IDEA

Rensselaer **Institute for Data**
Exploration and Applications

The New York City Tree Explorer: An Immersive Exploration of Open Government Data

Sankalp Chauhan (Sonny), John S. Erickson and James A. Hendler
Summer Undergraduate Research Program (SURP)
Rensselaer Polytechnic Institute, Troy, New York,12180

Introduction

- Our goal was to create a compelling interactive application on the Rensselaer IDEA Campfire driven by geographically-linked open government data
- The data for this project is freely avaialble from NYC Open Data [1], a portal with "open data for all New Yorkers"
- We utilized basic principles of linked data to filter data and efficiently integrate data visualization modes
- We utilized **mwshiny** [4], an RPI-developed framework for multi-window web-based R applications, to create an interactive app for the Campfire
- Our application can be easily extended to visualize and explore datasets that link geographic locations with other information resources and data assets

Figure 1: Tree Census Data via NYC Open Data

Materials and methods

- **Primary data;** Street Tree Data from the New York City **TreesCount! 2015 Street Tree Census,** conducted by volunteers and staff organized by NYC Parks & Recreation and partner organizations.
- **Data preparation:** Sub-divided into five datasets by NYC borough (Brooklyn, Bronx, Staten Island, Manhattan, Queens); extracted variables considered important in the study incl. unique tree identifer, latitude, longitude, borough, health, latin and english names of tree, and zip code of tree
- **Species linking:** Utilized unique latin names trees to create valid Wikipedia URLs. These were used in our Campfire app to display species data on an External Monitor
- **Location linking:** Utilized lat/log to create valid Google Streetview URLs. These were used in our Campfire app to display street-level panoramic views on the Campfire "Wall"
- **Geographic Tree Health Visualization:** Location and health data from the dataset were used to display colored dots representing tree health on a NYC street map. The **leaflet** R package [5], based on Open Streetmap, was used to create the interactive map interface on the Campfire "Floor"
- **Summary of Tree Health:** A histogram comparing the health of trees for the selected borough has been prototyped using ggplot2 and will be integrated into the final Campfire app.

About the Rensselaer IDEA Campfire

- The NYC Tree Explorer runs on the Campfire, a multi-user, collaborative, immersive computing interface [2].
- Campfire is a desk-height, 10-foot panoramic screen (the Wall) and floor projection (Floor) that users gather around and look into, maintaining contact with one another with no artificial or virtual barriers between themselves as they observe and engage with presentations and applications.
- Two large monitors adjacent to the Campfire complement the integrated Wall and Floor visualizations with appropriate content, enabling investigators to be fully immersed in their exploratory tasks.

Results and Discussion

- The results of the work include a Multi-Window Shiny R app presented on the Rensselaer IDEA Campfire
- The Campfire "Floor" presents a map overlaid with datapoints; each datapoint is colored by the health of the tree as shown below in Figures 2 and 5
- **Interaction:** When the User clicks on a datapoint the following changes occur:
  1. A panoramic view of the neighborhood of the tree appears on the Wall, showing the local environment of the tree
  2. A Wikipedia page providing information about the species of tree
  3. A Controller enabling the User to filter map based on the health of tree, zip code of tree, and/or species of tree
- Example summary results for the borough of Manhattan: honeylocust is shown to have the highest number of healthy trees, followed by Callery pear, then ginkgo
- **Possible applications:** The significance of knowing which tree species are healthiest helps people know which trees are most likely to be planted and which trees people would be better off planting in order to increase the amount of vegetation and wildlife present in the area
- Also knowing which boroughs of New York City have the most and least amount of trees is also helpful as it informs people of those boroughs that it may be important to plant trees and which boroughs need more trees in their area

Figure 2: The Rensselaer IDEA Campfire (Overview)   **Figure 3: NYC Tree Explorer "Controller"**

Figure 4: NYC Tree Explorer "Wall" View: Street-level panorama via Google Streetview

