

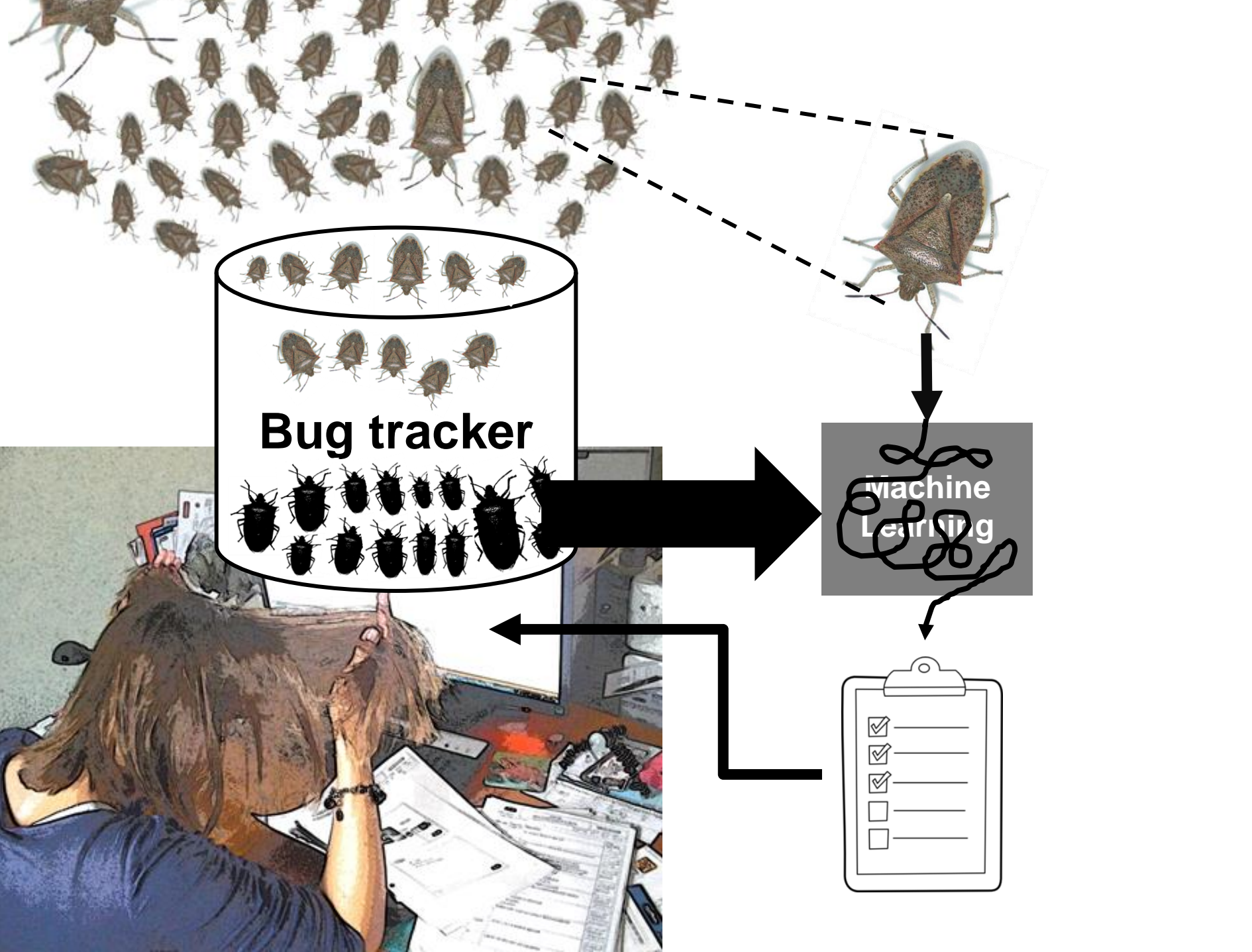


# Automation in the Bug Flow

- MACHINE LEARNING FOR TRIAGING AND TRACING

MARKUS BORG, LUND UNIVERSITY





**Bug tracker**

**Machine Learning**

- \_\_\_\_\_
- \_\_\_\_\_
- \_\_\_\_\_
- \_\_\_\_\_
- \_\_\_\_\_

# Markus Borg

Ph.D. Student

*Department of Computer Science*

[cs.lth.se/markus\\_borg](http://cs.lth.se/markus_borg)



