SYSTEMS ENGINEERING RESEARCH CENTER

PURDUE UNIVERSITY

Data science approaches to prevent failures in systems engineering

Sponsor: DASD(SE)

By
Prof. Karen Marais and Prof. Bruno Ribeiro
9th Annual SERC Sponsor Research Review
November 8, 2017
FHI 360 CONFERENCE CENTER
1825 Connecticut Avenue NW, 8th Floor
Washington, DC 20009

www.sercuarc.org

Project failures occur despite systems engineering best practices




Project delays, cost overruns, quality concerns, cancellations…

Why aren't these methods helping (as much as we hope)?

Several possible reasons…

① They rely on extensive data creation, collection, and tracking, which is hard to do

② We think they are not useful, and so they are not

Our core ideas:

① risk assessment based on the “real reasons” for systems engineering failures,

and

② augment existing data with about team assessments, Wisdom of the Crowd (WoC), to uncover problems and likely “real reason” causes

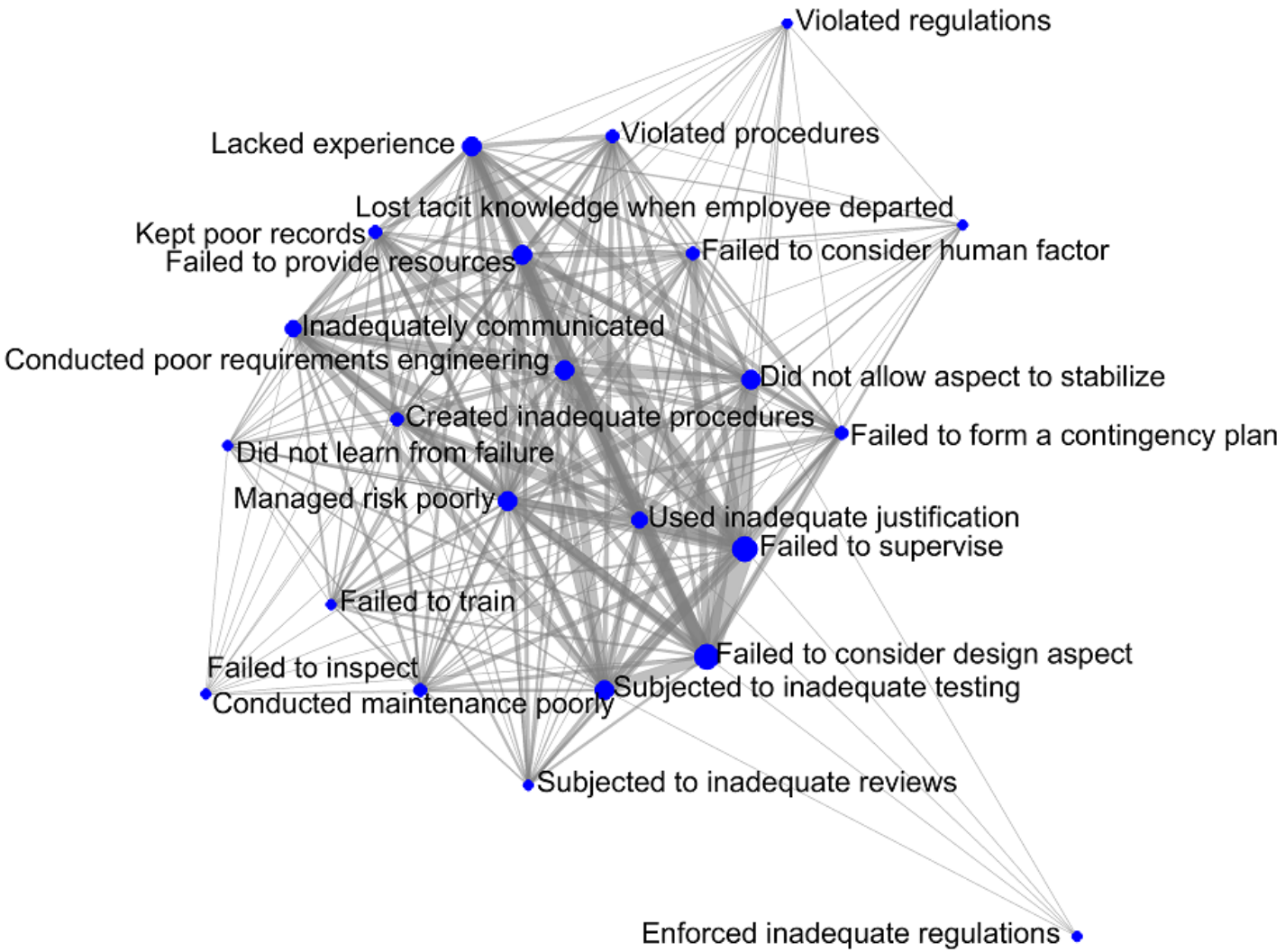
Most systems engineering failures do not involve black swans

