Empire AI gains momentum as officials visit potential sites at UB

By Janet Gramza

Venu Govindaraju, vice president for research and AI scientist at the University at Buffalo, describes the future Empire AI supercomputing center to be built at UB and shared by a consortium of New York research institutions to harness AI for the public good.

New York’s Empire AI consortium gained ground this week as members of the state’s think tank on artificial intelligence visited the University at Buffalo to view potential sites for the AI supercomputing hub to be housed on its North Campus.
Gov. Kathy Hochul created the Empire AI consortium of public and private research partners and funders in this year’s state budget, including designating two potential sites at UB for an AI supercomputing center to be shared by the group to advance AI research in New York.

Members of the consortium have been meeting mostly virtually for months to launch the $450 million project, whose mission is to give New York researchers the same state-of-the-art AI tools developed and used by Big Tech in order to harness AI “for the public good.”
Representatives of New York's new Empire AI consortium arrive at potential sites for the new AI supercomputing center to be housed at the University at Buffalo.

Janet Gramza/Buffalo News
UB elaborates on Empire AI in roundtable with state reps
UB and SUNY Chancellor John B. King were at UB's National AI Center for Exceptional Education Wednesday to share UB AI research with state legislators who launched Gov. Kathy Hochul’s Empire AI consortium.

Representatives of the seven founding partner institutions – the SUNY and CUNY systems, New York, Columbia and Cornell universities, Rensselaer Polytechnic Institute and the Simons Foundation – came to UB this week to look at two sites less than a mile apart, both fairly close to UB's Center for Tomorrow and the Flint Road entrance to North Campus.

The group of about 20 people also included UB officials, legal and corporate consultants and private funders including Tom Secunda, co-founder of Bloomberg LP, who donated $5 million to his alma mater, SUNY Binghamton, to participate in Empire AI research.
“It’s really important to get this going as soon as we can,” Secunda said. “AI is not sitting back and waiting.”

Secunda clearly favored the first of two sites the group looked at, a 1.8-acre area close to UB’s electrical substation and Flint Road solar arrays that’s easily accessible to construction vehicles and has room to expand beyond the 10,000 square feet anticipated for a one-story facility to house AI supercomputing hardware and infrastructure.

The only important utility that site would need is water for a cooling system, which is why the state also included a second site, a 1.3-acre property adjacent to the Baker Chilled Water Plant, as a potential option, said Kelly Hayes McAlonie, UB’s director of campus planning and design.

But while the second site has ample chilled water, the pipes that carry it run under the site and would have to be moved, said Tonga Pham, UB associate vice president for university facilities. Site 1 also is flat while Site 2 has a 6-foot elevation differential, Pham said. So besides moving water infrastructure, “We would have to construct a service road for that site,” she said.

McAlonie and Pham said it would be easier to bring water to Site 1 than to bring electricity and a service road to Site 2. But, Pham said, “this has to be a consortium-informed decision.”
Venu Govindaraju, UB vice president for research and an AI scientist who led the tour, said the Empire AI center will take a year or two to build. He estimated the building will cost $50 million to construct and the supercomputing infrastructure will cost $200 million to purchase and install.

Hochul allocated $275 million for Empire AI in her 2024-25 state budget, and has secured pledges of $125 million in corporate and foundation funding so far. Govindaraju said it will cost about $150
million to operate and maintain the computing center for a period of 10 years.

He said the Empire AI supercomputing center will be 20 to 40 times more powerful than UB’s existing Center for Computational Research, which is considered one of the leading academic supercomputers. That facility occupies about 4,000 square feet at the UB Center of Excellence in Bioinformatics & Life Sciences on the downtown medical campus. But Govindaraju said that doesn’t mean the AI supercomputing infrastructure will need 20 or 40 times more space.

“Computer racks are becoming more dense, so their square footage is shrinking,” he said. Going forward, computational power will increase while taking up less real estate, he said.
Krystyn Van Vliet, vice president for research and innovation at Cornell University, said visiting the potential sites with fellow consortium members “absolutely gives the project momentum.” “All of the organizations involved have been working together for many months now, on Zoom and phone calls and emails, so now with this rhythm of meeting together in Buffalo, as well as in New York City a couple of months ago, that’s where we are developing the relationships that make it feel real,” she said. The consortium has hired a consultant, Hangar, to design the facility depending on computing equipment and capacity, McAlonie said.
Group members said one concern they have is that the construction may compete with other state projects for the skilled trades workers needed to make it happen as quickly as possible.

Secunda said he would like to see the facility built in phases in a way that can get high-level AI research underway while the next phase is being built.

Besides the design and infrastructure, Van Vliet said the consortium also needs to map out a plan for allowing partners to solve computational problems of all sizes – from crunching climate data to designing cures for diseases.
“Once we plan what the facility is going to house, then you have to think about how are we going to share those computers to do a combination of very large jobs that take up a lot of resources, so-called large language model building and testing, as well as very important high-social impact problems that don’t use as much of the computing power at once,” she said.

“That’s a complex optimization problem, because you also want to make sure that the facility will be running 24-7,” she said. “That’s what will make it impactful and cost efficient.”