

Building data power

New initiative aims to connect central Indiana's data-sciences talent, resources

By Susan Orr
sorr@ibj.com

Fueled with a \$36 million grant from Lilly Endowment Inc., the Central Indiana Corporate Partnership has launched an initiative called AnalytiXIN to promote innovations in data science throughout Indiana.

Its strategy? Build connections between Indiana's manufacturing and life sciences companies and the university researchers who can help them use artificial intelligence and advanced data analytics to tackle big challenges like reducing a factory's carbon footprint or improving worker health.

"This is one way to ensure early that these kinds of critical collaborations are happening," said David Johnson, president and CEO of the Indianapolis-based Central Indiana Corporate Partnership. "It's an ambitious effort."

About half of the \$36 million will be used to hire university-level data-science researchers, some of whom will be based at 16 Tech in Indianapolis. The other half will go toward the creation of "data lakes," or large data sets built from information from multiple contributors. One of the data sets will be manufacturing-focused, while the other

“This is one way to ensure early that these kinds of critical collaborations are happening. It's an ambitious effort.

David Johnson, Central Indiana Corporate Partnership president and CEO

will contain health sciences data.

The ultimate goal for both parts of the effort: Build up Indiana's data-science resources, to bolster existing Hoosier

See next page

Continued from previous page

businesses and attract new activity in a fast-developing field.

“It makes a deliberate statement that this is a priority for us in Indianapolis,” said Tatiana Foroud, executive associate dean for research affairs at the Indiana University School of Medicine.



Foroud

The medical school is one of AnalytiXIN's three university partners, along with Purdue University and the University of Notre Dame.

Building data lakes

Foroud is involved in the creation of the health sciences data lake. The value of a data lake, she said, is that it contains a large amount of data, ideally from multiple sources, that can reveal patterns and insights that might not be evident in a smaller pool.

As an example, she said, a data lake might include things like anonymized medical records, genomic data, and geographic and demographic patient information. Taken together, this data might reveal that patients who live in certain areas are prone to certain health problems, for instance, or that patients treated with specific types of medicine tend to have better health outcomes than others.

“Data lakes allow you to ask those questions and be more confident in the response,” she added.

To date, Foroud said, Indiana University Health has committed to sharing some of its data, and her team is talking with both Anthem Inc. and Eli Lilly and Co. about contributing data as well. The team is also tackling a host of other issues, such as how to make the data both secure and accessible and how to structure it so people can find what they're looking for.

By the end of the year, Foroud said, the health-sciences project aims to have its data partnerships secured and to have identified the key health issues the data lake will focus on. Diabetes is one possible area of inquiry, she said.

The intent is to make the data lake available not just to AnalytiXIN partners but to others as well. Foroud said she thinks the health-sciences data lake could be of interest to a range of users both inside and outside Indiana, from tech companies and health care systems to life insurance companies.

AnalytiXIN

Purpose: Advance innovation in data sciences, including artificial intelligence and advanced data analytics.

Formed: September 2020

Headquarters: At 16 Tech, 1210 Waterway Blvd., for research, plus meetings, workshops or other events.

Partners: Indiana University, Purdue University, University of Notre Dame, Eli Lilly and Co., Anthem Inc., Cummins Inc., Energy Systems Network, Indiana University Health, OneAmerica Financial Partners Inc., BioCrossroads, Central Indiana Corporate Partnership, Conexus Indiana

Funding: \$36 million grant from the Lilly Endowment Inc., available over 2-1/2 years

Distribution

- ▶ About half will pay researchers from IU, Purdue and Notre Dame, some of whom will be based at 16 Tech.
- ▶ Most of the remaining funding will be used to develop one data set focused on life sciences and another on manufacturing. Data will be available to AnalytiXIN partners and others.

Source: Central Indiana Corporate Partnership

“I think you're going to start seeing people line up” to access the data, she predicted.

The manufacturing-focused data lake will include data from Columbus-based engine-maker Cummins Inc. and several other partners, said David L. Leach, executive director of analytics and artificial intelligence at Cummins. Leach declined to identify those other partners.



Leach

Cummins already uses data analysis and artificial intelligence for its own proprietary projects, Leach said. The data lake will focus on information that can help solve industry-wide challenges: how a manufacturer might reduce its carbon footprint, improve workers' health and safety or beef up cybersecurity.

“The more data, and the more diverse data, you get, the more powerful it is,” Leach said.

