Visualizing Data Structures in the Insight Engine: Information Processing Structure, Intelligence, Resilience

Hojung (Ashley) Kwon
114 S. Buchanan Blvd, Room A260, Bay 10, Durham, NC 27708. Duke University, USA
hk216@duke.edu

Bill Seaman
114 S. Buchanan Blvd, Room A260, Bay 10, Durham, NC 27708. Duke University, USA
bill.seaman@duke.edu

The Insight Engine uses an intelligent information processing structure in cooperation with a transdisciplinary database, a natural language API, and a series of visualization systems. The Insight Engine provides multiple avenues to data in different ways. This data visualization interface can also be personalized depending on users' preferences. The development of Micropeers enables a unique platform of collaboration.

The system works through intelligent information processing structures that are articulated through bissociation and poly-association. As mentioned in Arthur Koestler's Act of Creation, bisociation refers to the combination of two different entities into one. Poly-association, which is a term coined by Bill Seaman, associates multiple entities to enable intelligent linkages to be created across systems. These two concepts are crucial in creative thinking in many disciplines such as mathematics, physics, visual arts, poetics … etc. These processes are empowered through alternate interactive interfaces, which the user navigates through differing interactions with multiple modalities of the interface system, and are employed based on their preferences. Some of these are two-dimensionally oriented, while others are three-dimensionally oriented.

The central search avenue in the point cloud system are research papers. In terms of bisociation, the three axes in the space represent alignment among papers (x-axis), author relationalities (y-axis), and publication date (z-axis). Each point in the 3D space is a research paper, meanwhile in a system where many point clouds are overlaid, each set of intersecting points represent associations among papers, authors, or concepts discussed in the papers. On each point, along with the general information about it represented with the three axes, other information that the paper represents such as its keywords are potentially overlaid on top of it as multidimensional vectors. Depending on how much information about the paper we want to represent with the point, the number of dimensions of the vectors can vary. Users will be able to search for specific papers using these numerical intersection points related to a high-dimensional numerical intersection of the content such as cosine distance. This introduces a different way of visualizing intersections among research papers from the other two methods discussed in later parts of this paper.

Along with the 3D point cloud, a knowledge tree can be a way to represent associations among authors. Each author can be represented as a tree, with multiple papers making up the trunk and branches. The tree would be composed of the author's research papers and/or research papers that the author searched for and collected. Earlier papers are located near the bottom of the tree, composing its root, and more recent papers are located near the top of the tree, composing its leaves. Information about each author will be updated over a set time period so that the trees can "grow" by receiving and adding new publications about authors they represent. Users will be able to see this information by hovering the cursor over the tree. The interface of the Insight Engine will look like a forest composed of multiple trees. Text

will also apprar at the bottom of the interface in a more traditional manner. The distances among trees will be determined based on information gathered about authors that are represented in the Insight Engine; the trees of authors whose works are closely related to each others' in terms of research topics or authors who collaborated on the same paper will be located closer to each other. Users will be able to change categories based on where the trees are distributed in the forest and reorganize the trees into different formations. Also, there will be bridges between trees that represent collaborations among authors, bisociations and poly-associations. At the end of each bridge will be an element that composes a tree, namely a research paper. This connection can represent different natures of collaborations among authors; it may represent the research topic of one paper being inspired by that of the other or two papers discussing the same issue. Overall, the visualization will look like a forest with varying distances between trees and bridges connecting trees. Users will be able to search for specific authors and papers or multiple authors and papers to visualize associations among them.

Another method of visualization that involves a 3D space is through a wheel structure, which borrows its inspiration from Giulio Camillo's Memory Theater. The wheel is divided into different sections, each representing an academic discipline as well as potential 3D models, still images, videos, and sound. Each section will be composed of research papers that are relevant to the topic it represents. Users will be able to change the order of papers loaded into each section based on their interests. By default, in a series of concentric rings visible from the center of the wheel, papers that are more strongly related to topics represented by sections of the wheel will be closer toward the center.

The section will vertically extend toward the circumference of the wheel and include information about authors and research papers relevant to the discipline. As in the two other models described above, users can either search for an individual author or a paper or associations among multiple authors and papers. Associations among those elements composing each section of the wheel will be represented as lines connecting the sections. By hovering over a part of an association, users will also be able to see an automatically generated summary of it such as how the papers that are linked in the association are related to each other. These functionalities can be empowered by Natural Language Processing (NLP) and knowledge graph building APIs like Diffbot, which already have the functionality of extracting association among entities in a graph using NLP. In the end, the visualization will look like a spider web with vertical lines going out from the center of the wheel with horizontal lines connecting different parts of the vertical lines.