



IO Informatics Kicks Off Beta for Sentient 3.0 Software, Unveils PICC Parkinson's Partnership

Nov 19, 2015 | [Uduak Grace Thomas](#)

NEW YORK (GenomeWeb) – IO Informatics has begun beta testing a new iteration of Sentient suite, its enterprise software product that uses semantic web technology to aggregate and analyze data from multiple sources, ahead of a formal launch planned for the first quarter of 2016

Robert Stanley, IO Informatics' president and CEO, told GenomeWeb that a number of unnamed pharmaceutical companies began testing version 3.0 of the company's software in November as part of the beta and are providing feedback in terms of the usability and feel of the platform.

In addition to existing data integration and query features, the planned release features a newly redesigned software architecture that uses frameworks such as Angular JS and Apache Spark and microservices, offers new programming interfaces to the company's core technology, and also incorporates a data integration workflow complete with data modeling and integration tools. The release will also include a new so-called data decisions page, through which customers can request changes to their data and improvements to the user interface that make it more attractive.

Sentient can be accessed remotely but most customers prefer local installations of the software because of concerns about sensitive patient data, Stanley said. Pricing for the company's platform varies and can be more or less expensive depending on what sort of data customers want to integrate and what sort of services they require, among other factors. Prices range from under \$10,000 for a small installation to several millions of dollars for an enterprise solution that provides automated data harmonization and supports unified queries across a large institution's internal and external data resources — these costs cover local software installation or cloud-based hosting as well as any associated licenses and services.

Earlier this month, IO Informatics announced a partnership with the Parkinson's Institute and Clinical Center (PICC) to implement a customized version of the Sentient technology — branded as Parkinson's Insight — locally that integrates and aggregates data from various research and clinical studies focused on Parkinson's disease. Carolee Barlow, PICC's CEO, told GenomeWeb that the institute tapped IO's platform to deal with the daunting task of manually aggregating data collected from about 10,000 individuals who have participated in different studies conducted by the institute in the 25 years of its existence.

"We have this enormous wealth of information about each and every patient that's been at the institute, and yet we can't look at all of it to find new insights that could help us find a cure for this disease or find new insights about [a] particular patient which would help us do a better job taking care of [them]," she said.

Besides research into Parkinson's disease mechanisms and drug targets, the institute also runs clinical trials in partnership with other institutes and pharmaceutical companies, and it designs and runs its own internal clinical studies. All of that research and clinical study data is gathered and stored in different databases and formats, which makes it difficult to query across datasets, Barlow said. A clinical patient, for example, might participate in a drug trial, contribute data to a research study, and also donate skin cells for an induced pluripotent stem cell project all at PICC.

"That patient's got a lot of data about themselves spread throughout our entire institution, but nobody can go and get all of that information and put it right in front of them [and] say 'what is this patient teaching us about Parkinson's?'" she said. Adding further fuel to the fire is the fact that data from one clinical study may have been collected in a completely different format from a similar study done 10 years later.

Moreover, there are entirely different software packages used now that weren't available a decade ago and, similarly, some software tools that existed 10 years ago have gone the way of the dodo. For example, over the last 25 years, PICC has used two electronic medical record systems: MediNotes, which no longer exists, and eClinicalWorks, which it currently uses. Both of those tools collect and format data in different ways making it difficult to integrate the information from the two systems, Barlow said.

Parkinson's Insight uses Sentient's core semantic technology to aggregate and query information related to things like disease presentation and progression; image data; treatment information; genetic, protein, and metabolic information; and EMR data. Data is aggregated in such a way within the system that new datasets can be added on to existing integrations without breaking existing links between datasets, an improvement on traditional data warehouse technologies, which may not be as accommodating when new datasets need to be integrated with older ones, according to Stanley.

IO's semantics approach to data integration also sidesteps problems faced by existing approaches to data integration that rely on applying thesauri, vocabularies, and standards to connect siloed data or use lexical matching algorithms to try to match corresponding data elements, he said. But without the semantic layer, "you run into a lot of big problems." So, for example, a user trying to integrate a series of databases might find multiple matches for the letter K including Potassium, kilogram, or Kaposi sarcoma.

"You need to have the semantic layer so that you can say if it's related to something that says it's an element and a dietary supplement, then you know its potassium," Stanley explained. "By applying that semantic layer, which lets you do reasoning on relationships, you can do a much better job with integration than traditional methods."

PICC's iteration of the system has so far been used to integrate data collected as part of 70 out of about 100 Parkinson's-related studies and databases, and there are plans to integrate the remaining datasets moving forward — the reason for integrating only a subset of the data was to test Sentient's aggregation abilities, Barlow said. Specifically, the researchers worked to integrate data from patients that had submitted tissue samples, such as brain tissue, iPSCs, fibroblasts, and blood, for analysis as part of several different studies — the remaining datasets to be integrated come from patients that haven't donated tissue samples.

The information that has been aggregated so far has yielded some interesting insights that could have implications for diagnosing and treating Parkinson's, according to Barlow. By integrating the data, the researchers were able to connect tissues and samples to their original owners.

"Somebody could have donated their DNA as part of a study 15 years ago, and then when they passed away they donated their brain," she explained. Those bits of information were stored in separate databases that did not talk to each other.

"When we did the integration, we were able to identify how many patients we have that have various levels of specimens and how many of those we have brains from in addition to the clinical information," she said. Further analysis of the data revealed specific cases where patients presented with Parkinson's-like symptoms but did not actually have the disease.

"We realized that there were patients who were being [diagnosed with] Parkinson's disease, but when you looked at their neuropathology, there were no Lewy bodies" — abnormal protein aggregates that develop in nerve cells in Parkinson's cases, she said. Moreover, there are some genetic diseases that show similar symptoms to Parkinson's.

"If we were trying to study [changes in] those genes to look for potential therapies for Parkinson's disease, they wouldn't work because those genes, when they are abnormal, don't cause Parkinson's, so we would have been looking in the wrong spot for treatment," she said.

IO also developed some custom applications for PICC's iteration of the platform including capabilities for visualizing and exploring imaging data, Barlow said. They are currently working on ensuring that PICC's system complies with requisite industry standards for maintaining patient anonymity and security, she added. Moving forward, in addition to integrating new datasets, the company will develop more custom applications for PICC that will, for example, allow researchers to compare patients and identify individuals with similar profiles in their datasets, Stanley said.