Most failures result from rather prosaic and predictable white swans:

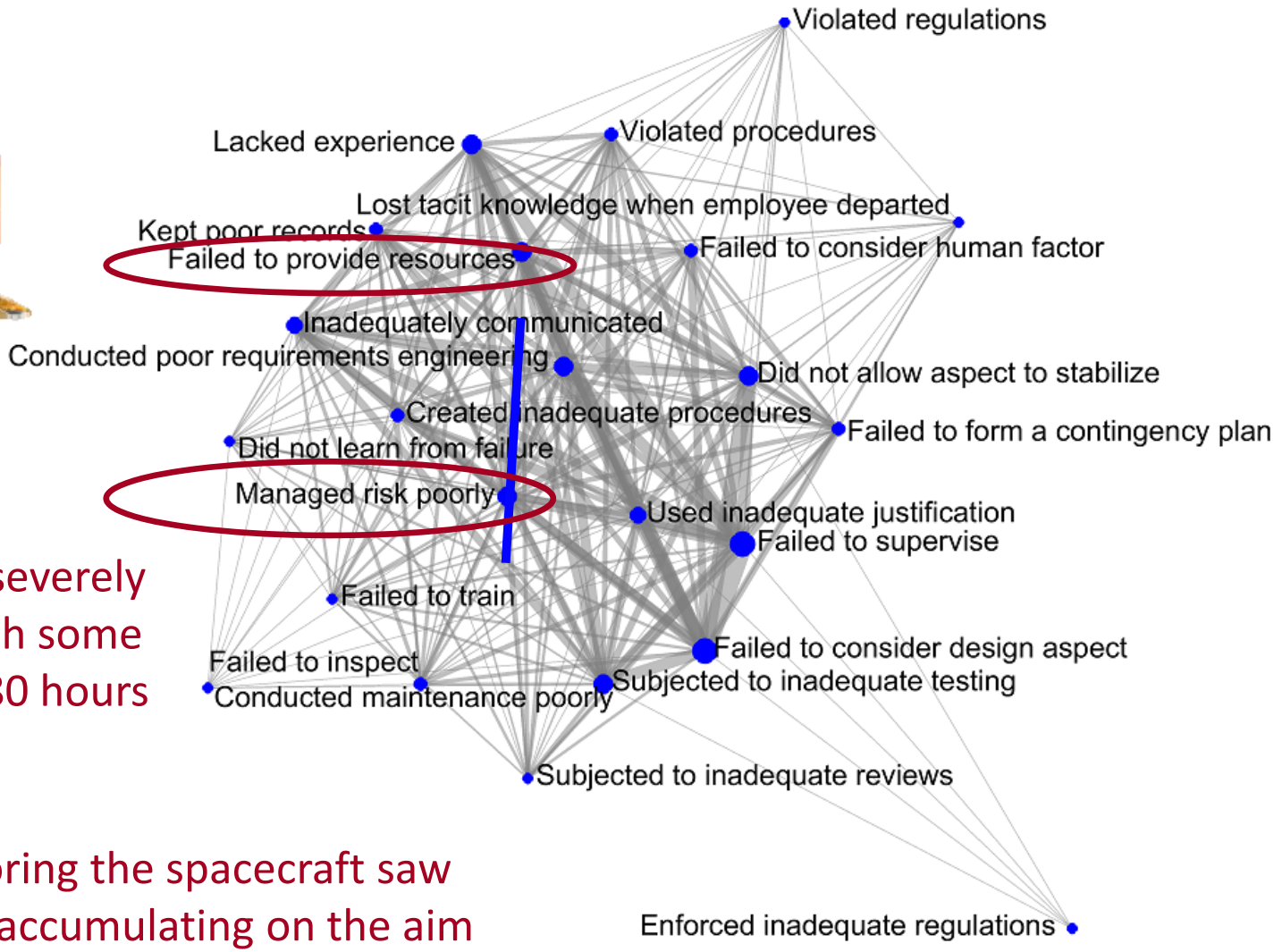
Lost tacit knowledge when employee(s) departed	Subjected to insufficient testing	Created deficient requirements
Failed to provide resources	Violated regulations	Failed to inspect
Used inadequate justification	Violated procedures	Subjected to inadequate reviews
Failed to form a contingency plan	Managed risk poorly	Kept poor records
Failed to consider systems factor	Created deficient procedures	Failed to supervise
Lacked experience	Enforced deficient regulations	Did not allow aspect to stabilize
Failed to consider human factor	Did not learn from failure	Failed to maintain

