Cloud Engineering Supports Spatiotemporal Data Analysis

Jack McNulty, Junior, Computer Science
Jeffrey Oltmanns, MSSE, AF Lt. Col.

RESEARCH QUESTIONS
1. How is multi-variate unstructured data useful to humanitarian efforts?
2. Do cloud-based solutions exist to engineer, visualize, and assess big data?

Using Amazon Web Services and Etegent Technologies’ NTellus Earth View, we can layer multiple datasets in one view. Polish reception points are places that are caring for Ukrainian refugees. Border crossing points are ones that refugees will cross to leave the country to safety. We can then draw conclusions on which route they took to get there.

Using NTellus’ Distance feature, we were able to choose two points on the map and find the distance between them. We found that the distance between the reception point and the border crossing was 1.5km away. Humanitarian initiatives could use this information to more effectively search for injured refugees or plan for logistical deliveries and security. Considering it takes approximately 15 minutes for the average person to walk 1km, we could extrapolate this information and determine how long it should take to walk from border crossing to reception point.

Using NTellus’ Distance feature, we were able to choose two points on the map and find the distance between them. We found that the distance between the reception point and the border crossing was 1.5km away. Humanitarian initiatives could use this information to more effectively search for injured refugees or plan for logistical deliveries and security. Considering it takes approximately 15 minutes for the average person to walk 1km, we could extrapolate this information and determine how long it should take to walk from border crossing to reception point.

Using the Apache Spark Library, I was able to relationalize the data and change it from a non-standardized format, to comply with Spatiotemporal Asset Catalogue (STAC). Relationalizing the data makes it so that it is no longer stored in a nested format. As seen in the above code, I was able to rename and recast the data types for each field to put it in a format that works with our tools hosted in the Commercial Innovation Center (CIC). Doing this allows the data to be hosted in a non-relational database that can be accessed by various tools that support NoSQL queries and visualize the data. The primary purpose of this was to ensure that the latitude and longitude fields were of type double, as they were initially presented in type string.