How Program Managers can exploit DevSecOps Pipeline of Pipelines to Inform Data Driven Decisions

NDIA

October 2023

Bill Nichols
Julie Cohen

Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213

Make Program Decisions!

Your SOCOM commander has learned that an adversary has unexpected capabilities.

You must **reprioritize** your **capabilities**.

Whiz-Bang Software for Sensor Fusion, which was added to the roadmap 18 months ago, is now **top** priority with a need of within the next 6 months.

Two capabilities that were to be the next priority, due in 3 and 9 months, are critical and must be completed as close to the original due dates as feasible.

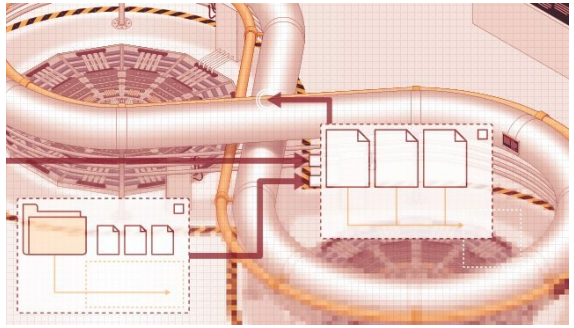
You need to know the following:

- How can the two next-priority capabilities and the new, highest priority capability be delivered without affecting the staffing?
- To deliver the Whiz-Bang capability within six months while maintaining the other two capabilities' schedules, **will more teams be required**? **How many**? **At what cost**?

What Do You Do?

How do you make informed, defensible decisions?

Agenda



Making defensible decisions with data

Automating data and analysis with ACE/PoPs

Indicators for Program Managers

Observations and Lessons

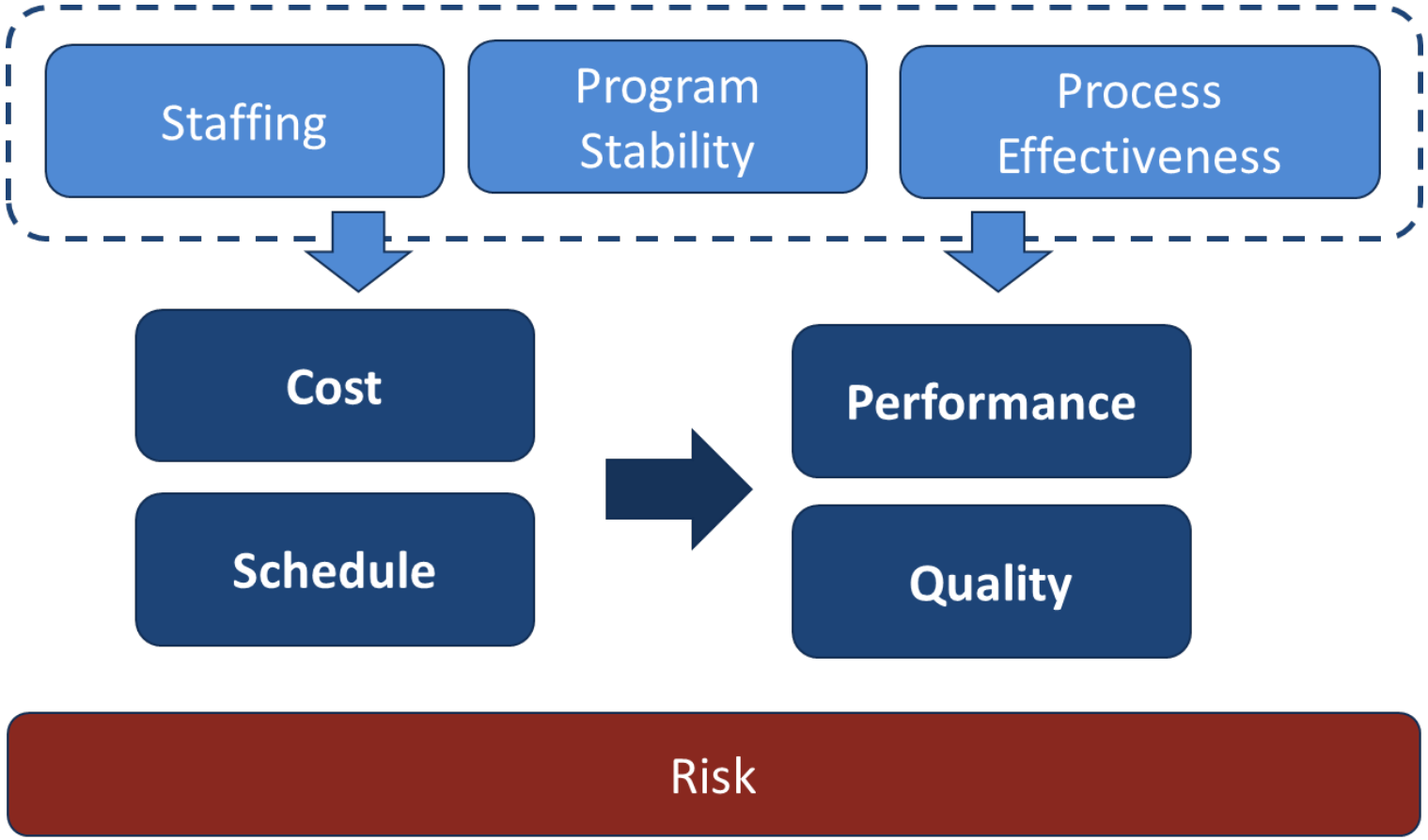
Call to Action

ACE/PoPs Overview

Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213

Start by getting the information you need

Information needs: What are your **targets**? What can you **manage**?



Automated Continuous Estimation for **Pipelines of Pipelines**

Automation drives continuous integration and delivery of software, but outpaces program control

To solve this problem:
Automate data collection
Model DSO systems with **Monte Carlo**, and provide continuous reporting.

- Determine status
- Project future events
- Provide evidence for corrective actions

Goal: Programs using DSO(DevSecOps) have constant access to information needed to monitor and control schedule and cost commitments.

Status and projection models should be available in real time.