**LUND**  
UNIVERSITY



Per Runeson





**The Challenge**  
**The Solution**  
**The Evaluation**

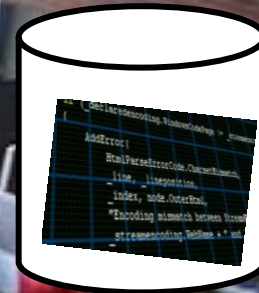
**Reqs. DB**



**Issue Repo**



**Code Repo**



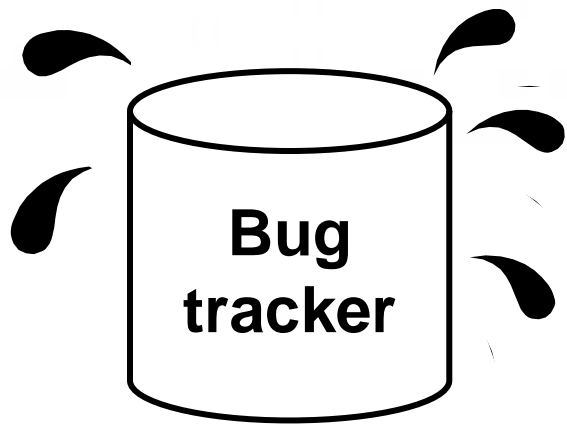
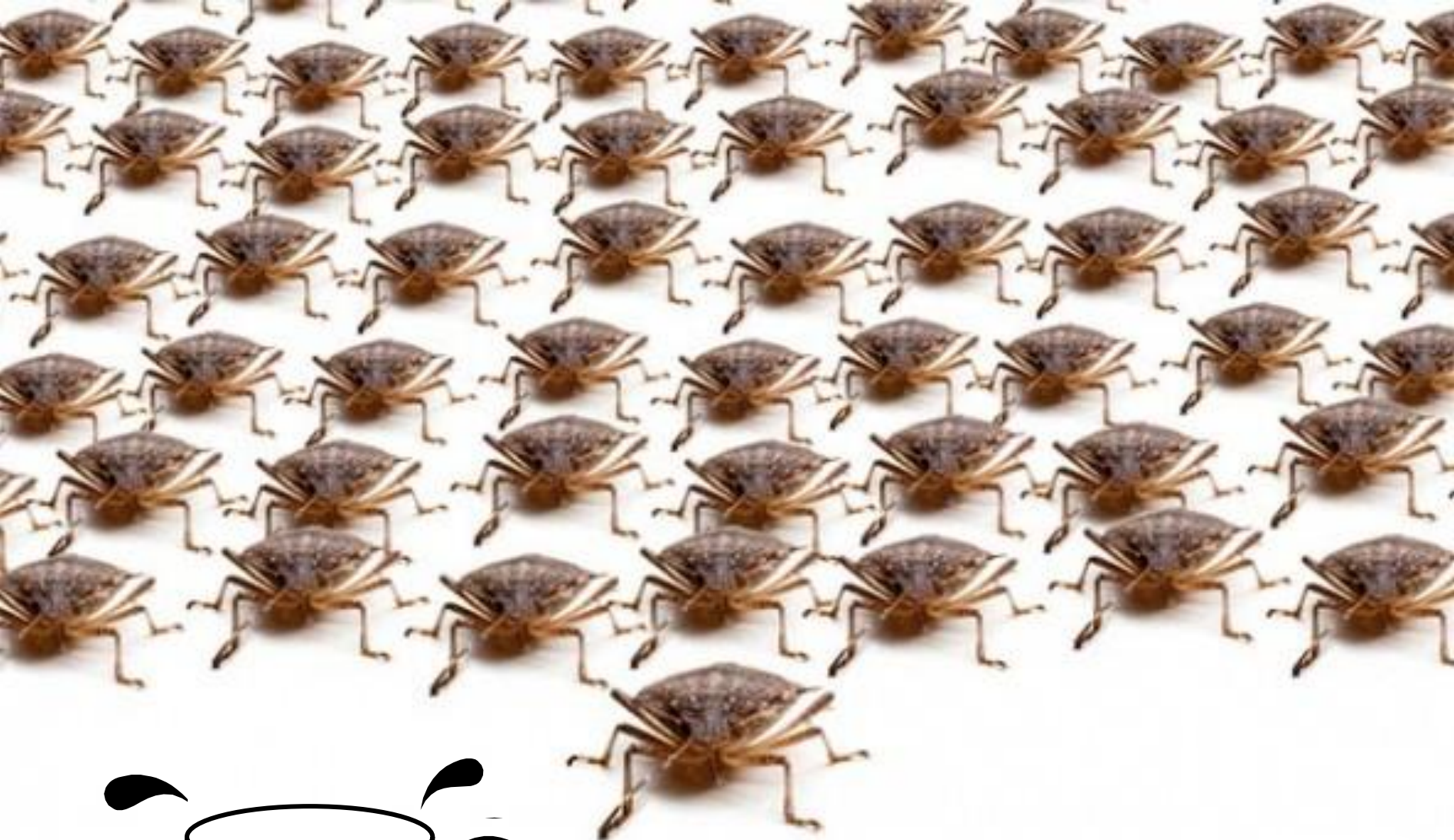
**Test DB**



**Doc. DB**



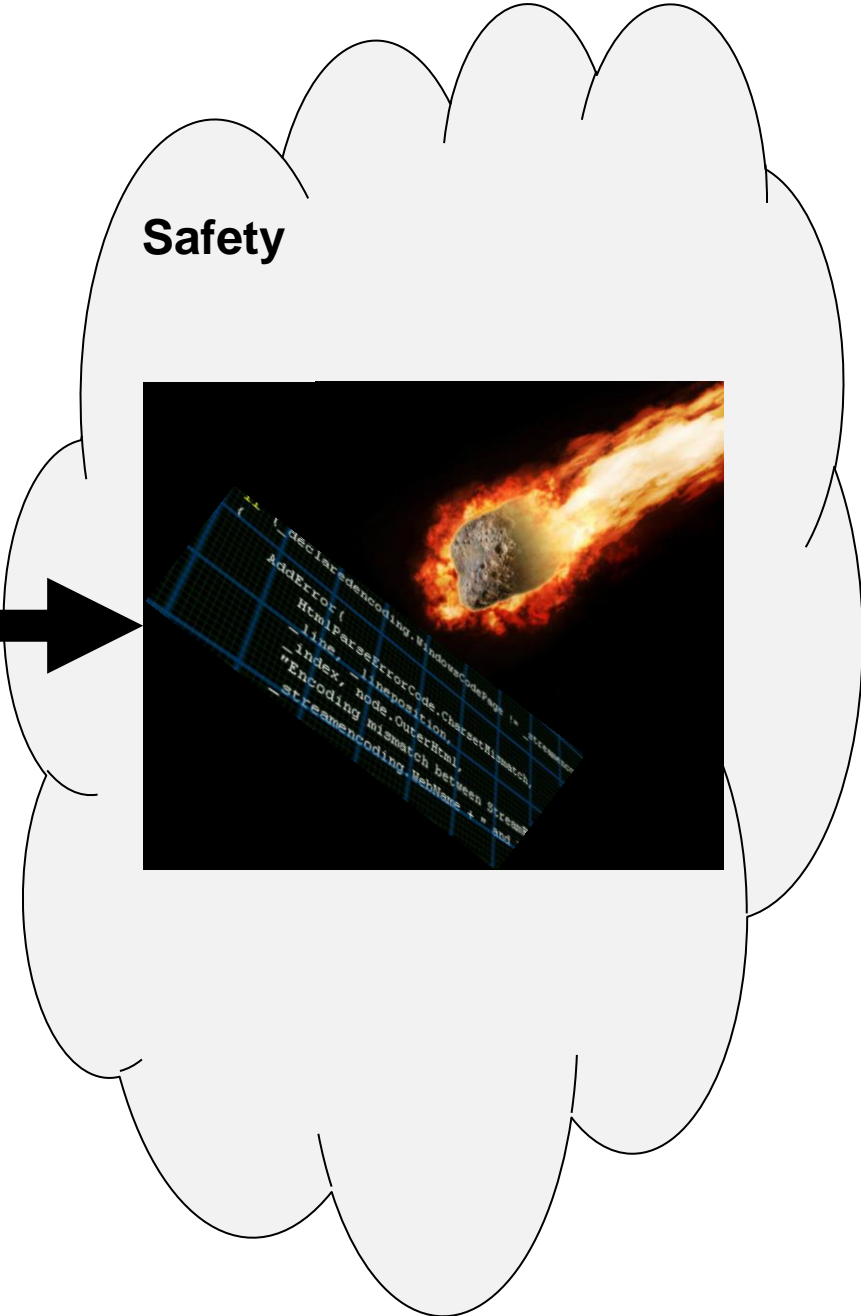
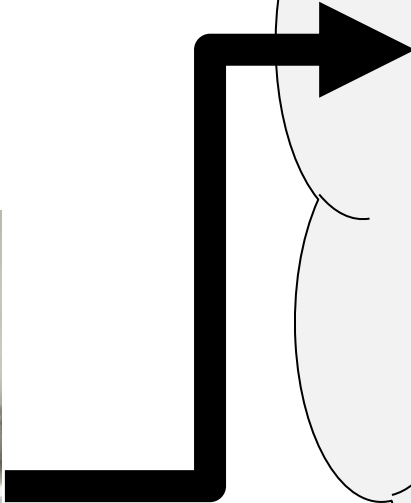
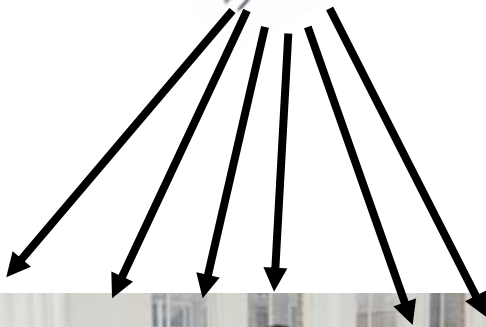








