Rebecca Lipton, American University
Armin Mehrabian, 610.2, Science and Exploration Mission Directorate

Constructing a Knowledge Graph & Applying Graph Algorithms to Draw Insights about GES-DISC Jira Tickets

May 10, 2 – 4 pm EST

ABSTRACT

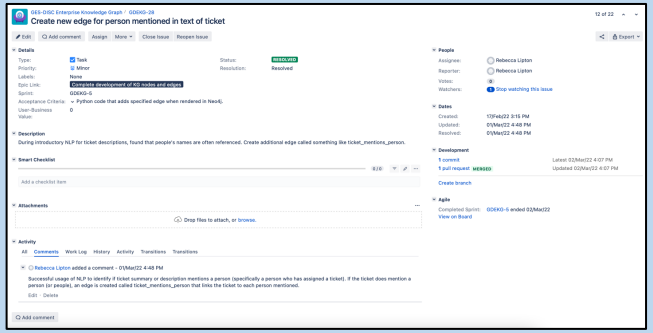
In order to assess the complexities of Jira tickets created by NASA Goddard Earth Sciences Data and Information Services Center (GES-DISC), it was beneficial to create a knowledge graph. The knowledge graph receives ticket data through the Jira API. The creation of a knowledge graph will help to answer high-level questions about internal structure, knowledge gaps, and team organization within GES-DISC. To work towards this goal, the knowledge graph was constructed in a dockerized Neo4j graph database. Once the graph had been created, graph algorithms were applied to answer high-level questions, such as exploring the role of staff in relation to projects, which qualities of a ticket contribute to the formation of communities within the graph, etc. To answer these questions, centrality and community detection algorithms were applied using Cypher querying language. The analysis of the results of the algorithms indicated that, as expected, certain individuals were more connected to some projects, while others were serving as hub nodes between two or more projects. Similarly, specific keywords are more likely to increase a Jira ticket's centrality in the graph. In terms of community detection, when tickets in a community have certain qualities, it is more probable for them to be grouped together. To best visualize which nodes had higher centrality scores or were grouped into certain communities, interactive graphs were created in Python using Plotly and Matplotlib. Ultimately, the project was successful in creating and deploying a knowledge graph to better understand the relationships between data in GES-DISC Jira tickets.

BACKGROUND

NASA Goddard Earth Sciences Data and Information Services Center (GES-DISC) utilizes Jira tickets to organize projects within the programming lifecycle. Adding the data from these tickets to a graph database allows for additional research to be done and insights to be drawn about the relationships between ticket properties.

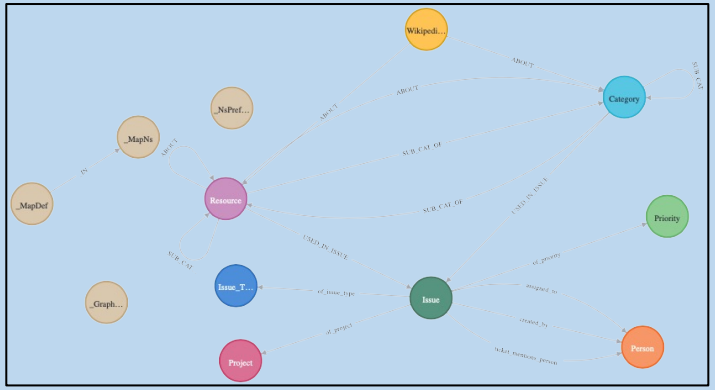
With that in mind, goals for this project included:

- Constructing and analyzing a knowledge graph of Jira tickets and personnel data for NASA GES-DISC to inform future team organization, knowledge gaps, and workload distribution
- Querying the finalized knowledge graph using Cypher and Python to draw connections between Jira tickets and personnel data
- Applying machine learning to derive and answer high-level questions about the knowledge graph

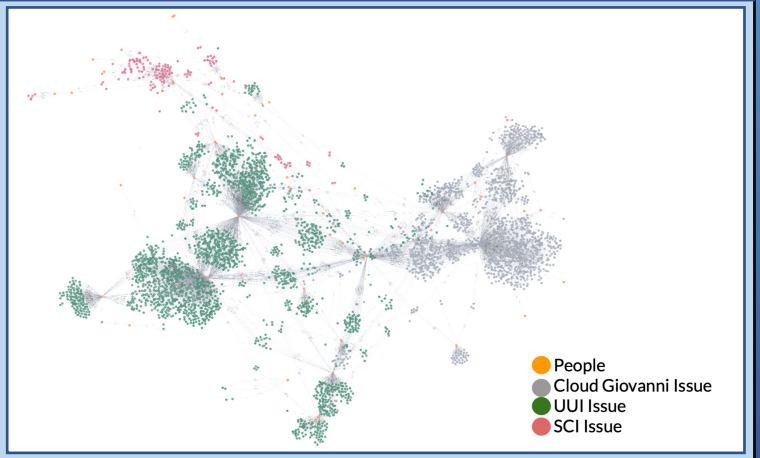


METHODS

- Employ Python, Neo4j, Py2neo, and Jira's API to create a knowledge graph of Jira tickets
- Design graph schema to support querying for anticipated user questions
- Conduct data cleaning using Pandas, NumPy, and natural language processing (NLP) in Python to build schema, nodes, and edges for a Neo4j graph
- Apply graph theory and graph algorithms using Cypher queries to derive and answer high-level questions about the knowledge graph
- Render visualizations in Python to better understand the output of graph algorithms



RESULTS

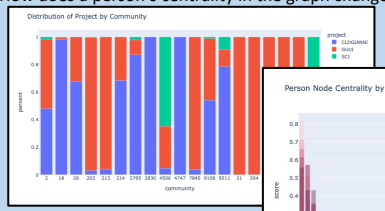


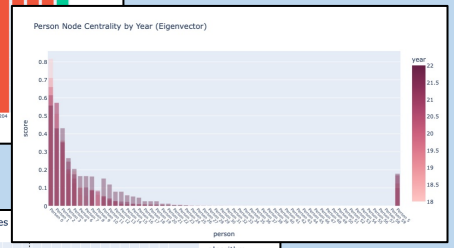
DISCUSSION & CONCLUSIONS

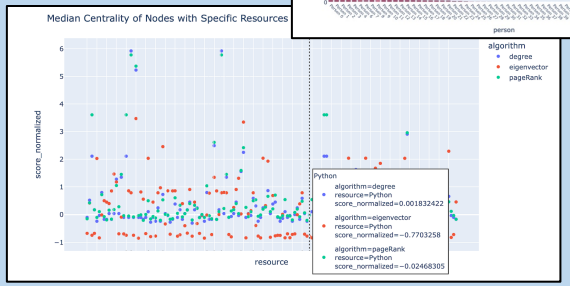
After rendering the knowledge graph successfully and applying graph algorithms to the knowledge graph in the form of Cypher queries, we are now able to draw insights from the graph.

At the beginning of the internship, some of the questions we were interested in assessing were:

1.) Are there any knowledge gaps?
2.) Does anyone work independently, and if so, who?
3.) Which individuals are more central in the graph?
4.) Can we identify which tickets more central individuals are associated with, and perhaps what skills they comprise?
5.) What skills or keywords determine a ticket's centrality level?
6.) How does a person's centrality in the graph change over time?







REFERENCES & ACKNOWLEDGEMENTS

- Armin Mehrabian: Thank you for being a fantastic mentor and supporting me in each of my learning endeavors.
- Fellow NASA Interns: Thank you for providing a sounding board for help with my project.
- NASA Goddard Internship Coordinators: Thank you for making this amazing opportunity a reality and guiding us throughout the way.