



Tech

## This all-knowing web tool will revolutionize the way students research

By Selena Larson

Jan 5, 2016, 2:24pm CT | Last updated Jan 5, 2016, 2:25pm CT

By: Selena Larson – January 5, 2016

As any college student can confirm, research papers are a colossal pain. When you start researching a topic, there's one tool everyone turns to first: Google. But Google serves up links upon links of information that may or may not be relevant to what you're looking for.

San Francisco-based startup Omnia wants to make it possible to both understand and visualize connections between documents and topics, whether or not they link to one another online. The company unveiled its semantic search engine at CES on Monday, and introduced a new way of searching publicly available data.

The platform searches meet ups, patents, Wikipedia, research papers, company websites, contact information and a slew of other data, and matches shared keywords to documents or topics your searching.

Unlike a traditional search system, which scours for keywords and headlines that seem to fit the words or phrase in the search bar, Omnia discovers related content that might not be linked, but draws connections between words or phrases that appear in content across the Web.

The screenshot displays the Omnity search interface. At the top, there is a search bar with the text "convolutional neural networks" and a "Discover" button. Below the search bar, the interface is divided into three main sections:

- Left Panel:** A list of search results. The top result is "Dynamically Configurable, Multi-Ported Co-Processor For Convolutional Neural Networks" with a patent number "Patent: US9442927". Below it are several other results, including "Cellular neural network", "Method And Apparatus For Using Convolutional Neural Networks In Speech Recognition", "Convolutional neural network", "Massively Parallel Processing Core With Plural Chains Of Processing Elements And Respective Smart Memory Storing Select Data Received From Each Chain", "Parallel computing", "Deep learning", "Stream processing", and "Artificial neural network".
- Center Panel:** A network visualization showing a central cluster of green nodes connected to a larger network of grey nodes. The nodes are connected by lines, representing relationships between documents.
- Right Panel:** A detailed view of the top search result. It includes the title "Dynamically Configurable, Multi-Ported Co-Processor For Convolutional Neural Networks", the date "January 31, 2010", the document type "Patent", the organization "NEC Laboratories America, Inc.", and an abstract. Below the abstract is a word cloud with terms like "kernel", "weights", "convolutional", "parallel", "processing", "memory", "architecture", "element", "parallelism", "CNN", "neural", "block", "convolver", "layer", "network", "instructions", "processor", "outputs", and "hardware".

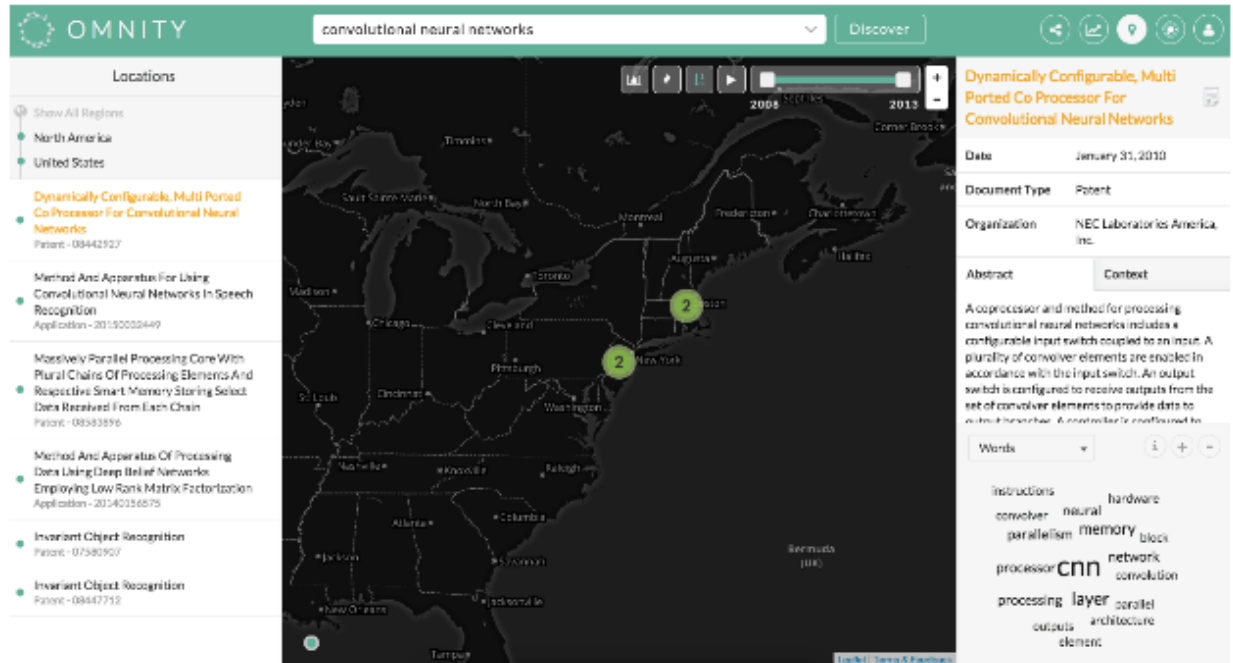
The Omnity logo is visible at the bottom center of the interface.

As CEO and co-founder Brian Sager explained to the Daily Dot, the algorithm takes all the rare shared words and creates an equation that filters through documents and publicly available data. It avoids the need to have keyword-based search or a taxonomy system that ranks websites in a list based on keywords or SEO.

The more rare the word, the higher it counts towards the page relationship rank. The software omits common words that are found in a variety of documents or papers and only searches those that have multiple shared words in common. Omnity then surfaces relevant documents or websites and literally draws connections between data that's related.

You're given a list of results that include primary documents labeled in green that are directly related to the search terms, and secondary documents in grey that share vocabulary with the primary document they nest under. A visualization shows how they're all connected.

The right-hand column includes a preview of your search result, where it came from, and a word cloud that shows the most popular words weighted by size according to their prevalence within the document. To see which documents contain a specific word, you can click on it to pare down your search results.



It's also possible to search entire documents—the tool will pull out keywords that appear more than others and use it as the basis for the search. For instance, Sager told me that when you input the text of a biography of Steve Jobs, Jobs's patents are among the data that will appear in search results.

While most of the search results made sense with a handful of queries I tested, they weren't all accurate. I tested the software on my story about coding programs in San Quentin State Prison, and it did show data regarding the prison and learn-to-code programs, but it also confused the last name "Houston" with the city in Texas and displayed information about programming meet ups in the area.

Omnity's free tier searches public data, so you'll see a lot of patents, Wikipedia pages, and group meet ups. A free student platform also searches open coursework and other research provided free to students and allows them to download three research reports per day. Paid tiers go deeper, and search for information that's found behind paywalls like case law and scientific research papers with subscriptions that cost up to \$400 per month and will be launching soon.

Sager and his team have worked on Omnity in stealth for four years, and finally launched publicly this week. The company is supported by grants from the National Science Foundation and has received an undisclosed amount of funding from private investors.

While education is an obvious market, Sager said enterprise uses could span across industries and include finance, medicine, law, and product development.

*Photo via CollegeDegrees360/Flickr (CC BY-SA 2.0)*