

- ❖ **Everybody** needs more compute and better software.
- ❖ So, I'll focus on other, more **hidden challenges**  
— i.e., treating **human time** as a **limited** resource.

# Code / Community / Culture

Deep Skies and a Vision for AI in the Cosmic Center at Fermilab

---

# The Deep Skies Lab and Community

2017: Josh Peek (STScI), Camille Avestruz (Umich)

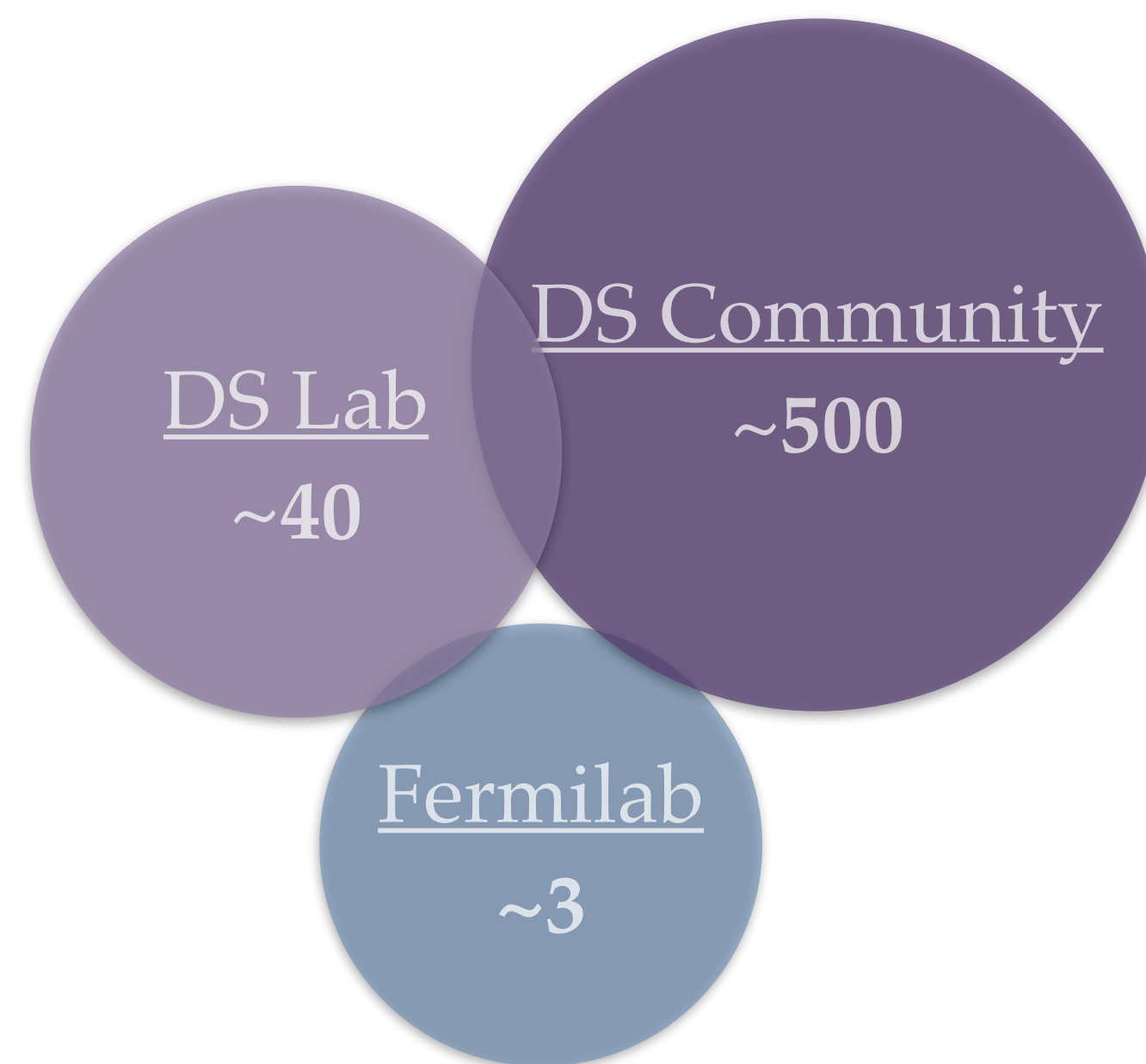
---

## ❖ The Community

- ❖ a meeting place
- ❖ Multi-institutional, transnational, ~4 continents

## ❖ The Lab is our research group

- Nord + Ciprijanovic
- ❖ Fermilab, UChicago, Cornell, UMich, STScI, NOIRLab, Riverside, UChicago Lab school, MIT, Pitt, +
  - ❖ High school to PI