- 1. Issue Assignment
- 2. Change Impact Analysis









# The Challenge

# The Solution

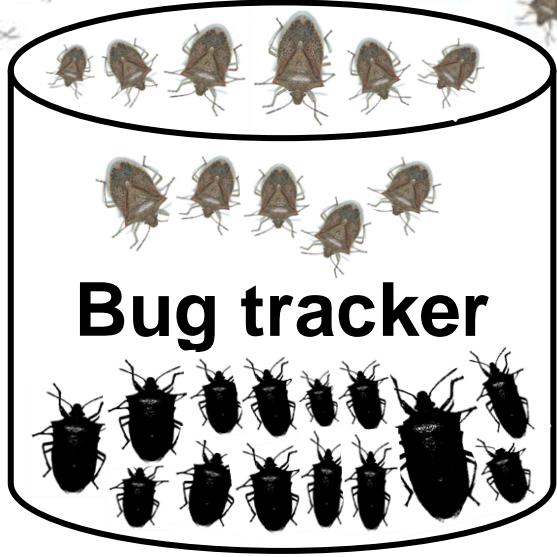
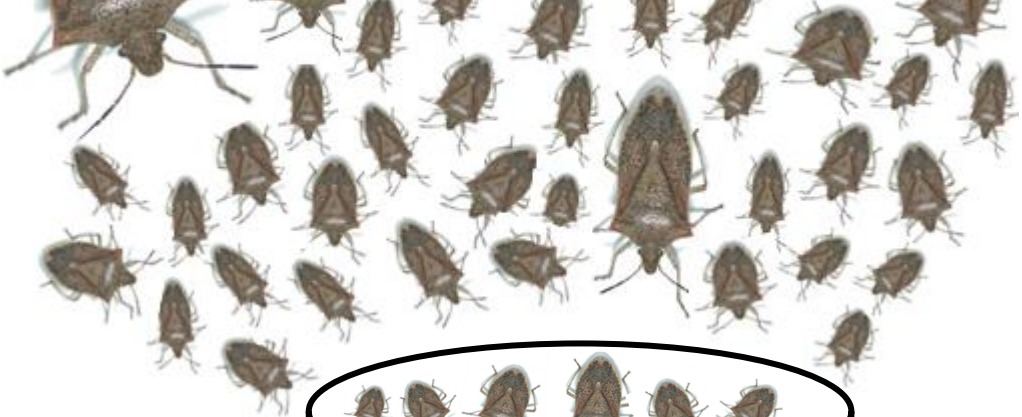
# The Evaluation



**LUND**  
UNIVERSITY

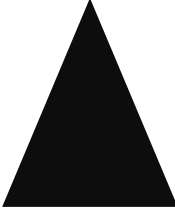


LUND  
UNIVERSITY

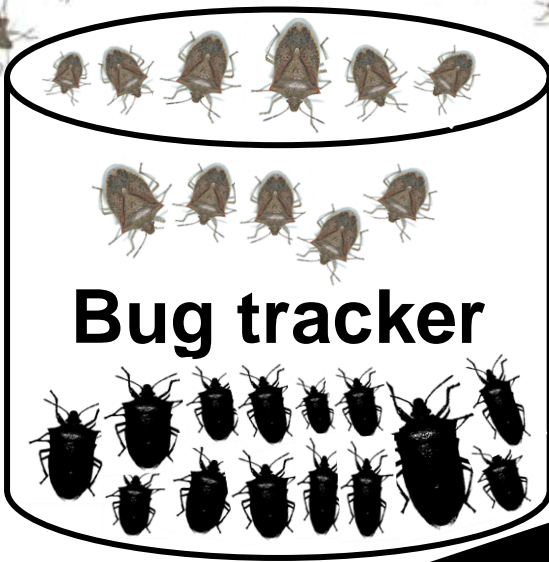
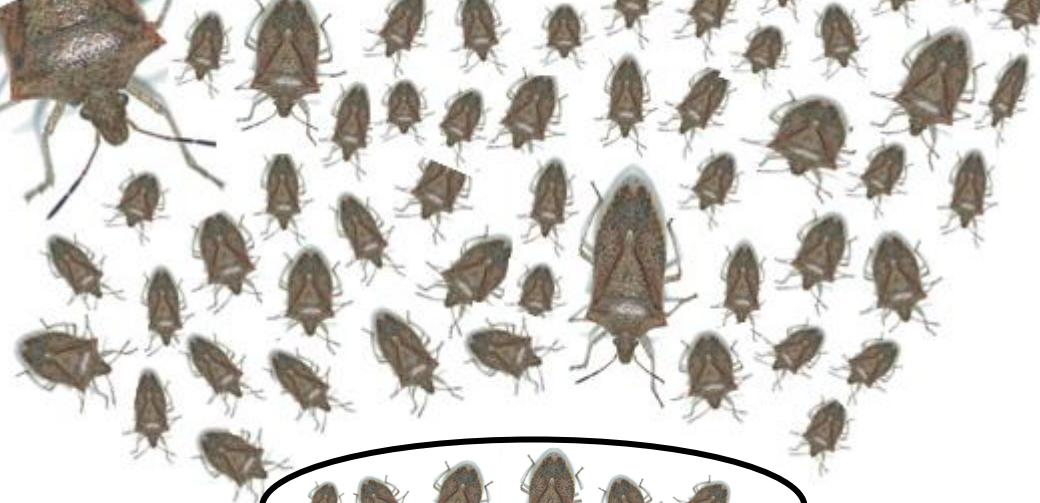


