

Apolo

Interactive Graph Exploration: Machine Learning + Visualization

Polo Chau (GT)

Niki Kittur (CMU)

Jason Hong (CMU)

Christos Faloutsos (CMU)



Carnegie Mellon

SAIC



Summary

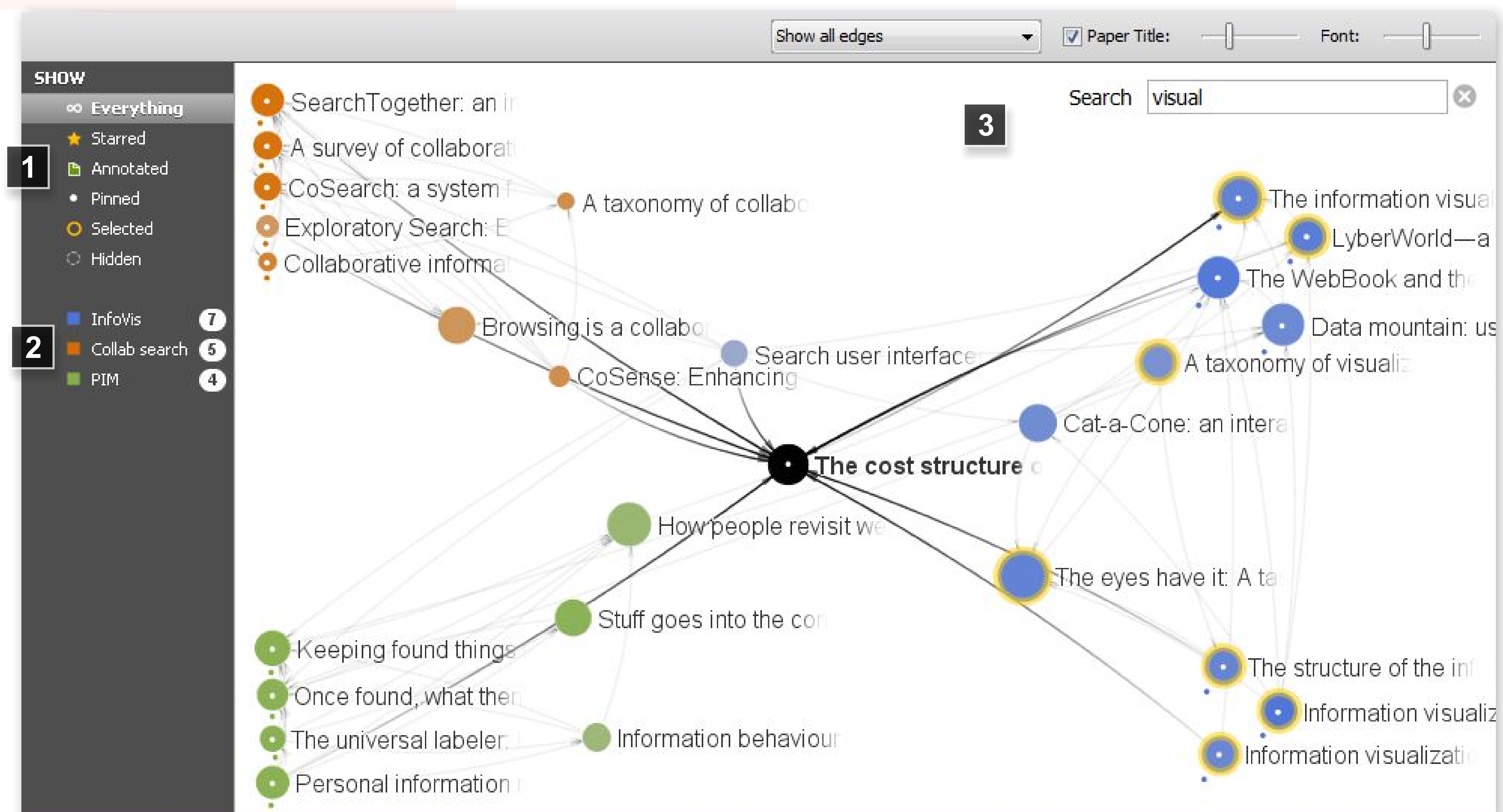
Apolo is a system that uses a mixed-initiative approach — combining **visualization**, rich user **interaction** and **machine learning** — to guide the user to *incrementally* and *interactively* explore large million-node graphs in real time. Here, we use a *citation network* as the example data.

Find Relevant Nodes with Exemplars.

The user specifies nodes of interest as exemplars, and Apolo computes other nodes' relevance, using the *Belief Propagation* machine learning algorithm that "spreads" relevance from the exemplars. The more relevance a node receives, the more relevant it is.

