

Does software testing need to be this way?

TOOLS, APPROACHES AND TECHNIQUES TO HELP TESTERS BE MORE EFFECTIVE

Rev: 28 Nov 2016

[Creative Commons License](#)

This presentation by Julian Harty is licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.

JULIAN HARTY

28th Nov 2016

Materials, logs, etc. remain the copyright of their respective owners.



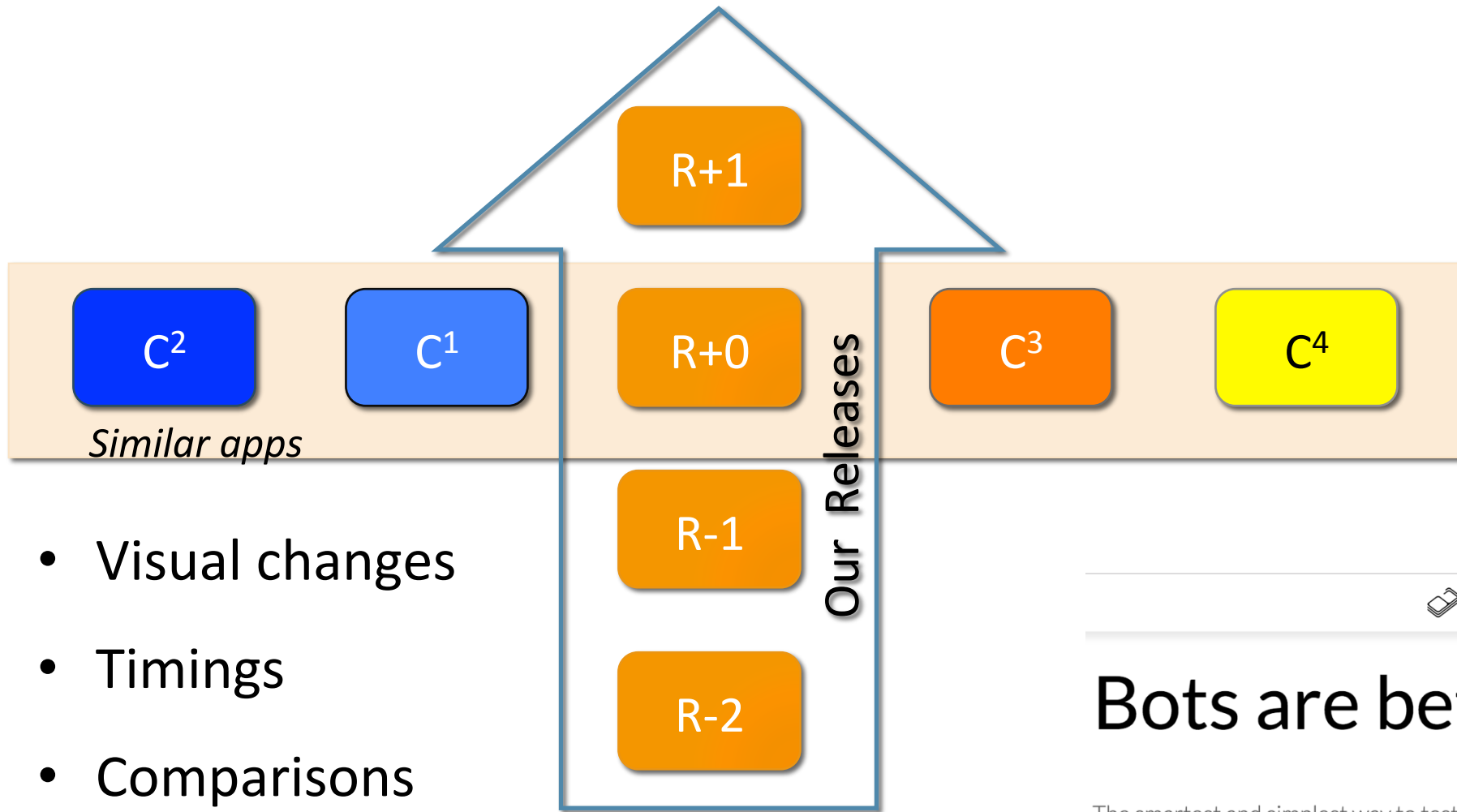
Provocative statements

- ☐ Are we trying to do better than today?
- ☐ Are we aiming for perfection (zero failures)
- ☐ Are we aiming for “good enough”
- ☐ Do we limit ourselves to what we “know”

The choices lead to very different outcomes/decisions

machines vs testers

AUTOMATED TESTING GETS STRONGER



- Visual changes
- Timings
- Comparisons

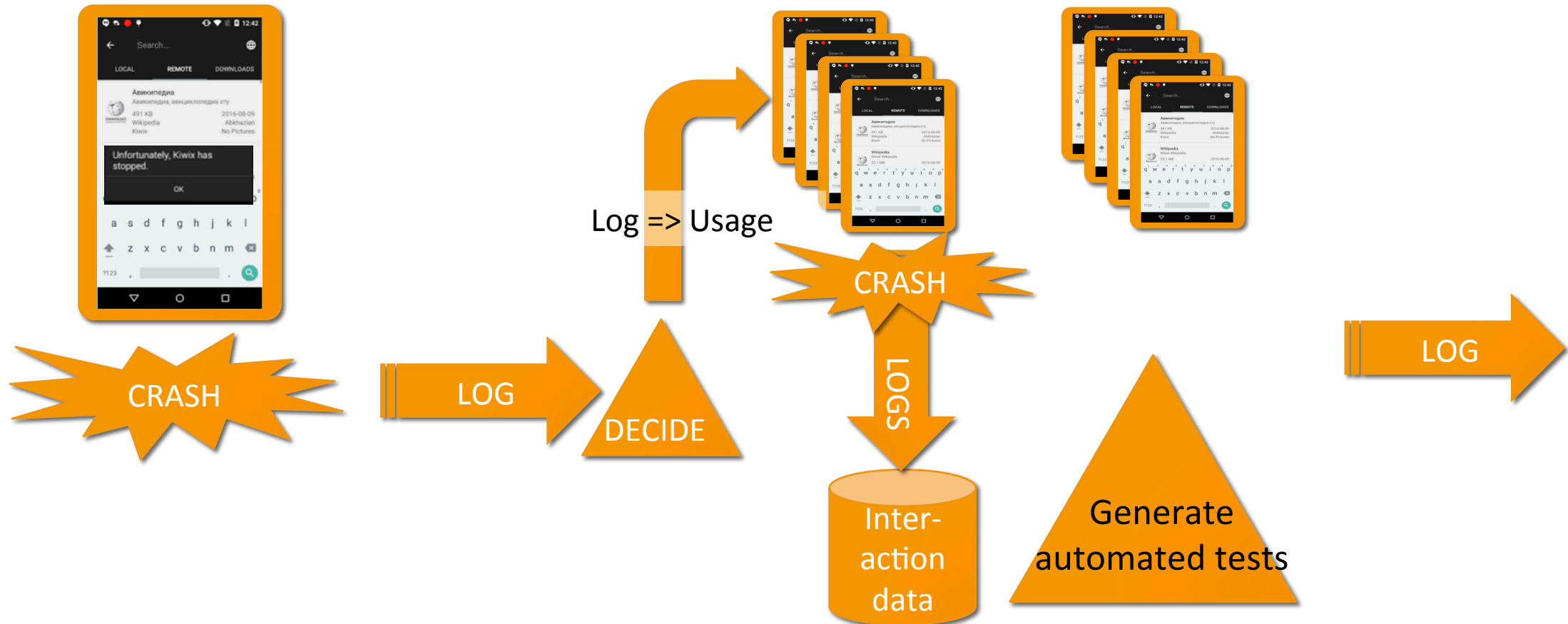


Bots are better

The smartest and simplest way to test. Intelligent bots tap, type, and swipe through your mobile app while measuring performance and tracking the user experience. No setup, no code, just results in minutes.

