Metrics2023: Measuring Success of CyberInfrastructure Projects

Brainstorming Session

1. How do we build trust in the data related to the metrics for success?

- ❏ Scoring model — multiple input sources, combinations of metrics.
- ❏ Transparency about sources of data, how analysis was done, caveats, reproducibility especially in different types of projects.
- ❏ Complementing with qualitative measurements — e.g., gathering feedback from direct engagement with specific users over time.
- ❏ Multiple kinds of trust — do we believe where it came from, do we believe it is correct, do we believe it is relevant…

2. What are the challenges related to gathering the metrics of success related to your projects?

- ❏ Some metrics are easy to get (e.g., number of downloads, social media mentions) but are not the best measure of what we really care about.
- ❏ Number of publications aren't a sufficient metric of how much science is being done per watt / core — publications aren't always equal, contributions to publications aren't always equal.
- ❏ Can't tell how much productivity is happening just by looking at system usage.
- ❏ If maximizing impact is the only metric used to design machines, it can result in machines that are barely usable.
- ❏ Collecting metrics can easily fall second to getting the actual work done.
- ❏ Long-term impacts are potentially massive but can be difficult to collect metrics for.
- ❏ Tendency to overvalue the metrics that we can easily get.
- ❏ If you depend on others to collect the data, have to incentivize getting it.
- ❏ Trouble getting management buy-in to collect metrics we aren't already collecting.
- ❏ It can take a lot of effort, e.g., experimental studies, even to determine which metrics should be gathered.

3. How are we currently using the data related to our project metrics, to inform or revise, our processes/actions/decision-making?

- Looking at metrics at greater frequency and more consistently helps reprioritize projects, redirect efforts and costs, see that initiatives are taking too long to start up and would tie up resources and people, shut down projects sooner.
- Recognizing users that need help, finding them, helping them.
- Determine which systems aren’t working well.

4. What are the considerations for developing or improving on the "responsible metrics" for assessing the CI projects and defining benchmarks for success?

- Different kinds of responsibility: social/ethical responsibility, academic responsibility, choosing the right solutions for the right problems, …
- Qualitative analysis.
- Not punishing lack of experience or lack of publications when evaluating researchers for allocations.
- When evaluating allocation proposals, how to work with the researcher to be more successful next time rather than flat-out rejection.
- Don’t cherry-pick the data to give the result you want.
- Have checklists and trainings on how to present and evaluate metrics when submitting proposals or reports and when evaluating these.
- Proxy metrics can be responsible or irresponsible, and they can hide the true metric you’re looking for.