**Bug tracker**

**Machine Learning**

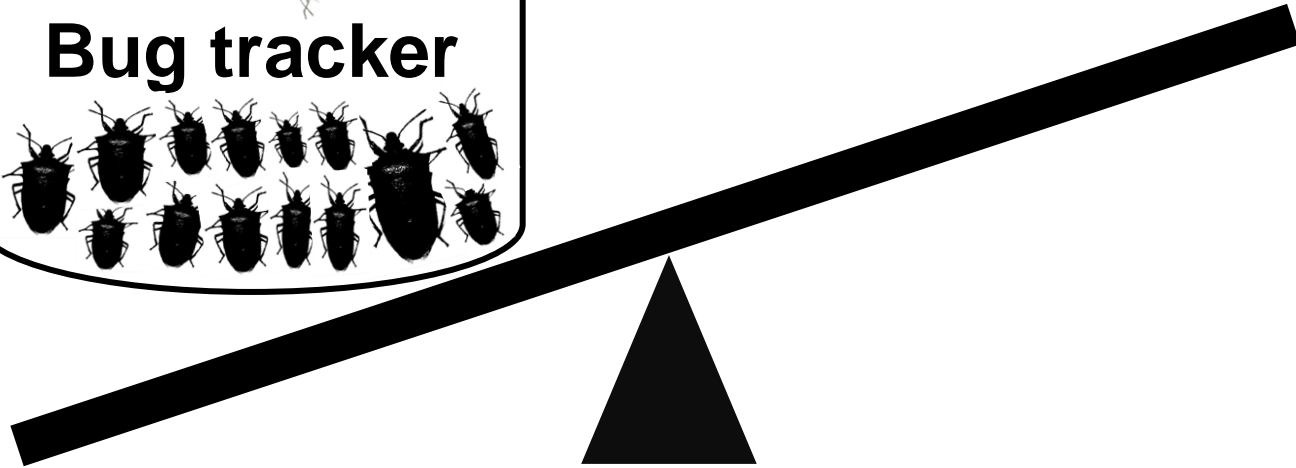


**LUND  
UNIVERSITY**



**Bug tracker**

**Machine Learning**



**LUND  
UNIVERSITY**

# Automated Issue Assignment

---

- Goal:

Useful tool deployable with minimum configuration effort

- Approach:

Bugs = textual data + basic metadata

Train classifiers on historical bug reports

Combine them using state-of-the-art ensemble learning



Leif Jonsson



LUND  
UNIVERSITY



# Automated Change Impact Analysis

---

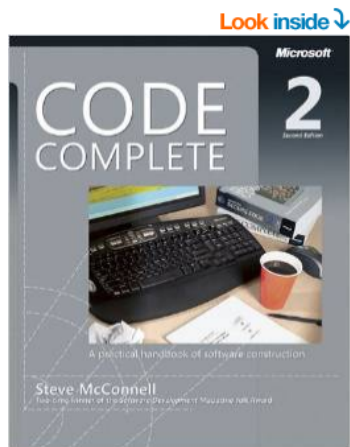
- Goal:

Intuitive tool to jump start analyses based on historical data

Faster + more accurate analyses compared to fully manual work

- Approach part 1: Mine the history
- Approach part 2: Recommend impact





Look inside

# Code Complete: A Practical Handbook of Software Construction, Second Edition

by Steve McConnell (Author) Paperback – June 19, 2004

★★★★★ 219 customer reviews

ISBN-13: 079-0145196705 | ISBN-10: 0735619670 | Edition: 2nd

## Buy New

Price: **\$31.72**

61 New from **\$28.00** | 56 Used from **\$18.95**

## Rent

Price: **\$22.00**

	Amazon Price	New from	Used from
Kindle	<b>\$24.79</b>	—	—
<b>Paperback</b>	<b>\$31.72</b>	<b>\$28.00</b>	<b>\$18.95</b>
Unknown Binding	—	<b>\$53.53</b>	<b>\$70.87</b>

Widely considered one of the best practical guides to programming, Steve McConnell's original CODE COMPLETE has been helping developers write better software for more than a decade. Now this classic book has been fully updated and revised with leading-edge practices—and hundreds of new code samples—illustrating the art and science of software construction. Capturing the body of knowledge available from research, academia, and everyday commercial practice, McConnell

[Read more](#)

Share

Rent **\$22.00**

**Buy New** **\$31.72**

Qty:  List Price: \$49.99  
Save: \$18.27 (37%)

**FREE Shipping** on orders over \$35.

**In Stock.**

Ships from and sold by Amazon.com.  
Gift-wrap available.

Yes, I want **FREE Two-Day Shipping with Amazon Prime**

[Sign in to turn on 1-click ordering](#)

**Want it tomorrow, Sept. 24?** Order within **9 hrs 43 mins** and choose **One-Day Shipping** at checkout. [Details](#)

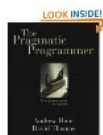
[Add to Wish List](#)

Flip to back



[See all 2 images](#)