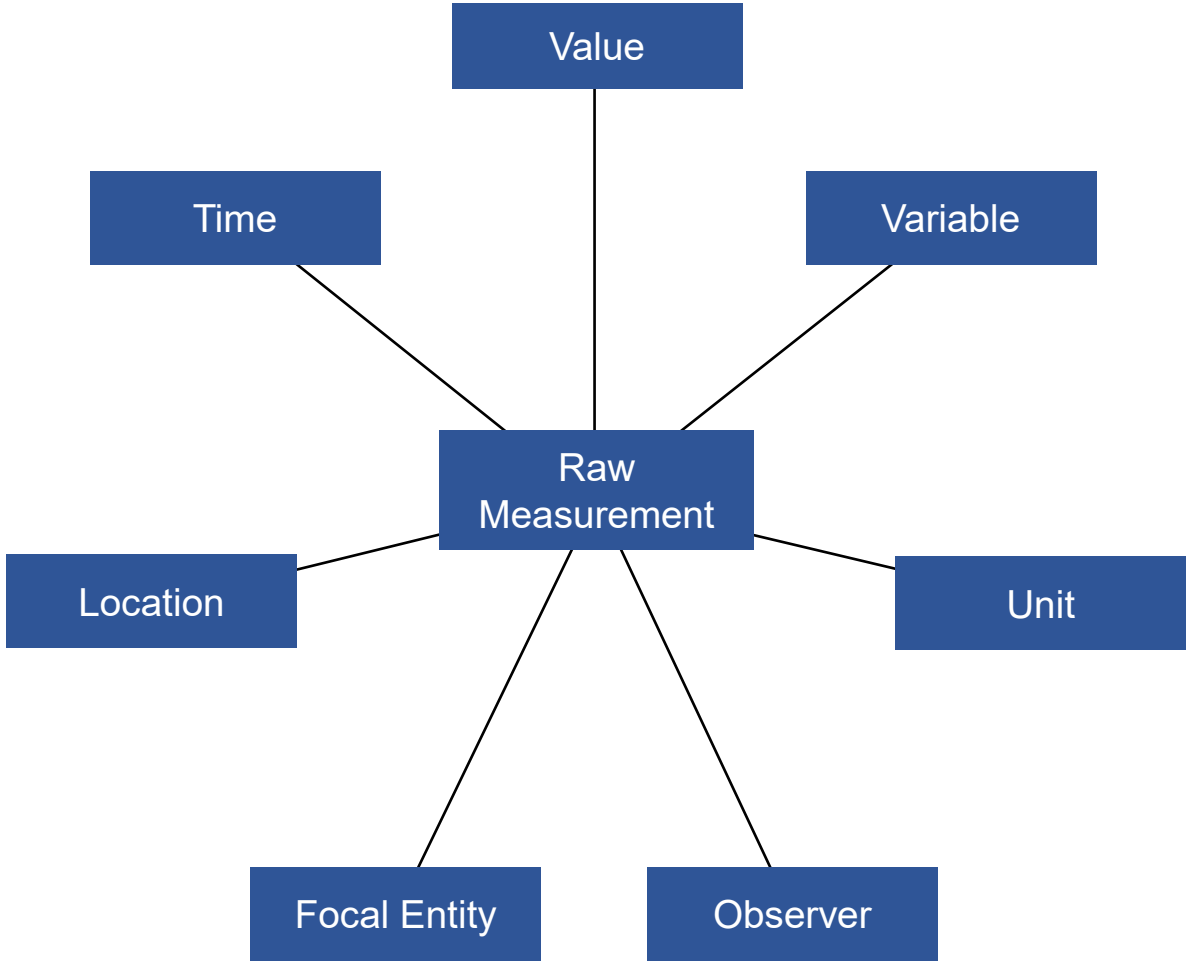
Model pipeline and pipeline-of-pipeline systems.

Automate data collection and Program Management Status Reporting for DevSecOps pipelines.

Directly collect data from DevSecOps pipeline tools

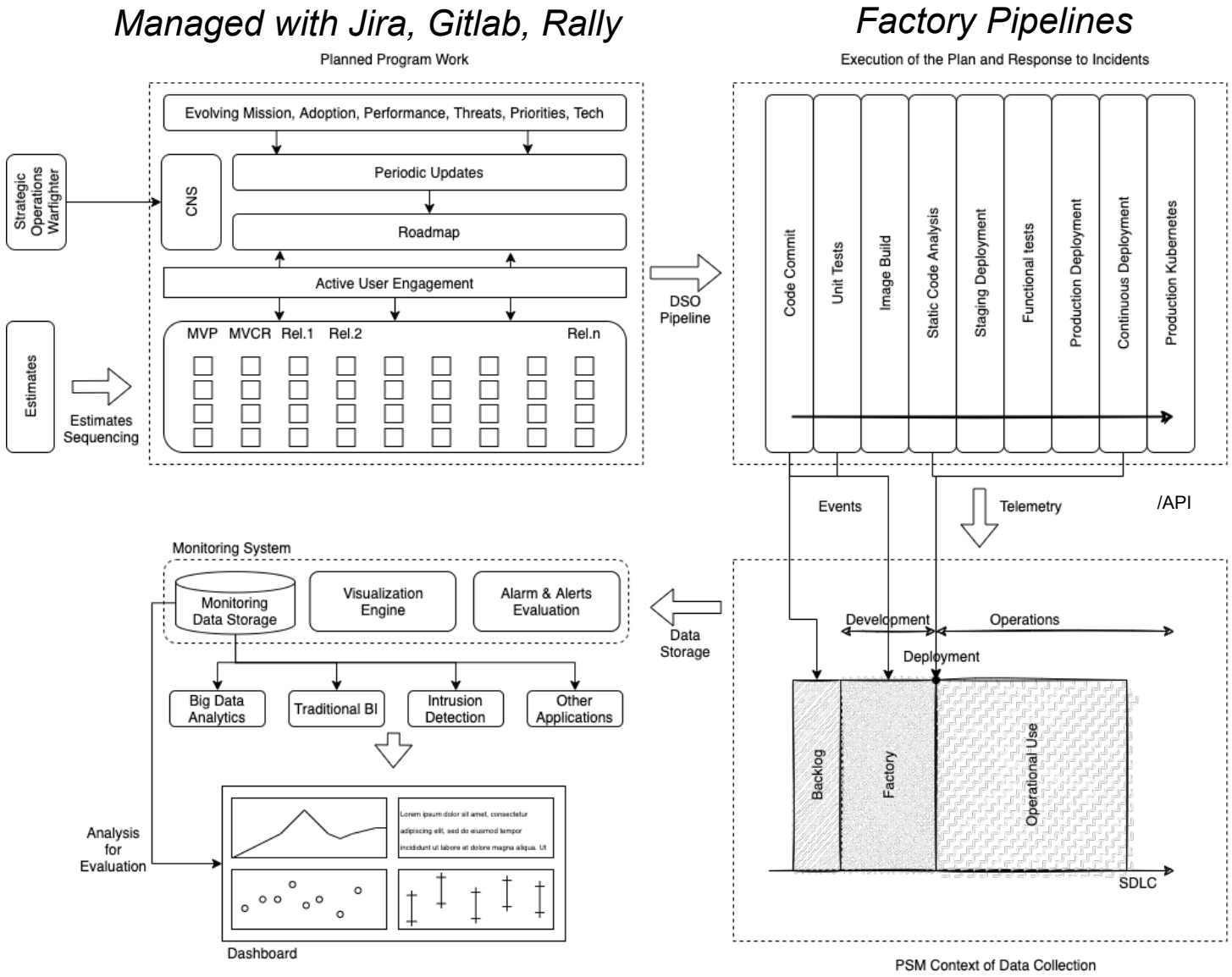
- Automate data collection, storage, and reporting
- Correlate data to project outcomes
- Present completion to-date and milestone predictions to Program Management in smart dashboards

Prototype DSO Measurement



- Measurement: *"A set of observations that reduce uncertainty where the result is expressed as a quantity." - Douglas Hubbard*

ACE/PoPs Model: Data Collection Context



Planned work includes the WBS, work packages, work sequencing, and estimates.