The goal is to have at least an initial version of the manufacturing data lake available by the end of the year, he said.

Nitesh Chawla, the Frank M. Freimann professor of computer science and engineering at Notre Dame, predicted that the data lakes will become a “state treasure” for Indiana. “This will really accelerate discovery and invention of new AI data methods.”



Chawla

“We are excited about what this opportunity means for the state of Indiana,” said Chawla, who is also founding director of the university's Lucy Family Institute for Data and Society.

Collaboration

Notre Dame plans to recruit faculty members from various disciplines as part of the AnalytiXIN effort. Those disciplines, Chawla said, could include data-driven manufacturing, analytics and cybersecurity, as well as a social scientist who studies the future of work.

Notre Dame, and the two other universities, will match the Lilly Endowment funding to pay the researchers' salaries, Chawla said, with the idea that the university would pick up the full cost at the end of the grant cycle.

At least one of the Notre Dame professors will be based in Indianapolis full time, he said. The university will also create a way for existing faculty members and students to participate, perhaps establishing Indianapolis-based student research projects or internships.

Johnson, the CICIP executive, said AnalytiXIN had its origins in a handful of separate conversations he had several years ago with CEOs in various industries. The executives all said they had to look outside the area for data-sciences research and development resources because they couldn't find the necessary expertise in central Indiana, Johnson recalled.



Johnson

From these conversations, he said, he grew concerned that locally based companies might start to “hollow out” if some of their highest-value activities moved elsewhere.

Indiana has some great universities working in the fields of data science,

See next page

Continued from previous page

Johnson said, but none of them is based in Indianapolis. That puts the city at a disadvantage to places like Atlanta and Boston, both of which are home to top-tier data-science schools.



Muro

The answer to that, Johnson said, is the AnalytiXIN offices at 16 Tech, which will provide space both for IU, Purdue and Notre Dame researchers and for industry partners interested in collaboration.

Efforts like AnalytiXIN could help Indianapolis boost its presence in the realm of artificial intelligence, according to one outside observer.

"We do see a lot of opportunities," said Mark Muro, a senior fellow at the Brookings Institution's Metropolitan Policy Program.

Opportunity abounds

On Wednesday, Washington, D.C.-based Brookings released a report called "The Geography of AI: Which Cities Will Drive the Artificial Intelligence Revolution?"

The report, which was partially funded by CICP, shows that the U.S. artificial intelligence industry is heavily

concentrated in a few key metro areas in California, mostly the "superstar region" of San Francisco and San Jose.

Indianapolis is among 87 metros that Brookings classifies as "potential adoption centers." These areas have some AI activity but not as much as in the 13 "early adopter" metro areas: New York; Boston; Seattle; Los Angeles; Washington, D.C.; San Diego; Austin, Texas; Raleigh, North Carolina; Boulder, Colorado; Lincoln, Nebraska; Santa Fe, New Mexico; and the California metro areas of Santa Cruz and Santa Maria-Santa Barbara.

Based on research activities at IU and Purdue, Bloomington and Lafayette-West Lafayette were also among 21 metro areas Brookings identified as hubs of federal AI research and contracting activity.

The Brookings report did not identify significant levels of AI activity in the country's remaining 261 metro areas.

"You really are looking at a potentially significant concentration and hub in Indiana," Muro said.

Central Indiana tech executives outside of the AnalytiXIN effort also see promise in the project.

"This has been talked about for a long time," said Christopher Day, CEO and co-founder of Indianapolis-based DemandJump Inc. "I'm so excited to hear that this is actually happening."

DemandJump, which Day founded in 2015, uses artificial intelligence and data

analytics to help companies improve their ranking in online search results.

Day said AnalytiXIN "has the potential to unleash a massive amount of entrepreneurial growth" because the data lakes and the brainpower will make central Indiana more attractive to entrepreneurs.

He also sees a great advantage to having data-science researchers based locally, working with local companies to solve problems. "When you centralize that brainpower, then people start asking questions of each other: 'What if we could do this? What if we could do that?'"

Lilly Endowment Inc. said in a written statement to IJB that the substantial project is worth tackling.

"The Endowment recognizes the complexities associated with an endeavor like this," the statement said. "However, we believe there is a far greater risk in not attempting to build a strong, more connected and entrepreneurial digital technology ecosystem in our state. The Endowment hopes that this grant will play a role in advancing that priority and, by extension, help to enhance the prosperity of Indiana."•



Day