RendezView: An Interactive Visual Mining Tool for Discerning Flock Relationships in Social Media Data

Melissa Bica[1] and **Kyoung-Sook Kim**[2]

[1]*Department of Computer Science, University of Colorado, Boulder, CO*
[2]*Artificial Intelligence Research Center, AIST, Tsukuba, Japan*

Summary

RendezView is an **interactive** data visualization framework to show flock **patterns and relationships** in complex data, particularly data from social media platforms such as Twitter. “Flock” refers to the clustering of similar data. The multiple interactive visualizations allow users to assess many levels of relationships in their data and easily detect patterns that would not be evident in a database alone. It is built in HTML5, CSS, and JavaScript.

Motivation

To develop a web-based interface for visual data mining[3] to extend the functionality of the previous Sophy[4] framework :

- Implement multiple, interactive data visualizations: geo-spatial, time, & topic data.
- Improve user interaction coupled with data mining processes: data filtering, selection, aggregation, etc.

Data Visualizations

The primary visualization is a **spatiotemporal 3D map**, implemented using Three.js[5] and GeoJSON[2] data. Geospatial data is represented in the X-Y (red/blue) dimensions, while temporal data is represented in the Z (green) dimension. Each cube mapped on top represents a matching row from the database.

The **word cloud** shows the frequency of keywords and hashtags used in conjunction with the searched on keyword. It updates to include the aggregate of all word frequency information when multiple boxes are selected. This was implemented using D3.js[1].

The **Sankey flow diagram** consists of nodes and links, where nodes are keywords and links are geospatial-temporal intersections between keywords. The width of the links shows the flow quantity, which is the measure of the intersection of the boxes in the 3D map. This is based on a D3 Sankey plugin.

Use Cases

RendezView can be used to investigate different social phenomena over a geographic region, such as information flow in disasters and other crisis events, patterns in work and hiring, or trends in political discourse, by iteratively using results from each visualization to inform subsequent interactions with the data. As an example, a social scientist can compare the flock pattern of job postings between the east and west coasts in the US by searching on the keyword “work.” The researcher looks at which keywords have a similar pattern, such as “apply” and “hire.” The Sankey diagram informs the next keyword search based on strong spatiotemporal co-occurrence frequency with the original.

