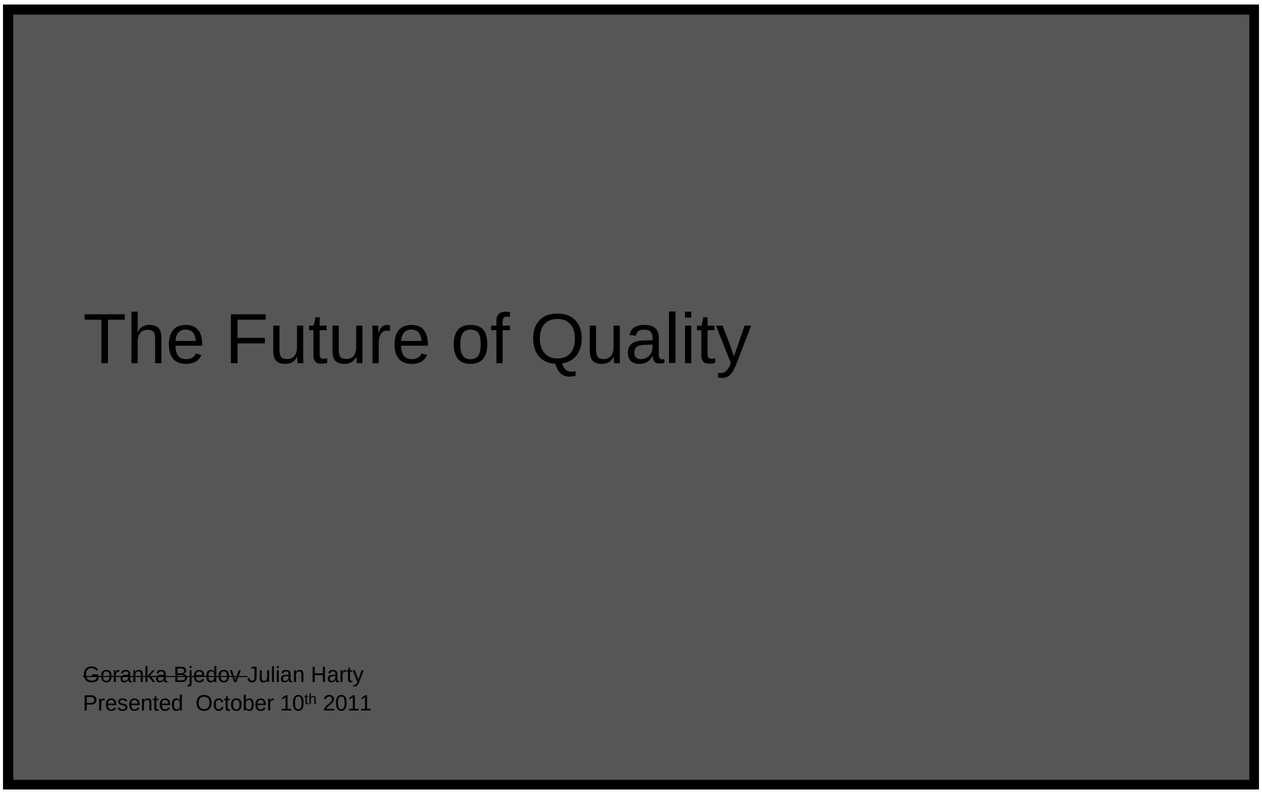
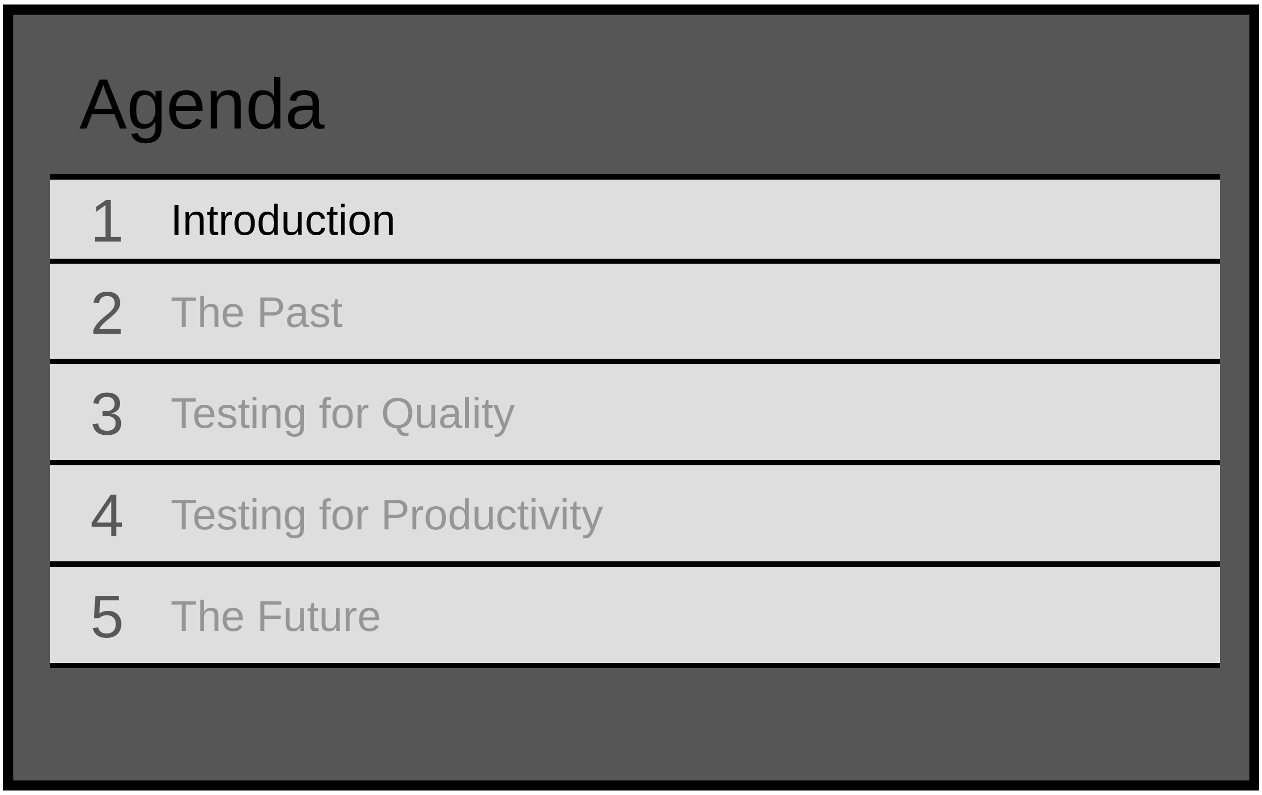