## Customers Who Bought This Item Also Bought



**The Pragmatic Programmer: From ...**  
Andrew Hunt  
★★★★★ (255)  
Paperback  
**\$37.44**



**Clean Code: A Handbook of Agile Software ...**  
Robert C. Martin  
★★★★★ (194)  
Paperback  
**\$38.85**



**Design Patterns: Elements of Reusable ...**  
Erich Gamma  
★★★★★ (366)  
Hardcover  
**\$45.76**



**The Mythical Man-Month: Essays on Software ...**  
Frederick P. Brooks Jr.  
★★★★★ (222)  
Paperback  
**\$30.26**



**Refactoring: Improving the Design of Existing ...**  
Martin Fowler  
★★★★★ (180)  
Hardcover  
**\$48.57**

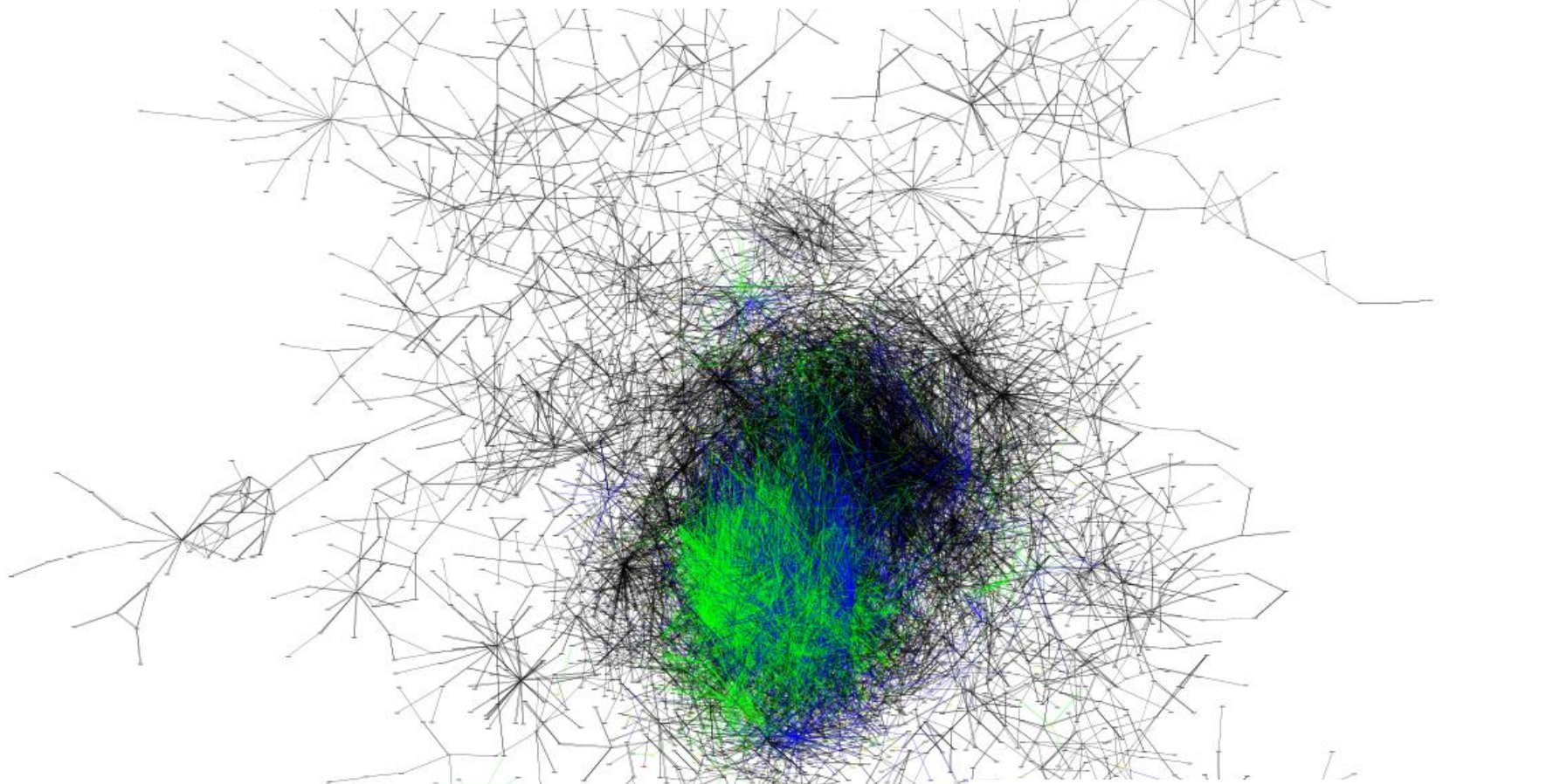


**Head First Design Patterns**  
Eric Freeman  
★★★★★ (431)  
Paperback  
**\$37.99**



**Programming Pearls (2nd Edition)**  
Jon Bentley  
★★★★★ (63)  
Paperback  
**\$31.75**

**Construct network of  
previously reported impact**



**Index textual data with**





**Calculate centrality measures**

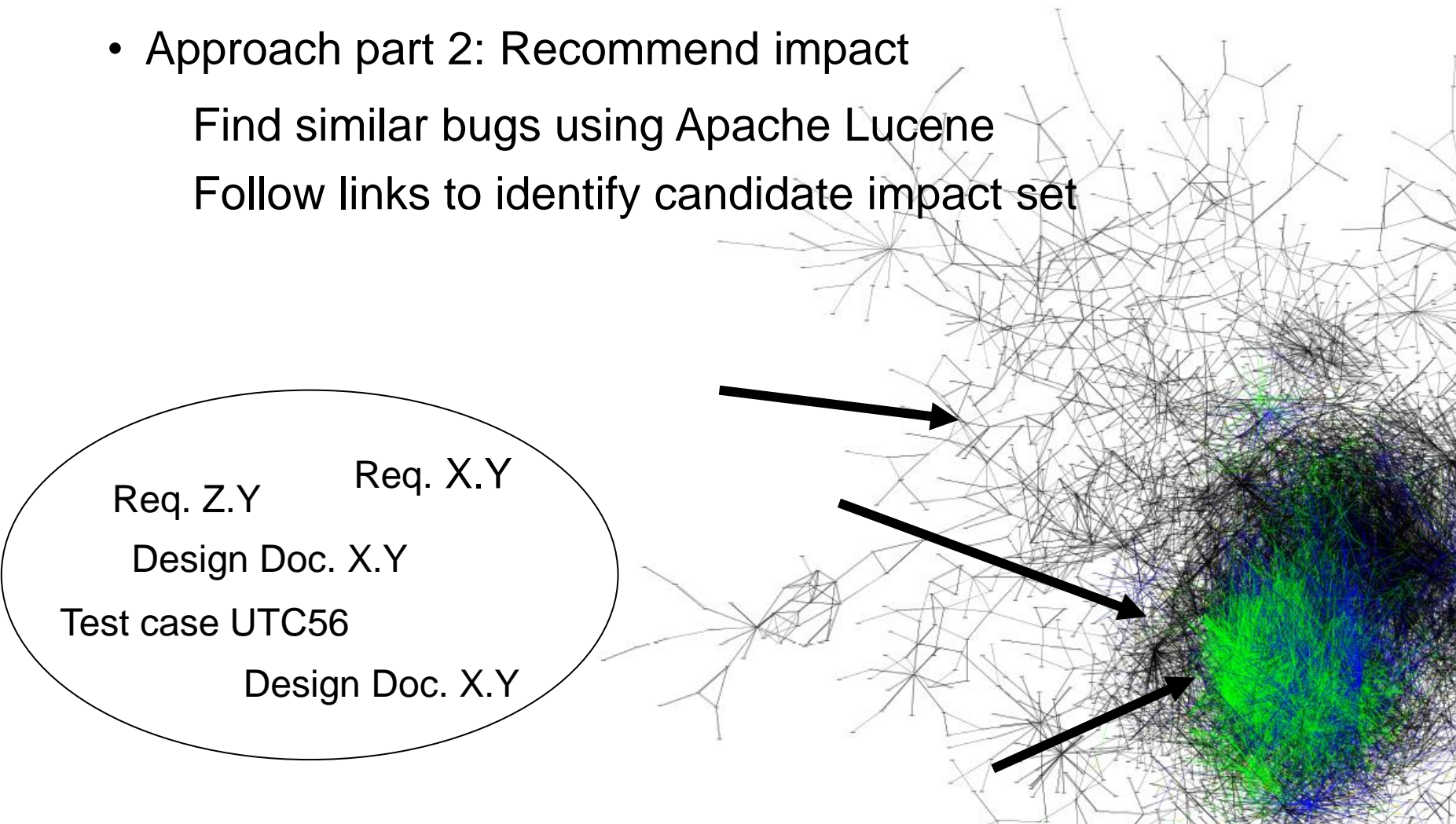
# Automated Impact Analysis

---

- Approach part 2: Recommend impact

Find similar bugs using Apache Lucene

Follow links to identify candidate impact set



Req. Z.Y      Req. X.Y  
Design Doc. X.Y  
Test case UTC56  
Design Doc. X.Y

# Automated Impact Analysis

---

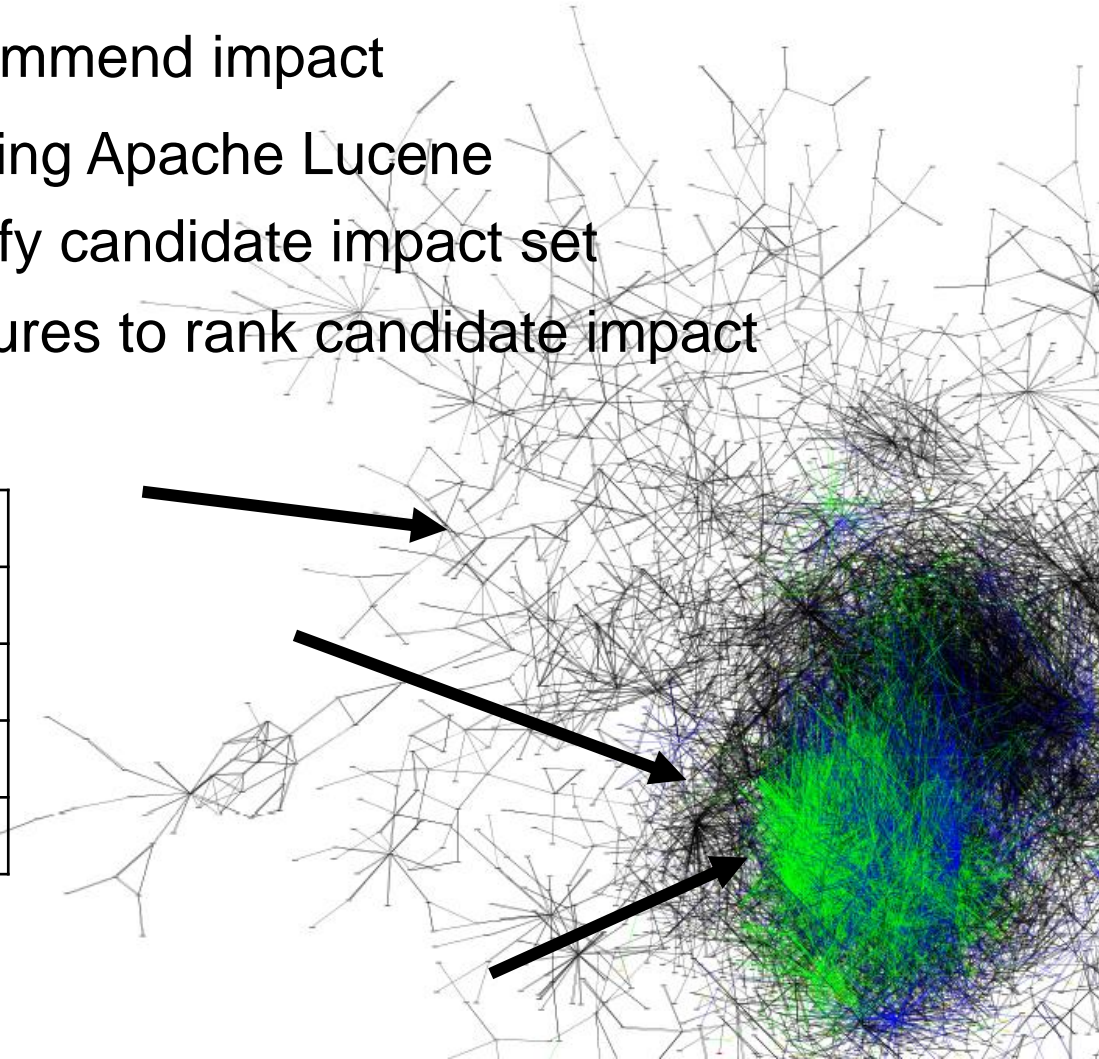
- Approach part 2: Recommend impact

Find similar bugs using Apache Lucene

Follow links to identify candidate impact set

Use centrality measures to rank candidate impact

1. Requirement X.Y
2. Design Document X.Y
3. Test case UTC56
4. Design Document X.Y
5. Requirement Z.Y



## Input Tracker Case

bug in memory reporting tool

Enter Tracker Case ID:

# 34533

Search

## WORKFLOW

1. Enter tracker case # to investigate
