Building the Bridge While Scientists Cross It

Suresh Marru

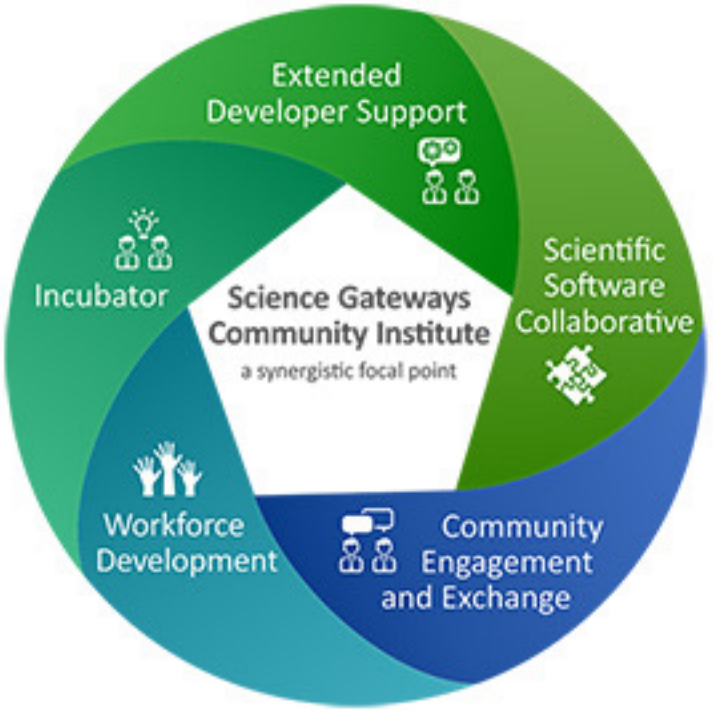
Georgia Tech
Center for AI in Science and Engineering




Congratulations!

