Software Quality: Moving Research To Practice

Ahmed E. Hassan

NSERC / RIM Research Chair
in Software Engineering of
Ultra Large Scale Systems
Queen's University, Canada

http://sail.cs.queensu.ca

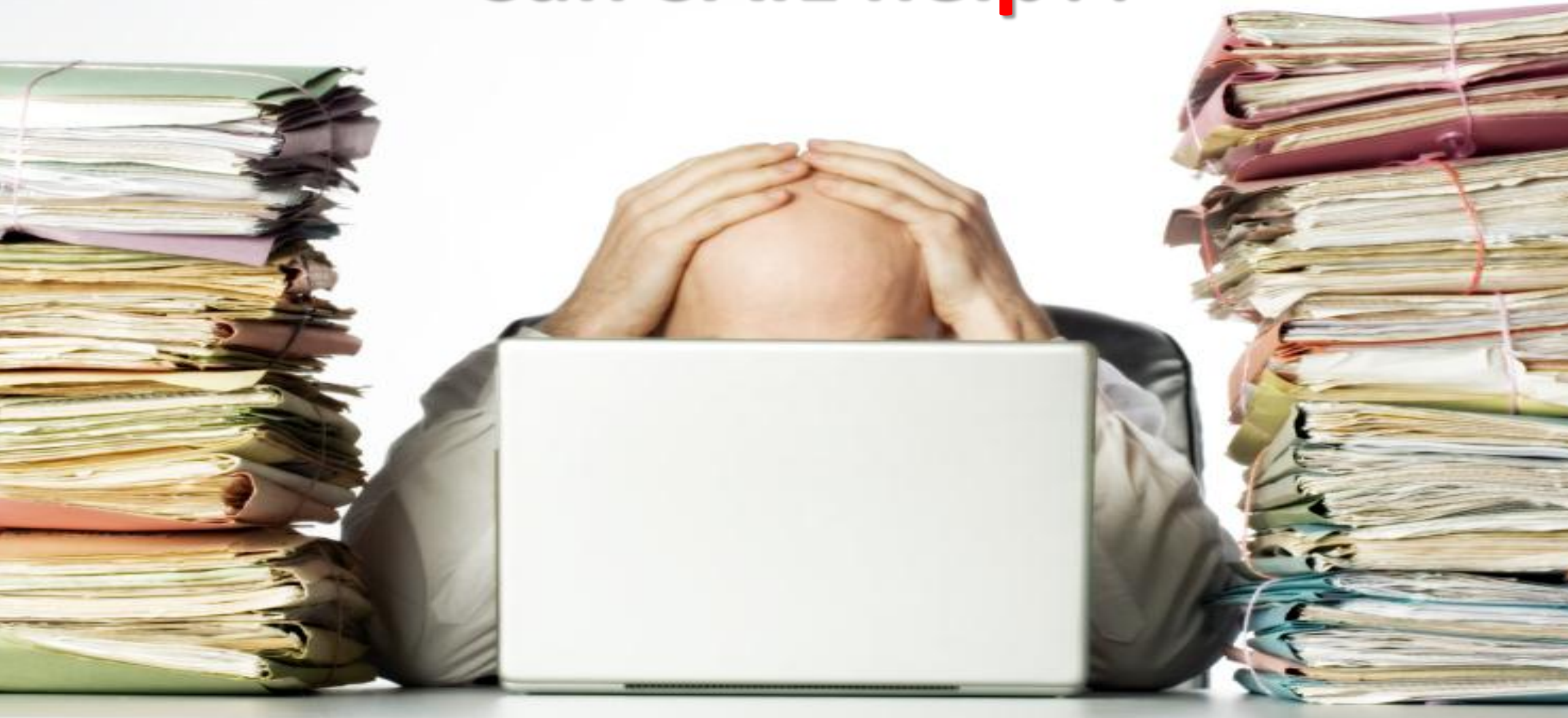
SOFTWARE ANALYSIS
& INTELLIGENCE LAB

Intelligence throughout the lifetime of a software system

From inception to production

http://sail.cs.queensu.ca

Can SAIL help?!



We got a paper for that!

