Telling More: a 2006 NASA SARP project (Executive Briefing)

Tim Menzies
Lane Department Computer Science & Electrical Engineering
West Virginia University
tim@menzies.us

Important disclaimer

- This is a very preliminary report.

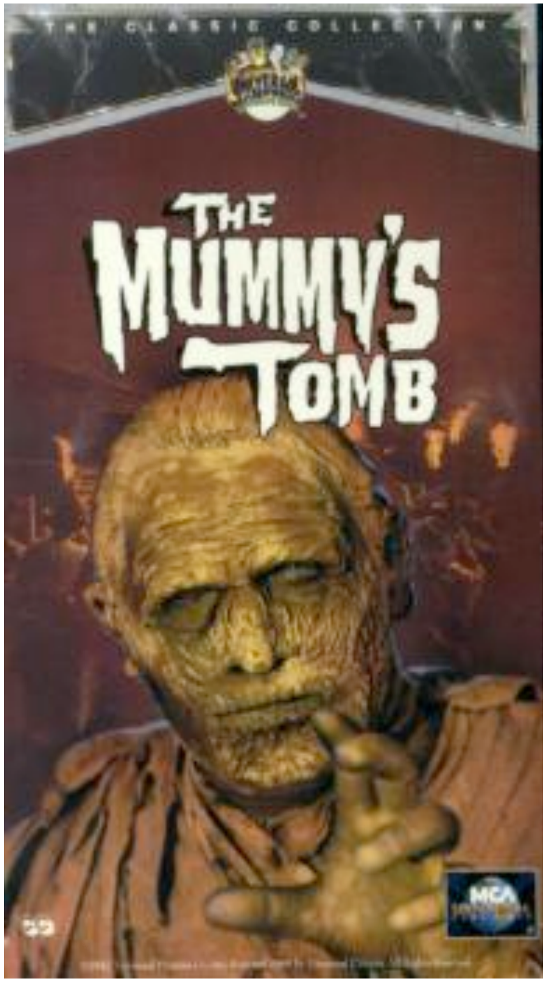
- While the planning&scoping team at Fairmont is currently reviewing this material…

  - … the views expressed here-in are the author's.
  - They do **NOT** reflect official NASA policy
  - They do **NOT** reflect the views of NASA civil servants.

Problem

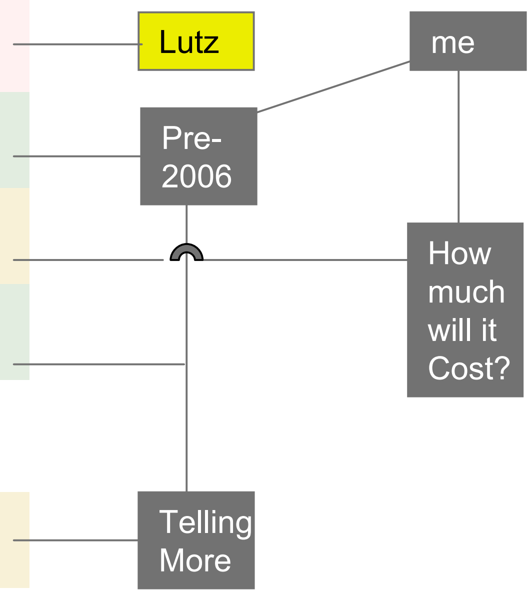
- NASA IV&V is in a unique position to review and comment on much of the NASA software enterprise.
  - We *see more*,
    but what have we *learned*?
  - What can we *tell*?

- NASA data = active repository?
- Or a data tomb?
  - write once;
  - read never;
  - buried;
  - doomed;