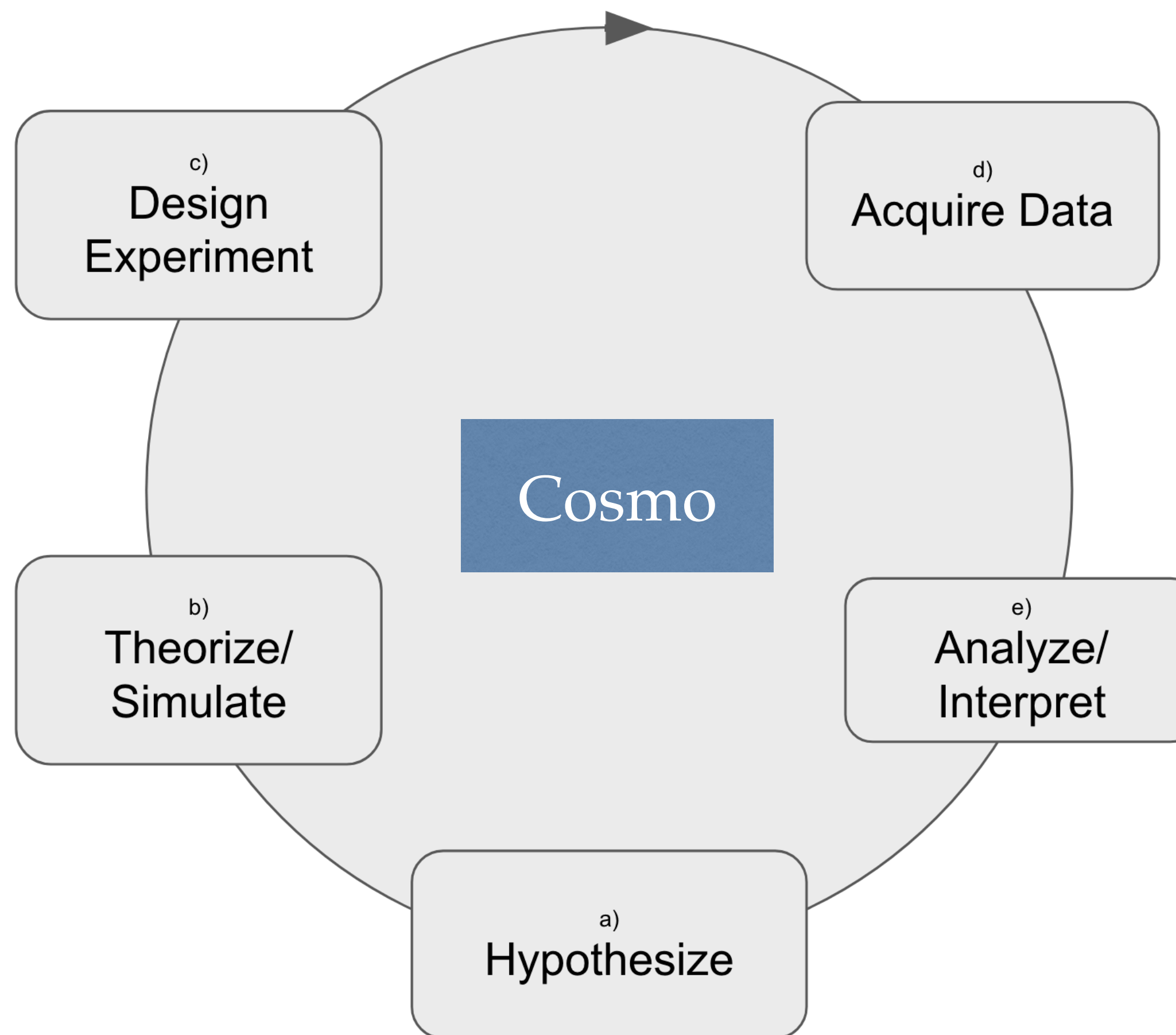


- ❖ **Brought AI to Fermilab Cosmic sector\***
- ❖ Created because we **needed a space ...**
  - ❖ ... to discuss AI + Cosmic
  - ❖ ... to have a healthy work environment\*\*

# DeepSkies at Fermilab: Science

---

- ❖ Tackle **every aspect** of the scientific cycle in a cosmic context.
- ❖ This includes **simulation, inference,** and **controls** performed both with standard **statistics** and **AI** tools.
- ❖ Development: AI Theory and Algorithms
- ❖ Applications in cosmic context.



# Barriers: Hardware

The “easy” ask.

*Approximate requests*

- ❖ Deep Skies Lab for the next 3-5 years:  
2 PI's + 5 postdocs + 25 students (HS, UG, Grad) = 30 people
- ❖ Not enough **just-in-time compute** for each user.