The Future of Quality

~~Goranka Bjedov~~ Julian Harty
Presented October 10th 2011



Agenda

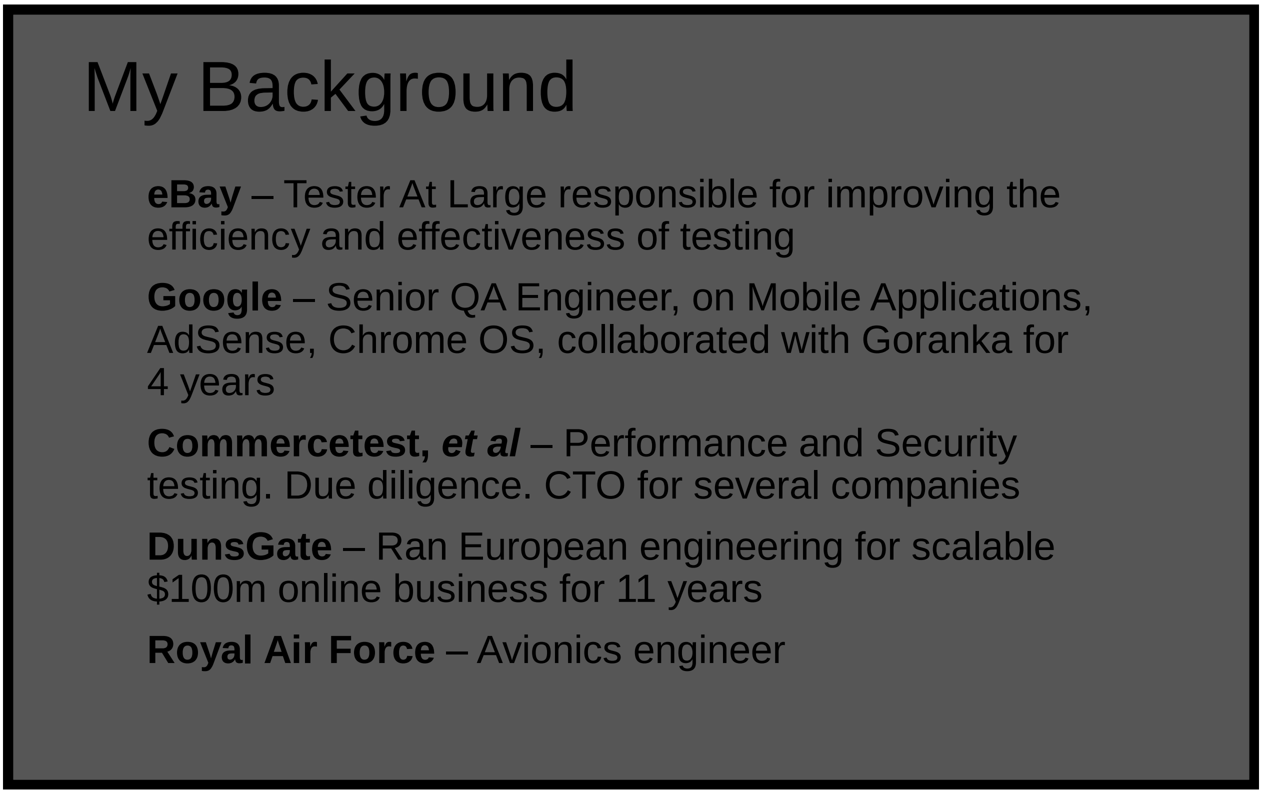
1 Introduction

2 The Past

3 Testing for Quality

4 Testing for Productivity

5 The Future



My Background

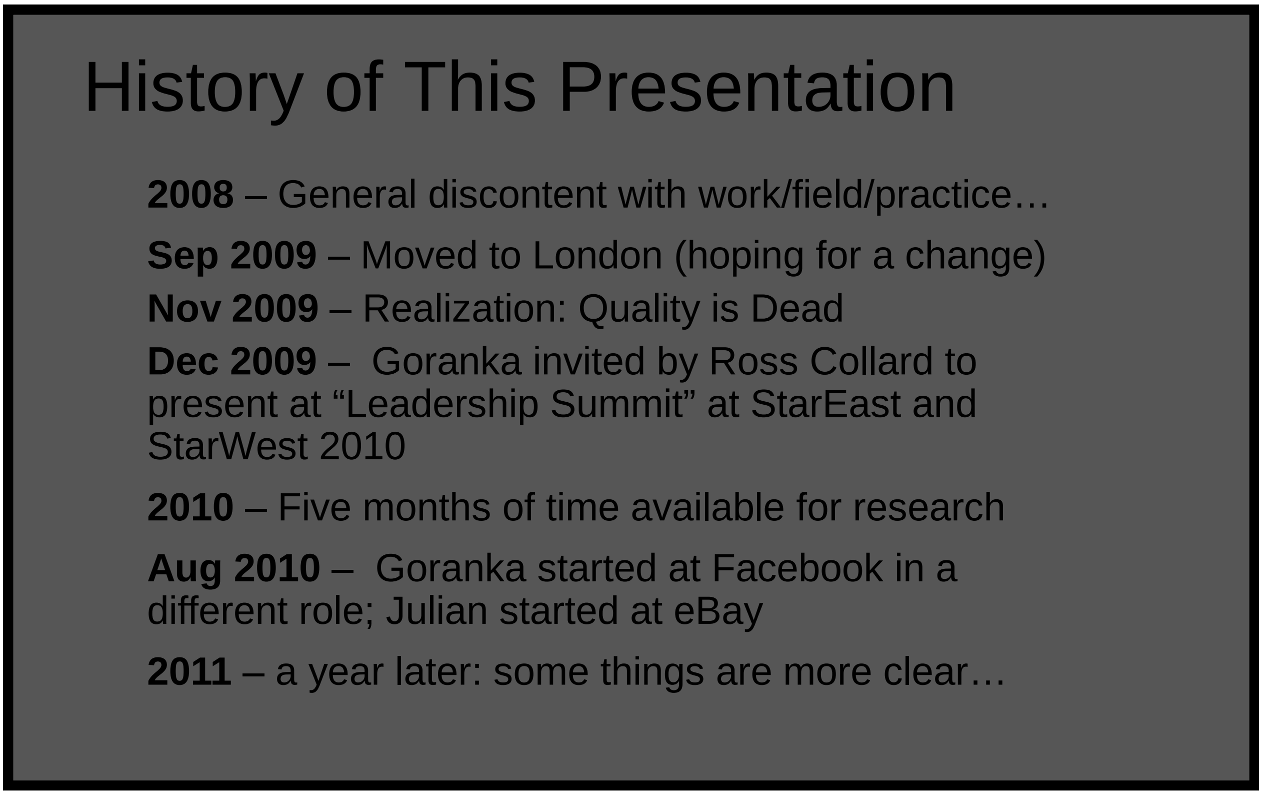
eBay – Tester At Large responsible for improving the efficiency and effectiveness of testing

Google – Senior QA Engineer, on Mobile Applications, AdSense, Chrome OS, collaborated with Goranka for 4 years

Commercetest, *et al* – Performance and Security testing. Due diligence. CTO for several companies

DunsGate – Ran European engineering for scalable $100m online business for 11 years

Royal Air Force – Avionics engineer



History of This Presentation

2008 – General discontent with work/field/practice…

Sep 2009 – Moved to London (hoping for a change)

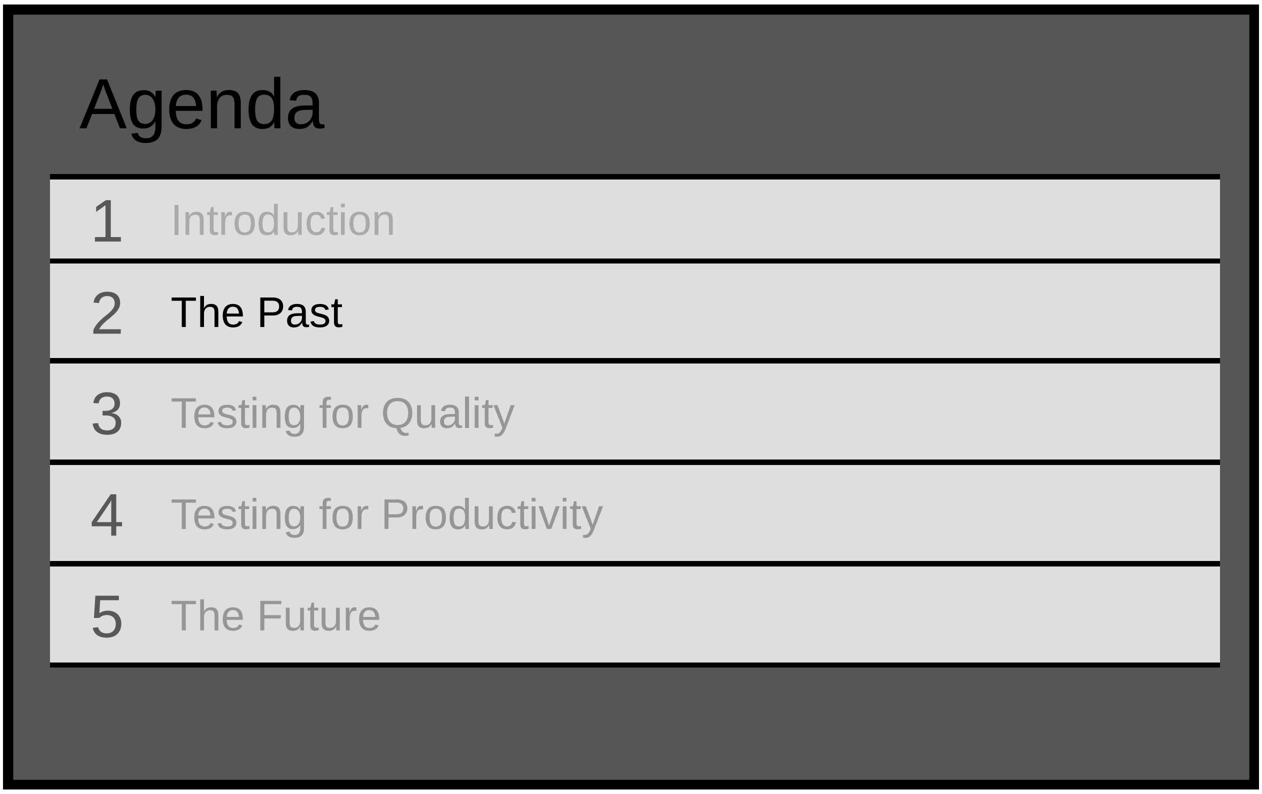
Nov 2009 – Realization: Quality is Dead

Dec 2009 – Goranka invited by Ross Collard to present at “Leadership Summit” at StarEast and StarWest 2010

2010 – Five months of time available for research

Aug 2010 – Goranka started at Facebook in a different role; Julian started at eBay

2011 – a year later: some things are more clear…



Agenda

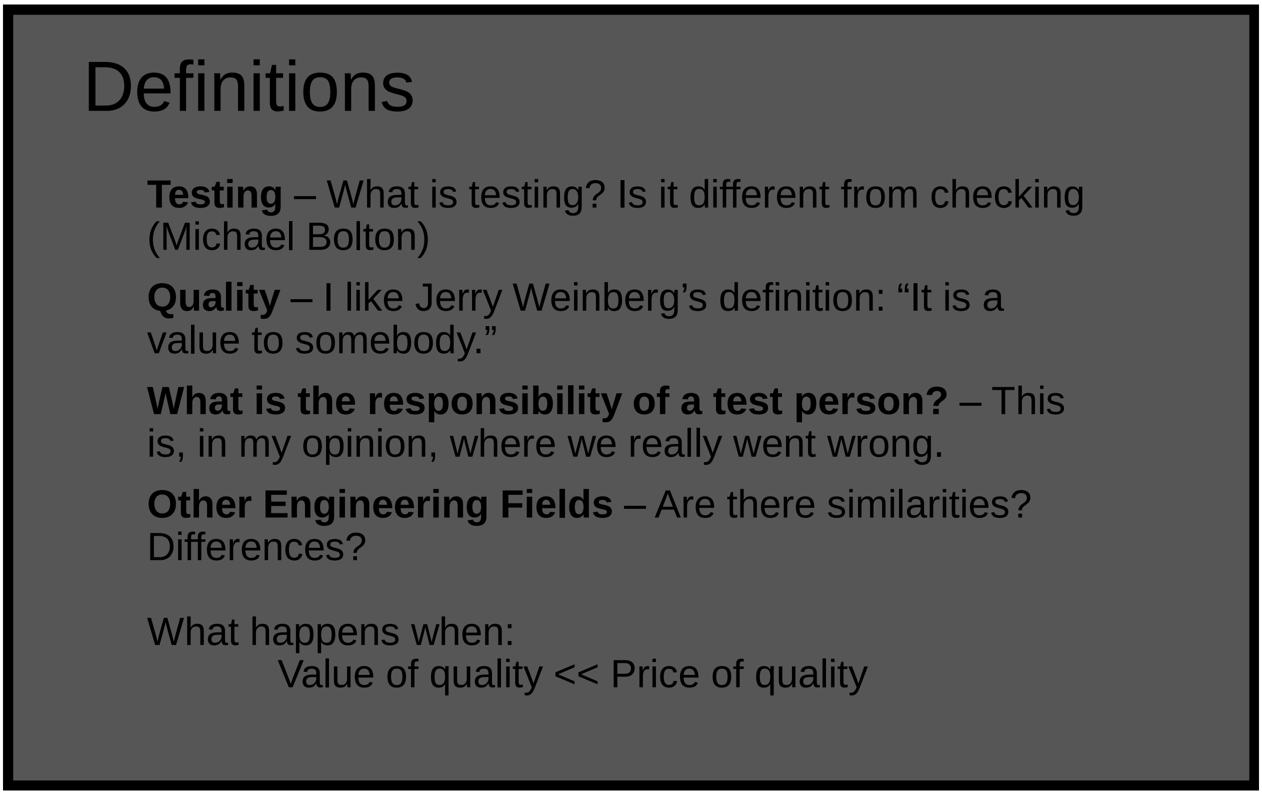
1 Introduction

2 The Past

3 Testing for Quality

4 Testing for Productivity

5 The Future



Definitions

Testing – What is testing? Is it different from checking (Michael Bolton)

Quality – I like Jerry Weinberg’s definition: “It is a value to somebody.”

