Statement of Intent – GTECH78803

The purpose of this document is to identify the background and other cursory information necessary for designing and building from scratch a reverse geocoder lookup for the New York City Department of Health and Mental Hygiene. This project will be used in part to fulfill the requirements of GTECH78803, the internship course for the Masters of science degree in Geoinformatics.

I am currently partaking in a part time internship position with the New York City Department of Health and Mental Hygiene (DOHMH) within the Information Technology division, in the bureau of Application Development and Database Administration. Within this bureau, the GIS administration for the agency handles tasks related to organization, maintenance, and updating of all servers, as well as maintaining all ArcGIS Online accounts related to the organization, all the while providing technical support to the few hundreds of employees about how to use GIS. While geographic tools are at their disposal, the DOHMH is in need of an application that relates latitude and longitude data to various political and social geometries that are defined, like United Hospital Funds and Community Districts. While there already exists an application that geocodes addresses from a CSV file to latitude/longitude points within the organization, enacting the reversal of this process is, however nonexistent. And it is also true that the Web App Builder for Portal within ArcGIS Online does this exact task, still the program is neither at the level of flexibility that the organization would like.

The goal of this project is to not only develop such a program from scratch, but also make it a flexible, easy to use, and secure internal application for the DOHMH. This program will then allow any user to upload a CSV file, relate the latitude and longitude points to the geometries of the user’s choosing, and returns a CSV file with the relevant information. In terms of security, all such processes will be done within the browser and saved on the cloud, thus allowing greater access to all users, while limiting access by saving said data onto any remote database. It goes without saying that almost all employees at the DOHMH handle sensitive data and information relating and referring to patients. This data cannot be allowed to be shared or handled without great urgency and care. Again, with the realization of this project, this application will process all data within the application, and therefore preserve the integrity and trust those patients entrust to us. In the form of website, this application will be developed that will be not only deployed, but also only accessible on the internal local server at DOHMH.

After having worked within the mainframe of DOHMH, after already utilizing many of the development servers and writing automation scripts of varying complexity, in it is well known the DOHMH contains the resources to build this program, especially in the computer language of Python. Now, given my year of programming in it, Python time and again showcases its functionality and flexibility. It is without doubt that Python is the ideal language for this task. In terms of the website coordinating with Python scripting, it will be developed using Flask: Flask is a micro web framework designed for simplicity both in compiling code and for users’ interface and is need of no specialized libraries or tools. As such, many of functionalities used in current web development, like form validation and upload handling, are not present in Flask’s library. Extensions, however, were developed to work with the core Flask framework. Now, while this can be a tedious process in the sense of needing to find and download all supplementary extensions, still there is a convince to this. Often, programs boggle down users with a plurality of tools and extensions. Using Flask for this project, the team would only download and utilize necessary extensions needed. This makes for higher efficiency in training software engineers, for they will be given all that they need, all the while keeping costs minimal.

The workflow for this application will be: users upload an Excel file to the browser and select which geometries they want referenced with their data; within the browser the file is then converted to a GeoDataFrame, while simultaneously the shapefiles that correspond to the geometries selected are uploaded as GeoDataFrames. After the uploading and conversation process is complete, this data is then into a list. Thereafter, a series of spatial joins are set in motion. The resulting GeoDataFrame will have have a column for each of the user-selected geometries where each value corresponds to the value in which the point intersects the polygon. This GeoDataFrame is converted to a regular DataFrame, and then exported as a CSV file for the user to download.

This internship project would be the culmination of my education within the Geoinformatics program. Much of my education has been focused on the more technical side of GIS (Free and Open Source GIS, Geocomputation 1 and 2, Spatial Databases) and having the opportunity to combine these aspects together and create a working reverse geocoder for the DOHMH would prove an ultimate learning experience. None of my in-school education has provided front end development for an easy user experience, and this project would allow me to dive head first into these aspects of development. . As the field of GIS moves toward development, I welcome the opportunity to embrace it, grow, and learn during the rest of my degree program as there is still much to be learned within development.