Uneven participation in discussions is worrisome

Participation of experienced developers is not a good sign

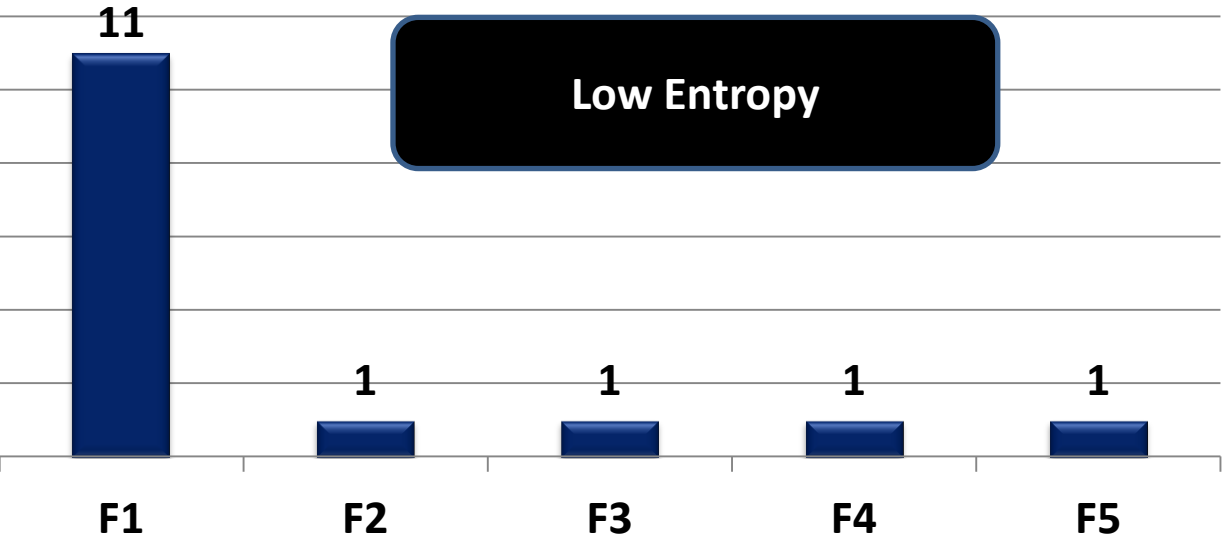
[ICPC 2010]
[EMSE 2011]

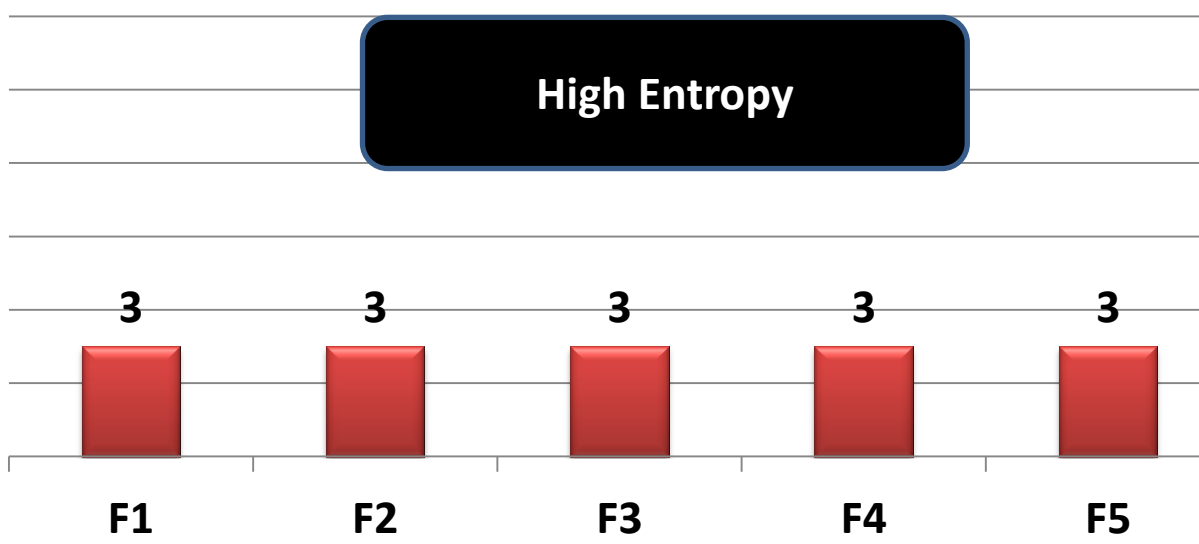
Missing to update code comments:
Not as big of a problem

Inconsistent comment updates is the problem!

[JSS 2011]

A Complex Code Development Process Impacts Code Quality





[ICSE 2009]

Building on Fresh Code is Dangerous!