What is the responsibility of a test person? – This is, in my opinion, where we really went wrong.

Other Engineering Fields – Are there similarities? Differences?

What happens when:
Value of quality << Price of quality

The Past

Packaged Software – testing/bug fixing cycle follows development which was preceded by design.

Cost of a Bug – bugs found and fixed in design are cheaper then those found and fixed in development, which in turn are cheaper then those fixed in the field.

Test Team Contribution – finding bugs.

Systems – relatively simple.

Code – relatively simple and traceable.

Development/Testing Tools – very primitive.

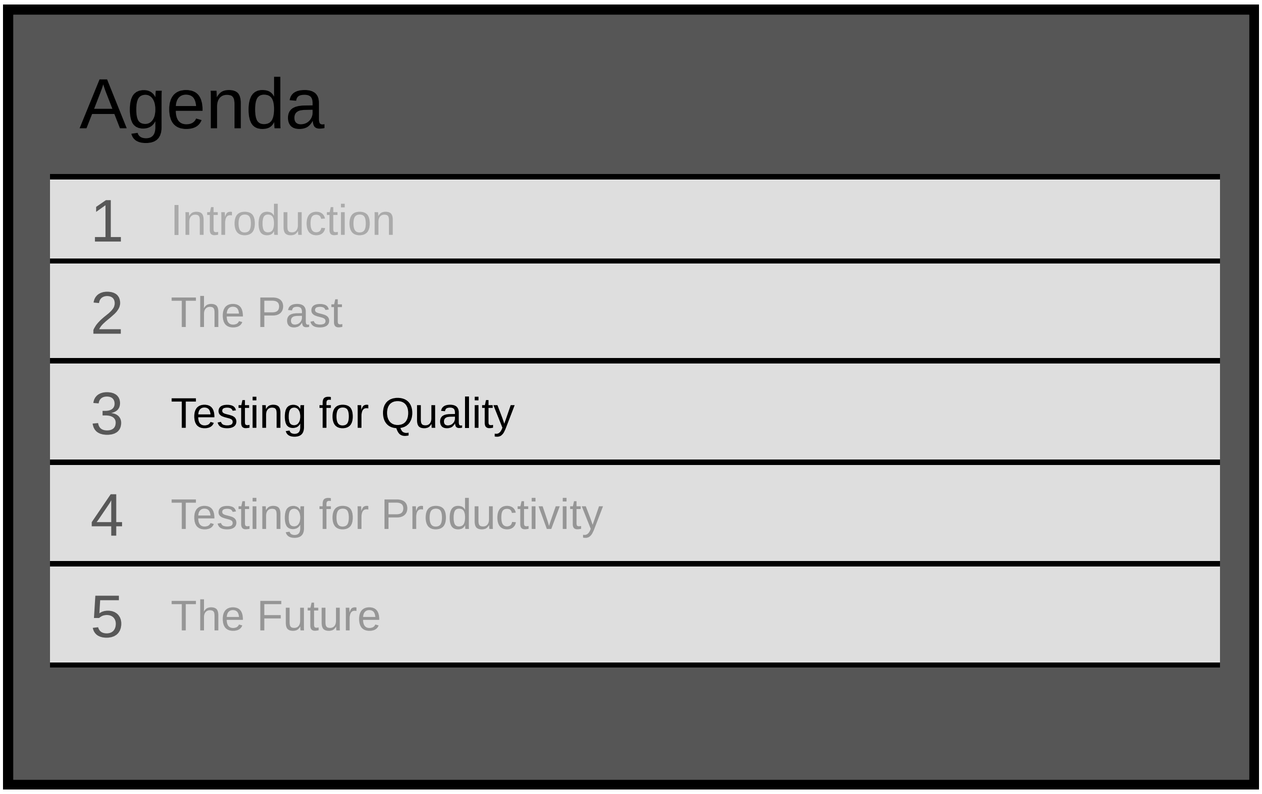
Testing Roles

Advocate for the Customer – it is our job to make sure nothing escapes.

Quality Consultant – it is our job to provide information on the quality of the software.

Quality Control – it is our job to stop bad products from being released.

IF WE BUILD OUR SOFTWARE WITH NO BUGS, WE CAN MAKE A 10% RETURN ON OUR INVESTMENT.
10%
Dilbert.com DilbertCartoonist@gmail.com
BUT IF WE DO A POOR JOB, WE CAN MAKE A 40% RETURN BY SELLING UPGRADES AND SERVICE.
2-3-11 ©2011 Scott Adams, Inc./Dist. by UFS, Inc.
BUT DON'T WORRY. WE ONLY HAVE THE BUDGET FOR A POOR JOB.
I CAN'T REMEMBER IF WE'RE CHEAP OR SMART.
PHEW!