2. Type or paste textual description.
3. Click 'Search'
4. Analyze recommendations, give feedback.
5. Conclude feedback by pressing 'Done'



## Conclude Feedback

Done

## Most Similar Tracker Cases in the Knowledge Base

OK?	IA	Links	Sim.	ID	Title
<input type="checkbox"/>			0.93	#45470	Sequence Test Framework bug
<input checked="" type="checkbox"/>	X	2	0.74	#30946	Compiler Statistics tool in CB
<input checked="" type="checkbox"/>			0.56	#23132	Reference handling bug in Shared Blob implementat
<input checked="" type="checkbox"/>	X	9	0.54	#28889	Debug-tool, write-trap for all memory pools which are
<input type="checkbox"/>	X	9	0.54	#30635	Debug-tool, write-trap for all memory pools which are
<input type="checkbox"/>			0.53	#20550	Non-critical memory leaks in AC800M Connect
<input type="checkbox"/>			0.52	#23558	Memory leakage during LEG session
<input type="checkbox"/>			0.52	#42822	Minor bug in CVarAccessItem::Find
<input type="checkbox"/>			0.49	#44345	Bug in MMS for EventNotification
<input type="checkbox"/>			0.47	#34911	Error in Memory Leaks OPC Server
<input type="checkbox"/>			0.47	#42821	Bug in the SFC editor
<input type="checkbox"/>			0.45	#22294	Not possible to dump memory on compact flash in H
<input type="checkbox"/>			0.45	#22223	Not possible to dump memory on compact flash in H
<input type="checkbox"/>	X	2	0.45	#41624	Error in Illegal memory access handling of PM891
<input type="checkbox"/>			0.44	#41867	Tool Routing Causes Large Heap Leakage and Cra
<input type="checkbox"/>			0.42	#21526	GSD Import Tool has several errors in diagnostics cc
<input type="checkbox"/>			0.42	#21357	GSD Import Tool has several errors in diagnostics cc