Talk Outline

Lessons from ~~building~~ nurturing, similar infrastructure communities



XSEDE

Extreme Science and Engineering
Discovery Environment

ACCESS

01 **Why dependable infrastructure matters**

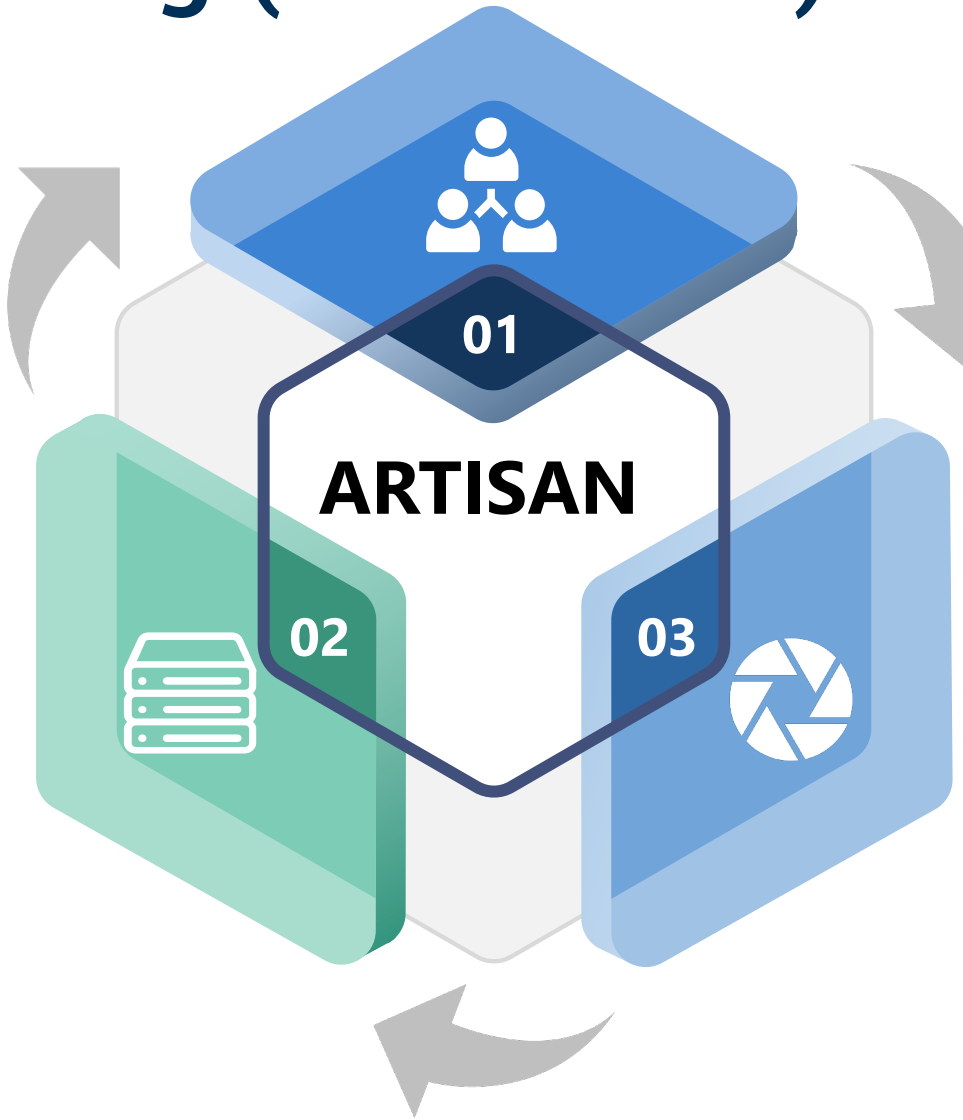
02 **Expectations Mismatch**

03 **Community Interest and Active Use**

04 **Evidence of scientific value**

05 **Questions worth asking early**

Georgia Tech Center for AI in Science and Engineering (ARTISAN)



Multidisciplinary Expertise

AI researchers, engineers, and domain scientists.

Advanced Computing

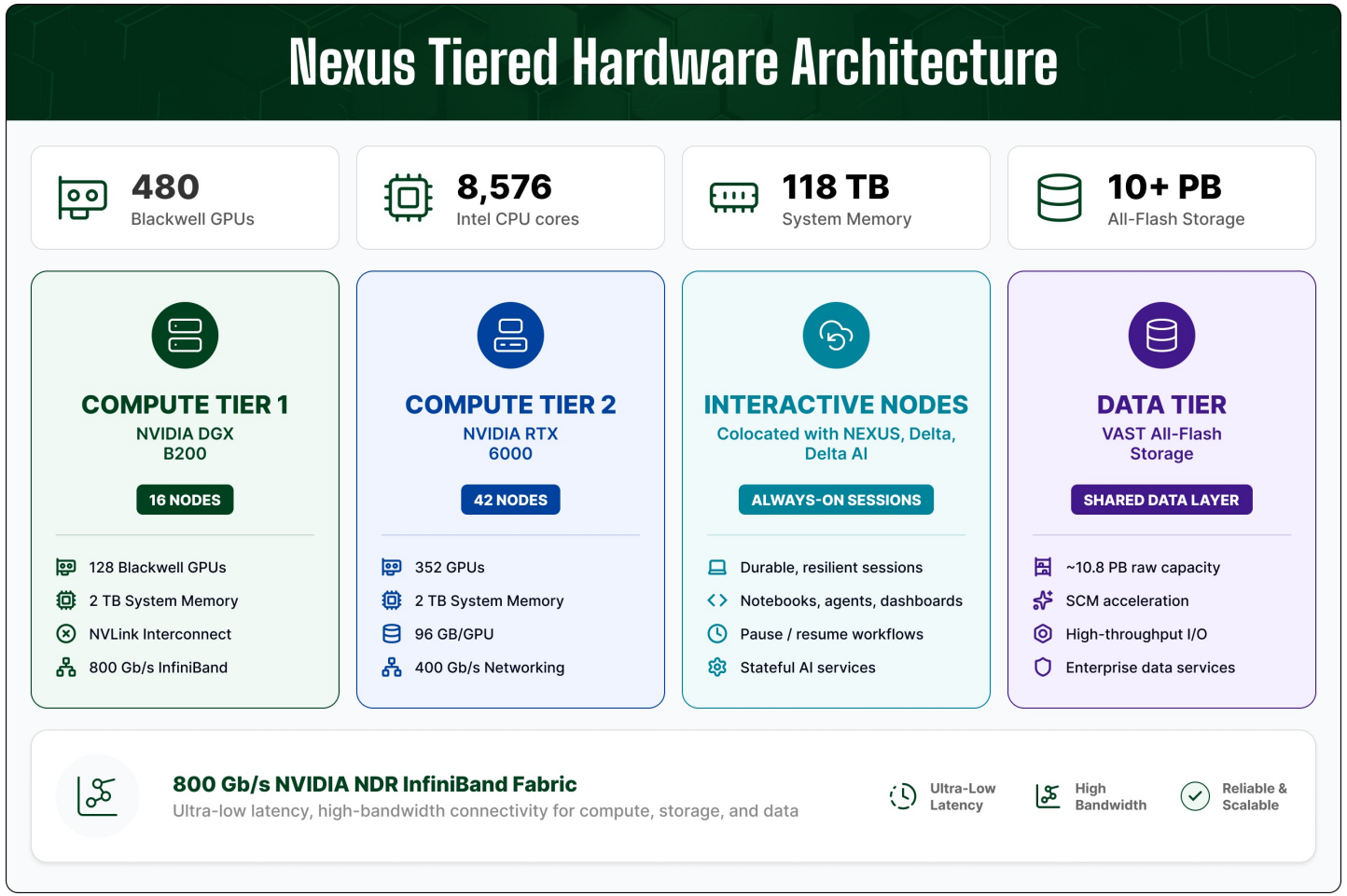
Next-generation systems enabling local-to-HPC scaling