Benefits & Importance of the research

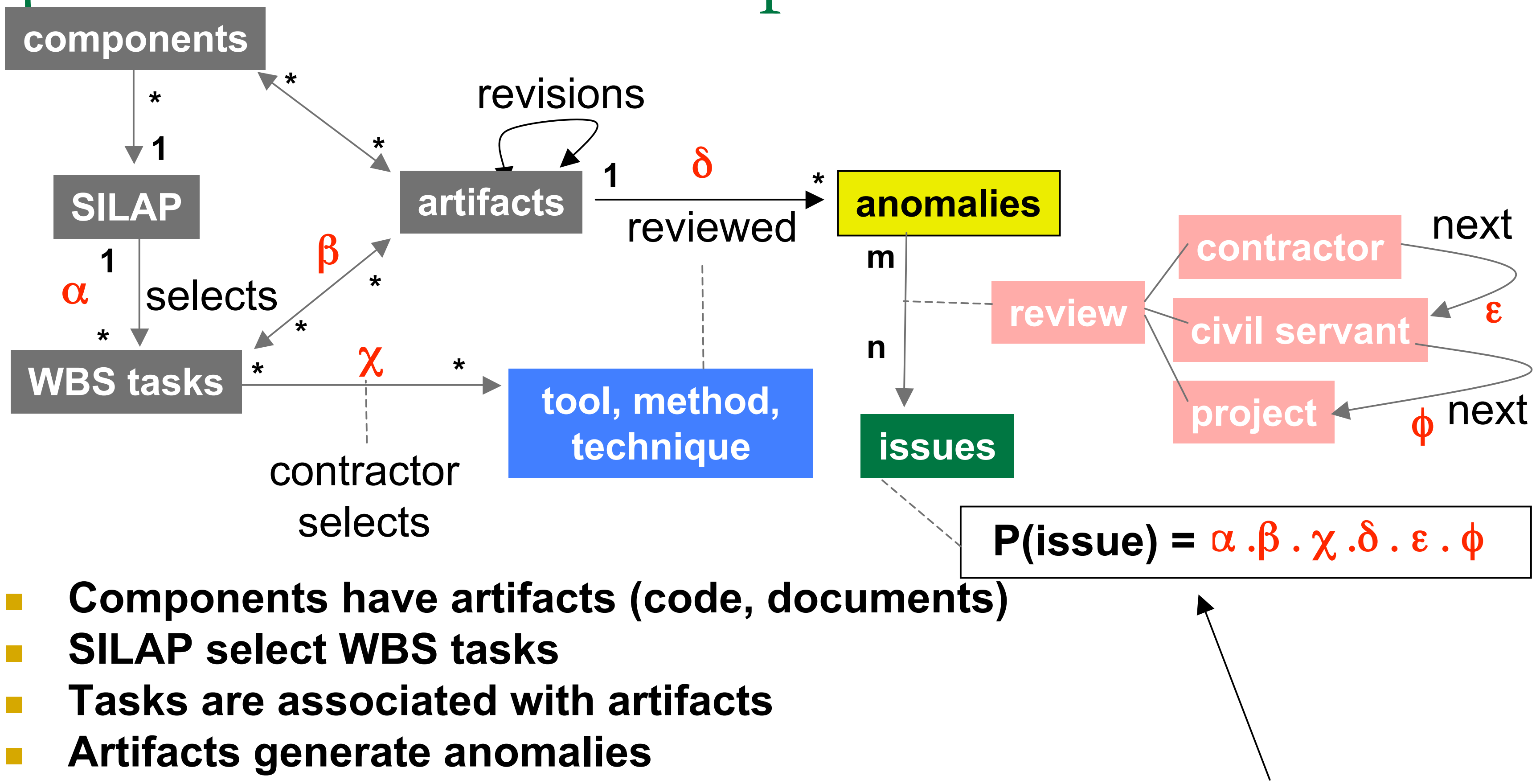
- If an organizations spends millions of dollars on data collection and archiving...
  - It should spend tens of thousands (at least) in analyzing that data.
- Surprises, found before by SARP projects:
  - Largest source of post-launch *deep space* anomalies is *ground systems* [Lutz,2004]
  - Common conflation of *severity* and *priority* in NASA defect logs [me,2002]
  - *Small* changes in data) *massive* changes in cost estimation [me,Hihn,2005]
  - Static code measures surprisingly good at predicting for issues [me, 2005]
- And the discoveries continue:
  - IV&V tasks often the *same*, despite processing different *projects* [me, today]
- What might we learn tomorrow?



Approach: Use AI

- Apply AI machinery to NASA data repositories.
  1. When data is plentiful, use *data miners*. (e.g. cost and defect estimation)
  2. When data is scarce and domain intuitions exist:
     - build a *what-if* simulator for those intuitions,
     - Monte Carlo the simulator,
     - goto 1
  3. When data and intuitions exist, use Bayesian belief nets;
     - e.g. Dabney, Fenton, etc etc
  4. When complex domain models exist, use semantic web tools to generalize from *here* to *there*.
     - Find and collect available data
     - Maintaining security and confidentiality requirements.
     - Match data sources to available machines
     - Apply the machines.