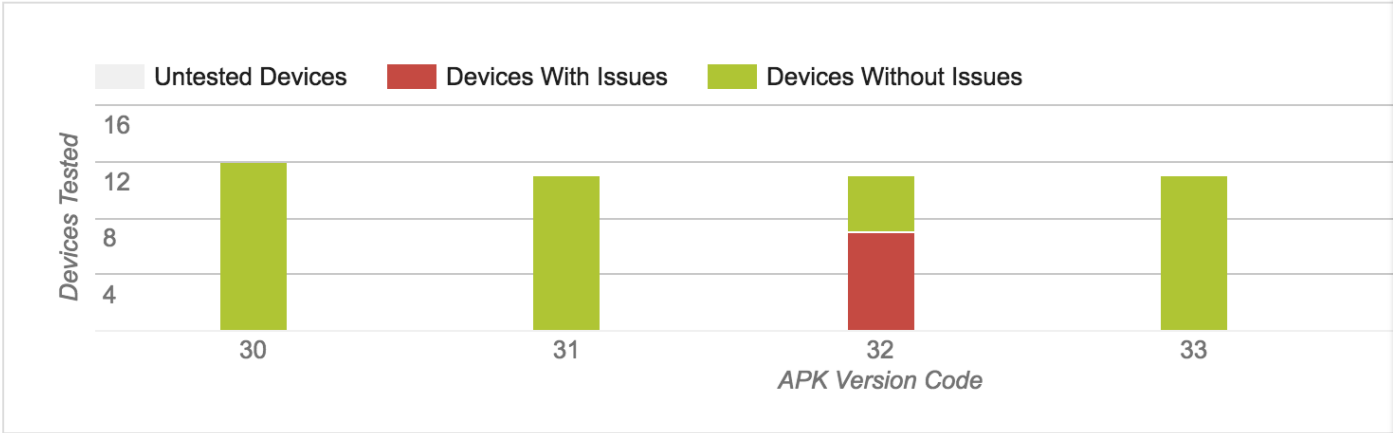
For example: <https://www.appdiff.com/>

Automatically reproducing crashes



monkey testing

HOW DOES MONKEY TESTING COMPARE?



TEST RESULTS FOR APK VERSION 32

Devices With Issues 7	Devices Without Issues 4	Devices Tested 11
--------------------------	-----------------------------	----------------------

DEVICES

	MODEL NAME	ANDROID VERSION	LOCALE	DESCRIPTION
⊘	Moto G (1st Gen)	Android 4.4	Hindi	android.view.InflateException: Binary XML file line #60: E
⊘	LG G3	Android 4.4	English	android.view.InflateException: Binary XML file line #60: Error inflating class Butt...
⊘	Nexus 5	Android 4.4	Arabic	-
⊘	Galaxy Note3	Android 4.4	English	android.view.InflateException: Binary XML file line #60: Error inflating class Butt...

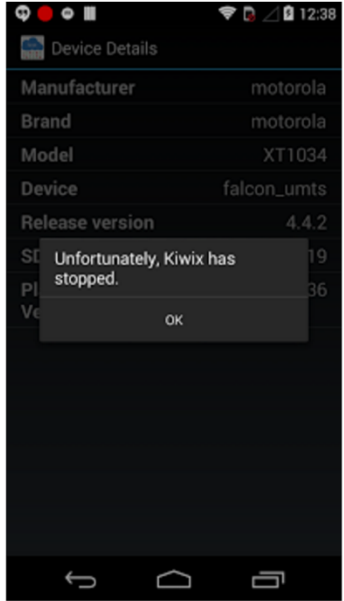
⊘ MOTO G (1ST GEN)

Test ID	1	Screen Size	720 x 1280
Test Duration	300 seconds	Screen Density (dpi)	320
Model Name	Moto G (1st Gen)	RAM	2048 MB
Manufacturer	Motorola	OpenGL ES Version	3.0
Android Version	Android 4.4	Native Platform	armeabi-v7a
Locale	hi_IN	CPU Make	Qualcomm
		CPU Model	MSM8626

▶ Issue: android.view.InflateException: Binary XML file line #60: Error inflating class Button

▶ LogCat

Screenshot



Video

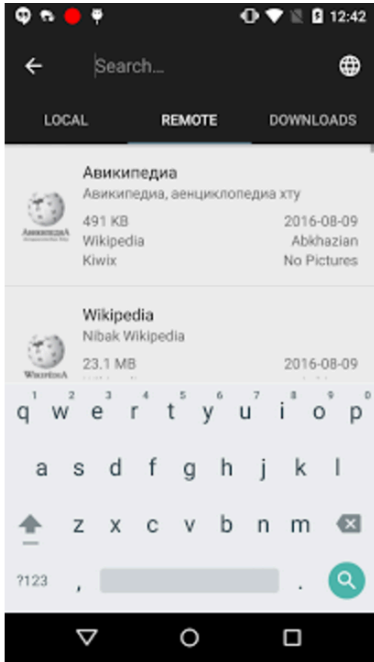


✔ NEXUS 5

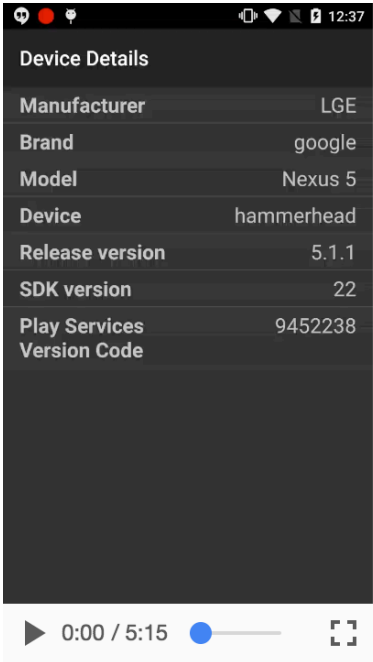
Test ID	10	Screen Size	1080 x 1920
Test Duration	300 seconds	Screen Density (dpi)	480
Model Name	Nexus 5	RAM	2048 MB
Manufacturer	Google	OpenGL ES Version	3.0
Android Version	Android 5.1	Native Platform	armeabi-v7a
Locale	en	CPU Make	Qualcomm
		CPU Model	MSM8974

► LogCat

Screenshot



Video



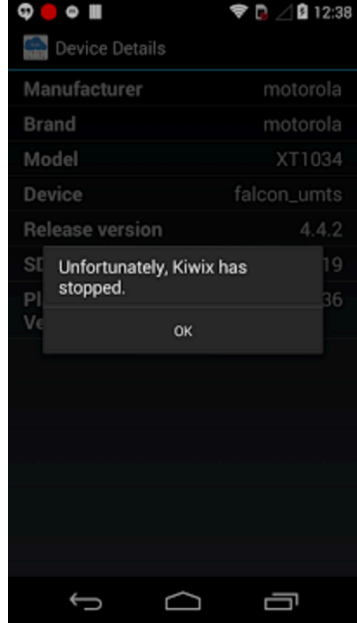
✗ MOTO G (1ST GEN)