## Tracker Case Details

Reference handling bug in Shared Blob implementation can cause crash in AC800  
 "We introduced a concept in SV4.0 called Shared Blob, which the aspect system developer can use to reduce the amount of memory needed when a lot of identical inherited aspects are activated at the same time. The unpacked blob is here only unpacked once and placed on the owning aspect. The inherited ASOs only have a reference to it. So far only AC800Connect (and our own Relative Name) has used it. We have now found a critical reference handling problem in this code. In some circumstances the DLL can be unloaded while there still are active shared blob objects. When the

## Reported Impact in Similar Tracker Cases

OK?	Conf.	ID	Type	Title
<input type="checkbox"/>	1	pa-fct-340	Requirement	SYSTEM VEF
<input type="checkbox"/>	0.64	sr-dgn-015	Requirement	SIL3 / CAT4
<input type="checkbox"/>	0.43	basichwlib	Hardware library	
<input checked="" type="checkbox"/>	0.32	sr-fct-005	Requirement	APPLICATION
<input type="checkbox"/>	0.21	s900ioci854hwlib	Hardware library	
<input type="checkbox"/>	0.19	3bse032067	Test description	FTTD MMU H
<input checked="" type="checkbox"/>	0.18	3bse032066	Test description	DTD MMU H
<input type="checkbox"/>	0.17	sr-fct-001	Requirement	
<input type="checkbox"/>	0.17	mmu-dof-030	Requirement	
<input checked="" type="checkbox"/>	0.17	mmu-dof-031	Requirement	
<input type="checkbox"/>	0.17	mmu-dof-032	Requirement	
<input type="checkbox"/>	0.17	mmu-dof-034	Requirement	
<input type="checkbox"/>	0.16	mmu-dof-009	Requirement	
<input type="checkbox"/>	0.15	sam-dof-410	Requirement	
<input type="checkbox"/>	0.15	sam-dof-430	Requirement	
<input type="checkbox"/>	0.14	sam-dof-420	Requirement	
<input checked="" type="checkbox"/>	0.13	dip-dof-004	Requirement	BUILD PROJ
<input checked="" type="checkbox"/>	0.13	3bse031917	Unspecified artifact	
<input type="checkbox"/>	0.13	s900ioci851hwlib	Hardware library	
<input type="checkbox"/>	0.13	3bse030861	Unspecified artifact	
<input type="checkbox"/>	0.11	3bse055868	Test description	
<input type="checkbox"/>	0.11	3bse048222	Test description	FTTD Access
<input type="checkbox"/>	0.09	3bse051937	Test description	FTTD Module
<input type="checkbox"/>	0.08	3bse032584	Test description	
<input type="checkbox"/>	0.08	sr-fct-003	Requirement	



**The Challenge**  
**The Solution**  
**The Evaluation**



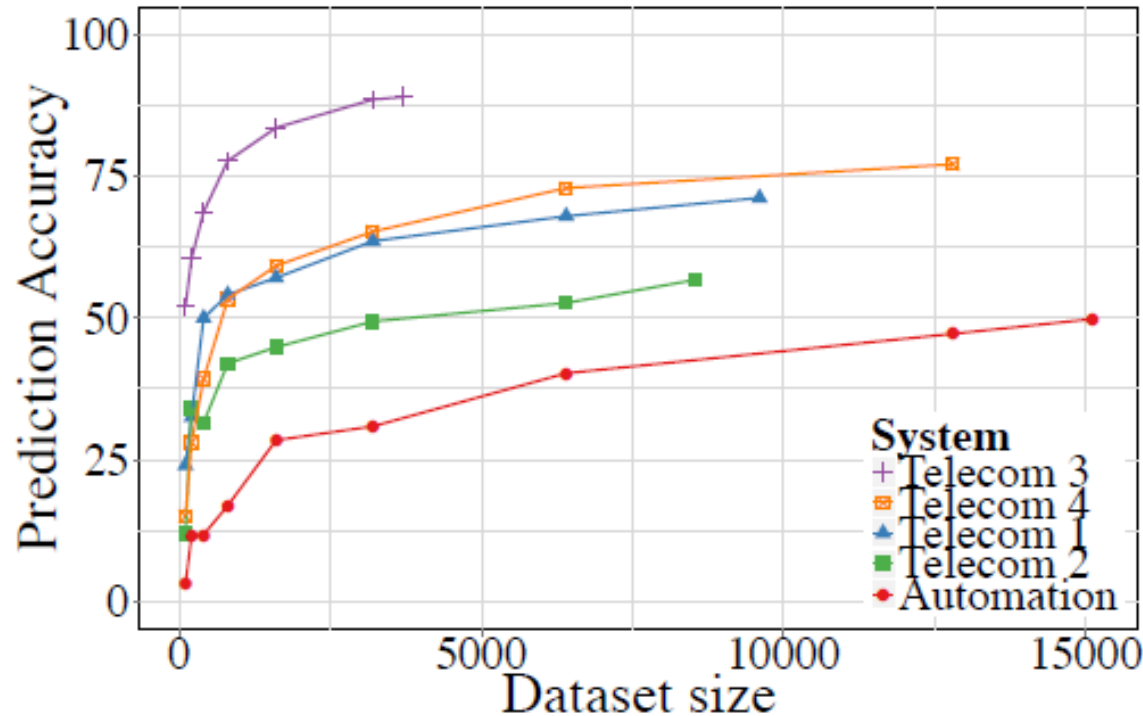
# Experiment: Issue Assignment

---

- Five large datasets from two companies
  - Telecom and Automation
  - 50.000+ issue reports
- 10-fold cross-validation and "replaying history"