Agenda

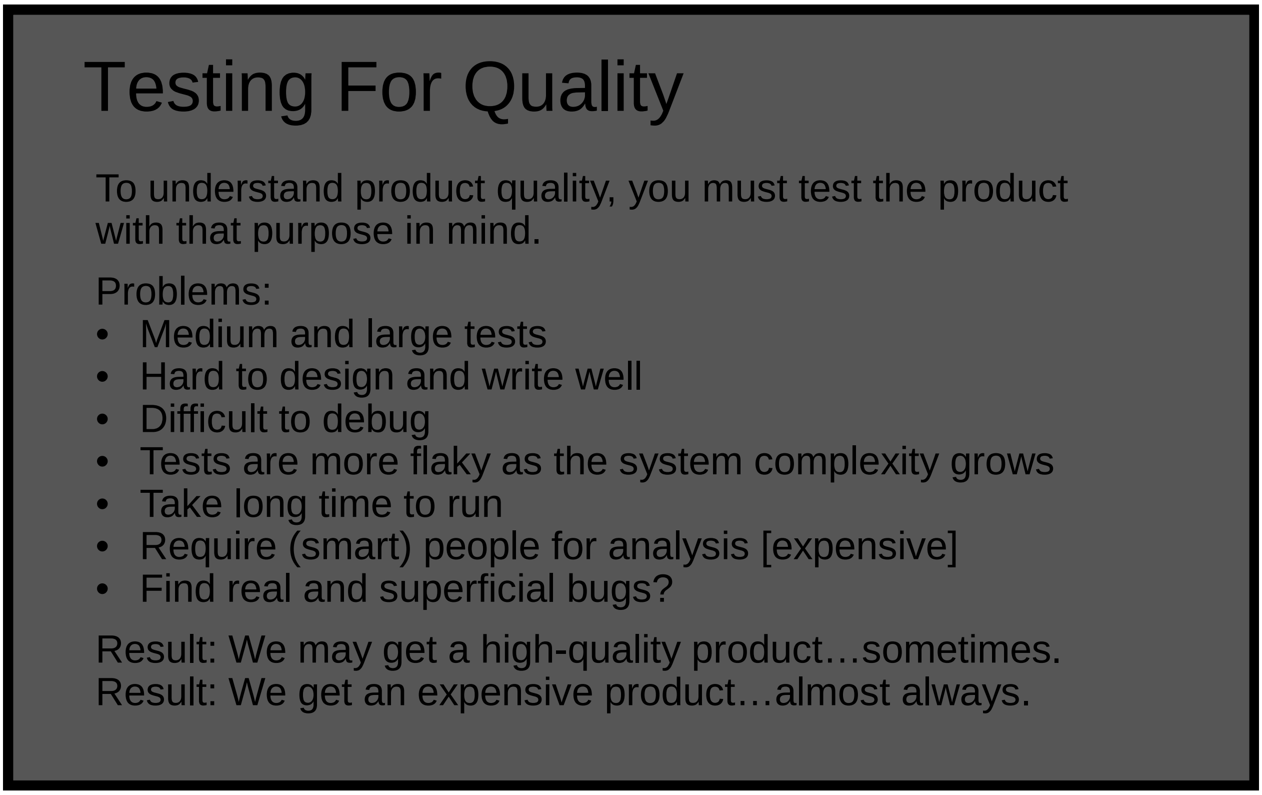
1 Introduction

2 The Past

3 Testing for Quality

4 Testing for Productivity

5 The Future



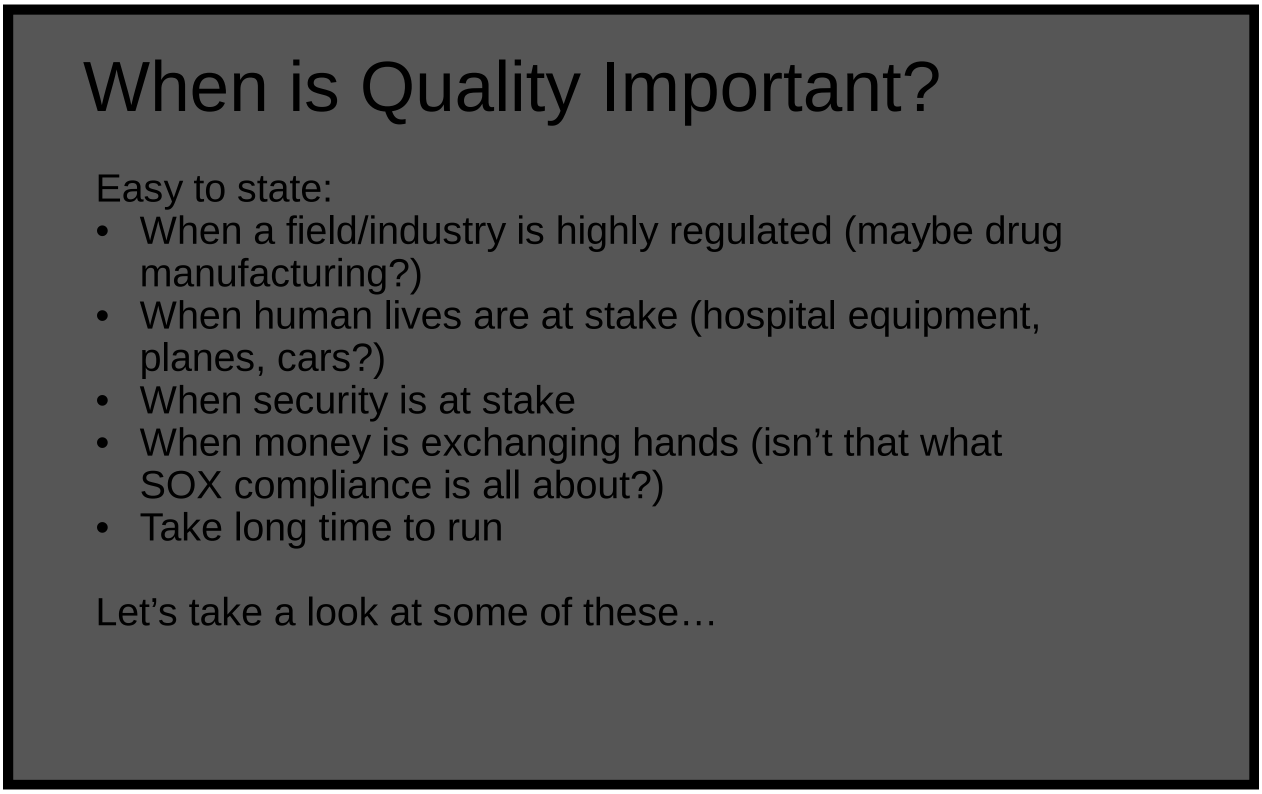
Testing For Quality

To understand product quality, you must test the product with that purpose in mind.

Problems:

- Medium and large tests
- Hard to design and write well
- Difficult to debug
- Tests are more flaky as the system complexity grows
- Take long time to run
- Require (smart) people for analysis [expensive]
- Find real and superficial bugs?

Result: We may get a high-quality product…sometimes.
Result: We get an expensive product…almost always.

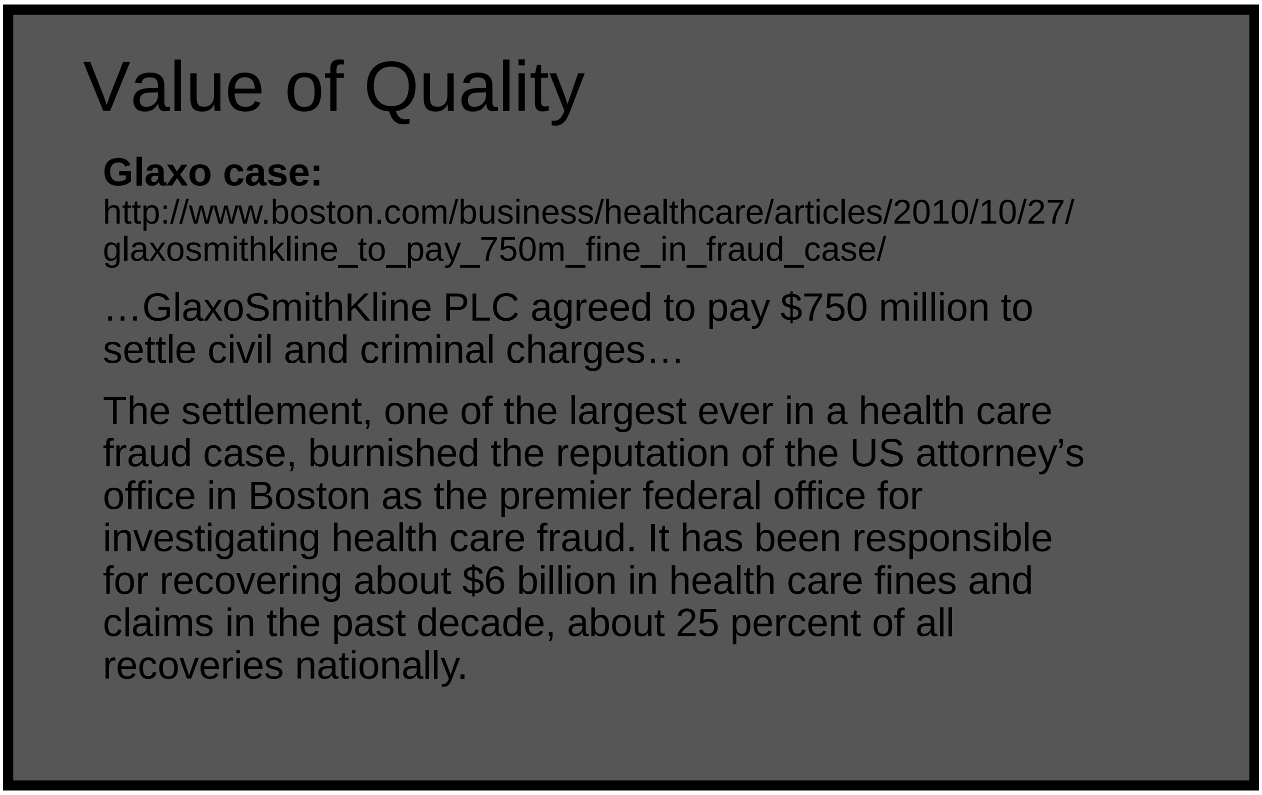


When is Quality Important?

Easy to state:

- When a field/industry is highly regulated (maybe drug manufacturing?)
- When human lives are at stake (hospital equipment, planes, cars?)
- When security is at stake
- When money is exchanging hands (isn't that what SOX compliance is all about?)
- Take long time to run

Let's take a look at some of these…



Value of Quality

Glaxo case:
http://www.boston.com/business/healthcare/articles/2010/10/27/glaxosmithkline_to_pay_750m_fine_in_fraud_case/

…GlaxoSmithKline PLC agreed to pay $750 million to settle civil and criminal charges…

The settlement, one of the largest ever in a health care fraud case, burnished the reputation of the US attorney's office in Boston as the premier federal office for investigating health care fraud. It has been responsible for recovering about $6 billion in health care fines and claims in the past decade, about 25 percent of all recoveries nationally.

Value of Quality, cont.

Airspeed inconsistency: http://en.wikipedia.org/wiki/Air_France_Flight_447
Do A330 aircraft have issues: http://www.spiegel.de/international/world/0,1518,766148,00.html

Not just aircraft: http://blogs.crikey.com.au/planetalking/2010/11/17/the-anatomy-of-the-airbus-a380-qf32-near-disaster/

Quote: Rolls-Royce had designed and was introducing a fix for the oil leak issues for this into the engines at its own speed. Qantas was left in the dark.

Value of Quality, cont.

It goes on and on:

- Amazon data center down April 21, 2011 http://www.bbc.co.uk/news/technology-13160929
- Gmail deletes user data for 150K users http://www.crunchgear.com/2011/02/28/storm-clouds-gmail-failure-reinforces-danger-of-becoming-too-cloud-dependent/
- Flikr deletes wrong paid accounts http://techcrunch.com/2011/02/02/flickr-accidentally-wipes-out-account-five-years-and-4000-photos-down-the-drain/
- ~~Medical devices (personal experience)~~
- Nodar Kumaritashvili, Feb 12 2010

Value of Quality, cont.

John Viega: "The Myths of Security" chapter 28,
"When Will We Get Rid of All the Security Vulnerabilities?"

2005 to today, average of 7000 vulnerabilities publicly disclosed

Quote: "But, if you're any other company (well, except for maybe a big financial company or a government agency), the cost of "doing it right" is far higher than the cost of only responding when someone does find a problem, and other companies are unlikely to be risking brand damage."

Some Entertainment

Cable bill - $16.4 million: http://www.daytondailynews.com/business/time-warner-charges-wright-patt-engineer-16-4-million-for-cable-1117224.html

While the dollar amount of DeVirgilio's billing ordeal makes it among the more egregious in recent memory, the kings of all billing mishaps have to be… Jon Seale, received notice that he owed a 17-figure sum that totaled almost 2,000 times the national debt: 23 quadrillion, 148 trillion, 855 billion, 308 million, 184 thousand and 500 dollars. The other, Josh Muszynski, was charged 23 quadrillion after buying a pack of cigarettes at a gas station.