Test ID	1	Screen Size	720 x 1280
Test Duration	300 seconds	Screen Density (dpi)	320
Model Name	Moto G (1st Gen)	RAM	2048 MB
Manufacturer	Motorola	OpenGL ES Version	3.0
Android Version	Android 4.4	Native Platform	armeabi-v7a
Locale	hi_IN	CPU Make	Qualcomm
		CPU Model	MSM8626

► Issue: android.view.InflateException: Binary XML file line #60: Error inflating class Button

► LogCat

Screenshot



Video



Pass

Pre-launch reports

Fail

DEVICES

All

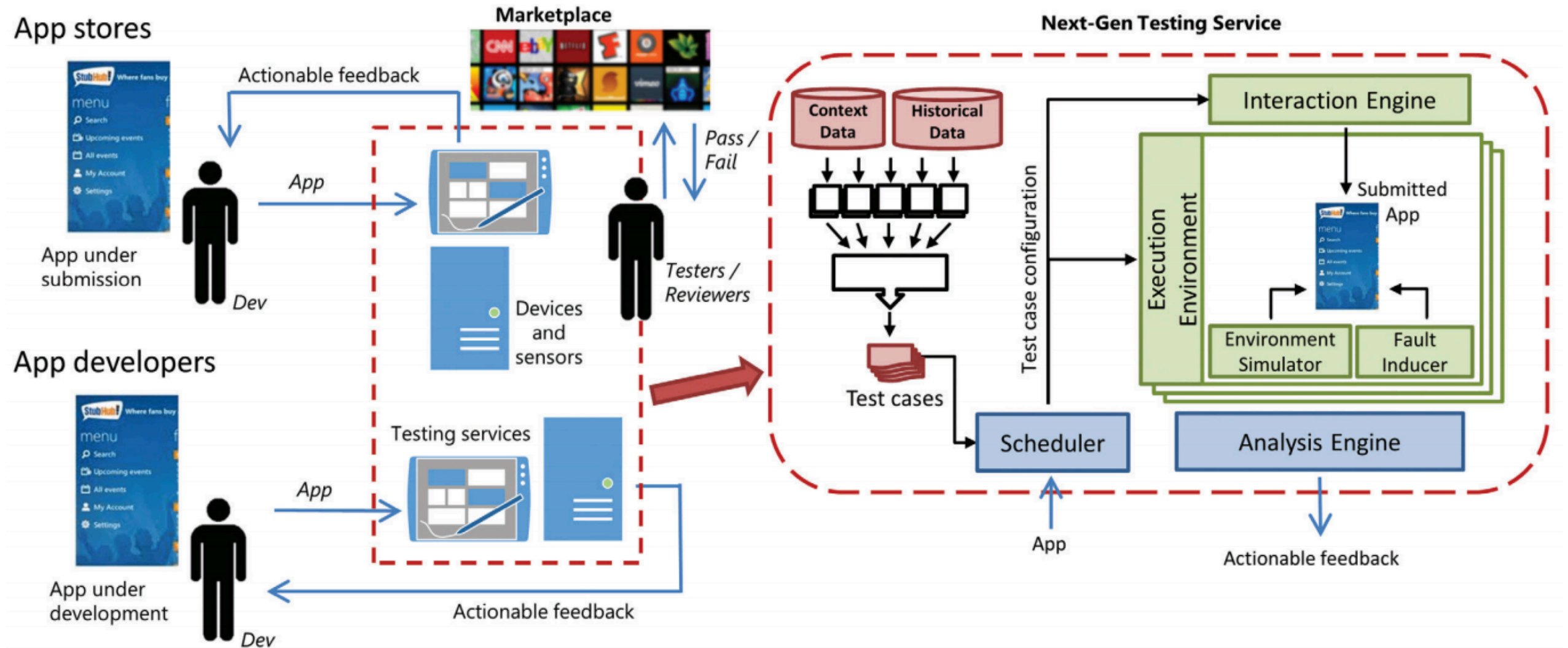
Issues

	MODEL NAME	ANDROID VERSION	LOCALE	DESCRIPTION
❌	Moto G (1st Gen)	Android 4.4	Hindi	android.view.InflateException: Binary XML file line #60: Error inflating class Butt...
❌	LG G3	Android 4.4	English	android.view.InflateException: Binary XML file line #60: Error inflating class Butt...
❌	Nexus 5	Android 4.4	Arabic	-
❌	Galaxy Note3	Android 4.4	English	android.view.InflateException: Binary XML file line #60: Error inflating class Butt...
❌	HTC One (M8)	Android 4.4	English	-
❌	Moto G (2nd Gen)	Android 4.4	English	-
❌	Moto X (2nd Gen)	Android 4.4	English	-
✅	Nexus 7	Android 5.0	English	-
✅	Nexus 9	Android 5.0	English	-
✅	Nexus 5	Android 5.1	English	-
✅	Galaxy S6	Android 5.1	English	-

Pre-launch reports: are there common factors?

data mining

MICROSOFT RESEARCH IN THE APP STORE



<https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/getmobile15.pdf>

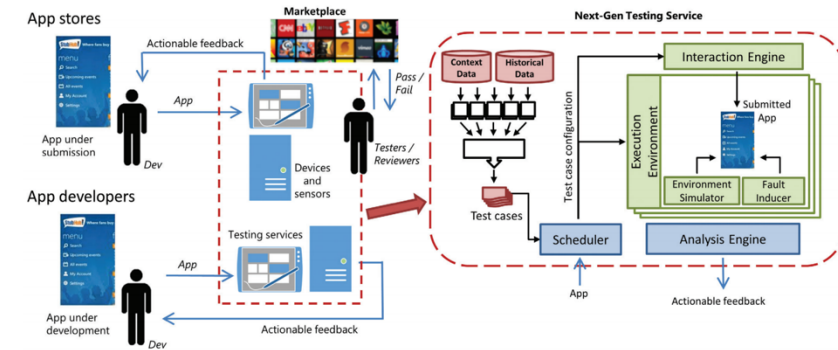
How to smash the next billion mobile app bugs

- Mined 25,000,000+ crash reports: found small number of causes e.g.
 - impatient user,
 - unexpected network response code.
- Uncovered 2,969 distinct faults;
 - 1,227 were previously unreported
- *Fully automated testing*

90% of bugs caused by
top 10 coding errors

<https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/getmobile15.pdf>

<https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/vanarsena.pdf>



Fault Inducing Modules

1. Web errors
2. Poor network conditions
3. Sensor errors
4. Invalid text entry
5. Impatient users

How to smash the next billion mobile app bugs

Is Carmen better than George?

TESTING THE EXPLORATORY TESTER USING HCI TECHNIQUES



Assessing exploratory testing using HCI techniques

before & after

MEASURING PROGRESS FOR A TDD DESIGN COURSE

Problem Summary - Big Bill's Tax Calculator

You see that big feller on the right?

That's Big Bill. He's part cowboy, part business mogul.

Big Bill's company, **Wild West Stores**, has a billing system that has been generating some *unusually big* bills of late.

Old Mrs. Smith was billed \$3,000 dollars sales tax for buying 5 items at the store.

Big Bill has had enough! He wants a quality sales tax calculator and wants YOU to program it.

Sales Tax Calculator Logic

Basic sales tax is applicable at a rate of 10% on all goods, except books, food, and medical products that are exempt. Import duty is an additional sales tax applicable on all imported goods at a rate of 5%, with no exemptions.

When you purchase items you receive a receipt, which lists the name of all the items and their price (including tax), finishing with the total cost of the items, and the total amounts of sales taxes paid. The sales tax is rounded up to the nearest 0.05 amounts.