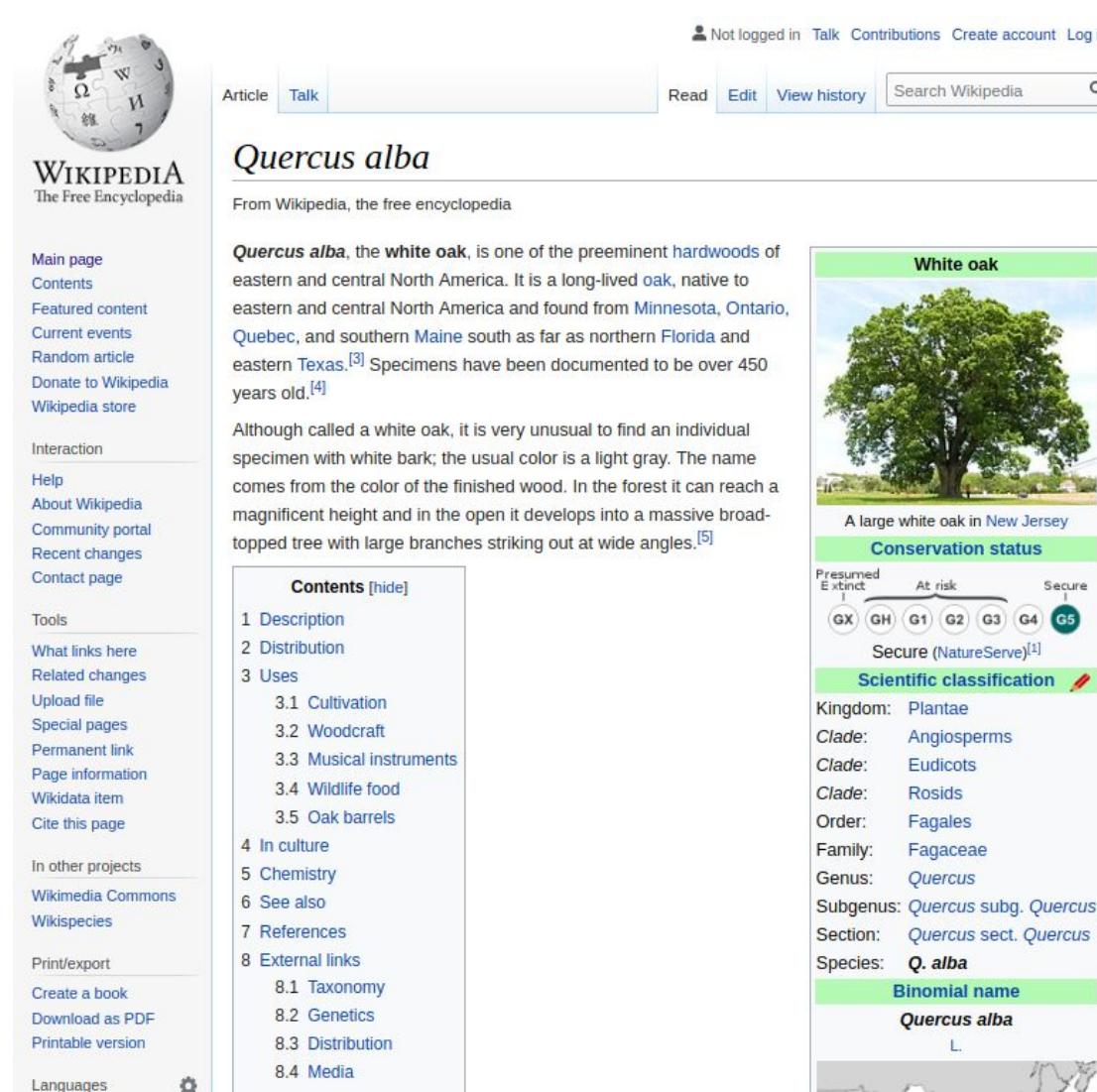**Figure 5: NYC Tree Explorer Campfire: Integrated "Wall" and "Floor"**