Agenda

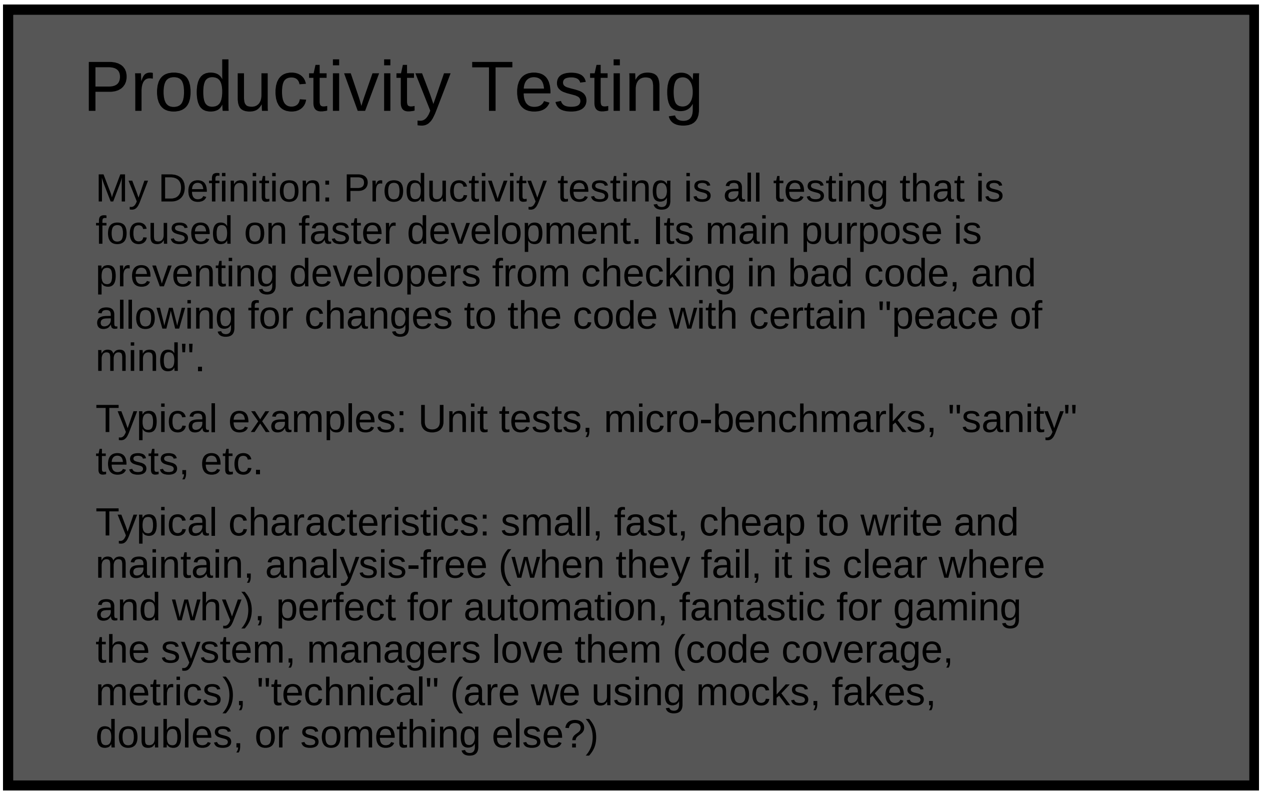
1 Introduction

2 The Past

3 Testing for Quality

4 Testing for Productivity

5 The Future

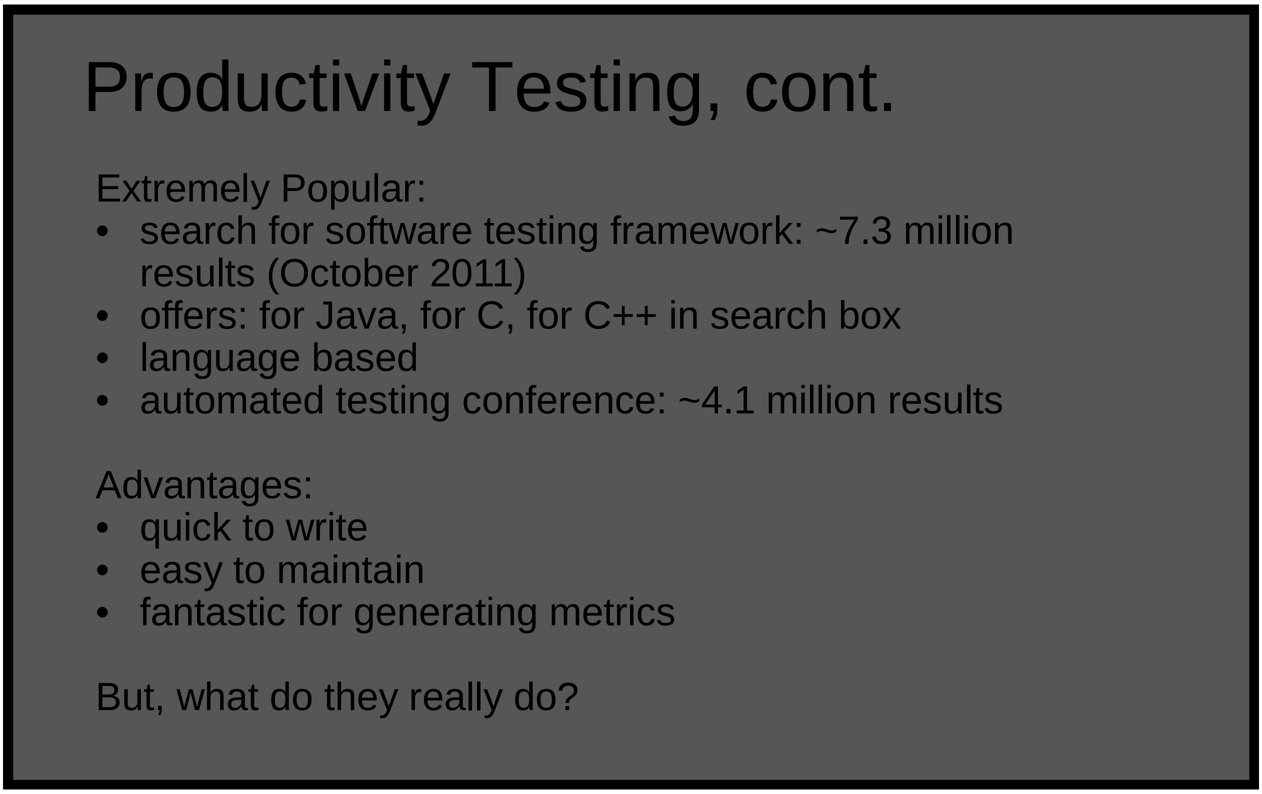


Productivity Testing

My Definition: Productivity testing is all testing that is focused on faster development. Its main purpose is preventing developers from checking in bad code, and allowing for changes to the code with certain "peace of mind".

Typical examples: Unit tests, micro-benchmarks, "sanity" tests, etc.

Typical characteristics: small, fast, cheap to write and maintain, analysis-free (when they fail, it is clear where and why), perfect for automation, fantastic for gaming the system, managers love them (code coverage, metrics), "technical" (are we using mocks, fakes, doubles, or something else?)



Productivity Testing, cont.

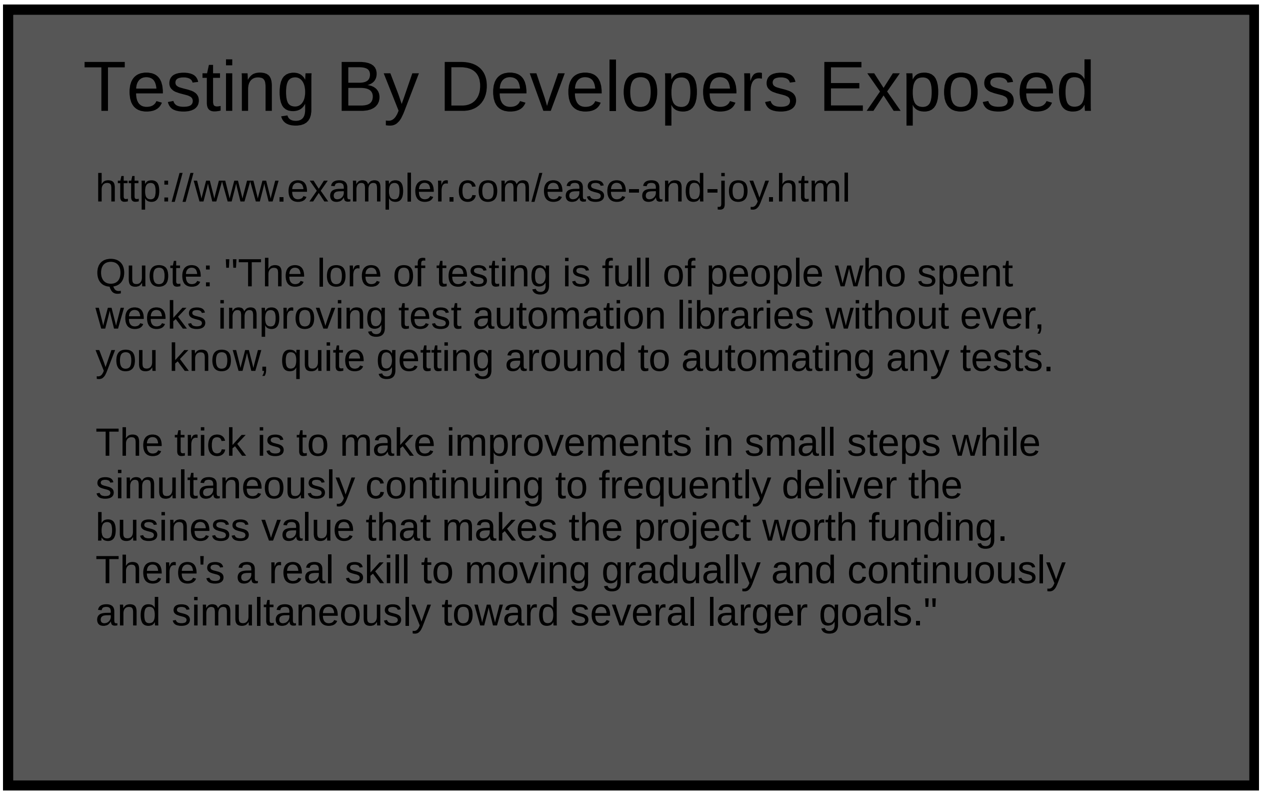
Extremely Popular:

- search for software testing framework: ~7.3 million results (October 2011)
- offers: for Java, for C, for C++ in search box
- language based
- automated testing conference: ~4.1 million results

Advantages:

- quick to write
- easy to maintain
- fantastic for generating metrics

But, what do they really do?