Diane Sorenson and **Karen Marais**, "**Patterns of Causation in Accidents and Other Systems Engineering Failures**," IEEE Systems Conference, April 2016, IEEE, Orlando, FL.

The failure cause network shows how these causes relate to one another



Consider the Mars Climate Orbiter failure



The project was severely understaffed, with some people working 80 hours per week.

The team monitoring the spacecraft saw that errors were accumulating on the aim point for the spacecraft, but did not investigate.

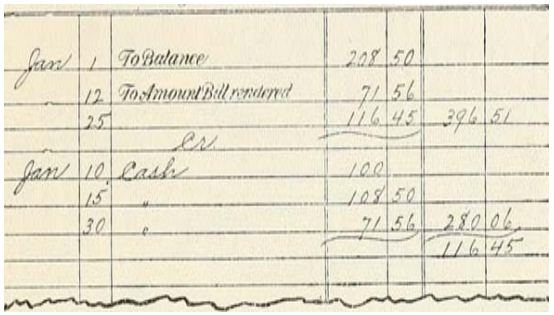
People knew something wasn't right…

Can we develop ways to get this information?

(without requiring ever more paperwork…)

Big data can help

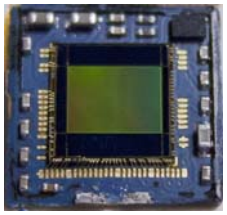
Expectation of Machine Learning/AI/Big Data

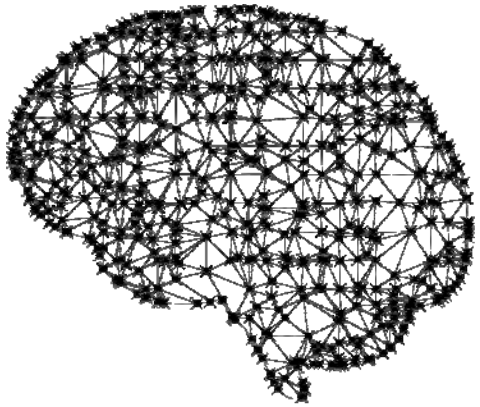


Human bookkeeping

+

Automated data collection (sensors)





Machine Learning / AI

=

insights

Prevent, forecast failures

Reality of Machine Learning/AI/Big Data

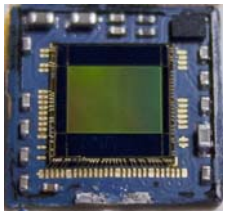
Estimated MTBF	~10 years? Check later with Emily	-
If sensor measurement is wrong do…	TODO: Describe our work-around. Need to revisit this later.	-

Data collection not high priority when project is in trouble

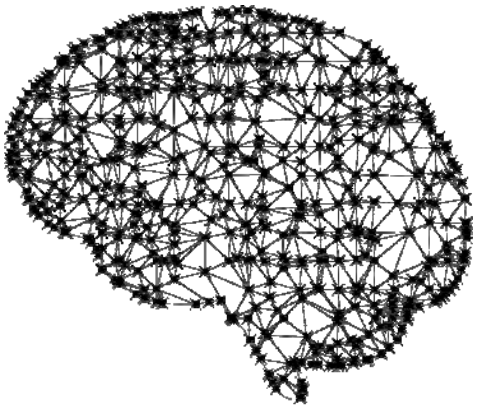
Human bookkeeping
(Low priority in crunch time)