- ❖ 2 Effort stages: Development, Production
- ❖ **Development:** 25xA100, >20GB memory, anytime in EAF
- ❖ **Production:**
  - ❖ **Simulation:** O(10) CPU for embarrassingly parallel runs
  - ❖ **Inference:** 25xA100, >100GB memory total, anytime in **Wilson**; +**Power9**
- ❖ **Total Storage:** ~500 TB

# Barriers: Software (1)

Code is our  
least-addressed  
systematic error.

- ❖ Tough: Getting code **spun up** on a resource (especially with GPUs).
- ❖ Tough: Becoming **familiar** with a given resource.
- ❖ Tough: Transporting code **between resources** (e.g., Laptop  $\rightarrow$  EAF  $\rightarrow$  Wilson).
- ❖ Tough: **Latest deep learning** versions of Torch/TF aren't always available.

These items can become seamless and require almost no effort from the (new) user: **documentation, automation.**

# Barriers: Software (2)

Code is our  
least-addressed  
systematic error.

- ❖ Most scientists (e.g., early-career) write code that is **not reproducible**.
- ❖ **Lacking skills** in clean code, versioning (git), packaging.
- ❖ Only major codes get **maintained** by the lab, so we need **CI/CD** another way.

- ❖ More skills **training**.
- ❖ More **automated workflows**.
- ❖ More **professional help**.

