the people who mentioned a land (geo-tagged social data).
  + Ranking of the areas where people mentioned a land (by number of mentions).

**People**

* **Where do people like to go?**
  + Ranking of places people in an area like to go to.
* **To what area these people belong to?**
  + Using geo-tagged social data to group people by their location.



Model

# Requirements

* **User Interface**
  + Easy to use user interface
  + Maps
  + Location Markers
  + Display Map/Locations Information
* **Data**
  + Social media data
  + Web data
  + Analysis of data
  + Categorization of data
  + Machine learning for clustering and classification
* **Development**
  + Dashboard with different statistical data and information
  + Data Extraction scripts
  + Web Application

# Possible Tools

* **Normal Web App**
  + Angular web framework for user interface
  + Some maps libraries such as leaflit.js
  + Python Django for server backend
  + NoSQL databases such as MongoDB and Postgress
* **ArcGIS & ArcGIS Online Web App**
  + ArcGIS Pro
  + Web App Builder
  + Geodatabase
  + Python Scripts
* **Data Extraction App**
  + Scrapping libraries for data extraction like (selenium, …)

# Similar Aspect Applications

* **Google Maps**
  + **Provides**
    - Maps
    - Location Data Visualization
    - Location Information
    - Location Search
  + **Lacks**
    - Geospatial Analysis
* **Google Places**
  + **Provides**
    - Detailed Location Information
    - Location Attributes (photos, reviews, ..)
    - Location Search
  + **Lacks**
    - Geospatial Analysis
* **Human Behavior On Social Media (By Human Dynamics For Mobile Age - HDMA)**
  + **Project**
    - Determine urban land-use patterns in **Beijing**, **China** using 9.5 million geotagged social media messages from social media platform **Sina-Weibo** for six months in the urban core areas of **Beijing** and compared them with 385,792 commercial points of interests (POI) from **Datatang**, a Chinese digital data content provider. To estimate urban land-use types and patterns, the team created a grid measuring 400 x 400 meters to divide the urban core areas into 18,492 cells.
  + **Methods & Tools Used**
    - Using K-means clustering algorithm, they identified seven types of land-use clusters in Beijing: residential areas, university dormitories, commercial areas, work areas, transportation hubs, and two types of mixed land-use areas.
    - Text mining, word clouds, and the distribution analysis of POI were used to verify the estimated land-use types successfully
    - Several computer programs to automatically or semi-automatically collect social media data from Twitter, Sina-Weibo, Google Places, and Reddit.
    - The data was saved in MongoDB, a NoSQL database.
    - The data was also exported from MongoDB to the ArcGIS platform to create point and kernel density maps, among other analyses.
    - The center has also created two software toolsets to analyze and display the data it collects from social media platforms.
      * The first is SMART Dashboard, a search tool for geotagged social media messages. It monitors and aggregates the dissemination of information related to changes in social behavior and provides insight into how a local population is responding to an event or situation.
      * The second is GeoViewer, a web-based mapping app that visualizes the results of the geotagged social media analyses performed by HDMA. Its geospatial functions are easy to use and include the ability to display hot spot and cluster data layers.
  + **Reference**
    - <https://www.esri.com/about/newsroom/arcnews/human-behavior-on-social-media-is-big-data-and-gis-makes-it-actionable/>

# Title