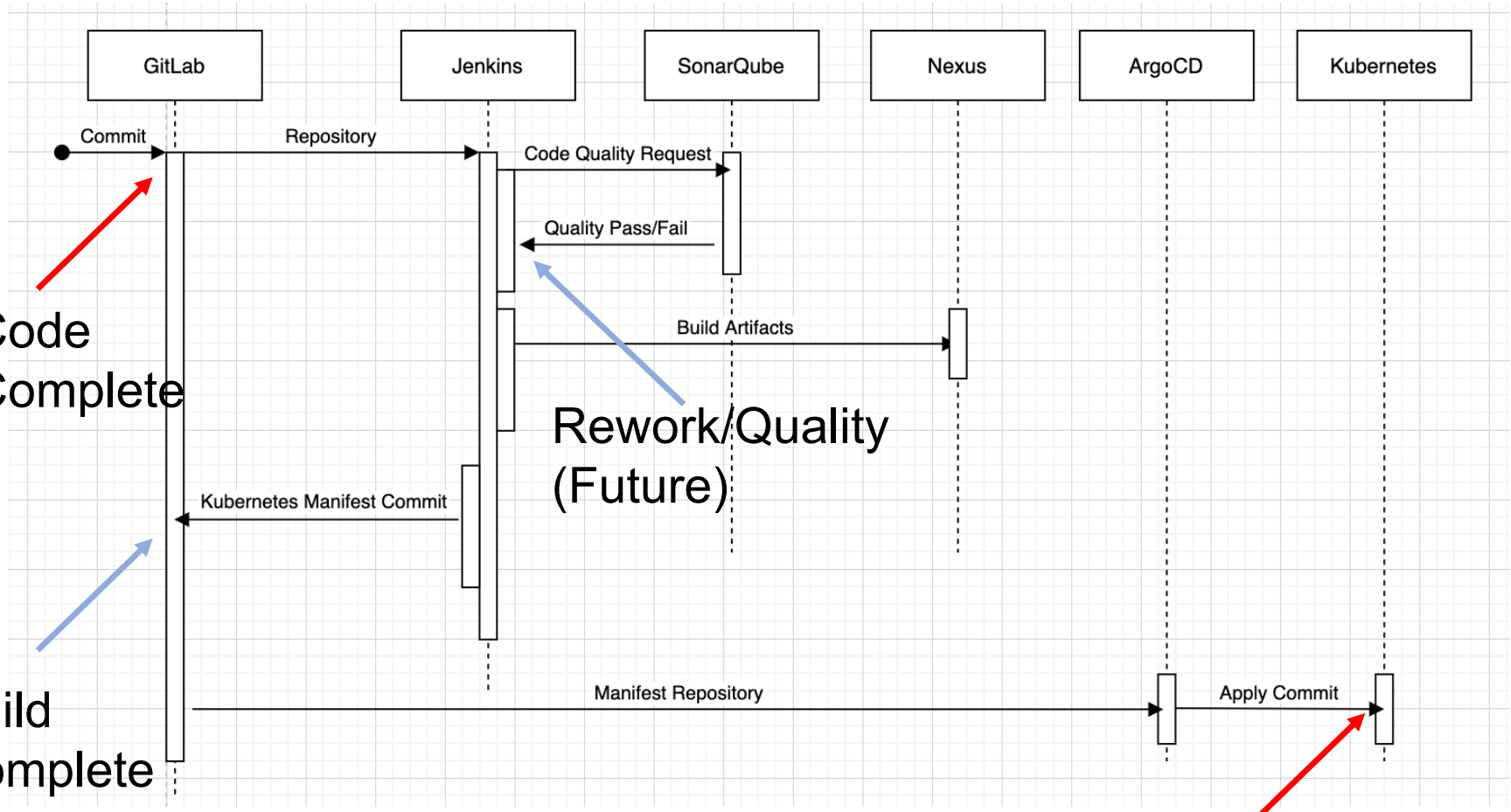
Work packages **execute** plan development stages. Tools trigger events (time stamps, package labels).

Data is collected and **transformed** for storage.

The **warehouse** loads the data and provides the interface for analysis and dashboards.

Track events through the Pipelines

Extracting metrics https://youtu.be/u96OFTXgr0g



Code Complete

Rework/Quality (Future)

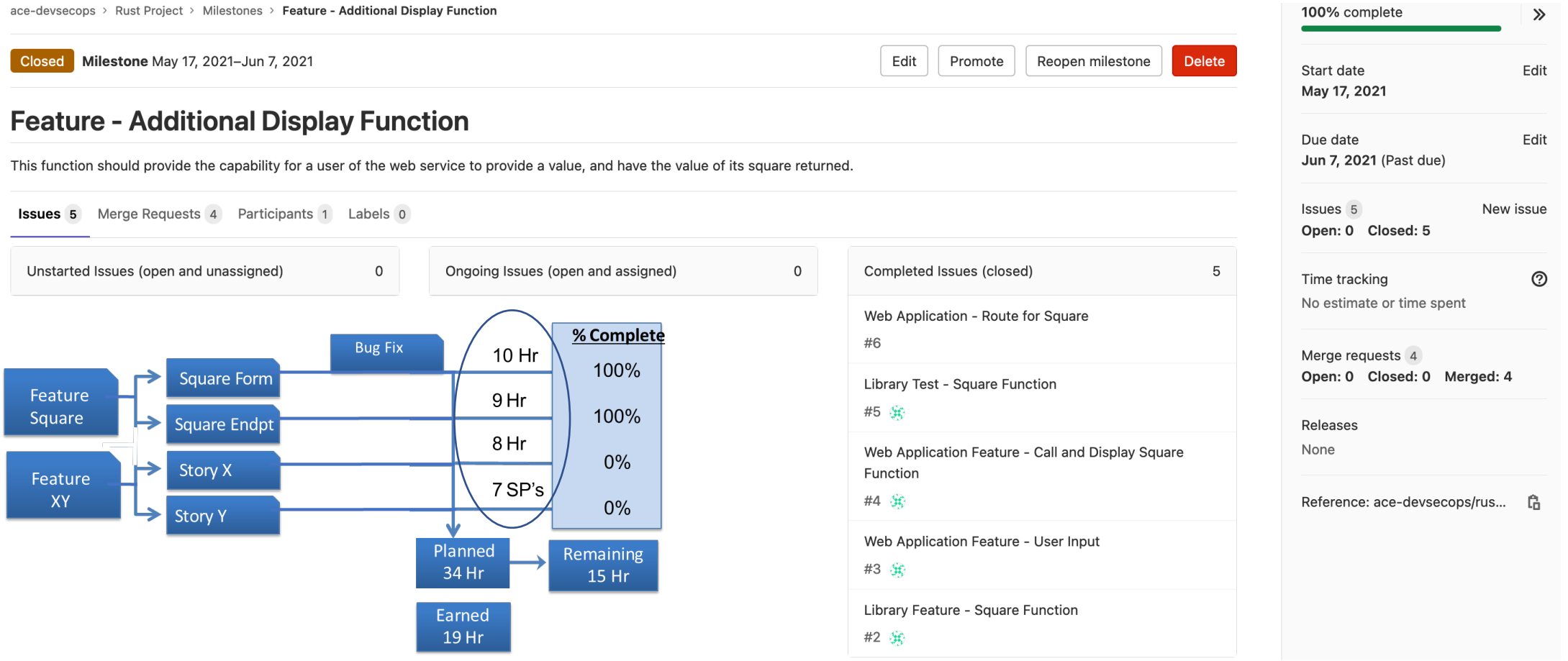
Build Complete

Deployed = Done

- Date is collected from key events.
- The data specification is on the following slide.
- Use Labels to connect WBS, RoadMap, and Backlog to work packages.
- Lead Times
- Estimated Dates
- Actual Times