Write an application that prints out the receipt details for these shopping baskets...

Input	Output
1 book at 12.49 1 music CD at 14.99 1 chocolate bar at 0.85	1 book: 12.49 1 music CD: 16.49 1 chocolate bar: 0.85 Sales Taxes: 1.50 Total: 29.83
1 imported box of chocolates at 10.00 1 imported bottle of perfume at 47.50	1 imported box of chocolates: 10.50 1 imported bottle of perfume: 54.65 Sales Taxes: 7.65 Total: 65.15
1 imported bottle of perfume at 27.99 1 bottle of perfume at 18.99 1 packet of headache pills at 9.75 1 box of imported chocolates at 11.25	1 imported bottle of perfume: 32.19 1 bottle of perfume: 20.89 1 packet of headache pills: 9.75 1 imported box of chocolates: 11.85 Sales Taxes: 6.70 Total: 74.68

Note: Document the necessary assumptions and email them to us.



Problem Summary - Mini Mausi's Meeting Assistant

You see our Mausi on the right?

That's Mini Mausi, manager @ a large company.

Mini Mausi usually wants to meet her team for very important meetings.

After spending hours trying to find a suitable meeting time, she usually realizes that the chosen meeting slot is no longer available.

Mausi has had enough!

She wants an easy, efficient way to find a possible meeting time with her team and wants YOU to build it.

Mini Mausi has high expectations from you and the new system.

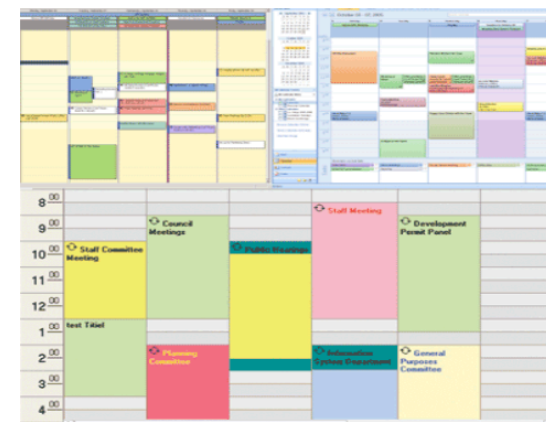
She also tends to get *a little cranky* when she is disappointed.

So please do your best to build a solution that will make old Mausi nice and happy.



Meeting Assistant Instructions

While trying to set up a meeting with her team members, Mini Mausi has a hard task of scanning through everyone's calendars on her little monitor to find a common, empty meeting slot.



Example Report

“Good improvement compared to the pre-assessment. Looks like you've at least applied some TDD principles, but there is a large scope for improvement.”

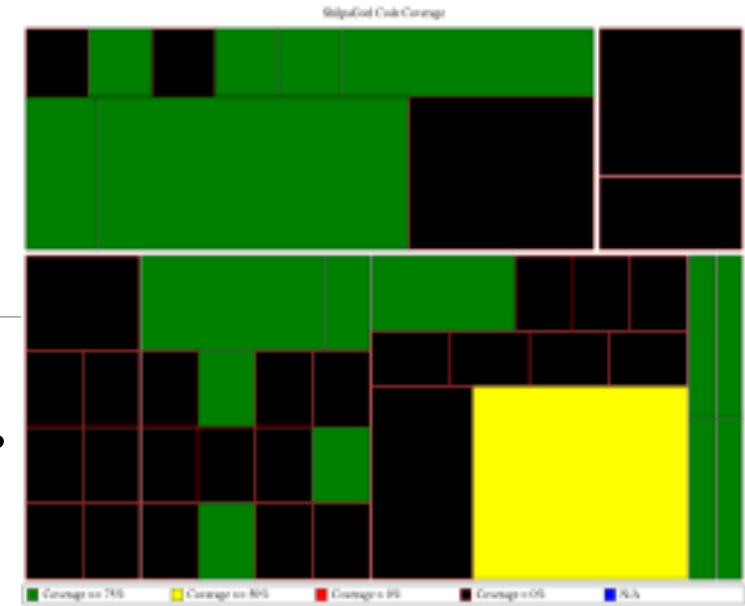
“Code Smells”