Testing By Developers Exposed

http://www.exampler.com/ease-and-joy.html

Quote: "The lore of testing is full of people who spent weeks improving test automation libraries without ever, you know, quite getting around to automating any tests.

The trick is to make improvements in small steps while simultaneously continuing to frequently deliver the business value that makes the project worth funding. There's a real skill to moving gradually and continuously and simultaneously toward several larger goals."

It is Usually Done Badly

Check any talks/papers/references. People talk about:

- Number of tests (Relevance?)
- Coverage
- Execution time (Fast = Good)
- Automated generation of tests/data
- Testing tools/harnesses/frameworks etc.
  e.g. http://www.gtac.biz/agenda

Interestingly, I cannot find any meaningful information on:

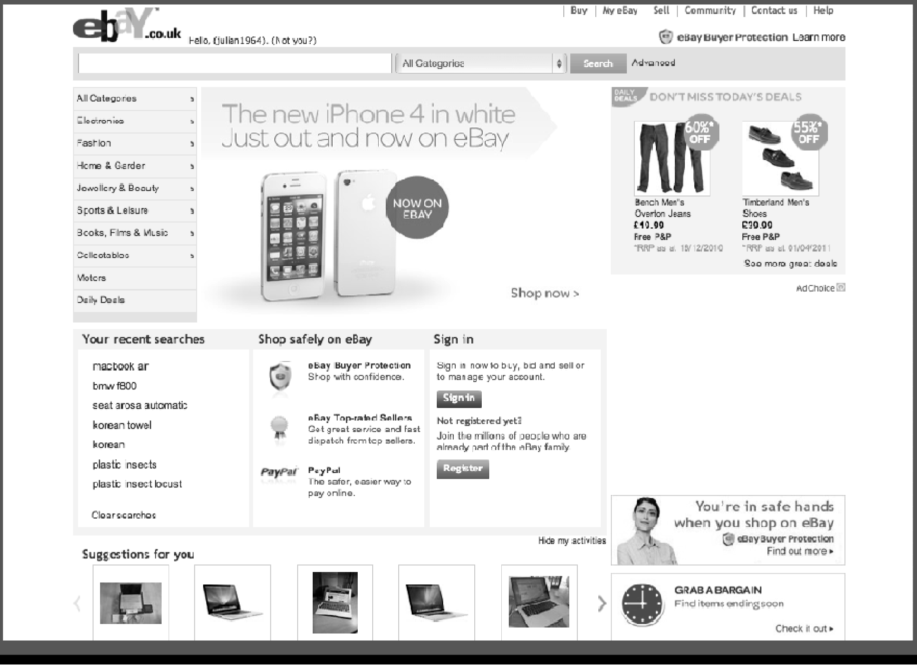
- What do your tests do?
- Who analyzes them?
- Do they save or waste money?

Over-confidence in the test results

- Automation as servant, not master, of our software delivery:
  - Inaccurate results
  - "Beware of Automation Bias" by M.L. Cumming[1]
- Automated tests and checks miss more than they find

[1] http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.91.2634&rep=rep1&type=pdf

What we think we are testing



Hello, (julian1964). (Not you?)

The new iPhone 4 in white
Just out and now on eBay

DON'T MISS TODAY'S DEALS

Bench Men's Overton Jeans
£10.99
Free P&P
*RRP as at 18/12/2010

Timberland Men's Shoes
£30.00
Free P&P
*RRP as at 01/04/2011

Your recent searches

macbook air
bmw f800
seat arosa automatic
korean towel
korean
plastic insects
plastic insect locust

Clear searches

Shop safely on eBay

eBay Buyer Protection
Shop with confidence.

eBay Top-rated Sellers
Get great service and fast dispatch from top sellers.

PayPal
The safer, easier way to pay online.

Sign in

Sign in now to buy, bid and sell or to manage your account.

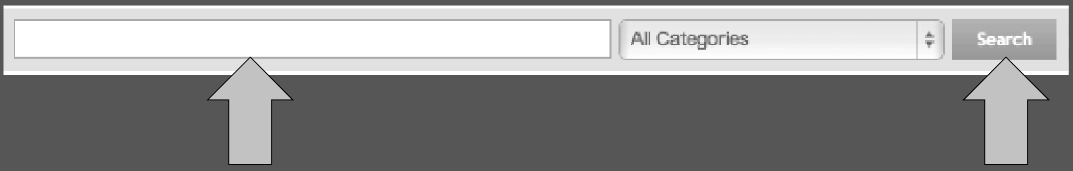
Not registered yet?
Join the millions of people who are already part of the eBay family.

Suggestions for you

You're in safe hands when you shop on eBay

GRAB A BARGAIN
Find items ending soon

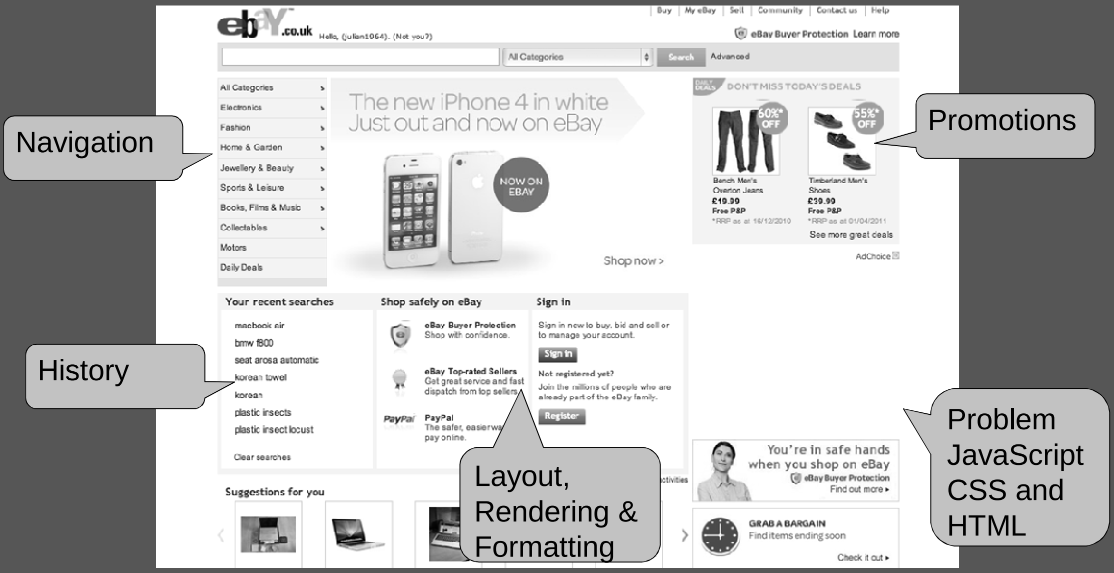
What our automated test actually interacts with…



Enter text here

‘click’

(Some of the) things our automated tests may miss…



Navigation

History

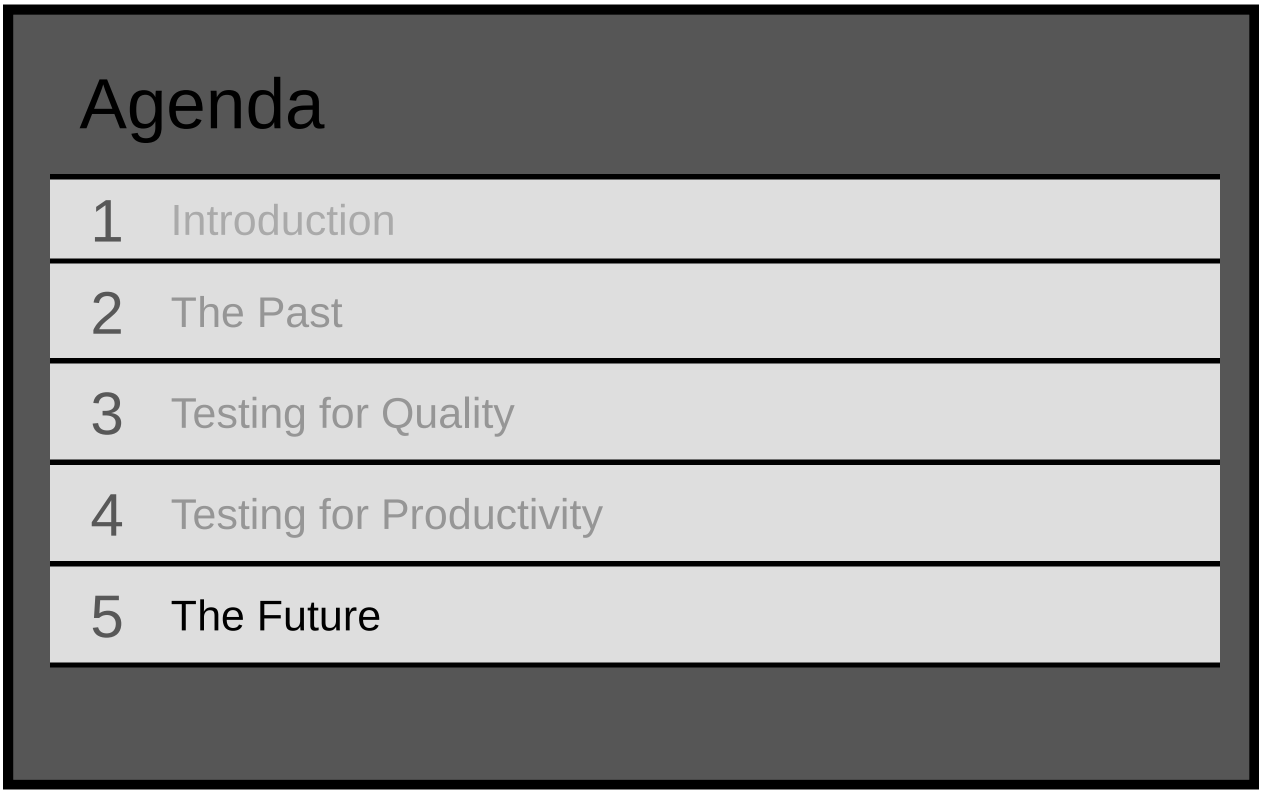
Layout, Rendering & Formatting

Promotions

Problem JavaScript CSS and HTML

Beware of poor-quality automated 'tests'