Track From the Roadmap Through the Pipeline

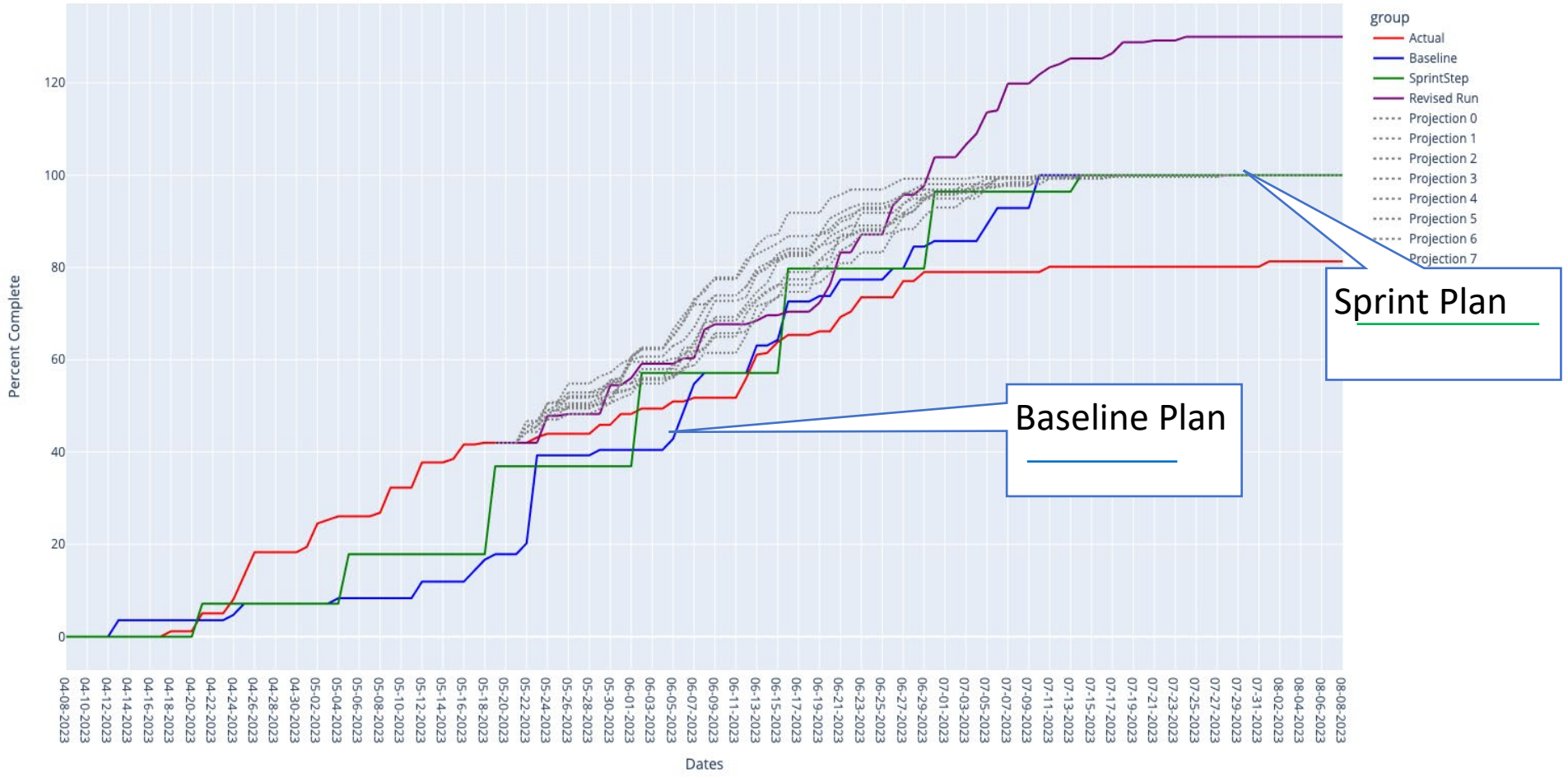
Work Completion https://youtu.be/X-R1mIZ3sPk



