

Title of Presentation: (Interactive Exploration of LEHD, A Case Study in Knowledge Discovery)

Question your work tried to answer	Developing a tool for knowledge discovery by visualizing and exploring large and complex datasets such as LEHD
Local Employment Dynamics data sources used	<input type="checkbox"/> OnTheMap <input type="checkbox"/> QWI <input type="checkbox"/> Industry Focus <input checked="" type="checkbox"/> Raw data files from CD or VRDC <input type="checkbox"/> Other: _____
Other data sources used	
Software/ data processing tools used	JavaScript HTML Three.js (for WebGL) Source code is available at: https://github.com/nebeleh/EVA
Brief description of methodology (if someone wanted to do a similar analysis, how should they approach it?)	<ul style="list-style-type: none"> • The server side: includes data acquisition, data cleaning, data labeling, and transforming original LEHD files into compressed binary files accompanied by relevant metadata. • The client side: includes an Ajax interface for downloading the data from the server and putting it into memory as JavaScript Array Buffer. Then uses Three.js to transform this data into final values used in the GPU (e.g. X, Y, Z, color). HTML is used to provide interaction mediums such as sliders, etc.
Benefits of methodology/ data	Facilitates knowledge discovery by directly visualizing a large dataset and also through using intuitive and responsive navigation methods which represent different projections of the data
Drawbacks/problems with methodology/data	Current system requires a client/server architecture with access to the Internet. Also, further statistical analysis is required to confirm the observed hypotheses.
Anything else?	
Who and how to contact for more information:	Saman Amraii – samirpou@andrew.cmu.edu