+

Automated data collection
(sensors)



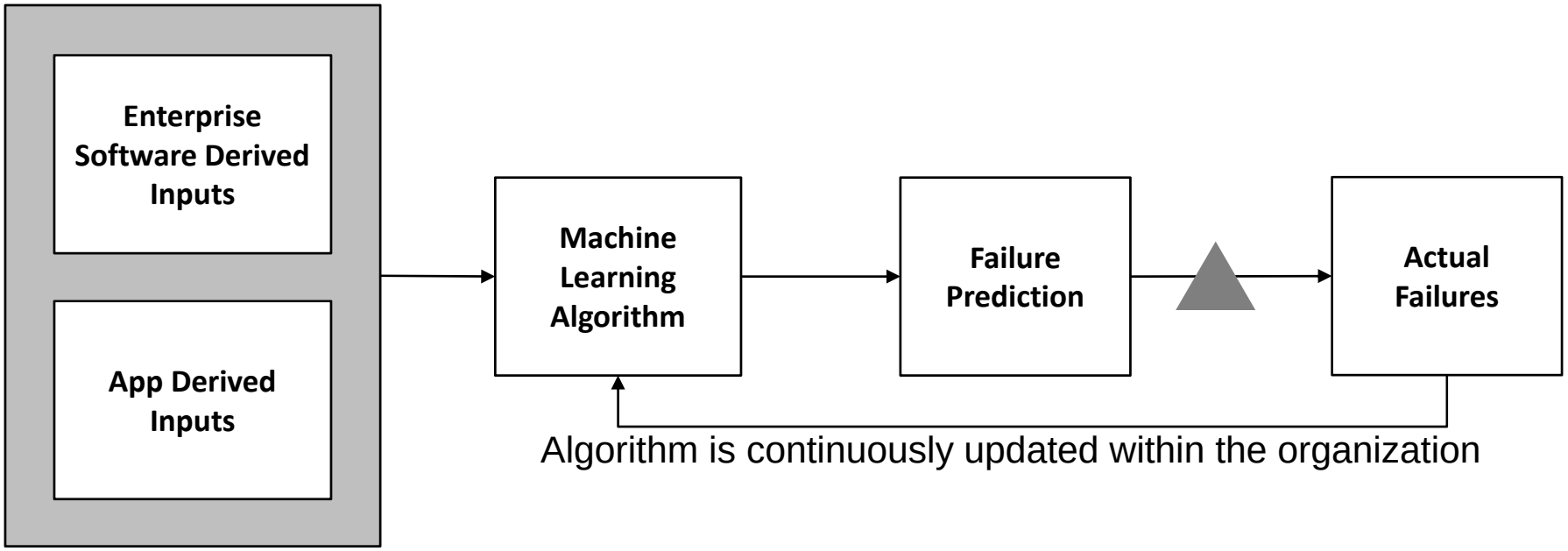
Low priority: faulty, incomplete when needed most