# Barriers for Cosmic in AI: Time, Energy, Culture

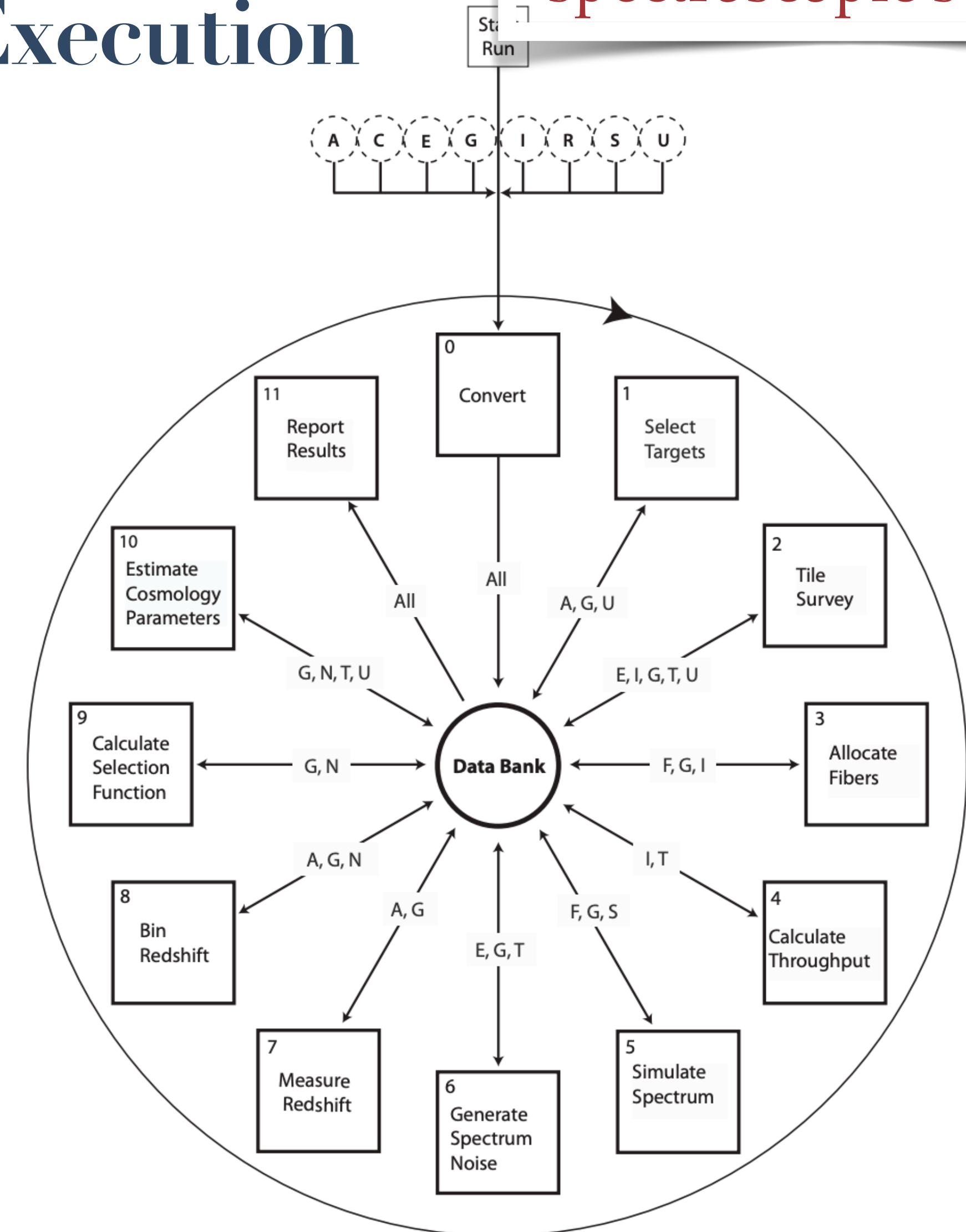
---

- ❖ **Inefficient** management of meetings and projects.\*
  - ❖ Unreasonable **bureaucracy** is literally **destroying our reputation** with col
  - ❖ Scientists **opposed** to exploring the potential of AI in cosmic.
  - ❖ Lack of **time** to understand AI.
  - ❖ Lack of time to understand AI **proposal ecosystem**.
  - ❖ Fermilab Cosmic is **struggle** to find a leadership role in the lab ecosystem
  - ❖ Too much **AI hype**.
  - ❖ **Wasting** non-science effort.
  - ❖ **Lack of accountability** causes trauma and stress.
- ❖ More skills **training** in management and meetings.
  - ❖ **Push back** on site office.
  - ❖ Reduce **useless bureaucracy**.
  - ❖ **Stop allowing** people to create a hostile work environment.

# Addressing the Struggle: Automated Experiment Design and Execution

SPOKeS is a multi-fidelity simulation of spectroscopic surveys.

- ❖ Our Cosmic center could perform automated and precise experiment planning and execution?
- ❖ And we could become known for it.





# Addressing the Struggle: **Proposal writing**

---

---

- ❖ **Remove** the non-scientific tasks of proposal-writing.
- ❖ Give people **time to ponder** what is valuable for cosmic-AI.
- ❖ Give people **time to learn** about how DOE / NSF design AI / Science FoA.

# Addressing the Struggle: **Value expertise and time**

---

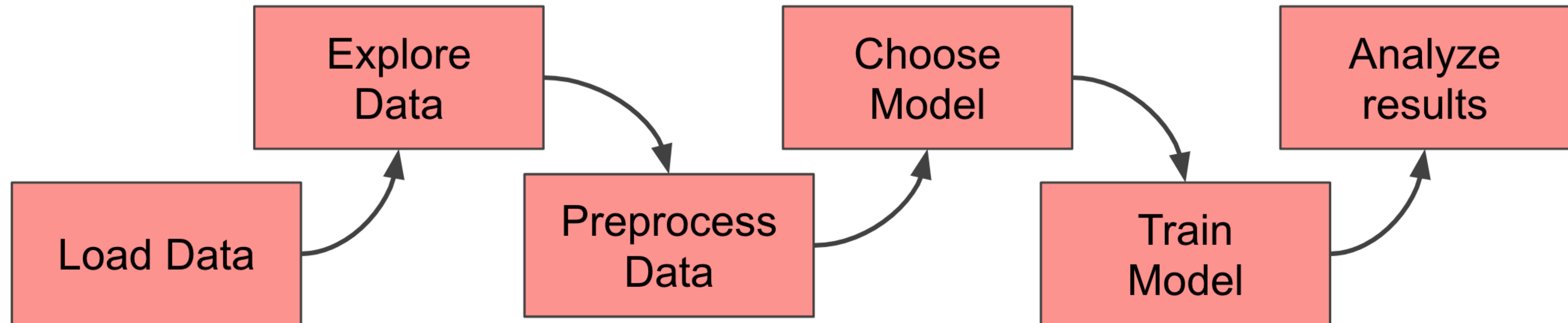
- ❖ The institution should let us do our jobs, and stop wasting our time.

Things like Comet.ML do some of this,  
but it's too superficial for science.

## Addressing the Struggle:

# Clean code and workflows: Deep Skies ToolBox

- ❖ CI/CD; Extensible; Auto-Diagnostics; Auto-Hyper-Optimization; Explainability; not reliant on commercial products;



# Addressing the Struggle: Ethical AI and Goodwill

---

- ❖ Scientists enjoy **goodwill** in society, because we can claim we are on the right side of truth.
- ❖ If we **betray** this, we risk **exhausting** that goodwill — e.g., facial recognition, weapons / DoD, corporate partnerships, over-hype.
- ❖ We need an ethics-informed approach to our computing, AI, and partnerships.

# Summary

---

- ❖ **Code** can be good, clean, reproducible.  
Let's make this part of our work culture.
- ❖ Can compute be made **simple to access and use**?  
Documentation is good, but trivial use is better.
- ❖ Let us do what we were actually hired to do:  
**science**, not trauma  
**science**, not bureaucracy  
science, **and** management?
- ❖ Plan for the **long term**, think new thoughts.