AI, Data & Software

Open, interoperable frameworks

Nexus: A National Open Science AI Infrastructure

- $20M NSF funding to build and operate Nexus.
- Catalyzing interdisciplinary AI4Science across computing, engineering, life sciences, climate, materials, and health.



NEXUS

Evolution: Serving Science by Getting Closer to the Researcher Workflow

LEAD

Ground-up Gateway infrastructure

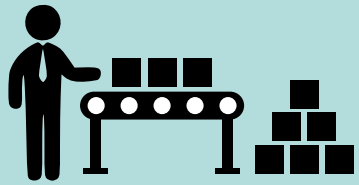
Wrap each science application as a web service



Airavata

Generalized Gateway Framework

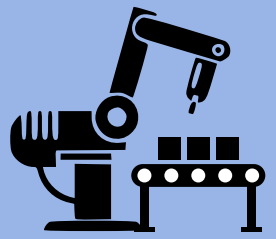
Evolving Architecture with high available as-a-service model.



SciGaP

Gateway Factory

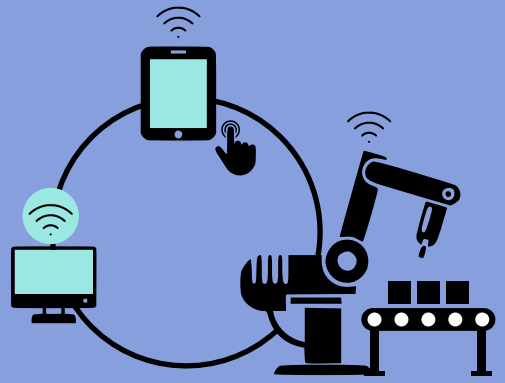
Integrated 70+ compute resources into 50+ Science Gateways from multiple disciplines.