1. `AssertTrue(true);`
2. No/Inadequate/Too many Asserts
3. Poor code design
4. Falsehoods (False positives / negatives)



Agenda

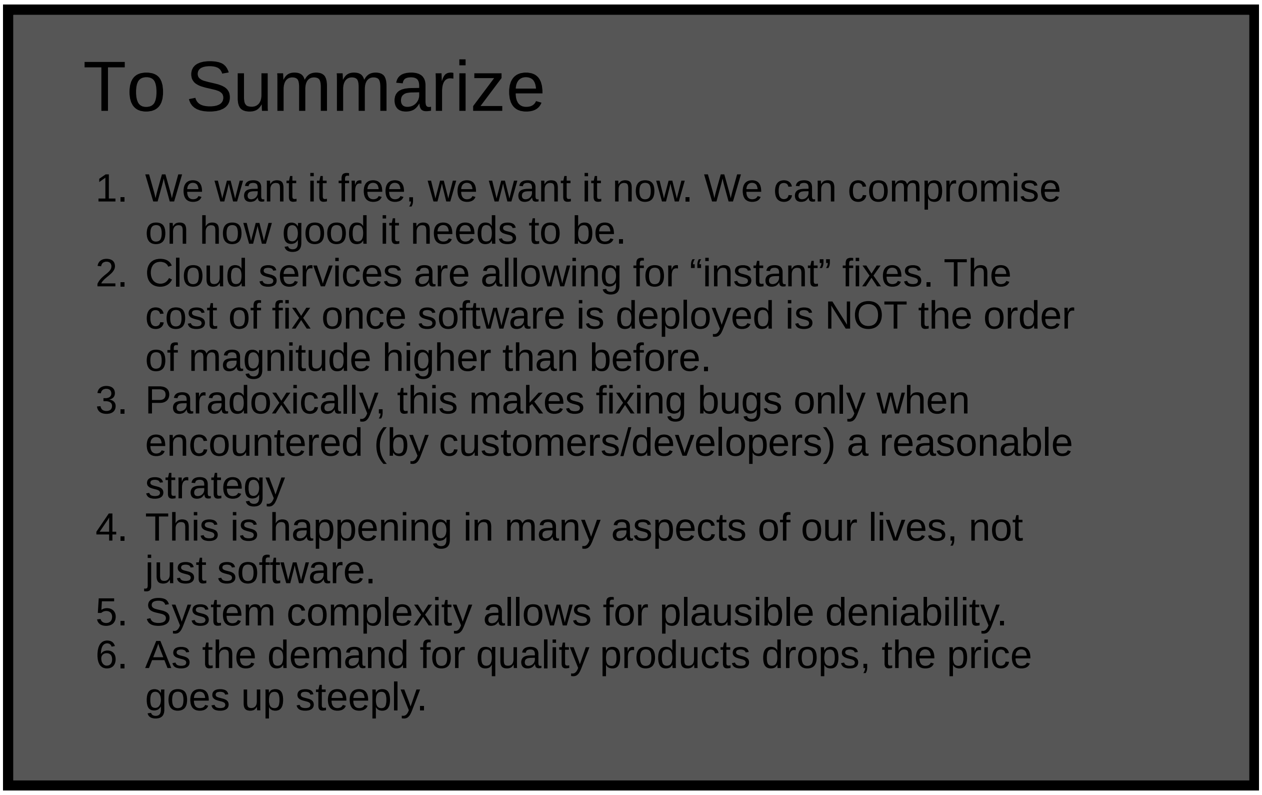
1 Introduction

2 The Past

3 Testing for Quality

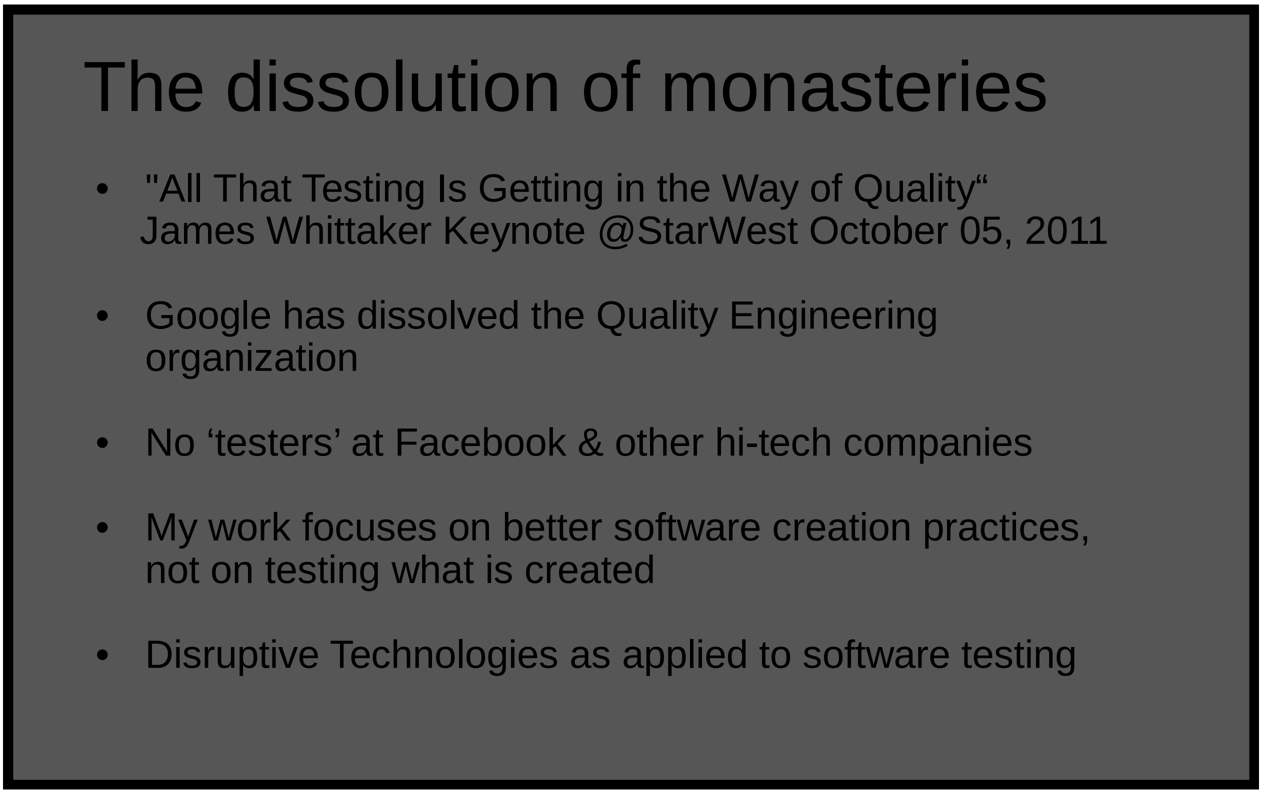
4 Testing for Productivity

5 The Future



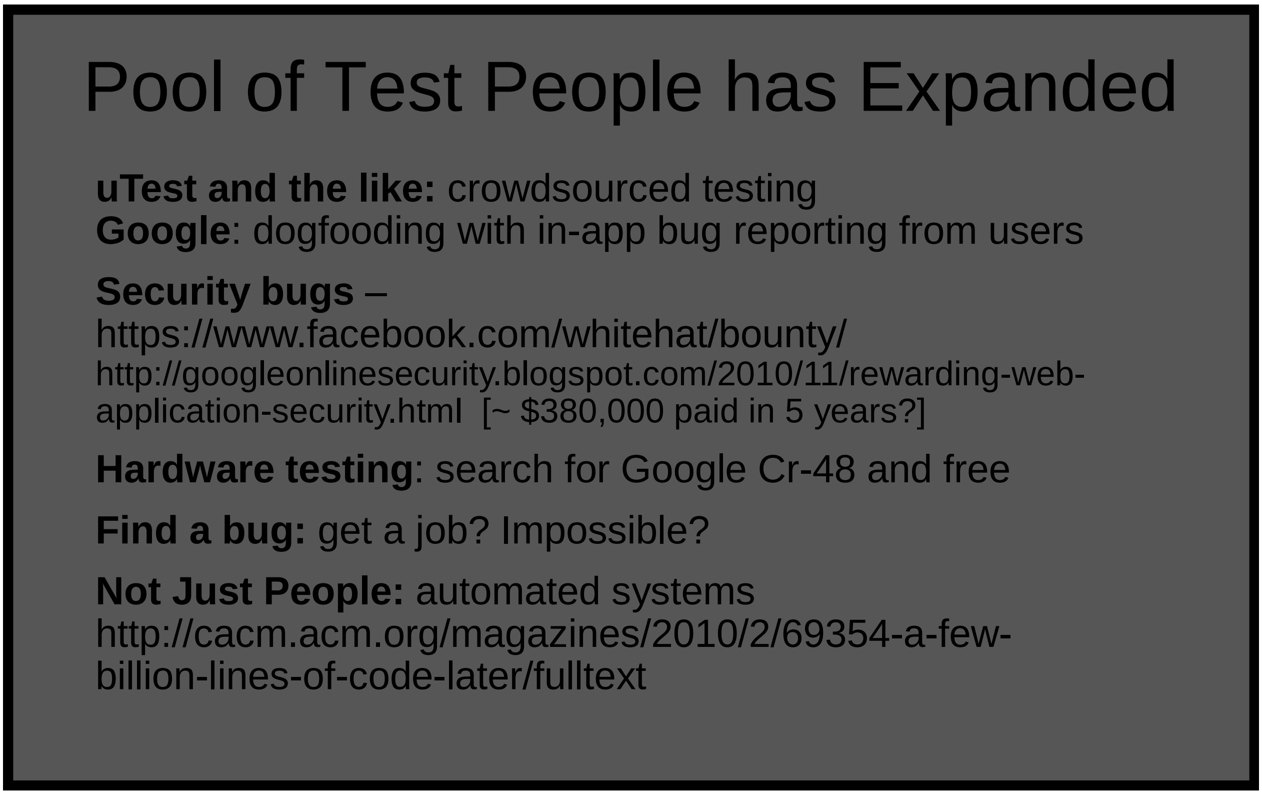
To Summarize

1. We want it free, we want it now. We can compromise on how good it needs to be.
2. Cloud services are allowing for “instant” fixes. The cost of fix once software is deployed is NOT the order of magnitude higher than before.
3. Paradoxically, this makes fixing bugs only when encountered (by customers/developers) a reasonable strategy
4. This is happening in many aspects of our lives, not just software.
5. System complexity allows for plausible deniability.
6. As the demand for quality products drops, the price goes up steeply.



The dissolution of monasteries

- "All That Testing Is Getting in the Way of Quality“
  James Whittaker Keynote @StarWest October 05, 2011
- Google has dissolved the Quality Engineering organization
- No ‘testers’ at Facebook & other hi-tech companies
- My work focuses on better software creation practices, not on testing what is created
- Disruptive Technologies as applied to software testing



Pool of Test People has Expanded