RendezView Interface

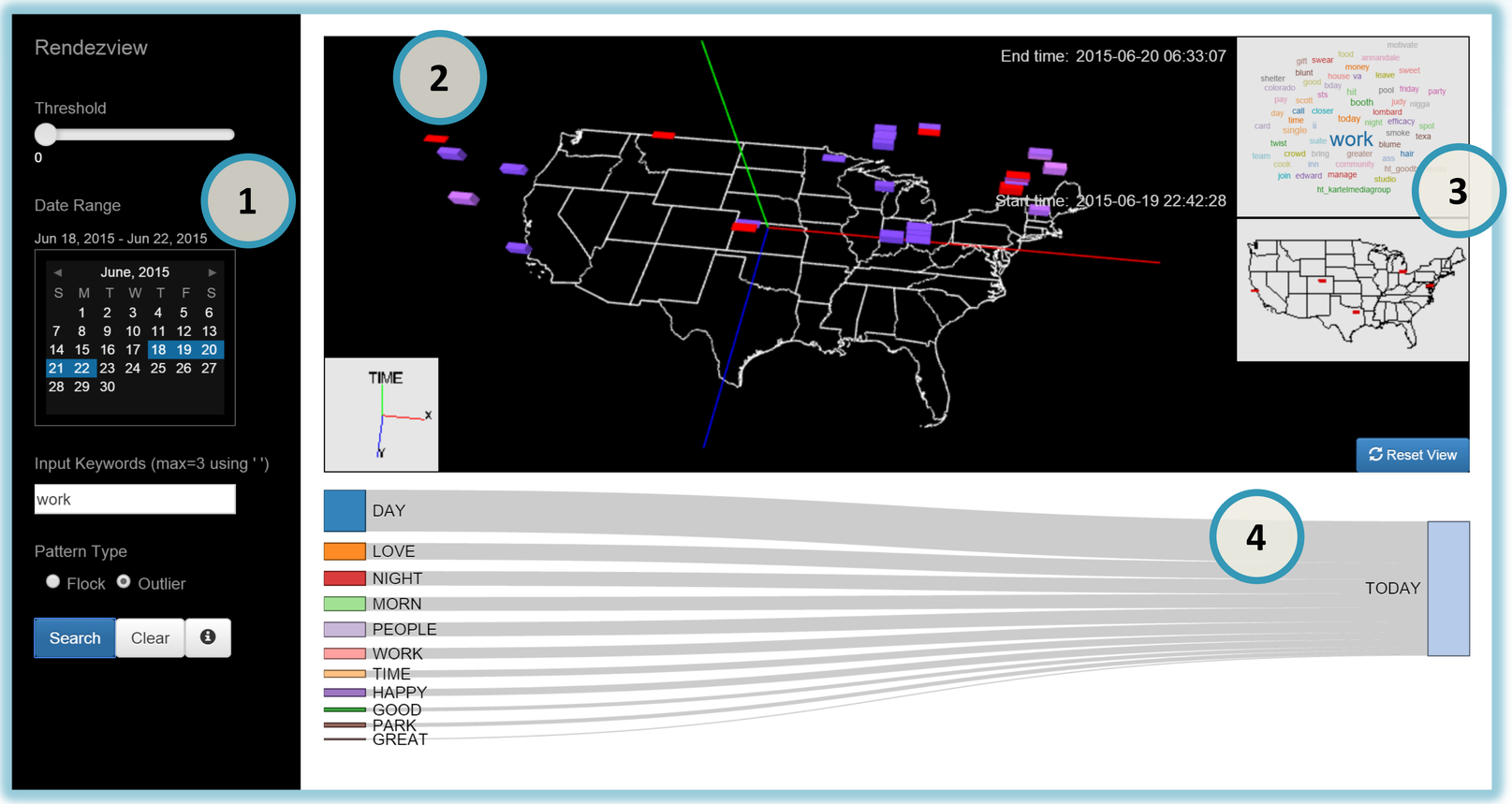

Figure 1. The above screenshot shows the UI layout, design, and functionalities of the RendezView framework. **(1)** The user inputs their options for **filtering** the data to display. **(2)** Database entries that match the selected filters are represented on the 3D map as **cubes**. Each cube’s position, dimensions, and color correspond to geospatial, temporal, and/or topic metadata. **(3)** Additional visualizations appear when boxes are clicked on. These include a **word cloud**, a 2D map to more clearly show geospatial dimensions, and labels displaying the time range of the selected data. **(4)** A Sankey diagram shows **flock relationships** between keywords. The user can click on links between keywords to display geospatial temporal intersections of those keywords on the 3D map.

Interactive Visual Data Mining

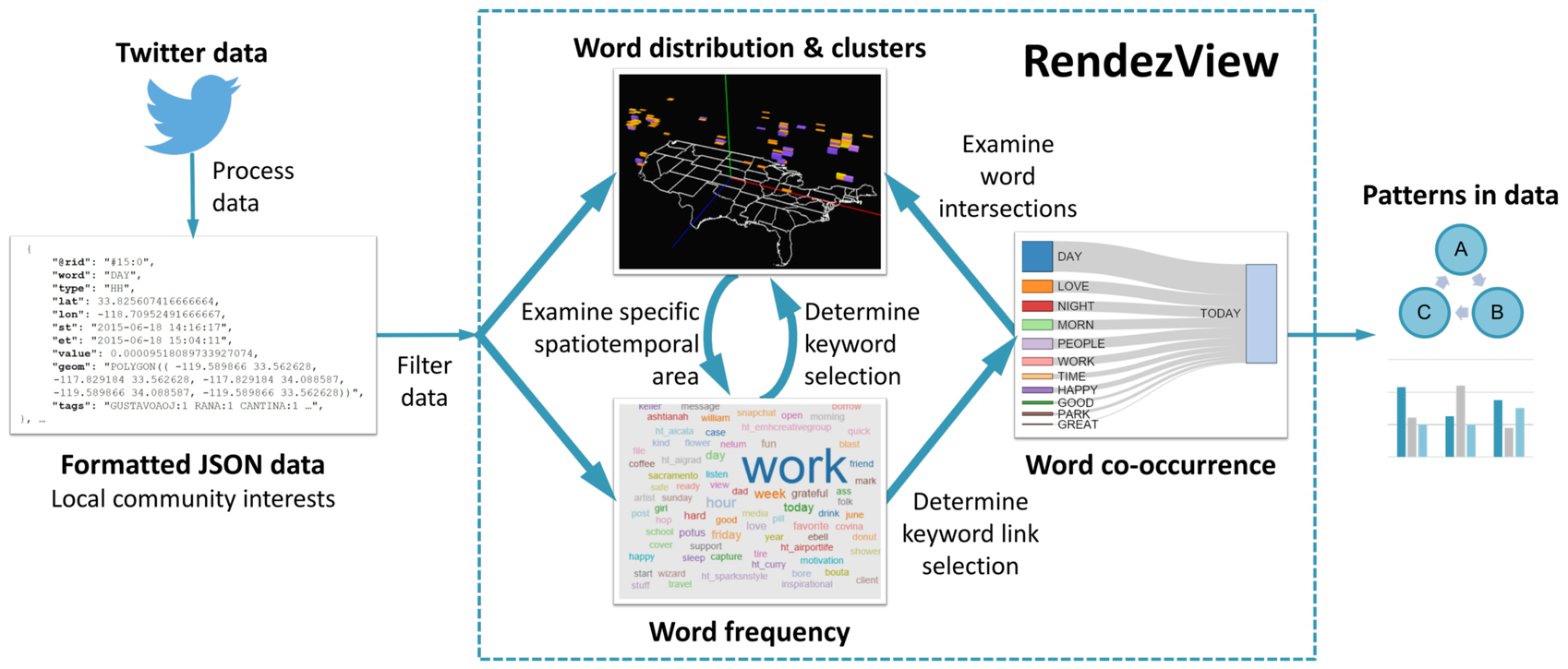

Figure 2. RendezView is a tool for *visual data mining*: the process of detecting patterns within big data using visualizations. This diagram shows how data is represented and the interactions between the visualizations.

Future Improvements

- Connection to live database.
- Sankey diagram functionality and interaction with 3D map, including option to view link widths according to various aggregation types.
- Front-end performance improvements.

References

[1] D3: http://d3js.org/
[2] GeoJSON: http://geojson.org/
[3] Keim D. Information Visualization and Visual Data Mining. 2002.
[4] KS Kim, H Ogawa, A Nakamura, I Kojima. Sophy: a Morphological Framework for Structuring Geo-referenced Social Media. 2014.
[5] ThreeJS: http://threejs.org/

Acknowledgments

The authors of this project thank the Partnership in International Research and Education (PIRE) program, the Open Science Data Cloud (OSDC), the National Institute for Advanced Industrial Science and Technology (AIST), and the NSF for their support. Additionally, we thank Dr. Bob Grossman, Dr. Maria Patterson, and Dr. Jason Haga for their guidance.