1. Magic Number
2. Dead Code
3. Duplicated code
4. Long Method
5. Speculative Generality
6. Feature Envy...

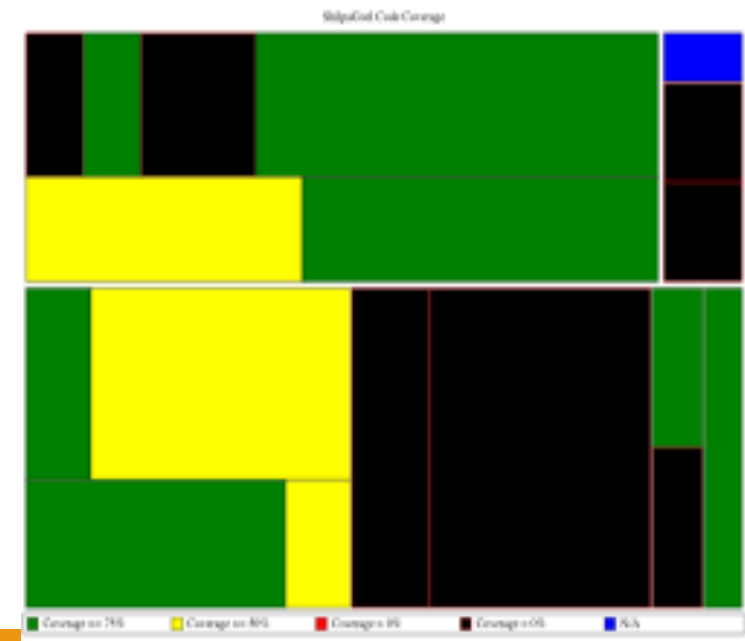
Human insights & interpretation

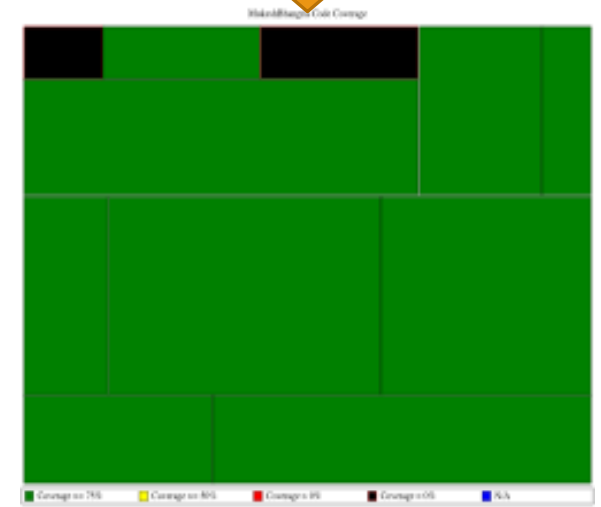
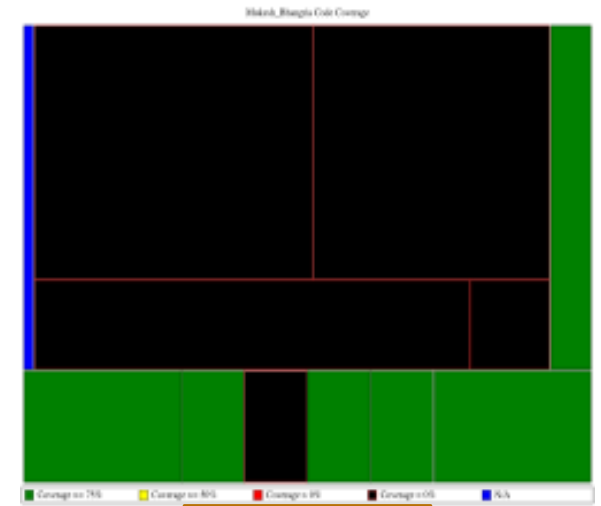
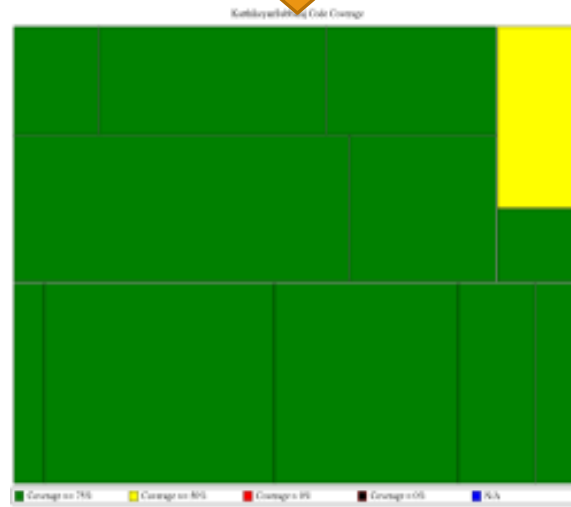
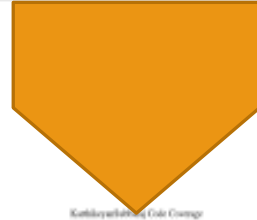
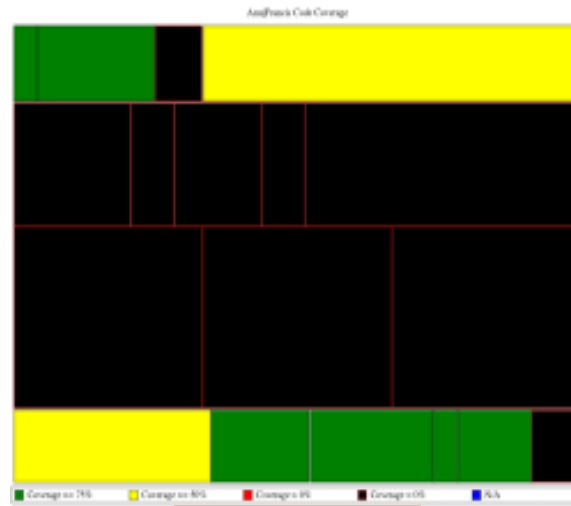
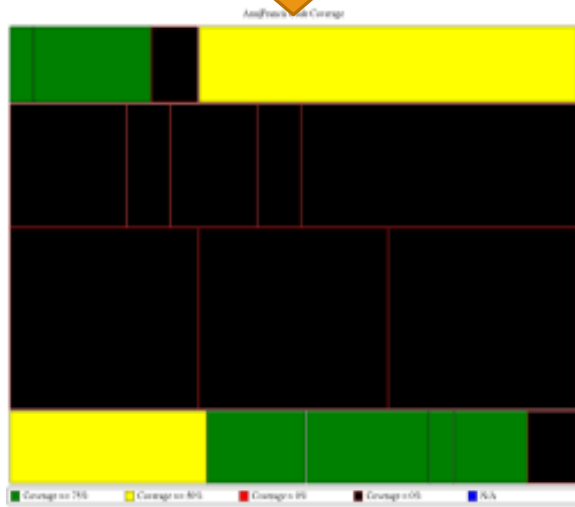
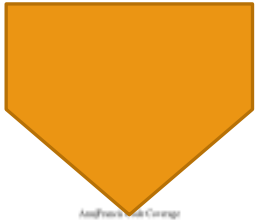
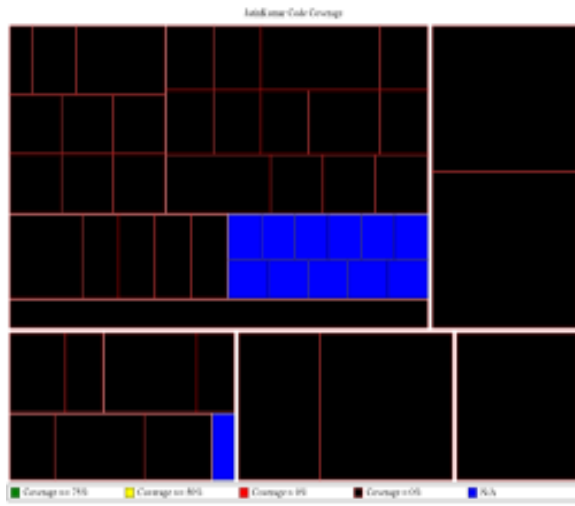
1. Has written one test using mock objects, but only partially tests the core logic.
2. Test has a huge setup highlighting problem with the design (tight coupling).
3. Since we are testing from a higher point, the test does not communicate the intent very well.

Pre
47%



Post
54%





Test Coverage improvements by participants

guidance

THE PROMISE OF GUIDES FOR BETTER TESTING

Absent, Access, Batch, Sequence, Clone, Concurrency, Configuration, Confluence, Constraint, Continuity, Corrupt, Customize, **Data**, Delay, Dependency, Feedback, **Function**, Hierarchy, History, Integrity, Interaction, Invoke, Lock, Loop, **Operations**, Parameter, **Platform**, Prerequisite, Priority, Refresh, Repetition, Resource, Restore, Roll back, Rule, Saturate, Scale, Share, Sort, State, **Structure**, Synchronization, Temporary, Time, Timing, Trace, Void...

SFDPOT

List reordered from: http://www.testingreferences.com/functional_testing_heuristics.php and combined with SFDPOT (devised by James Bach)

<http://curioustester.blogspot.hu/2009/10/power-of-heuristics-and-mneumonics.html>

<http://www.satisfice.com/tools/htsm.pdf>

Heuristics can help you focus your testing

Create your account

First Name



Last Name



Email



Password



Repeat Password



Company Name

☐

I AGREE WITH TERMS AND CONDITIONS

Enter the arena

Testgyltices

FOR THE KUNG FU TESTER IN YOU

Sign in to your account

Username



Password



Enough Talk. Let's fight.

[Forgot password?](#)

Social Sign In

f

g

Example of an online tool to help you focus your testing

Decide what to test and where

Setting the scope and focus of the testing

WHAT FEATURES WOULD YOU LIKE TO TEST?

x SEARCHING

x USABILITY

x FILE HANDLING

x EFFECT OF CHANGING THE SETTINGS

Not listed above? Enter features you want to test separated by commas

DEVICE TO TEST ON

x Samsung Galaxy Ace

x LG Google Nexus 5

x ASUS Nexus

BROWSER TO TEST ON

|

NETWORKS TO TEST ON

x Wifi

x 3G

OPERATING SYSTEMS TO TEST ON

x Android

CARRIER TO TEST ON

x Vodafone

WHAT IS YOUR TEST FOCUS?

x FUNCTIONALITY

x UX

Pick testing models (heuristics)

Heuristics can guide our testing

1

2

3

Environment

Model Selection

Test Ideas

SEARCHING

USABILITY

FILE HANDLING

EFFECT OF CHANGING THE SETTINGS

MODELS

☐ CRUSSPIC ?

☒ HICCUPPSF ?

☒ RIMGEA ?

☐ FAILURE ?

☐ CCD IS EARI ?