[JSS 2011]
[ICSM 2009]

Uneven participation in discussions is worrisome

Participation of experienced developers is not a good sign

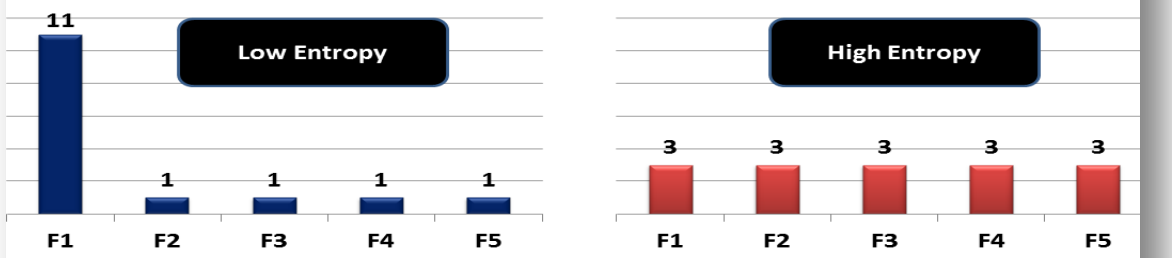
[ICPC 2010]
[EMSE 2011]

Missing to update comments: Not as big of a problem

Inconsistent comment updates is the problem!

[JSS 2011]

A Complex Code Development Process Impacts Code Quality



[ICSE 2009]

Building on Fresh Code is Dangerous

[JSS 2011]
[ICSM 2009]

Interesting

Makes sense!

Data Noise

Bias exists in the best industrial data sets

[WCRE 2010]

Inconsistent Results

Analysis level matters (files vs. subsystems)

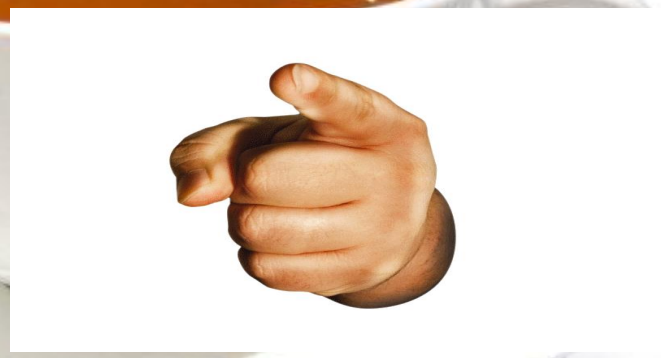
[ICSM 2010*2]

Why no adoption?


Cute!

How can we move our research to practice?!