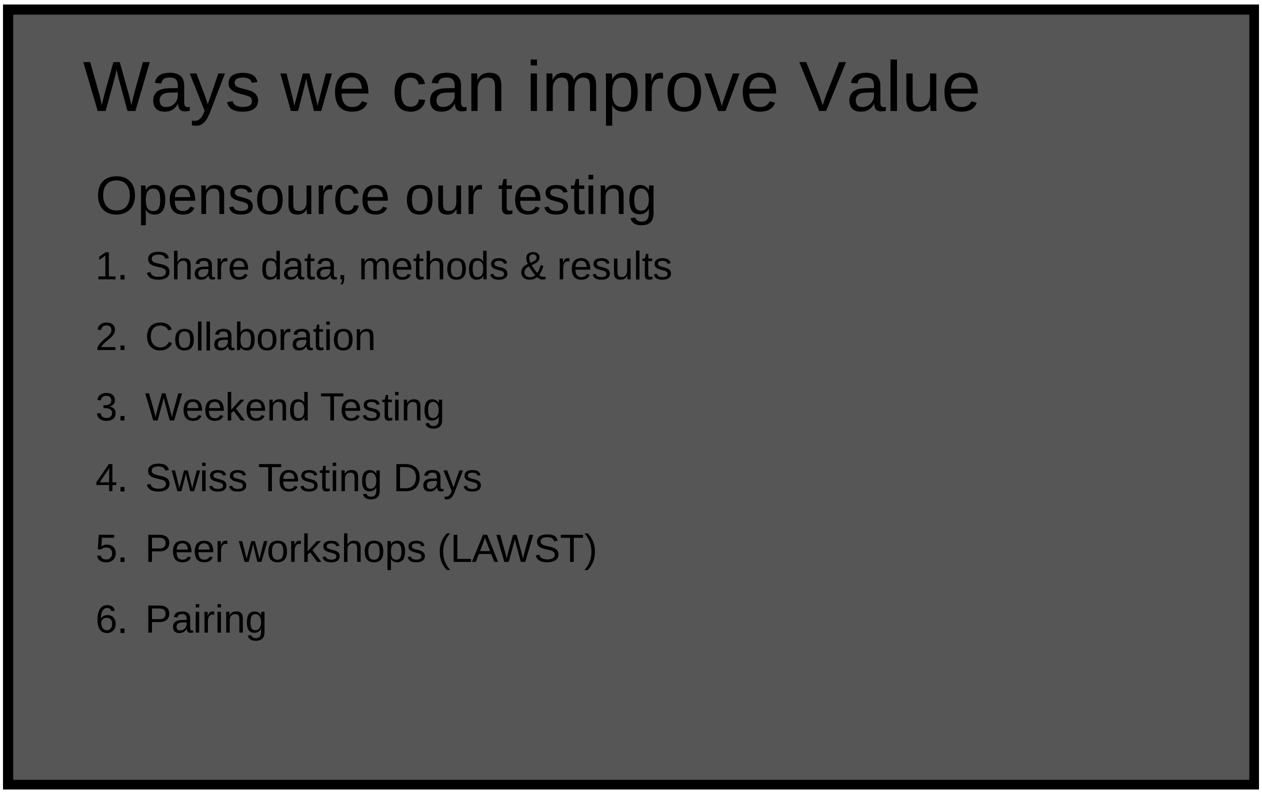
uTest and the like: crowdsourced testing
Google: dogfooding with in-app bug reporting from users

Security bugs –
https://www.facebook.com/whitehat/bounty/
http://googleonlinesecurity.blogspot.com/2010/11/rewarding-web-application-security.html [~ $380,000 paid in 5 years?]

Hardware testing: search for Google Cr-48 and free

Find a bug: get a job? Impossible?

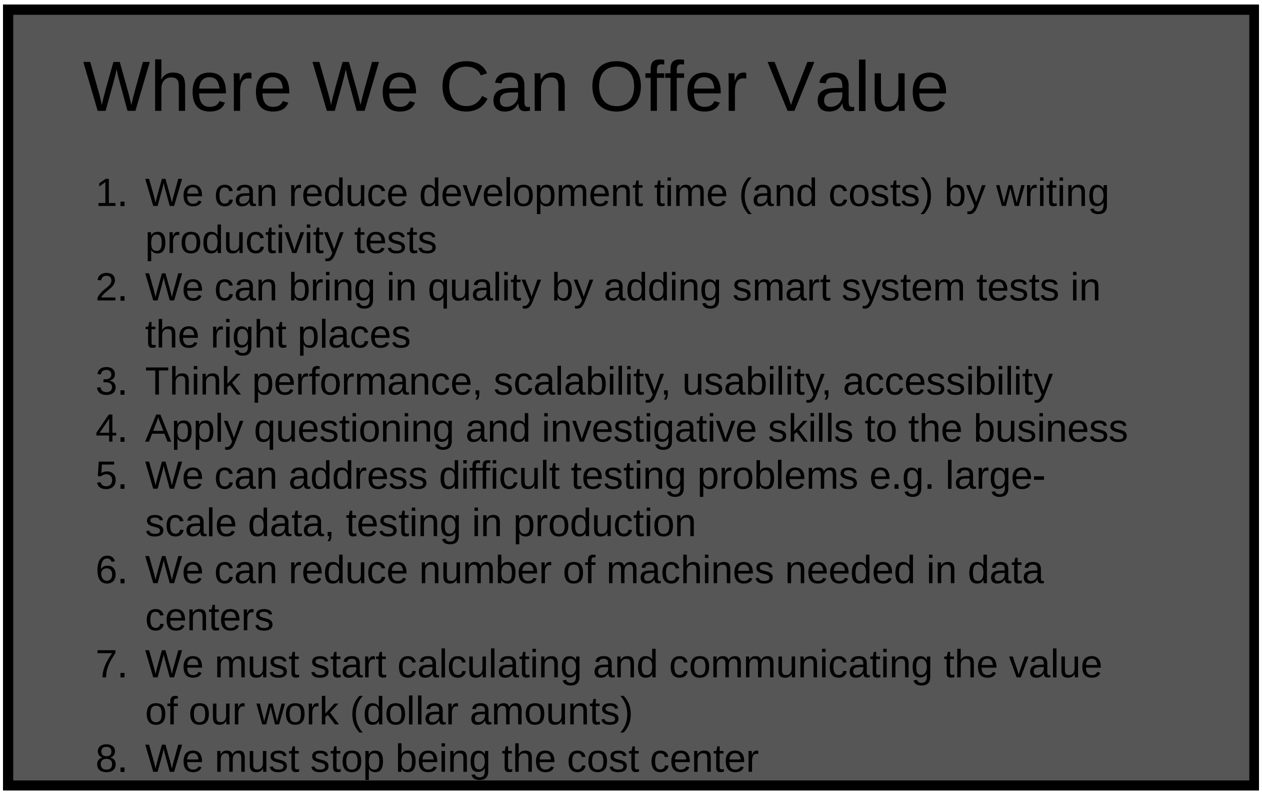
Not Just People: automated systems
http://cacm.acm.org/magazines/2010/2/69354-a-few-billion-lines-of-code-later/fulltext



Ways we can improve Value

Opensource our testing

1. Share data, methods & results
2. Collaboration
3. Weekend Testing
4. Swiss Testing Days
5. Peer workshops (LAWST)
6. Pairing



Where We Can Offer Value

1. We can reduce development time (and costs) by writing productivity tests
2. We can bring in quality by adding smart system tests in the right places
3. Think performance, scalability, usability, accessibility
4. Apply questioning and investigative skills to the business
5. We can address difficult testing problems e.g. large-scale data, testing in production
6. We can reduce number of machines needed in data centers
7. We must start calculating and communicating the value of our work (dollar amounts)
8. We must stop being the cost center

Q&A

julianharty@gmail.com
jharty@ebay.com