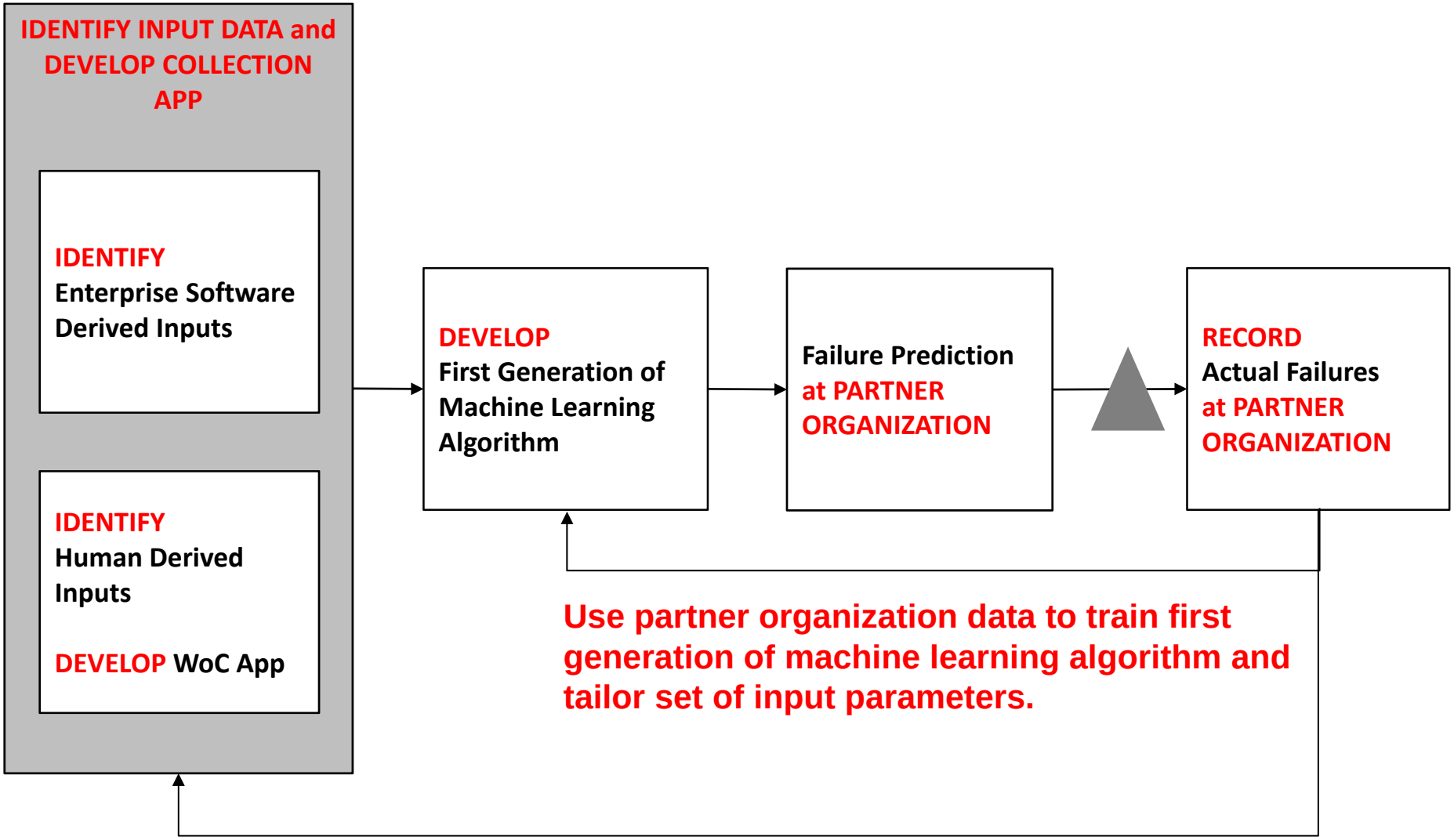
Machine Learning / AI

=

We propose a tool that uses both existing data and Wisdom of the Crowds to help predict failures