Timely
Explainable
Assignable

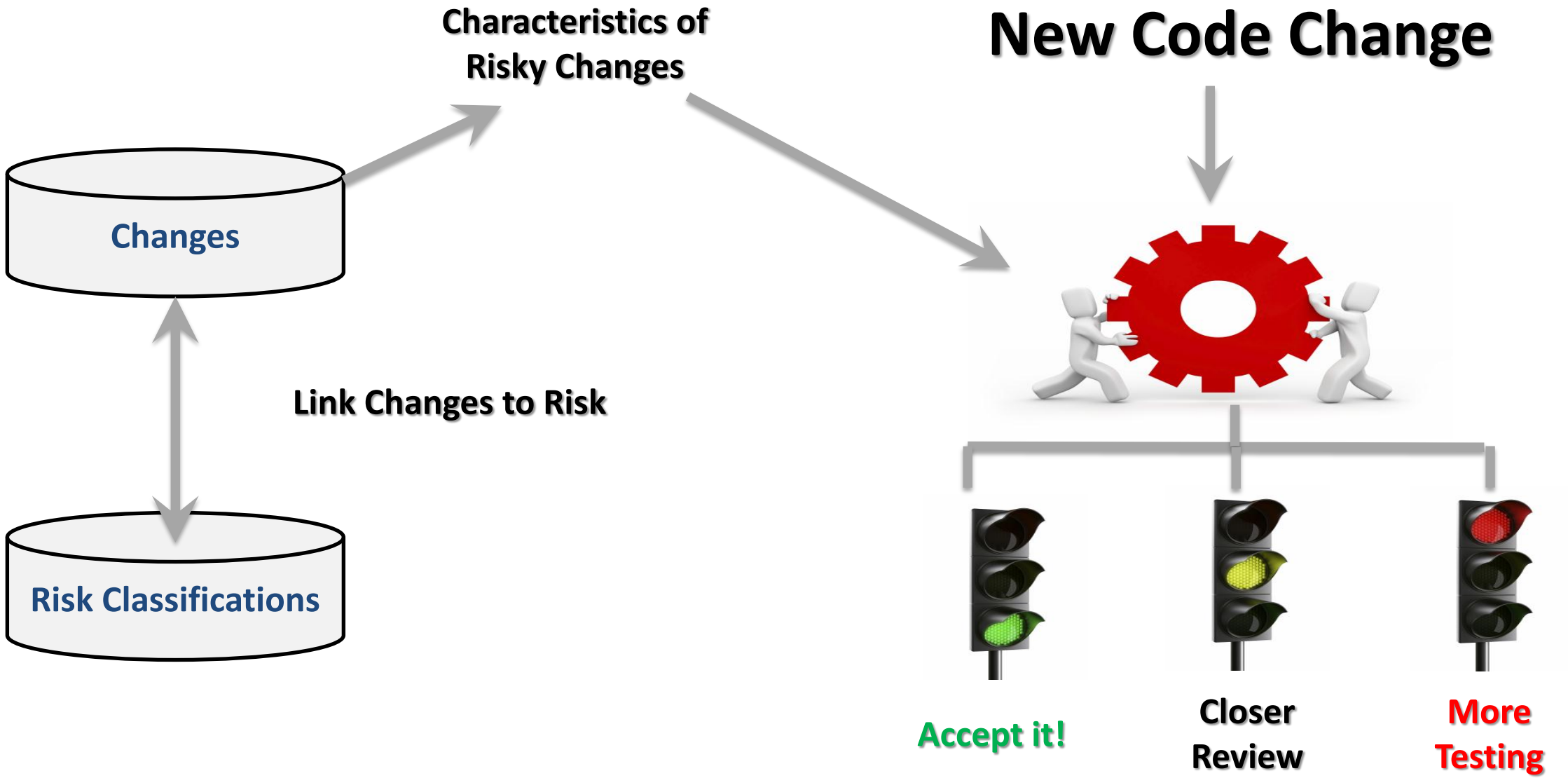


Two examples of industrial adoption

Risk Analysis of Code Changes

Automated Analysis of Load Tests

How risky is a code change?



As accurate as deve

BONUS:
Provides rationale for risk
(e.g. complex change,
too spread out)

Automated Risk

One Year of Changes

0.84

Developer Risk

450 developers
60 teams

[FSE 2012]

Timely
Explainable
Assignable

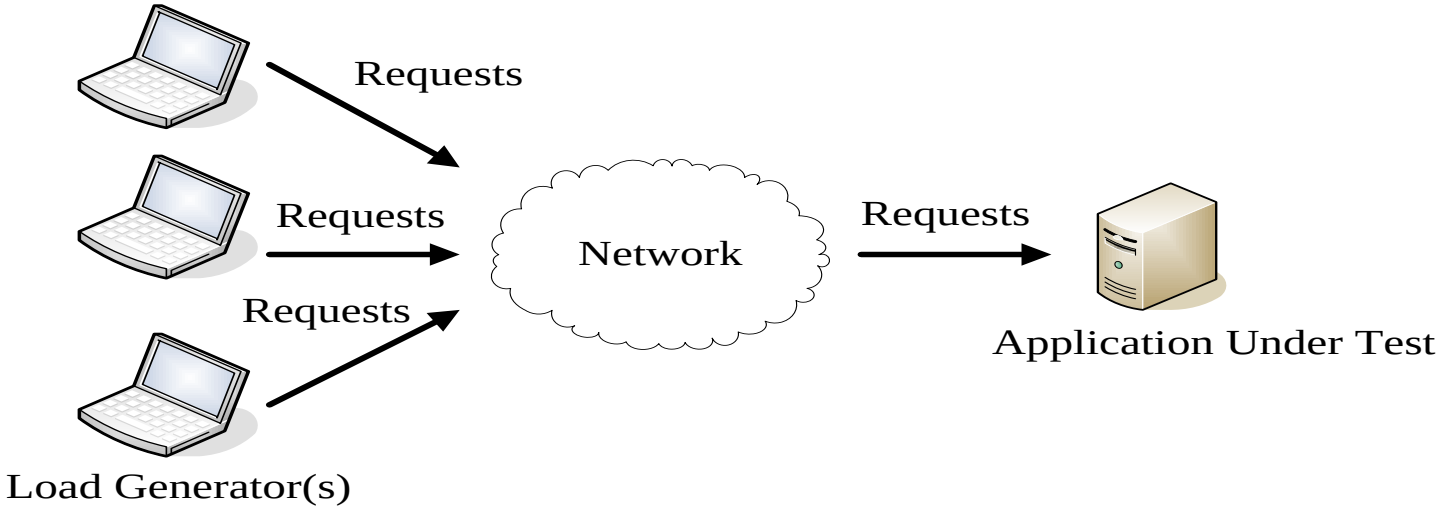
Two examples of industrial adoption

Risk Analysis of Code Changes

Automated Analysis of Load Tests

Most field problems for large systems are *rarely functional* instead they are *load-related*