☐ MUTII ?

☒ SPIES ?

☐ SEED NATALI ?

☐ CRUMBS ?

☐ STMPL ?

☐ CIDTESTD ?

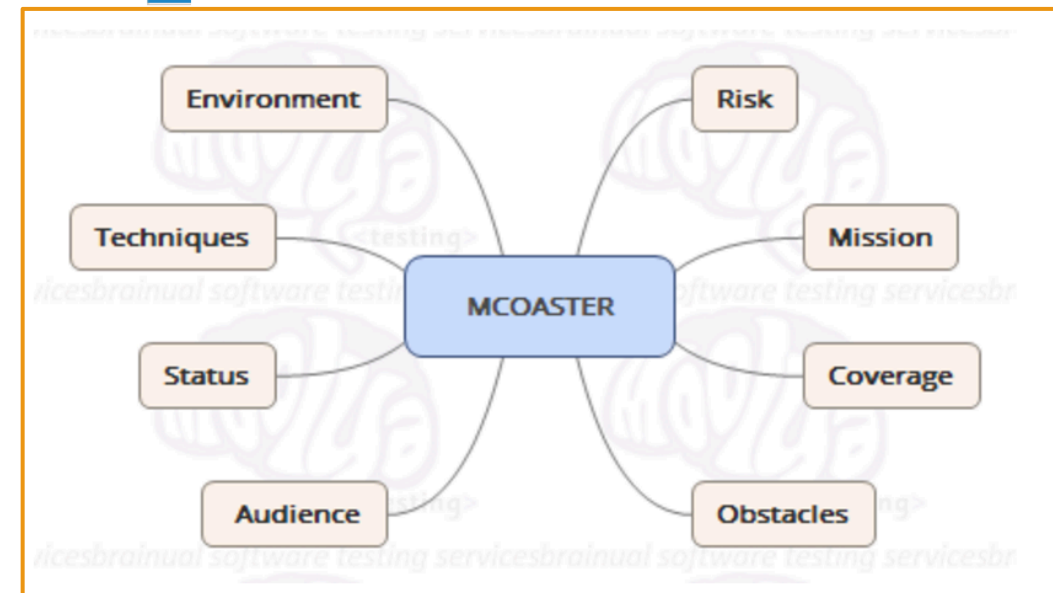
☐ SACKED SCOWS ?

☐ FCC CUTS VIDS ?

☐ FIBLOTS ?

☐ RCRCRC ?

☒ I SLICED UP FUN ?



1

2

3

EnvironmentModel SelectionTest Ideas

SEARCHINGUSABILITYFILE HANDLINGEFFECT OF CHANGING THE SETTINGS

ICCUPPSF (History,Image,Comparable Product,Claims,User Expectations,Product,Purpose,Standards And Sta

MGEA (Replicate It,Isolate It,Maximize It,Gen

ATES

Audi

Platform,Function,User Scenarios,Network)

COP FLUN ON (Communication,Orient

s,Guidelines,Updates,Notifications)

SPIES (Speacters,Pages & Content,Integrations>Error Messages,Special Formats)

HEURISTIC(EAS)

ORACLE

COVERAGE TYPE

CLASSIFICATION

SPIES

wiki.kiwix.org/searchterms

FUNCTIONALITY

Basic Flow

Focus on Special Characters & Content

SPIES

Combining what with how to test
Help focus our testing and save time



PROBLEM

NO PROBLEM

SHOW ALL

Export CSV

50 ▼

Search...

HEURISTICS (IDEAS) ▴ ▾

🔍 Filter HEURISTICS (IDEAS)

ORACLE ▴ ▾

🔍 Filter ORACLE

DEVICES ▴ ▾

🔍 nexus

ENVIRONMENT ▼

PROBLEM? ▴ ▾

🔍 Filter PROBLEM?

cop flung gun

Development Team

ASUS Nexus



SELECT



cop flung gun

Development Team

ASUS Nexus



SELECT



cop flung gun

Development Team

LG Google Nexus 5



SELECT



cop flung gun

Development Team

LG Google Nexus 5



SELECT



SPIES

Platform, File system limits

ASUS Nexus



SELECT



Clear progress reporting
Also helps tester(s) keep track

TEST RUNS



1

BUGS YOU FOUND



02:20:15

MINUTES



0

RISKS IDENTIFIED

BUG ID ?	ISSUE TYPE ?	FEATURE ?	WHAT ?	WHY ?	HOW ?	PERCEIVED PAIN ?	STATUS ?
BUG-2770	BUG	EFFECT OF CHANGING THE SETTINGS	App crashes	App crashes when user selects the Settings option	Select the Settings option from the drop-down on Android 4.x	HIGH	Problem

VIEW FULL BUG REPORT +

Track progress, risks and time taken

Some benefits of analytics data

No longer limited to what “we” think “they” need/do/etc.
Discover how the app behaves across virtually all the population

Lower cost of operations

Real usage can help drive our testing and analysis; brings realism to our testing

Guiding testing using data & analytics

Examples of what would be helpful

- Create new models based on what seems to suit/work for the tester.
- Data on which guides have been effective for similar apps.
- Guidance to test around a bug to detect the extent of the bug e.g. OS release, previous release of the app, device type, screen size, etc.
- Testing guides for each heuristic/constituent.

Trends,
Defect Reports

Engineering Activity,
Benchmarking, Testing

Extrapolation

Past

Present

Future

Information

What's happened?
(Reporting)

What's happening now?
(Alerts)

What will happen?
(Forecasting)

Insight

How and why did it happen?
(Factor analysis)

What is the next best action?
(Recommendation)

What's the best/worst that can
happen?
(Modeling / Simulation)

Software quality models,
bottleneck analysis

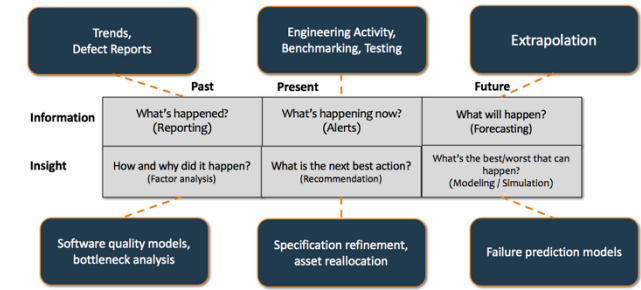
Specification refinement,
asset reallocation

Failure prediction models

<http://research.microsoft.com/pubs/136974/foser-2010-buse.pdf>

Analytics for Software Development

Testing based on analytics data



TESTING BASED ON INFORMATION

Popularity & volumes

Locales

User-flows, activities, etc.

Crashes

Aims include

Fast reproduction

1st-hand learning in controlled env.