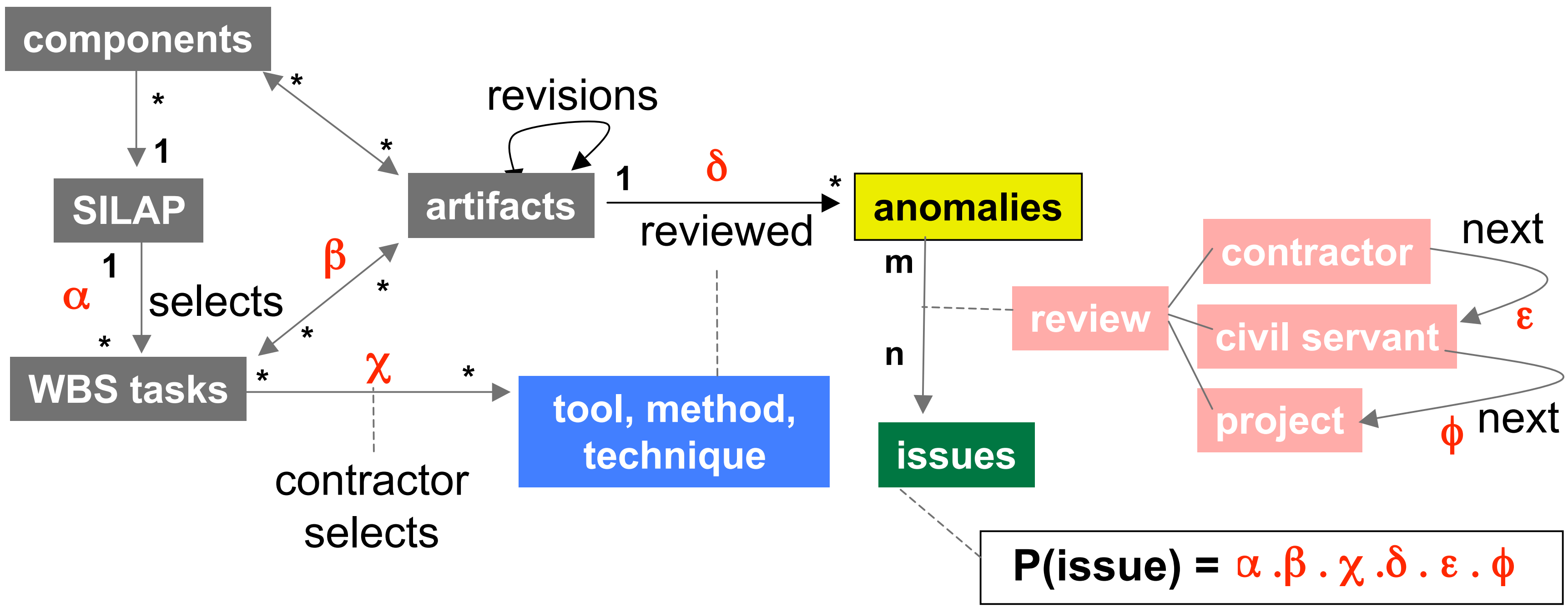
Approach (2): find critical parts of the business process



- **Components have artifacts (code, documents)**
- **SILAP select WBS tasks**
- **Tasks are associated with artifacts**
- **Artifacts generate anomalies**
- **Filters reject bogus anomalies**
- **Projects accept issues**

So, what do we know about these distributions?

Accomplishments



Working, left to right

- 2006: learn α

Input:

- MB1: 500 CSCI (a.k.a. sub-systems)
- Use these as inputs to SILAP; determine:

Output:

- what tasks we are doing most/least
- Is there a difference in tasks selected based on project type?

Accomplishments

```
DT3= use of defect tracking
CL3= CMM level
US3= use of standards
EX3= experience
HS2= human safety
```

```
DT3 <= 1: group2: 150
DT3 > 1
| CL3 <= 4
| | US3 <= 2: group0: 170
| | US3 > 2
| | | DT3 <= 2
| | | | EX3 <= 2: group3: 25
| | | | EX3 > 2
| | | | | HS2 <= 3: group0: 14
| | | | | HS2 > 3: group3: 6
| | | DT3 > 2: group0: 42
| CL3 > 4: group1: 19
```

tests on extreme values

Good news: many projects use elaborate defect tracking tools

- Surprise: IV&V WBS task selection NOT determined by…
  ???
  - "orbit": ground system, earth orbit, transfer, ground ops on different planet
  - "profile": type of science
  - "prime": who built the CSCI
- Rather, what distinguishes NASA software projects is project's willingness to
  - Reflect on its own process
    - E.g. use of standards
  - Let others reflect on their process
    - E.g. use of defect tracking tools
  - The development **process** is more important that the project **goal** (at least, for the purposes of IV&V)

Cornerstone values : what most separates the current sample of NASA projects seen at IV&V
- Over **half** the SILAP variables do not appear here
- Spend more time making certain that the above variables are **scored correctly**
- So there is a **standard type of NASA project** currently getting IV&V?
  - If that type changes, then should current IV&V practices change?