Load Testing



Mimics multiple users repeatedly performing tasks for hours or even days

Produces GB/TB of data that must be analyzed

Automated Verification of Load Tests

Fact:

Load testing repeatedly executes the same scenarios

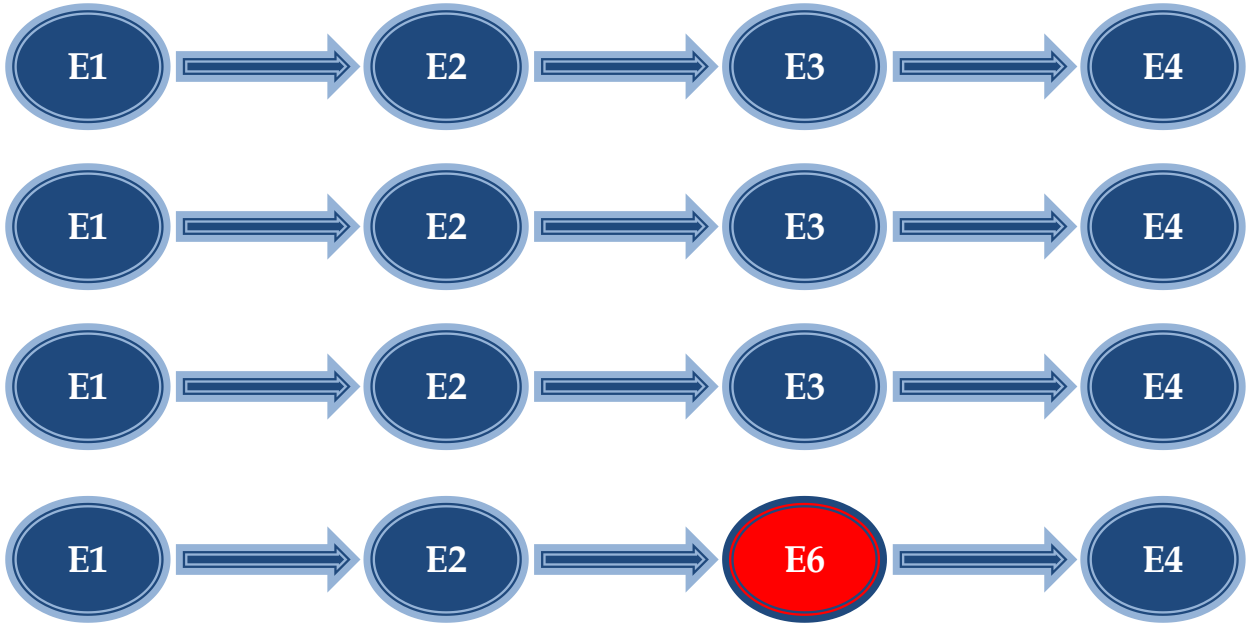
Intuition:

Dominant behaviour indicates normal operation

Minority behaviour is likely problematic

Automated Verification

- (E2, E3) are always together:
  - (acquire_lock, release_lock)
  - (open_inbox, close_inbox)
- If we see (E2, **E6**) this might be a problem