TESTING BASED ON INSIGHTS

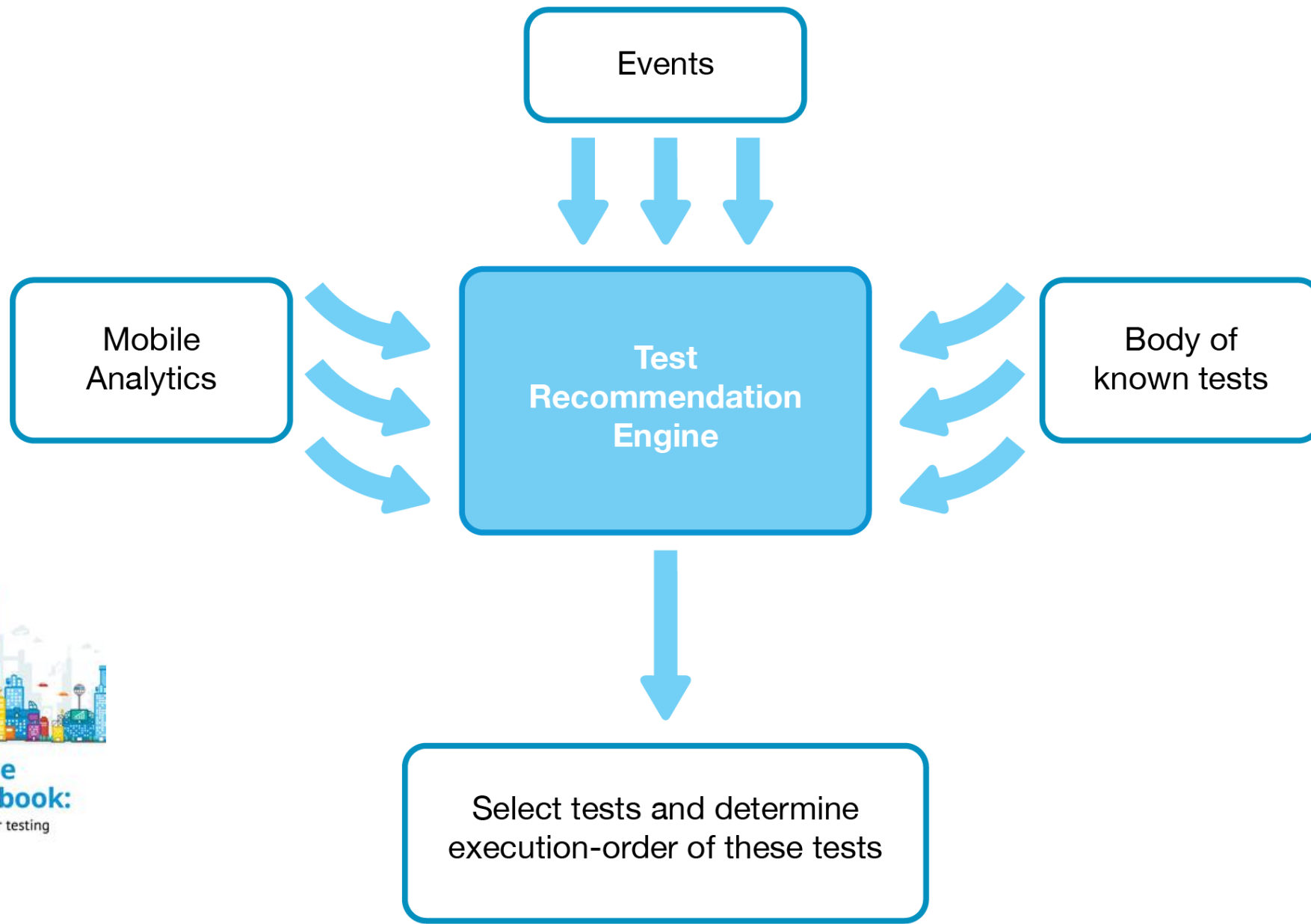
Delta's (rate of change)

Crossing thresholds

Anomalies

Aims include:

Maximising insights, agility, and ability to adapt & respond



The Mobile Analytics Playbook:
A practical guide to better testing

Julian Harty

Antoine Aymer

My Ph.D Test Recommendation Engine

And yet...

Tension between:

Automation Complacency & Automation Bias

Ask

What are the key skills we want people who use the systems to have and to hone

How can we devise the system to help them do so?

The character and the goals of the work, as well as the standards by which it is judged, are shaped by the machine's capabilities

'The Glass Cage should be required
reading for everyone with a phone'
Jonathan Safran Foer

The Glass Cage

Who Needs
Humans
Anyway?

Nicholas Carr

people can add discernment

ROLES FOR HUMANS

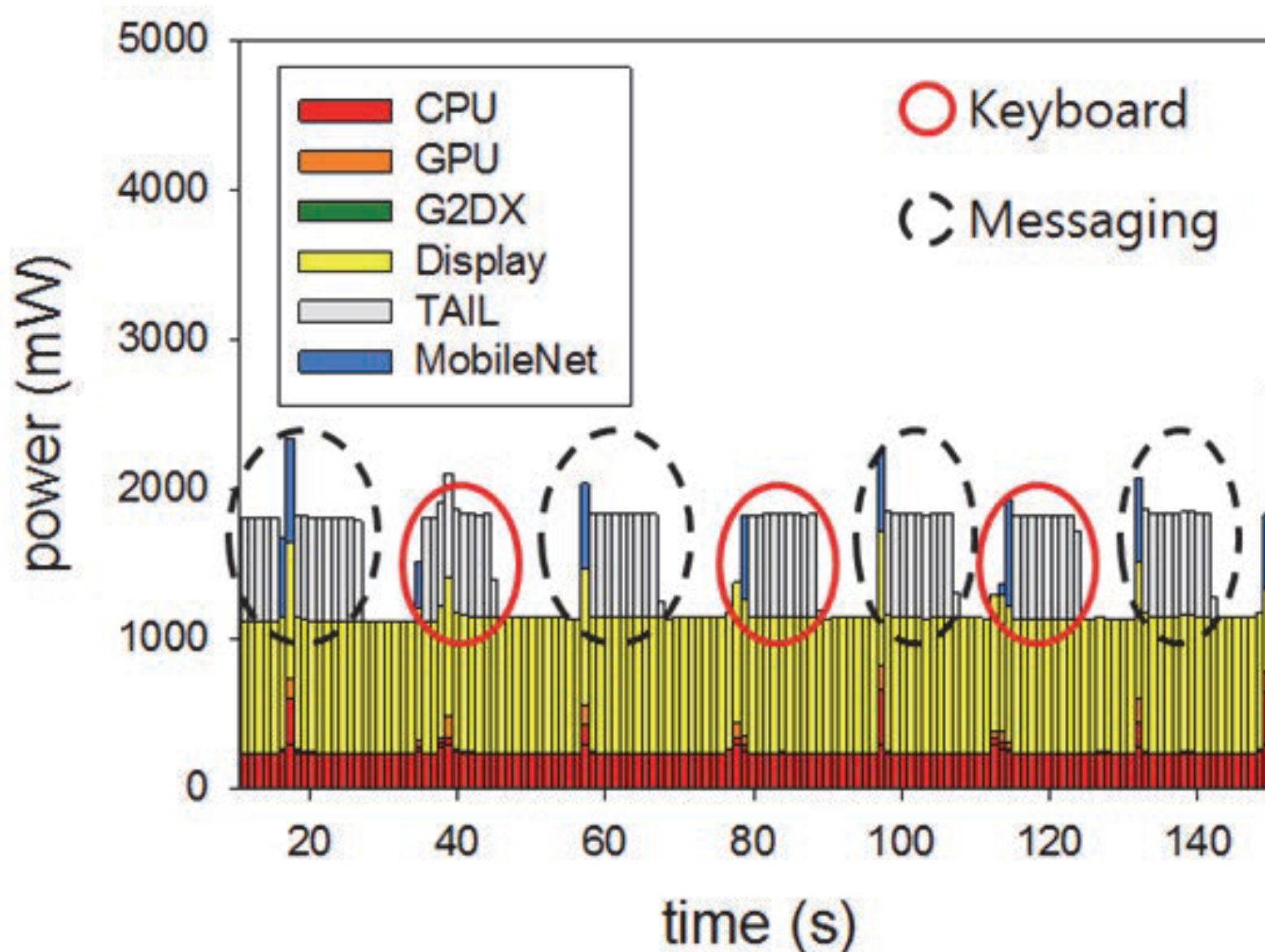


Figure © ACM

From: User Interaction-based Profiling System for Android Application Tuning

Correlation? Causation? Appropriate?