# Experiment: Issue Assignment

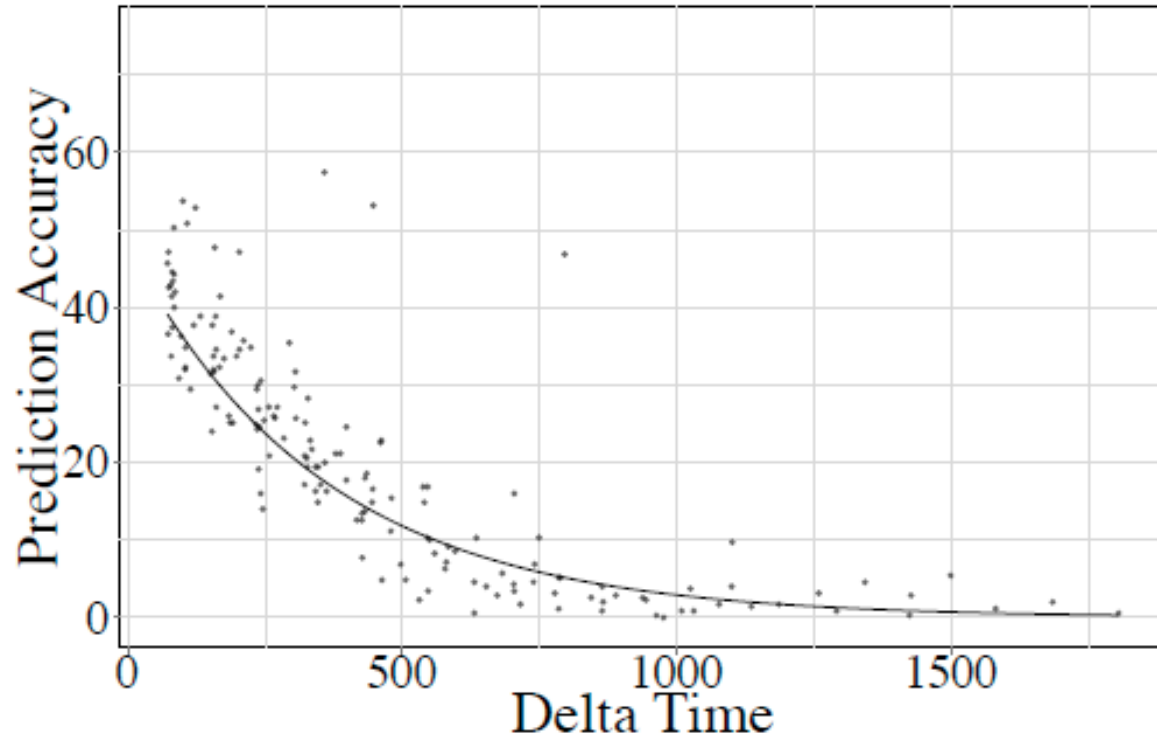


- Prediction accuracy in line with human activity
  - But instantaneous!
- At least 2.000 bug reports in the training set



# Experiment: Issue Assignment

---



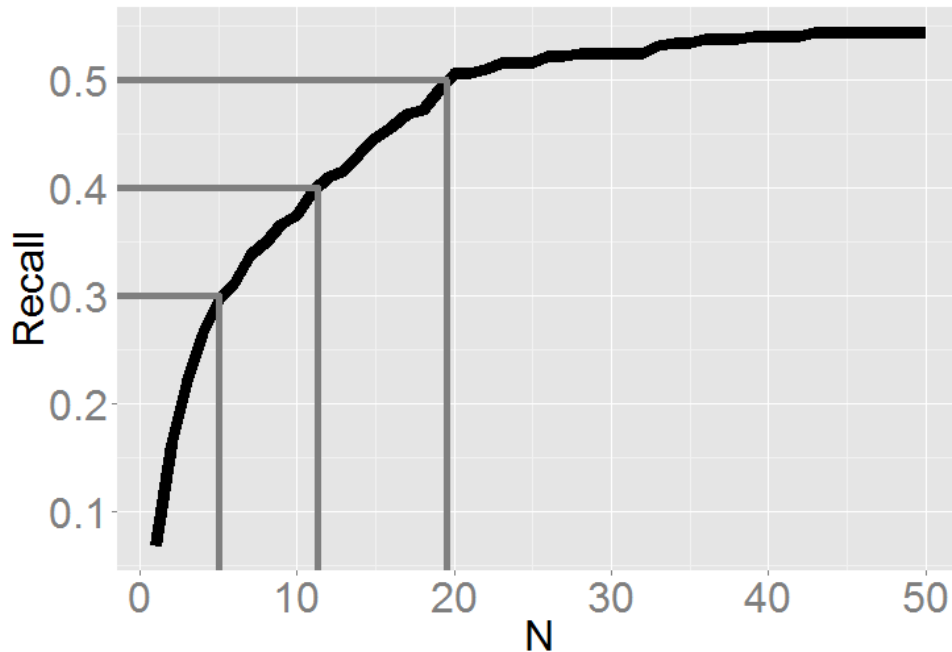
- Warning! Some systems need fresh training data



# Experiment: Change Impact Analysis

---

- Experiment with historical impact
  - Training set: 8 years, Test set: 2 years



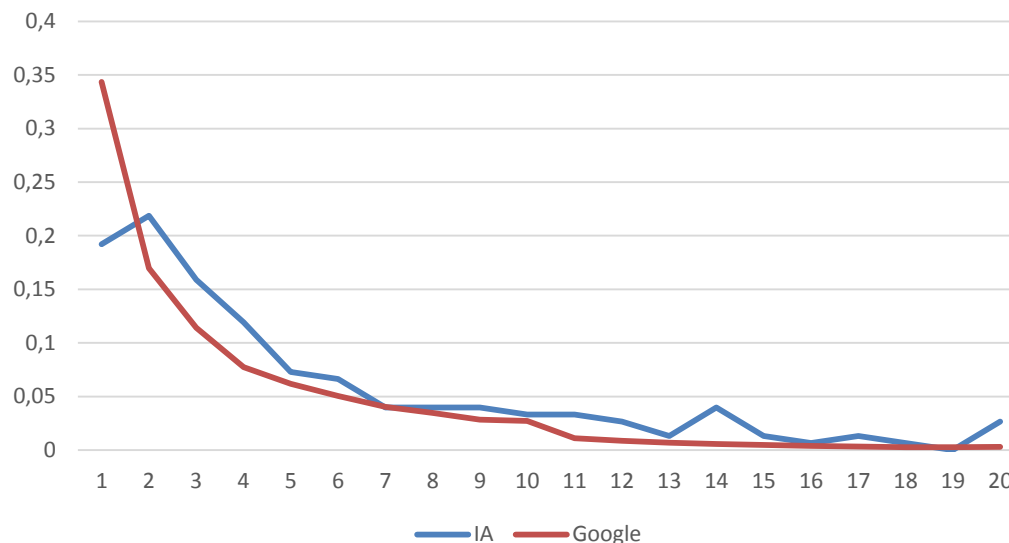
# Case Study: Change Impact Analysis

---

- Industrial case study
  - Two units of analysis: Team Sweden & Team India
  - Tool deployed in March 2014 & August 2014
  - Interviews and user log files



Click Distribution, top-20 hits





# Conclusion





*Embrace your bugs!*

Machine learning can:

- Guide maintenance
- Recommend impact

Potential more:

- Severity prediction
- Resolution times
- Noise filtering

# Thank you!

markus.borg@cs.lth.se

cs.lth.se/markus\_borg

 @\_Troddel\_



## PHOTO CREDITS

### **Brown stink bug**

- Marlin E. Rice

### **Isopods**

- Omoshiro Aquarium  
- Flickr: littoraria, coda

### **Cubicles**

- Flickr: templetonelliot, ifl, danburgmurmur

### **Eightball girl**

- Flickr: mobilestreetlife

### **Evaluate**

- Flickr: theideadesk

### **My wife**

- My wife



**LUND**  
UNIVERSITY





LUND  
UNIVERSITY