Cybershuttle

End to end Science workflow

Integrate resources at all scales, assist with full science life cycle



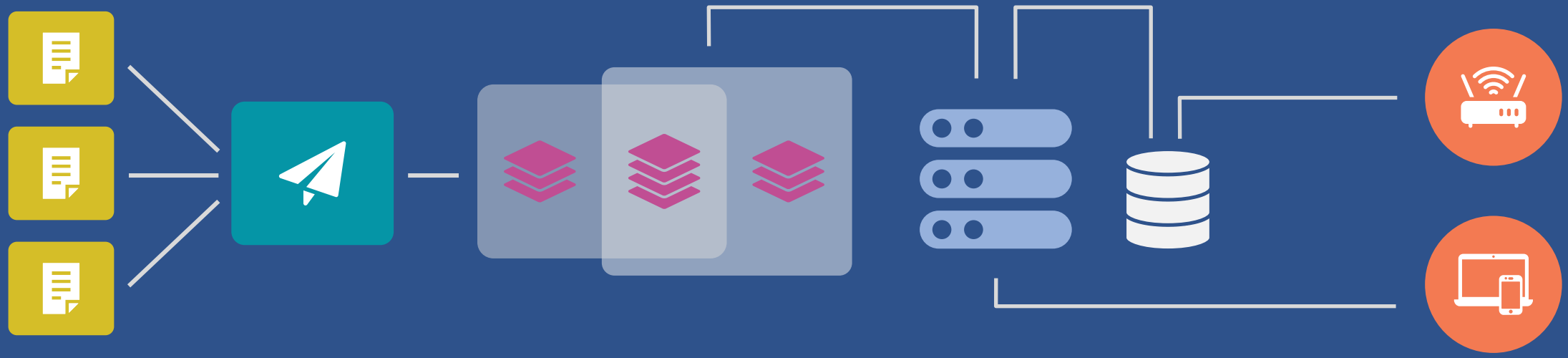
Cybershuttle enabled “**Science Workflow**”

Ingest data from public or authenticated repositories (and Analyse them locally)

Execute user-authored or available community applications/workflows

Integrate with Resources at all scales and minimize movement of data

Analyse results remotely or locally and publish when ready



Each of these steps are interactive, iterative, “recreatable” and sharable within a controlled collaboration and published eventually.

BRIDGE is a Category II IDSS Cyberinfrastructure
From Established Platform to National Operations

Current State

…transition of an established system from a smaller-scale or prototype setting to national-scale production and operational quality…

Next 3 years

- A service they can rely on for real research
- A usable path through the scientific data lifecycle
- Clear access, documentation, training, and support