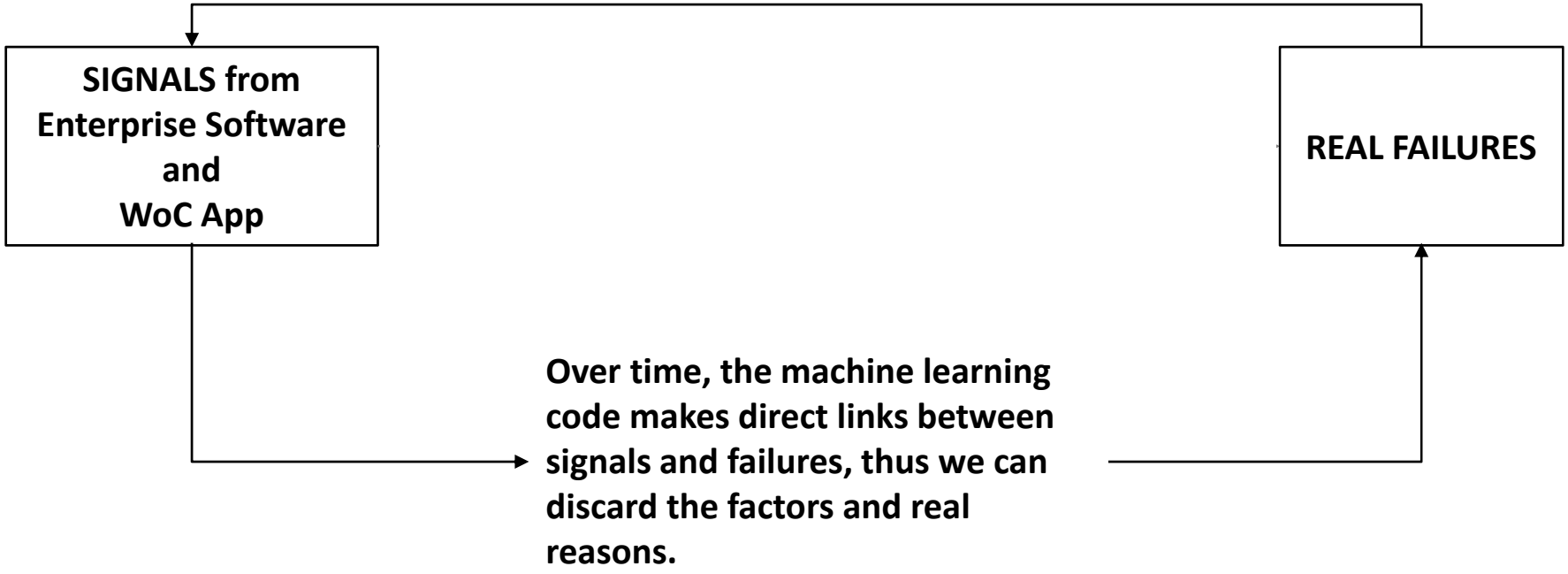


How will we get there?

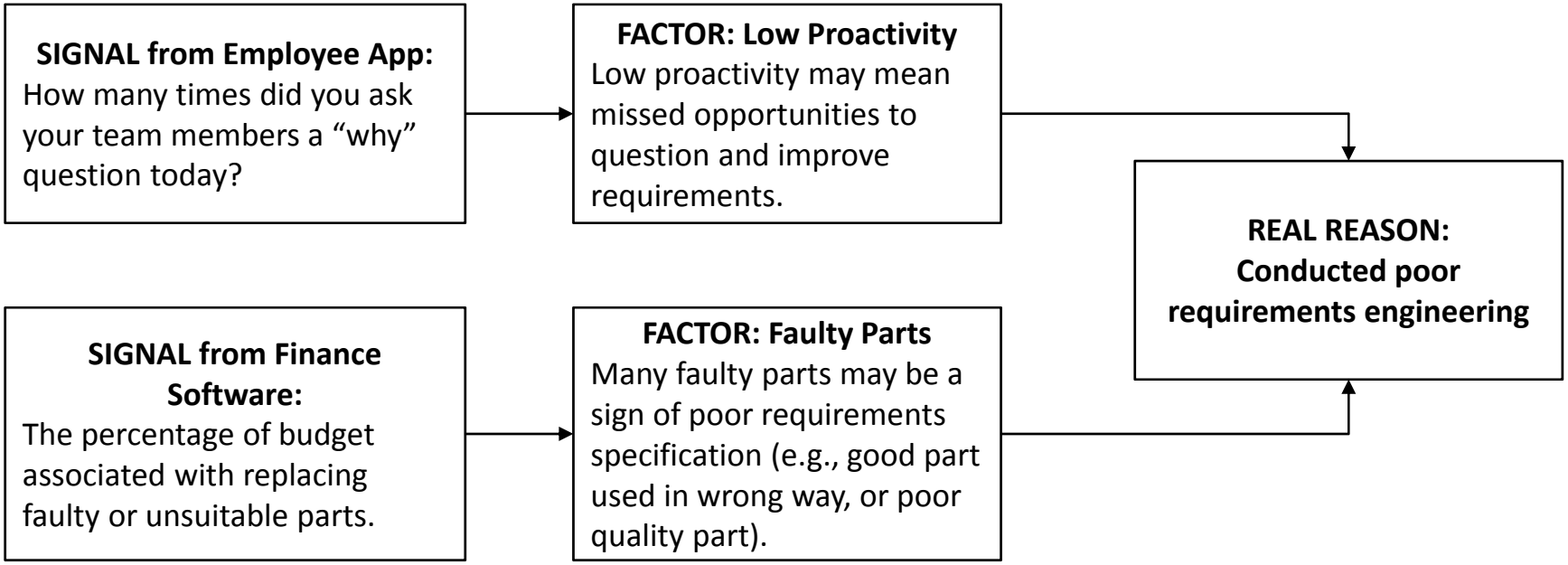


Identifying input data

Use student and partner organization data to train first and second generations of machine learning algorithm and tailor set of input signals.