August Home Adds IFTTT
Integration to Smart Lock

LEARN MORE ►

IFTTT

August Home and Nest
Integration

LEARN MORE ►

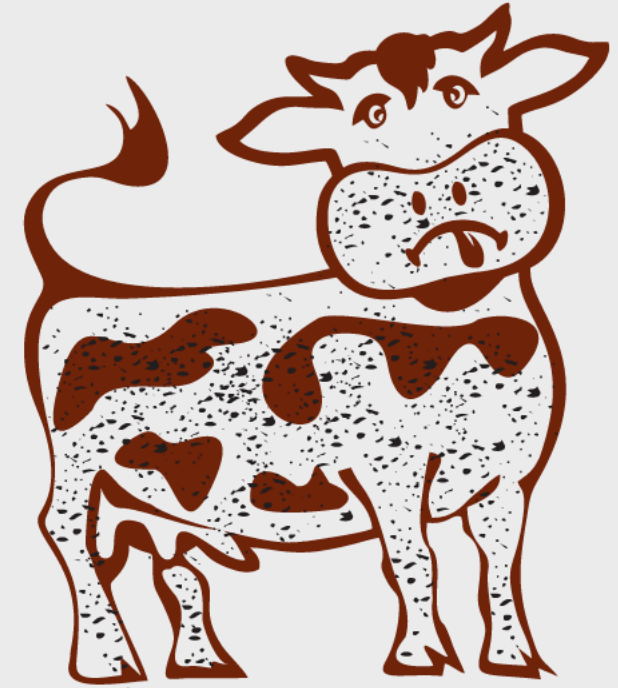


Research: 30 Million U.S.
Households to Add Smart
Home Technology

LEARN MORE ►



<http://august.com/>



DIRTY COW

<https://dirtycow.ninja/>

Smart Objects

Hard problems?

Cyber attack: hackers 'weaponised' everyday devices with malware to mount assault

Hundreds of thousands of devices such as webcams and DVRs were infected with malicious code to create a so-called 'botnet' to target leading sites



<https://www.theguardian.com/technology/2016/oct/22/cyber-attack-hackers-weaponised-everyday-devices-with-malware-to-mount-assault>

ACTi: admin/123456 or Admin/123456
American Dynamics: admin/admin or admin/9999
Arecont Vision: none
Avigilon: Previously admin/admin, changed to Administrator/<blank> in later firmware versions
Bosch: None required, but new firmwares (6.0+) prompt users to create passwords on first login
...
Samsung Electronics: root/root or admin/4321
Samsung Techwin (old): admin/1111111
Samsung (new): Previously admin/4321, but new firmwares require unique password creation

<https://ipvm.com/reports/ip-cameras-default-passwords-directory>

Smart Objects

How can we help users be smart too

learning from medicine

HIGH VALUE CARE

“High Value Care improves health, avoids harms, and eliminates wasteful practices.”

Better testing improves the health of the software and the team, avoids harm and eliminates wasteful practices.

<https://www.acponline.org/clinical-information/high-value-care>

“1-Minute Consults, called Smart Testing”

Lightweight, fast assessments can help improve our practices.

[http://www.ccjm.org/?id=105745&tx_ttnews\[tt_news\]=369817&cHash=25e67e4e823dd22a9e41a94a7cd30202](http://www.ccjm.org/?id=105745&tx_ttnews[tt_news]=369817&cHash=25e67e4e823dd22a9e41a94a7cd30202)

[2] Perhaps can we learn from medicine...

Quotes from medical articles	Possible application to testing software
<p>“The Smart Testing series examines scenarios in which diagnostic tests are commonly ordered in the absence of supporting data.</p>	<p>Evidence of the results of our testing, including assessments and retrospectives can help improve our testing.</p>
<p>The articles use evidence-based recommendations to clarify when testing is appropriate.</p>	<p>Focus, also, on discouraging testing that may be excessive or harmful.</p>
<p>Authors are encouraged to include specific examples of testing that may be excessive or harmful.”</p> <p>http://www.ccjm.org/corporate-links/journal-info/instructions/smart-testing-instructions.html</p>	
<p>[3] Perhaps can we learn from medicine...</p>	

Quotes from medical articles

“ACP defines High Value Care as the delivery of services providing benefits that make their harms and costs worthwhile.

Value is not merely cost.”

"The Institute of Medicine has estimated that \$765 billion is wasted annually in the United States on care that provides no value to patients,"

<https://www.acponline.org/acp-newsroom/smart-testing-tackles-chief-culprits-of-wasted-care-unnecessary-imaging-procedures-and-diagnostic>

Possible applications to testing software

Only embark on implementing and using stuff where the benefits far outweigh the potential harms and costs.

[3] Perhaps can we learn from medicine...

Quotes from medical articles	Possible application to testing software
“High Value Care improves health, avoids harms, and eliminates wasteful practices.”	Better testing improves the health of the software and the team, avoids harm and eliminates wasteful practices.
“1-Minute Consults, called Smart Testing”	Lightweight, fast assessments can help improve our practices.
<p>“The Smart Testing series examines scenarios in which diagnostic tests are commonly ordered in the absence of supporting data. The articles use evidence-based recommendations to clarify when testing is appropriate.</p> <p>Authors are encouraged to include specific examples of testing that may be excessive or harmful.”</p>	<p>Evidence of the results of our testing, including assessments and retrospectives can help improve our testing.</p> <p>Focus, also, on discouraging testing that may be excessive or harmful.</p>
“ACP defines High Value Care as the delivery of services providing benefits that make their harms and costs worthwhile. Value is not merely cost.”	Only embark on implementing and using stuff where the benefits far outweigh the potential harms and costs.

Perhaps can we learn from medicine...



Old systems are no longer fit for purpose
They stagnate and become obsolete

A cyclist in a red and black jersey is riding a road bike at night. The rear light of the bike is glowing brightly with a red and orange light, creating a lens flare effect. The background shows a dark landscape with some distant lights under a twilight sky.

INTRODUCING THE ICON REAR LIGHT

FROM SEE.SENSE

[LEARN MORE](#)

Smart Systems

Bicycles, Streetlights, ...

Where Humans can focus

- Questioning the use of automation to understand the purpose, implementation & limits
- Helping devise the testing of & for the new systems, concepts, processes
- Hone deep understanding to perceive where we might go awry
- Providing the human intuition &
- finding ways to improve the human to computer to human communications

Questions?

julianharty@gmail.com

Credits

Pradeep Soundararan of Moolya

Naresh Jain of Agile FAQs

Andy Glover: Cartoon Tester

Isabel Evans

Further Reading

<http://www.testingeducation.org/BBST/testdesign/>

Using Software Testing to Move Students from Trial-and-Error to Reflection-in-Action

<http://wrestt.cis.fiu.edu/about-wrestt-com> vetted Learning Objects in software testing.

Automation Bias in Intelligent Time Critical Decision Support Systems , M.L. Cummings

Heads-Up Displays (HUD)

- James Whittaker <https://testing.googleblog.com/2009/07/plague-of-blindness.html>
- Alan Richardson
<http://blog.eviltester.com/2010/03/a-poor-mans-testing-head-up-display.html>
- David Burns <http://tech.smartfocusdigital.com/2009/12/testers-heads-up-display.html>