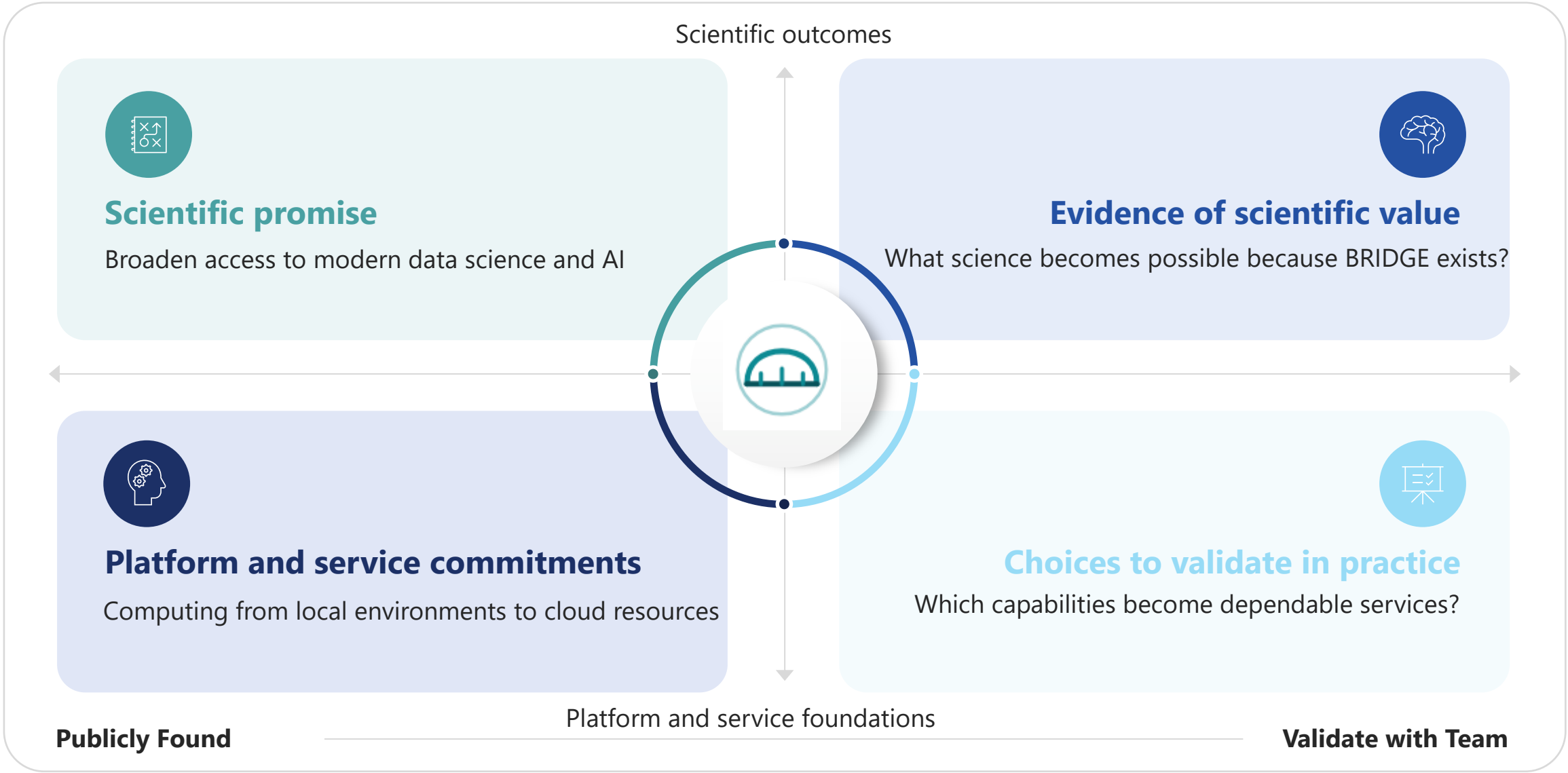
Renewal/Cat I proposal

01 Evidence that the resource enables reproducible scientific outcomes

02 Stable operation while the platform continues to improve

03 Identify the science that would not happen without BRIDGE

My Reading of BRIDGE's Public Promise





Georgia Tech
Center for AI in Science and Engineering

Zero-Sum Game

Which capability in BRIDGE will be stable, reliable, and repeatable in 3 years?

Which scientific journeys deserve stability promise?



01

VS Code or code IDE's

Probably you can tolerate a bleeding edge IDE

02

Laptops

May be a "cool" laptop can excite us over a boring robust one.

04

A compiler for scientist?

Do you expect real science to happen on Bridge or prefer cool demos?

03

Java Compiler

Would you deploy your Scala projects on an evolving compiler?

I will share my experience

I learnt this the hard way











LINKED
ENVIRONMENTS
FOR ATMOSPHERIC
DISCOVERY











Anticipating Tornados

- Project LEAD
  - Using technology to dynamically adapt to weather conditions
  - Decrease false alarm rate
  - Dramatically improve probability of detection



LEAD Accomplishments

"The stretch goal for LEAD was to begin ushering in this paradigm change. After 6 years, 415 papers and presentations, a mention of LEAD by Bill Gates at Supercomputing 2007, 9 PhD's awarded, 14 Master's degrees, and 1 Bachelor's degree; and 5 TB of user generated forecast and analysis results, we can safely say that LEAD has succeeded in making substantial inroads towards advancing this paradigm change."

– LEAD final Report to NSF.

Feedback: a promising platform but is it an Essential Infrastructure?

This is not useful for me, but many people in my community will use it

If we stop funding you now, what science will not happen?

Domain scientists, in the end, were like, “Thank you for showing what's possible, but sorry, we could not rely on it.”

Enabling Science vs Innovating Capabilities



Georgia Tech
Center for AI in Science and Engineering

Q&A

I look forward to learning from what you build.



01 **What will scientists be able to do with BRIDGE?**

02 **If BRIDGE disappears tomorrow, what science would no longer happen?**

03 **What leaderboard will each of you look at each day?**