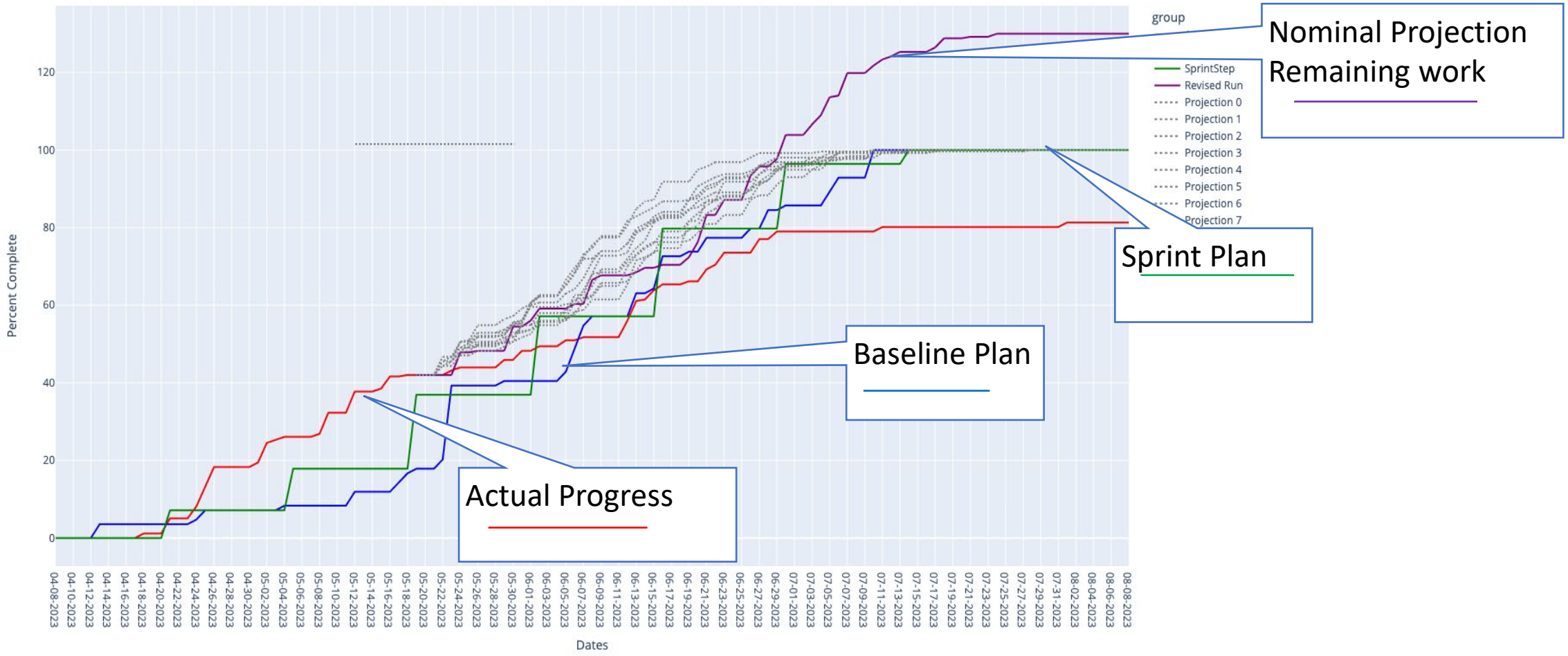
Indicators

Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213

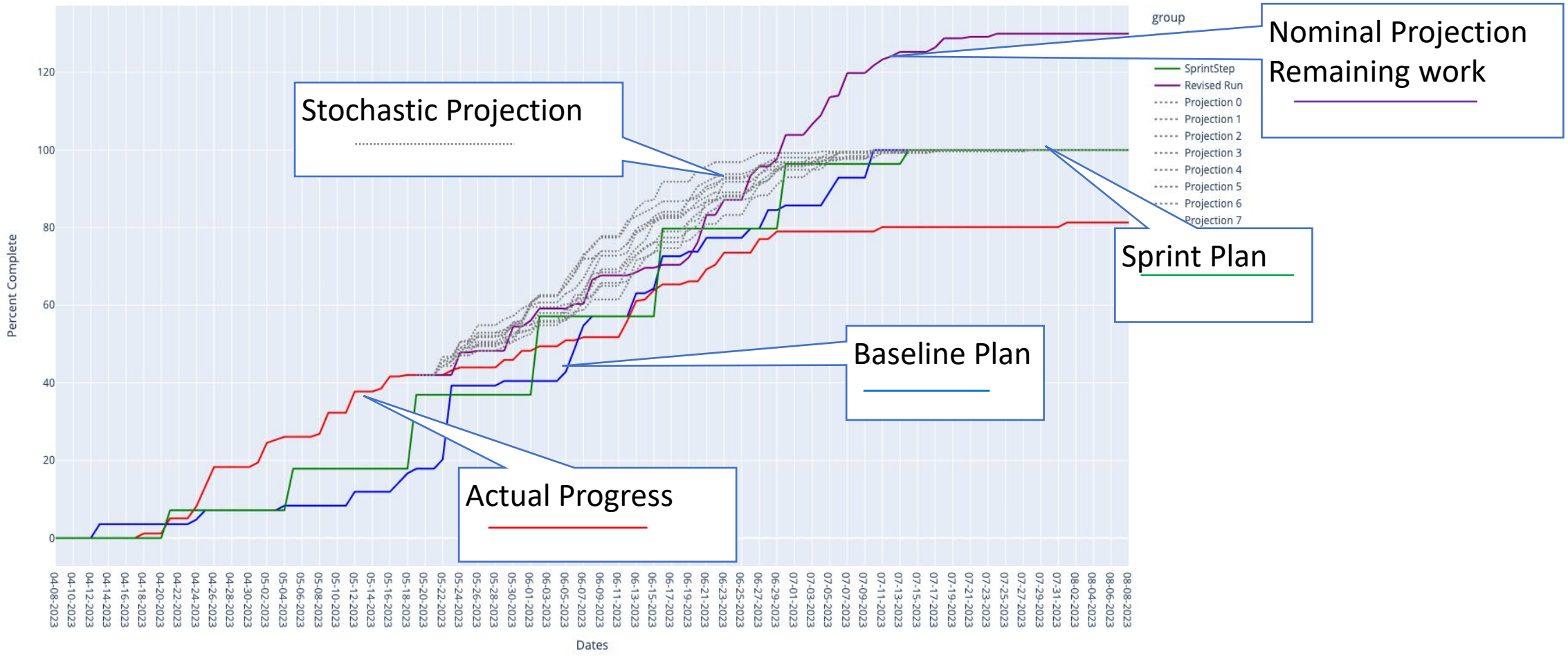
Analysis: What is% Completion, When will we Finish?



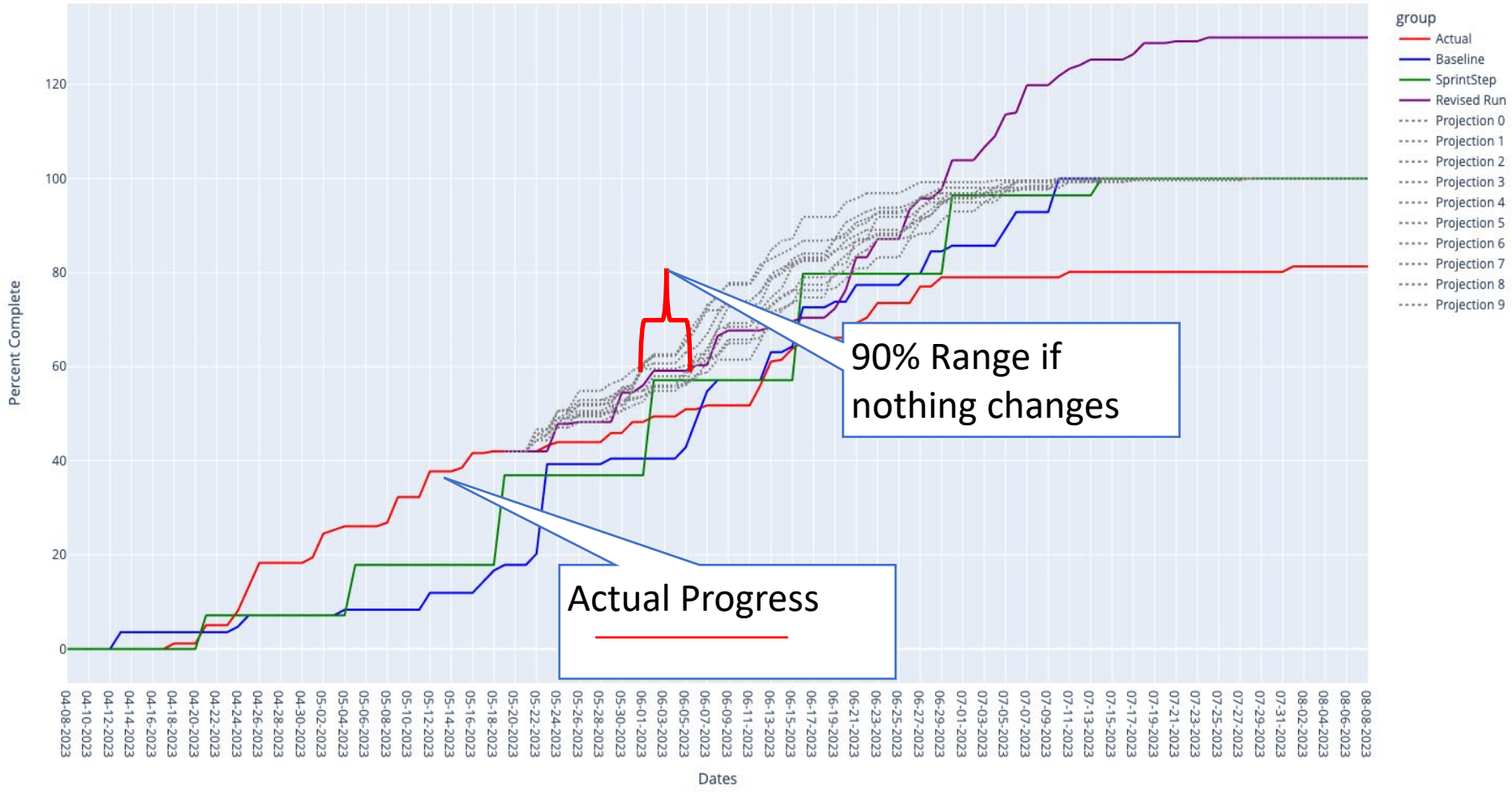
Analysis: What is% Completion, When will we Finish?