Next steps

SARP research working closely with NASA business

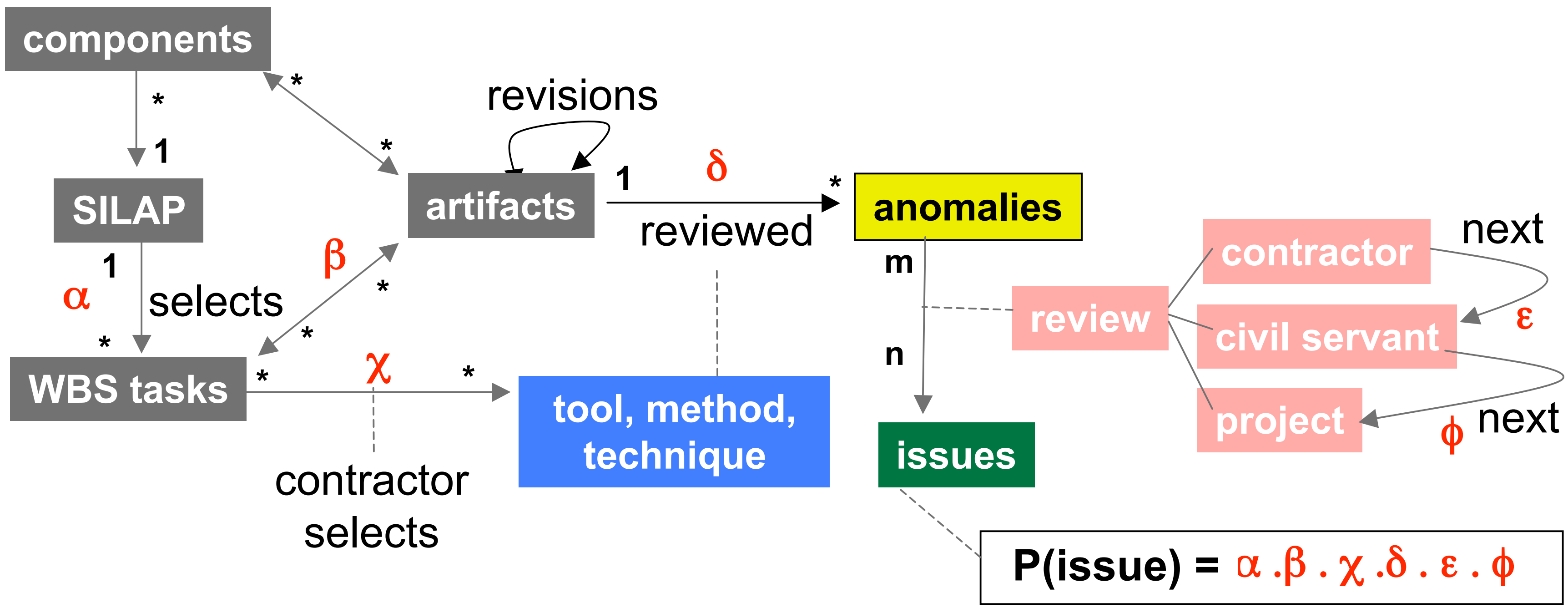
WVU Research team ↔ NASA IV&V Planning & scoping team

- More data mining

Can AI/ data miners simplify/ optimize any of that discussion?

- Review/change SLP 9-1 (the WBS)
- revising SILAP doco
- identified areas for ?change:
  - right tasks selected by scores?
  - are factors the best selectors?
- Is the criteria sufficient/ correct?
  - Map factors directly to tasks?
  - Study planned vs actual to find a “best” or most common architecture breakout?
- Minimum set of tasks needed to add IV&V value?
  - defining sets of "common" tasks for specific types of functions?
- ?? break code analysis out into
  - tool execution only?
  - tool execution + review of results?
  - full-up code inspection?
- Etc etc

Next steps (general)



Working, left to right

- 2006: learn α
- 2006+: learn the rest
- Determine how to optimize IV&V task selection