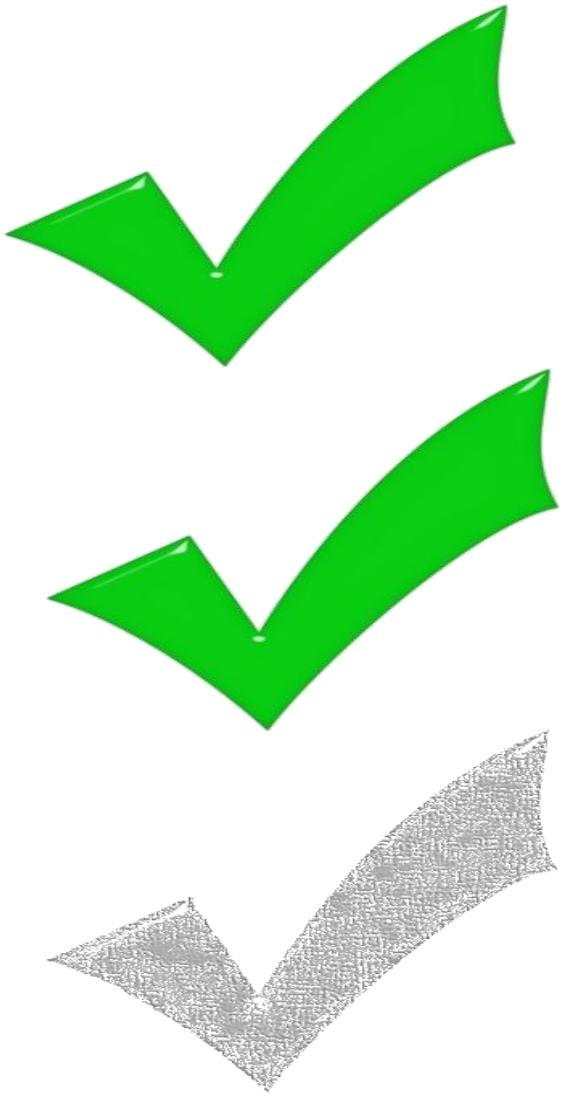
Report for Dell DVD Store

#	Z-Stat	Kinds	Min	Max	Total	Event
E_{13}	79.61	3	358	87,528	89,322	SessionID=19420, Entering purchase for simple quantity queries
E_6	39.96	2	1	14,393		
E_{19}	34.73	2	317	16,273		
E_{22}	20.65	2	1	3,857		

Freq	Sample	Details (Sort by Freq)
87,528 (98%)	ds2logs.txt 688 ds2logs.txt 689	E_{13} --> SessionID=19420, Entering purchase for simple quantity queries E_{14} --> SessionID=19420, Initial purchase, update cart
1,436 (<1%)	ds2logs.txt 2,484 ds2logs.txt 2,488	E_{13} --> SessionID=16242, Entering purchase for simple quantity queries E_{13} --> SessionID=16242, Entering purchase for simple quantity queries
358 (<1%)	ds2logs.txt 10,020 ds2logs.txt 10,021	E_{13} --> SessionID=13496, Entering purchase for simple quantity queries E_{15} --> SessionID=13496, Finish purchase before commit

99.99% reduction in viewed log lines
with a precision of 56-100%

Timely
Explainable
Assignable



Uneven participation in discussions is worrisome

Experienced developers talking is not a good sign

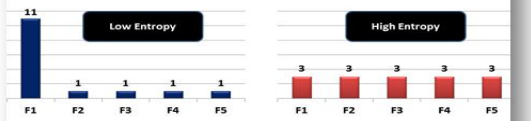
[ICPC 2010]

Missing to update comments: Not as big of a problem

Inconsistent comment updates is the problem!

[JSS 2011]

A Complex Code Development Process Impacts Code Quality



[ICSE 2009]

Building on Fresh Code is Dangerous

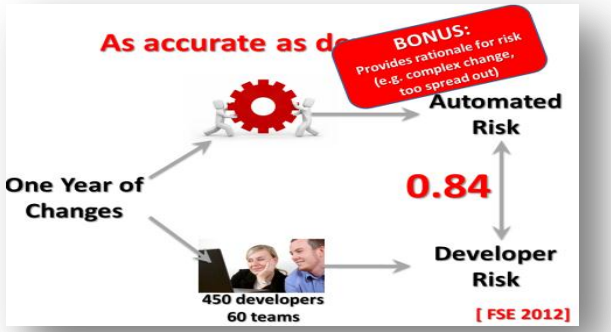
[ICSM 2009]

RISK!

Timely
Explainable
Assignable

As accurate as d

BONUS: Provides rationale for risk (e.g. complex change, too spread out)



[FSE 2012]

Two examples of industrial adoption

Risk Analysis of Code Changes

Automated Analysis of Load Tests

Automated Verification

- (E2, E3) are always together:
  - (acquire_lock, release_lock)
  - (open_inbox, close_inbox)
- If we see (E2, E6) this might be a problem

SOFTWARE ANALYSIS & INTELLIGENCE LAB

http://sail.cs.queensu.ca

http://MSRConf.org

http://promisedata.org