How does input data relate to the real reasons?



Wisdom-of-the-Crowd Information

Accurate Wisdom-of-the-Crowd Predictions from Incomplete Pictures

An expert with a complete view and understanding of the entire process may be able to give a reasonable assessment of potential problems and delays.

As projects become more complex, this assessment is increasingly hard and dedicated experts are no longer able to have a complete view.

(Wisdom of the Crowd)
Can we use the assessments of non-experts with partial views to train neural networks to learn to

a. Predict success and failures using non-experts with incomplete (possibly biased) information?

b. Ask relevant questions to these non-experts to help make the data richer to better predict success and failure?

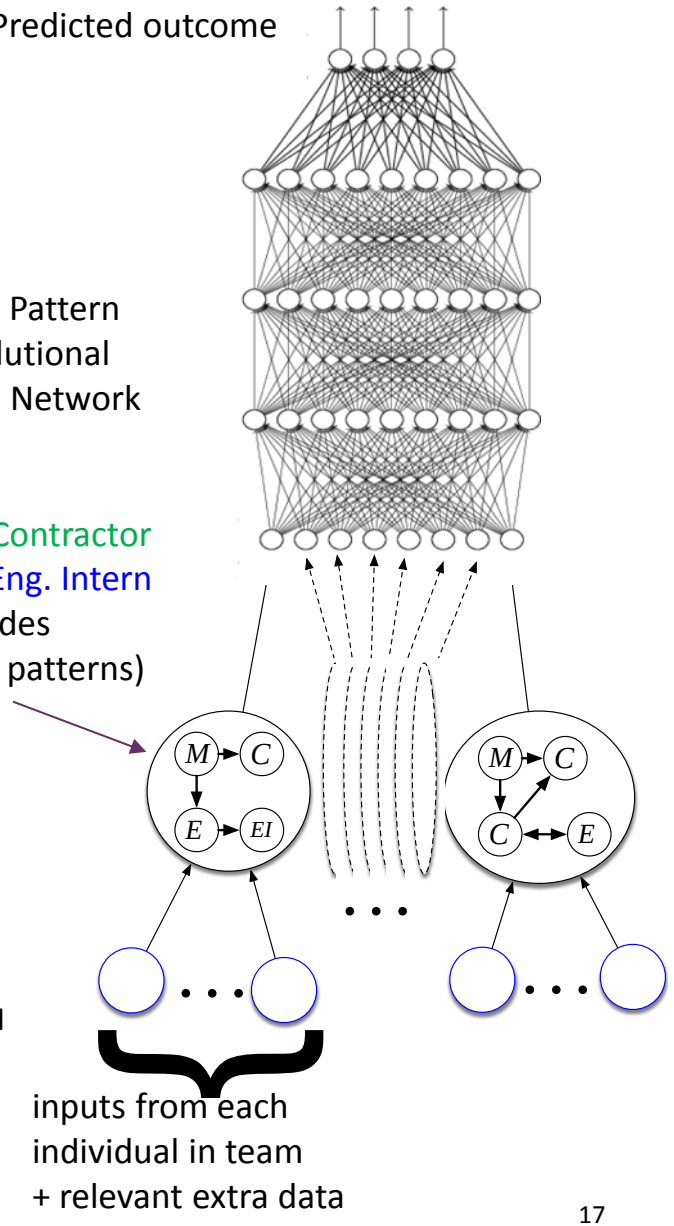
Predicting Outcomes from WoC Inputs

Hypothesis: Non-expert opinions and their relationships can help predict project outcomes

Approach: Use our newly developed Sparse Pattern Convolutional Neural Network (SPCNN), to learn dynamic relational dependencies between group actors, their opinions, and project outcomes.