Create & Modify Groups.

Organize your representation about the data visually and interactively. Apolo supports groups, and you can spatially arrange them any way you want. What's more — each group can have its own exemplars. They enable Apolo to find relevant nodes for all groups at the same time.



The Apolo user interface displaying citation data around the article *The Cost Structure of Sensemaking*. The user has created three groups: **Information Visualization** (InfoVis), **Collaborative Search**, and **Personal Information Management** (PIM). A color is automatically assigned to each group. A node is assigned the color of the group it most likely belongs to; saturation indicates "belongness". Each exemplar has a color dot below it. **1**: The Filter Panel provides filters to control the type of nodes to show; the default filter is "Everything", showing all kinds of nodes, except hidden ones. **2**: The Group Panel lets the user create, rename, and delete groups. **3**: The Visualization Panel where the user incrementally and interactively builds up an understanding and a highly personalized visualization of the network. Articles whose titles contains "visual" are highlighted with yellow halos.

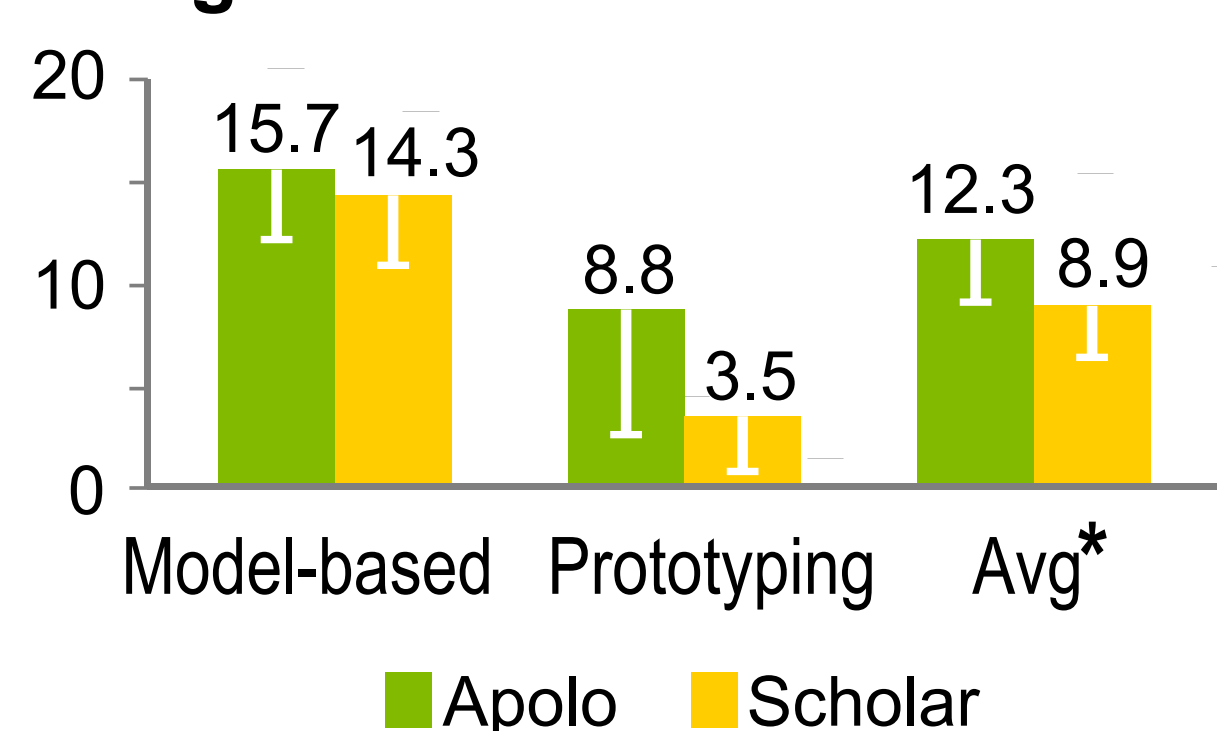
Cognitive, Visual & Algorithmic Scalability.

Visualizing the full million-node graph is rarely useful. Apolo finds and visualizes a small set of the most relevant nodes and edges. When you are ready for more, ask Apolo to show it. Apolo's scalable algorithm processes millions of nodes and edges in real time.

Evaluation

In a between-subjects study, 12 participants used Apolo and Google Scholar to find relevant new papers to update two sections of an existing survey paper. Two expert judges rated papers found by Apolo to be statistically significantly more relevant. Subjective feedback of Apolo was also positive.

Judges' Scores for Each Section



In Making Sense of The Literature, Which Software Seemed More ...