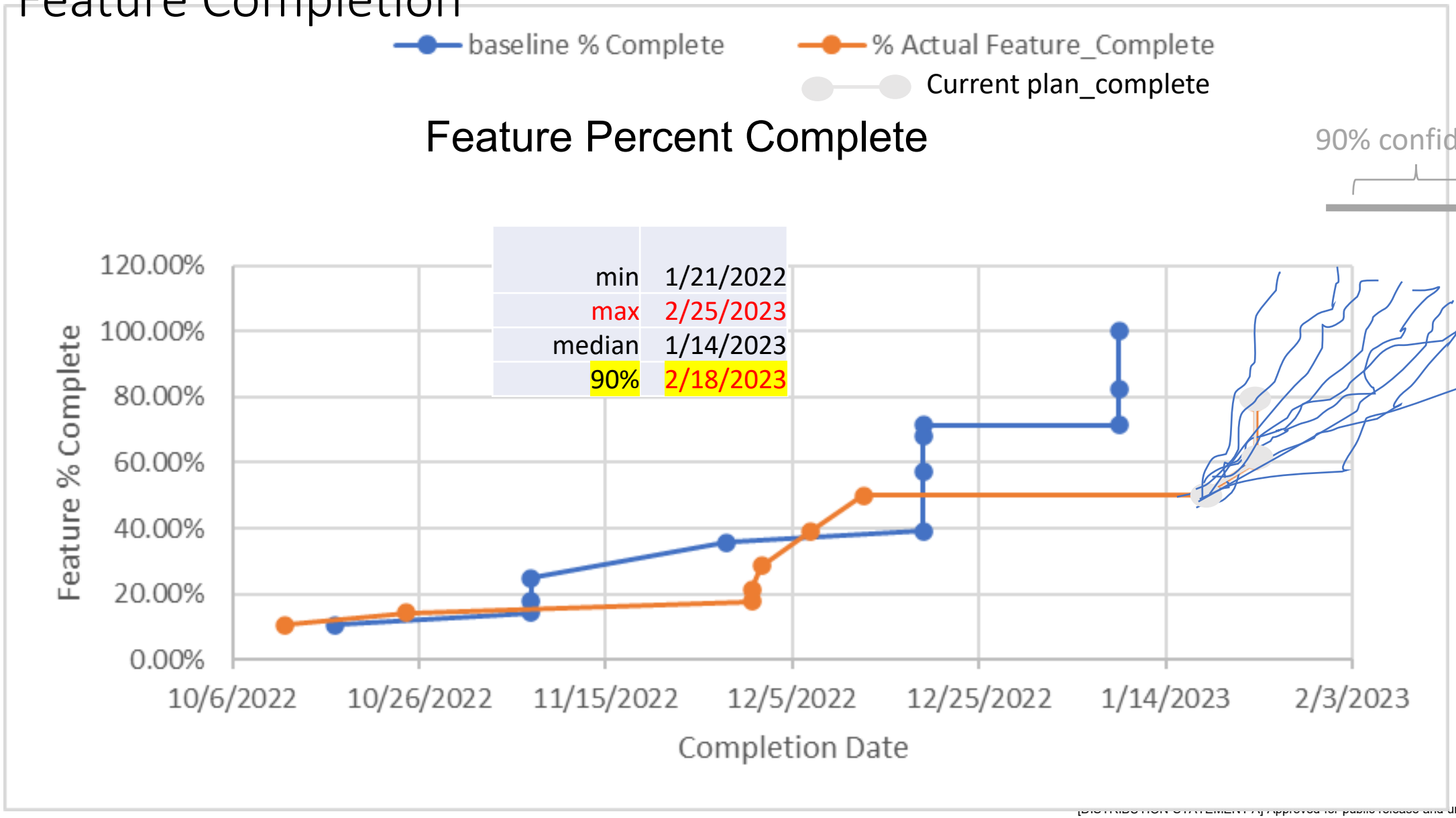
Analysis: What is% Completion, When will we Finish?



Analysis: Did Something Change?

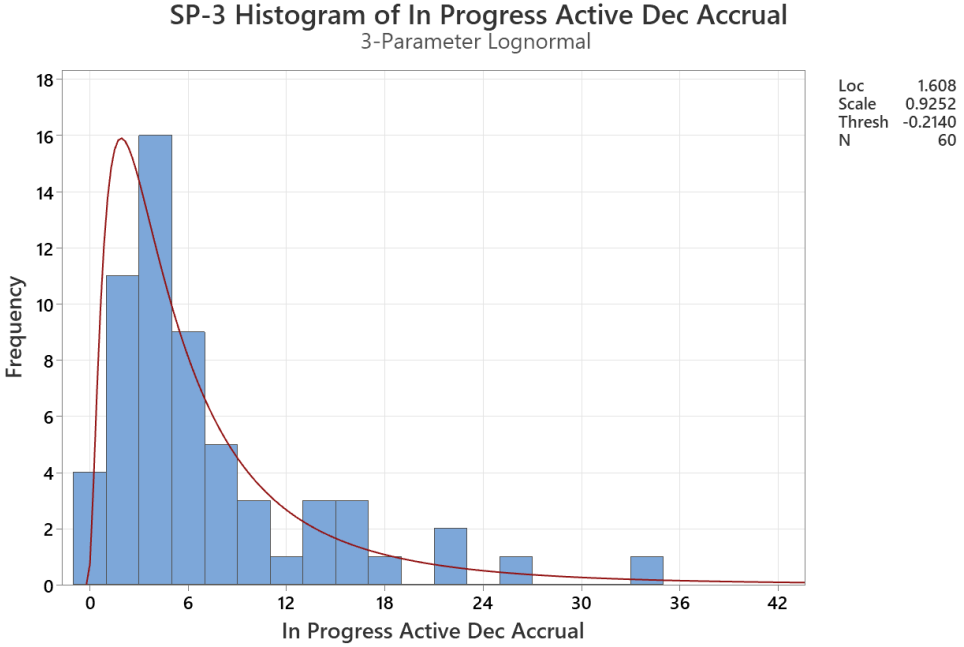


Feature Completion



Measure Estimation Accuracy and Precision

Calibrate for each pipeline.
Determine that work is predictable **within a range**.
Estimate all work: program planning, road mapping, program increments, sprints.
Measure the work as it passes through the planning, backlog and pipeline.
Compare estimates with actuals.



While the estimation is stable, use the ranges to make future predictions.