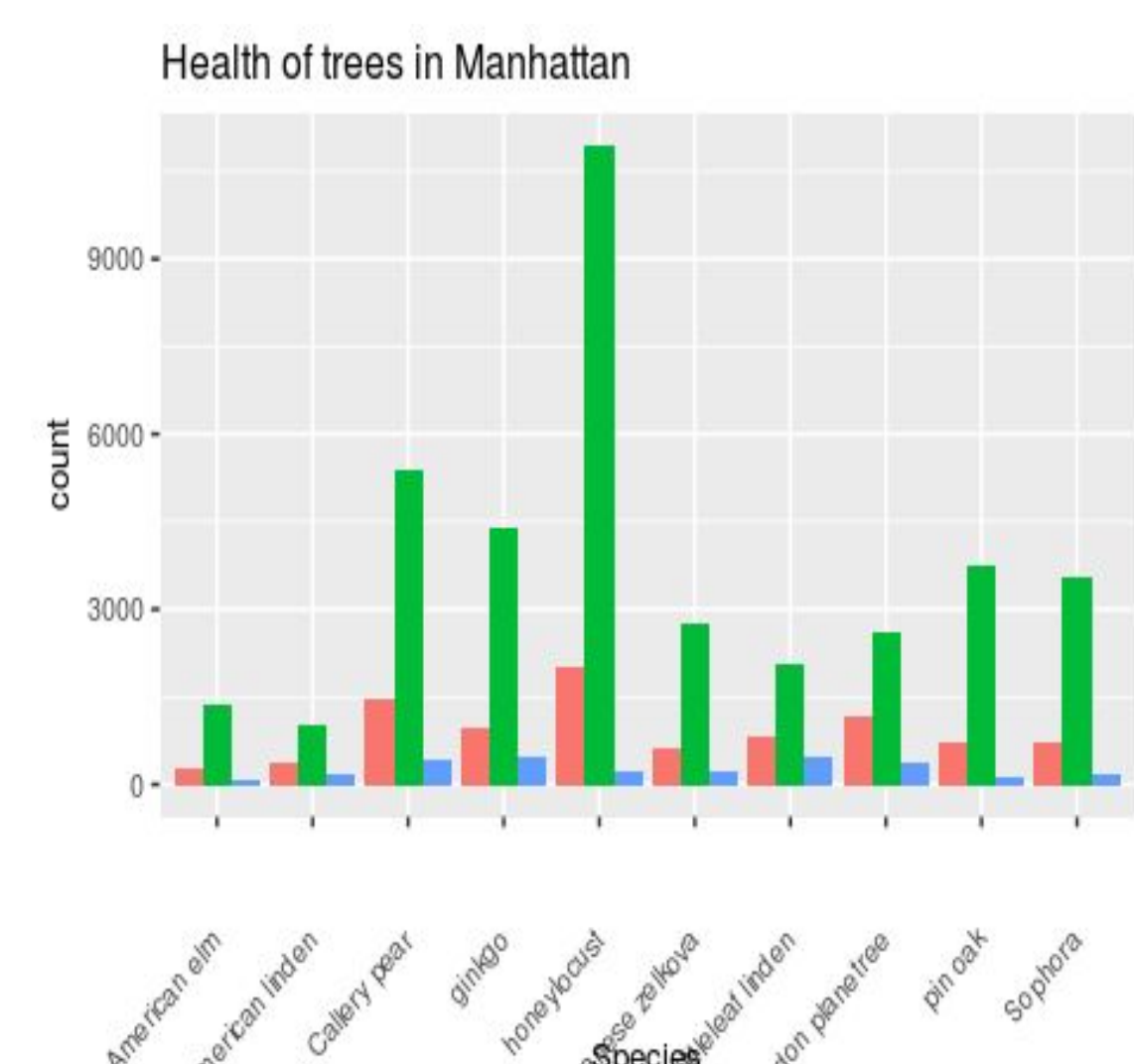**Figure 6: NYC Tree Explorer: Species Details (via Wikipedia)**

Figure 7: NYC Tree Explorer: Summary of Tree Health

Conclusions

- Open government data can be valuable as a basis for creating relevant and useful applications for communities
- Using open government data from NYC, we created an app that anybody from the city can understand. Our app can be used to help inform people about the conditions of the trees near where they live
- In this project we learned the power of linked data for simultaneously accessing and displaying as much information about the trees as we can. From the NYC we constructed URIs (uniform resource identifiers) to link to Wikipedia pages and physical location-based web APIs such as Google Street (the street-level panorama)
- The IDEA Campfire is a great environment for simultaneously showing different aspects of our application and data

Future work

1. Finish integrating Tree Health Summary histogram view into main Campfire app (External monitor)
2. Implement per-borough data loader (Controller)
3. More fully implement principles of Linked Data [6]
4. Experiment with other geographially-linked open datasets!

References

1. NYC Open Data Portal. https://data.cityofnewyork.us/
2. The Campfire: A novel multi-user, collaborative, immersive computing interface. http://bit.ly/empac_campfire
3. R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.
4. Hannah De los Santos, John Erickson and Kristin Bennett (2019). mwshiny: 'Shiny' for Multiple Windows. R package version 1.1.0. https://CRAN.R-project.org/package=mwshiny
5. Joe Cheng, Bhaskar Karambelkar and Yihui Xie (2018). leaflet: Create Interactive Web Maps with the JavaScript 'Leaflet' Library. R package version 2.0.2. https://CRAN.R-project.org/package=leaflet
6. Christian Bizer, Tom Heath and Tim Berners-Lee, (2009) Linked Data - the story so far. International Journal on Semantic Web and Information Systems, 5 (3), 1-22. (doi:10.4018/jswis.2009081901).

Acknowledgments