Input: WoC team member assessment of project health, WoC assessment of potential personal issues, team structure, traditional indicators

Output: Predicted outcome of project milestones

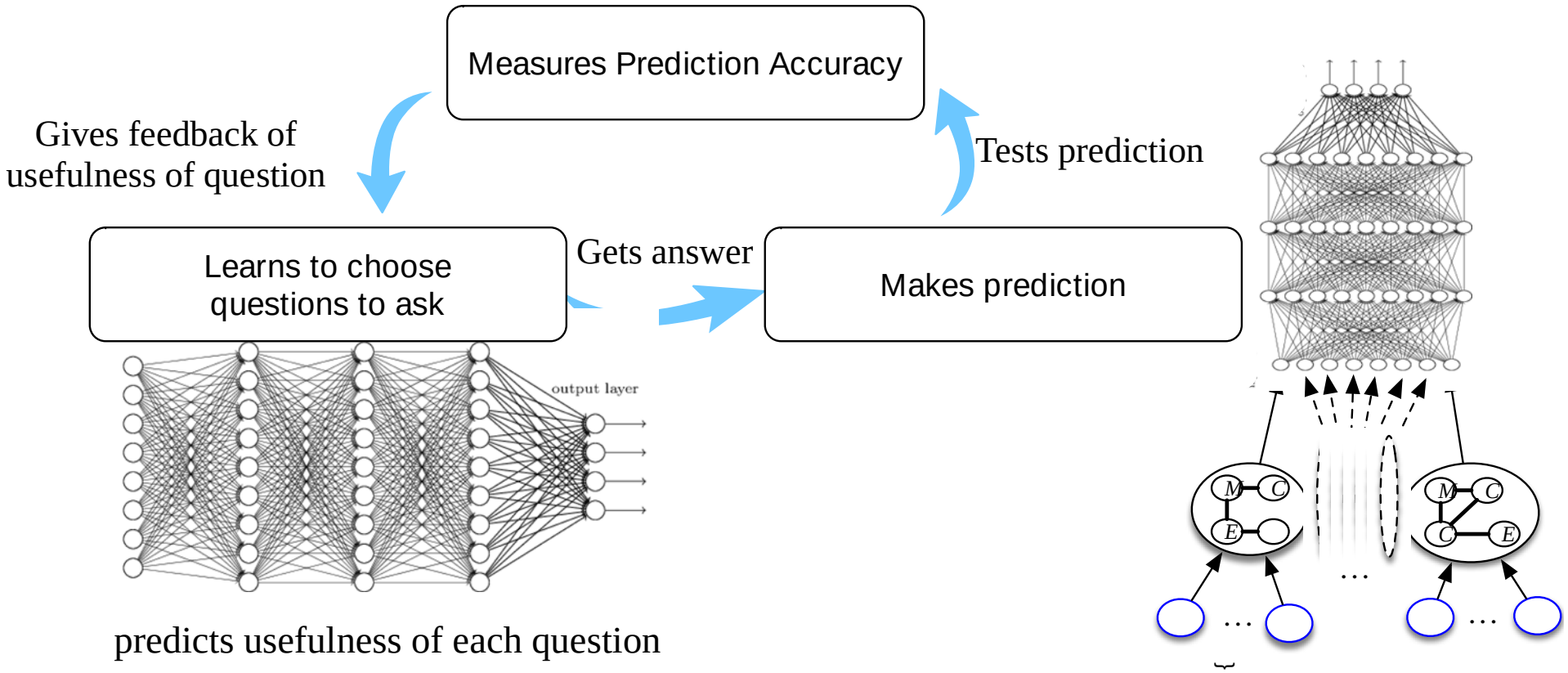


Meng, C., Sekar, C., Ribeiro, B., Neville, J., 2017, **Predicting Subgraph Evolution in Heterogeneous Dynamic Networks**, *(preprint)*.

Yang, J., Ribeiro, B., Neville, J., **Should We Be Confident in Peer Effects Estimated from Partial Crawls of Social Networks?** AAAI Conference on Web and Social Media (ICWSM), 2017

Active Learning + Contextual Bandits

- Problem: Too many questions we would like to ask
  - —Must limit the number of questions to ask (avoid subject fatigue)
  - —Which questions/answers most correlate with outcome?
    - o Active learning approach to learn which questions to ask



Vision for Product Development: **Year One** and **Future**