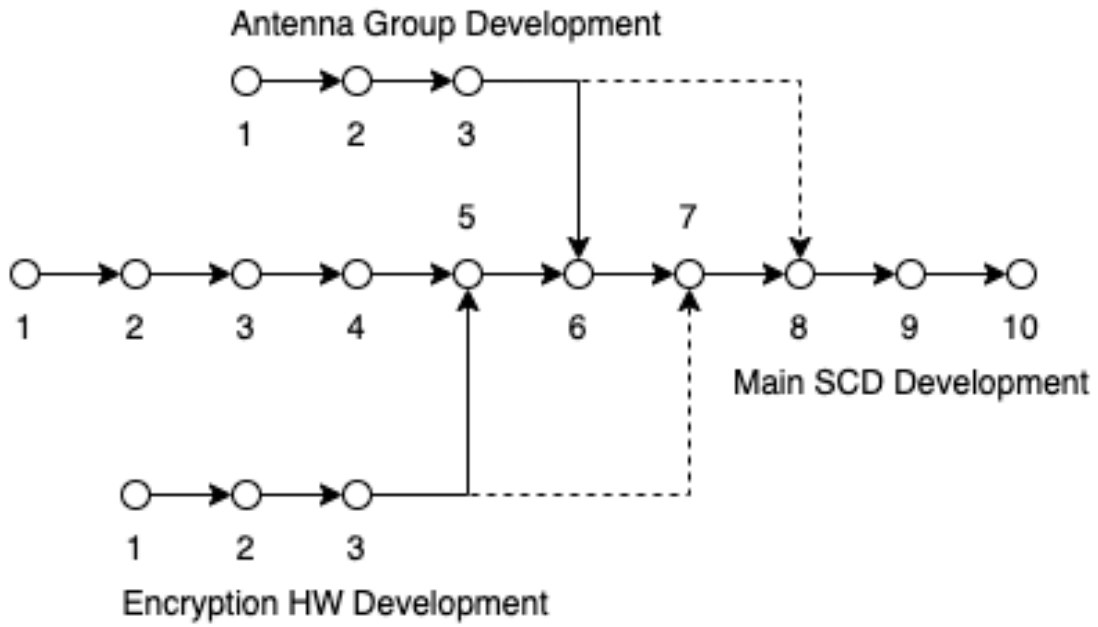
Statistics

Total Count	Mean	SE Mean	StDev
60	7.185	0.883	6.839

Minimum	Q1	Median	Q3	Maximum
0.335	2.583	4.599	8.965	33.44

Mode	IQR
8	6.382

Pipeline of Pipelines PoPs Workflow Network Example



Model a fictitious device that captures characteristics of a real project dependencies between hardware and software capabilities.

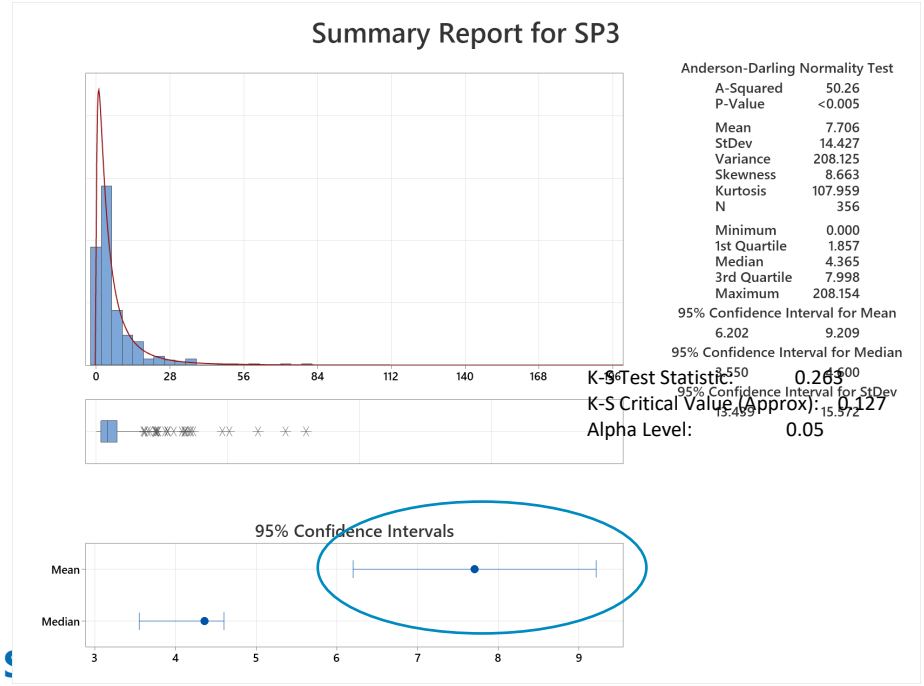
Each pipeline produces **dependencies** used to model schedule, cost, and technical performance risks resulting from production variation, accumulated variance, and rework.

Each pipeline has it's own personality and parameters.

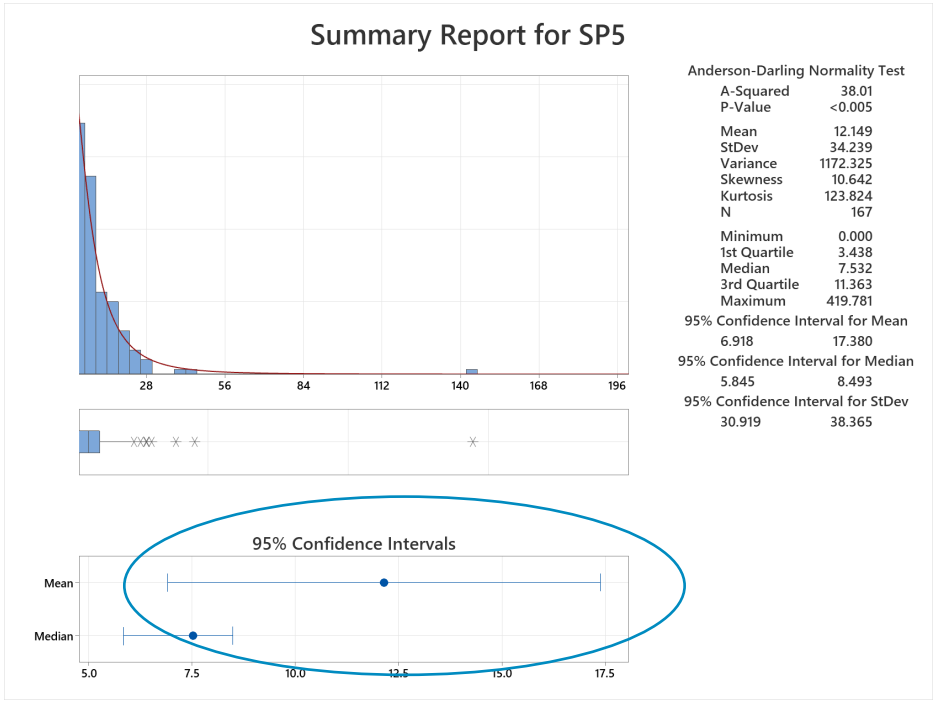
Each category of work item has distinct parameters.

Report dependencies in **range of days**.

Stories are **known unknowns, with** predictable distributions



K-S Test Statistic: 0.263
K-S Critical Value (Approx): 0.127
Alpha Level: 0.05

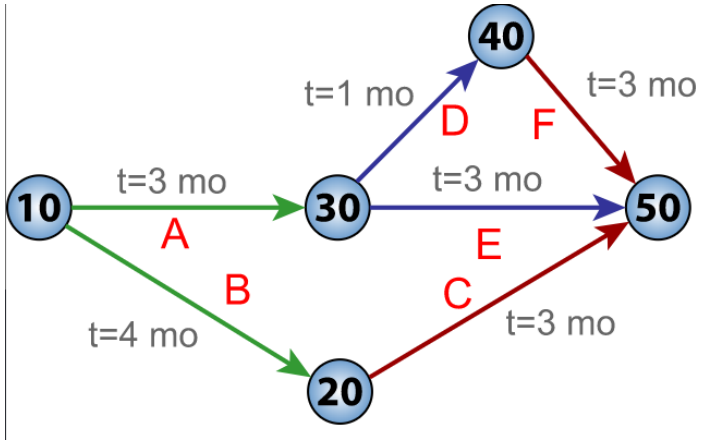


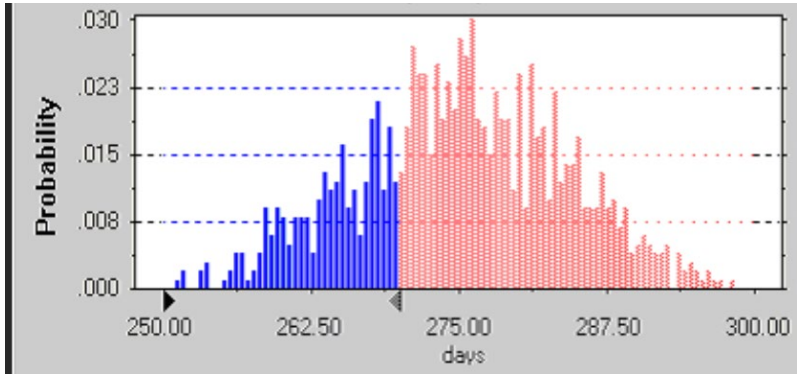
Variable	Count	Mean	SE Mean	StDev	Min	Q1	Median	Q3	Max	IQR
SP3	356	7.70	0.765	14.427	0.000	1.857	4.365	7.998	208.2	6.1
SP5	176	12.15	2.65	34.24	0	3.44	7.53	11.36	419.8	9.92

Different Distributions, but
Overlapping confidence intervals
On the Mean,
Distinct on Median

Q:Why are some outliers?
Q:Can we parameterize the distribution?

Extend Multi-Pipelines with Networks





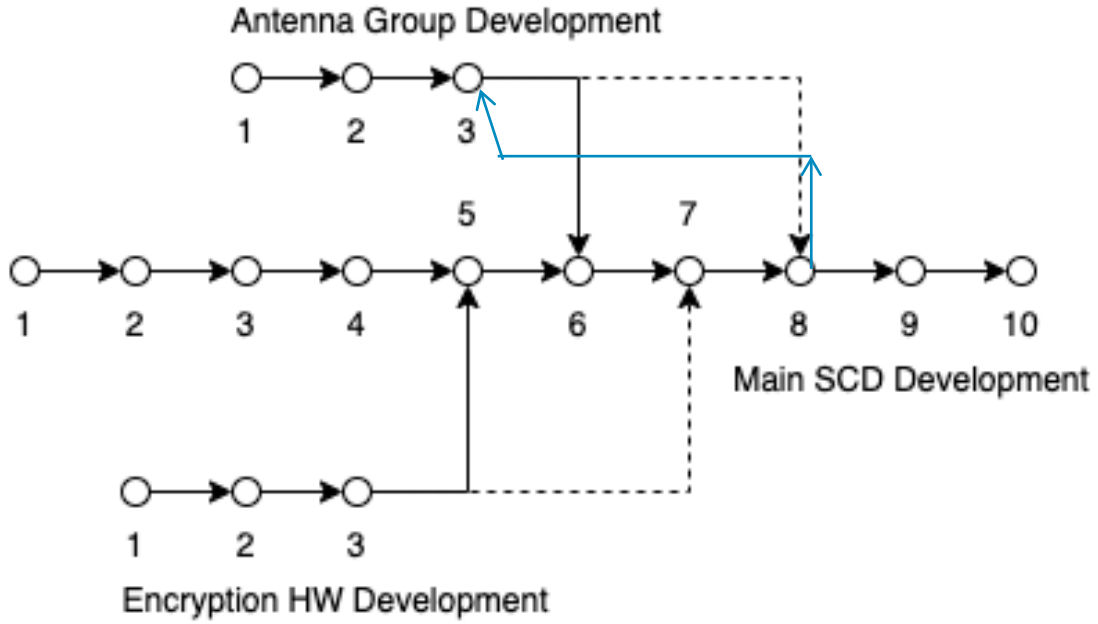
Approach

- Trace work item through development steps
- Identify blockers and integration points
- Probability of completion date

Data from DSO pipeline and other sources

- Product state node structure (capability based WBS, product dependencies, workflow)
- For each Pipeline obtain empirical data for
  - **Effort Rate** and **variation** (by skill?)
  - **Production Rate and variation** by **work type**
  - **Primary work** and **Rework** by activity
  - Defect Rates and fix latencies (build, test)

Pipeline of Pipelines PoPs Workflow Network Example



Each produce dependencies used to model schedule, cost, and technical performance risks resulting from production variation, accumulated variance, and rework.

Pipelines have backward loops. Sometimes with themselves.

We model likelihood and latency in these loops.

Little's Law assumptions are strongly violated except for **some** linear pipeline segments.

Typical Flow Metrics do not accommodate rework, merges, or multiple entry points

Observations and Lessons

Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213

Lessons Learned

Carnegie Mellon University Software Engineering Institute

Measurement tools are siloed, making it hard to work together.

Different types of work need to be accounted for (product, bugs, research …). They have different characteristics, estimation accuracy, and outputs.

Averages alone don't support information needs high priority changes or statistical modeling. We need distributions.

We need more specific lead time measures for process steps and baselines for zero rework lead times. (total time until test completes is a candidate quality proxy).

Typical flow metrics don't appear to apply to the pipeline-of-pipelines because of branching and other assumptions violations.

Measuring staff availability remains unresolved.

ACE/PoPs Links for Program Managers

Carnegie Mellon University Software Engineering Institute

Program Managers—The DevSecOps Pipeline Can Provide Actionable Data
https://resources.sei.cmu.edu/library/asset-view.cfm?assetid=890538

https://insights.sei.cmu.edu/blog/actionable-data-from-the-devsecops-pipeline/

Why Your Software Cost Estimates Change Over Time and How DevSecOps Data Can Help Reduce Cost Risk
https://insights.sei.cmu.edu/library/why-your-software-cost-estimates-change-over-time-and-how-devsecops-data-can-help-reduce-cost-risk/
https://insights.sei.cmu.edu/library/actionable-data-in-the-devsecops-pipeline/

How to get started
Getting Started with ACE/PoPs
https://resources.sei.cmu.edu/library/asset-view.cfm?assetid=890665

Coming
Starter packs

Next Steps

Increase number of GQIM indicator Templates

Improve projection precision

Experience Package

Reports

- Life Cycle Measurement Gap
- Non-Functional Requirement
- Capability Based Estimation Literature Review

Prepare for a Transition, search for transition partners

Stay out of the **SWAMP** (SoftWare Analysis Metrics Pool)

Automation requires precise definitions.

Every measurement has precise meaning in a known context.

Every metric supports an information need.

Disciplined work decomposition (WBS) connected to product deliverables

Categorize and estimate work items

Automation support for workflow (e.g. Jira/GitLab) and technical implementation (DevSecOps)

Consistent workflow with start and finish

Minimize humans in the loop for data collection

Call to Action

Would you benefit from continuous updates to status and projections?

Are you using DevSecOps tool chains, issue trackers, and workflow management?

Can you share process data and discuss results?

Will you participate in our quarterly research review?

We can help!

- Share our Program Management Measurement White Paper
- Specify information, data, and displays for your program management
- Recommend approaches and tools to get started
- Evaluate your results for effectiveness

ACE/PoPs Team

Carnegie Mellon University Software Engineering Institute

William Nichols
Principal Engineer
Telephone: +1 412.268.1727

Brent Clausner
DevOps Engineer

Luiz Antunes
DevOps Engineer

Julie Cohen
Senior Engineer

Email: info@sei.cmu.edu

ACE/PoPs Team

Carnegie Mellon University Software Engineering Institute

Brigid O'Hearn
Senior Engineer

Chris Miller
Senior Researcher

Anandi Hira
Senior Researcher

Michael Bandor
Senior Engineer