The 4th Industrial Revolution and Smarter Testing

Stuart Reid PhD, FBCS

(stureid.test@gmail.com / www.stureid.info)

Contents

- **The 4th Industrial Revolution**
  - Robots and Factories
  - Digital Disruption
  - Technology & User Expectations
- **Smarter Testing**
  - Random Testing & Fuzz Testing
  - Crowd Testing & A/B Testing
  - Testing using Artificial Intelligence (AI)
- **Questions?**

4th Industrial Revolution

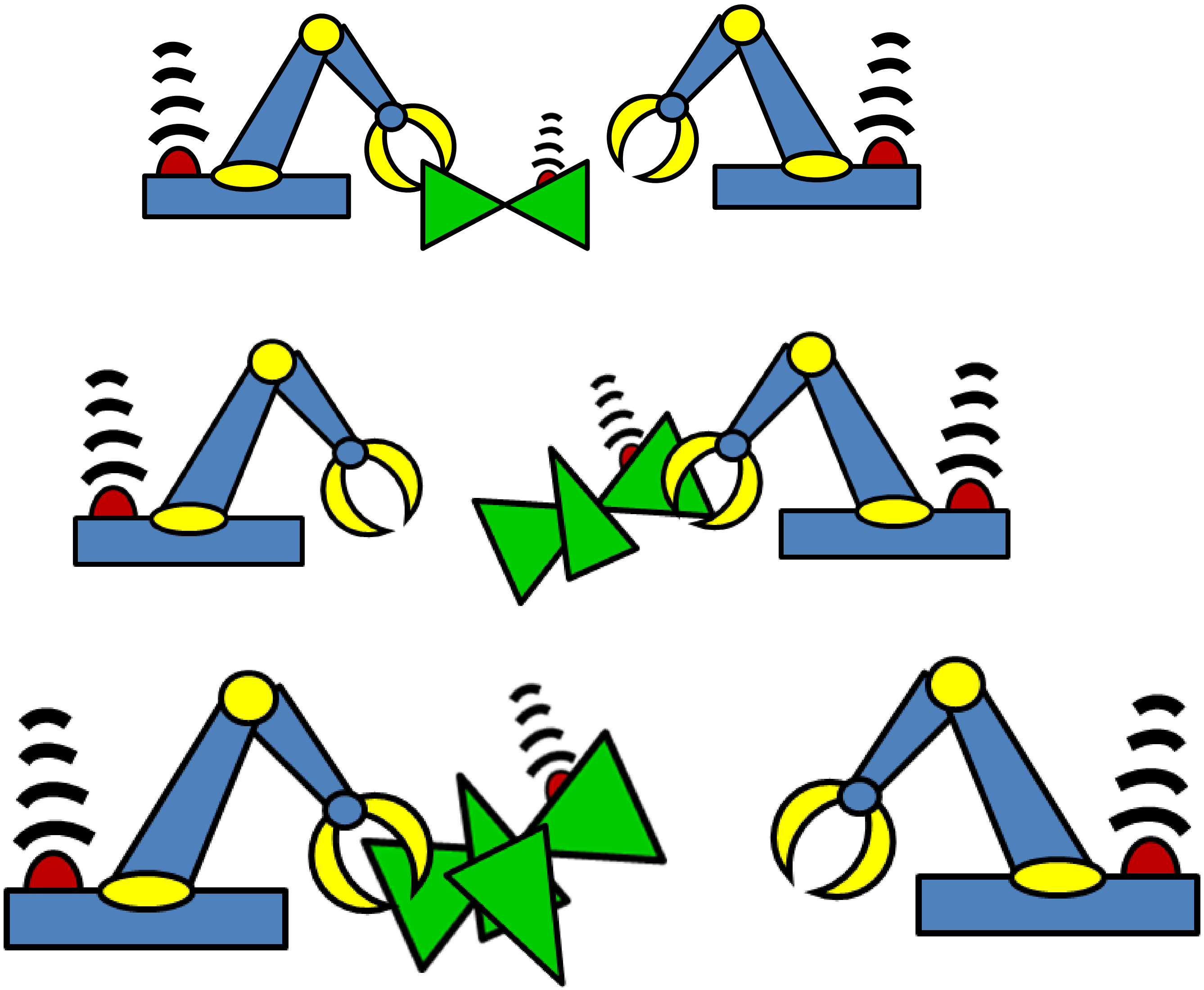
Industrial Revolutions

1. Factories & Steam Power (~1760)

2. Steel & Mass Production (~1850)

3. Electronics & IT (~1970)

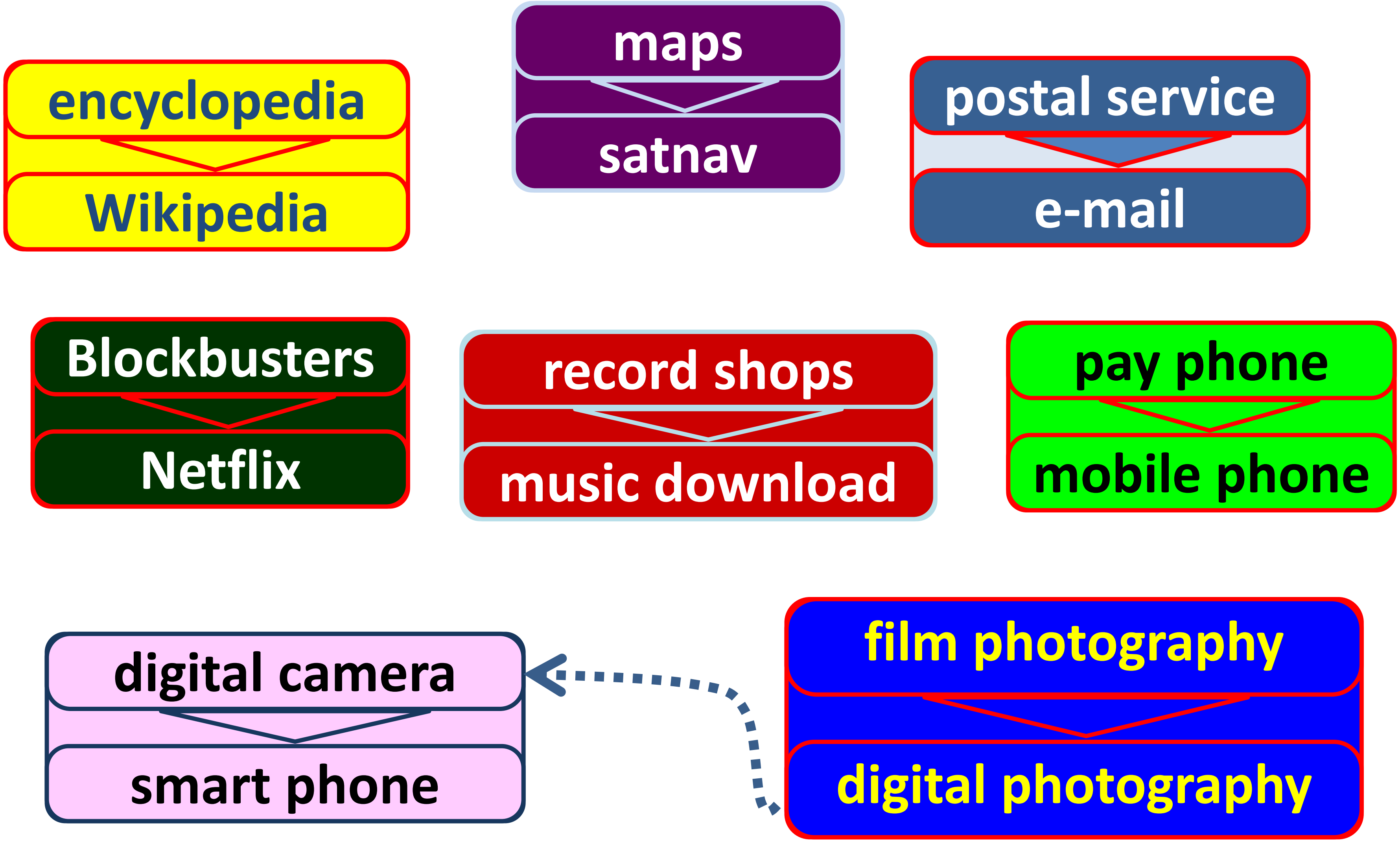
4th Industrial Revolution / Industrie 4.0



Robots - Industrial Level Disruption

- **Industrial Robot Production up 15% in 2015 from 2014**
- **Between 2015 and 2018, it is estimated that about 1.3 million new industrial robots will be installed in factories around the world**
- **China is the biggest market at 27% in 2015**
- **China/Foxconn**
  - May 2016: Foxconn factory reduces employee strength from 110,000 to 50,000 thanks to the introduction of robots
- **Korea**
  - #1 robot density in non-automotive industries with 365 robots installed per 10,000 employees in 2014
  - the world's 4th largest robot market, increased by 16% to about 24,700 units in 2014
- **Automotive industry is the biggest robot market**

Digital Disruption – Examples



encyclopedia
Wikipedia

maps
satnav

postal service
e-mail

Blockbusters
Netflix

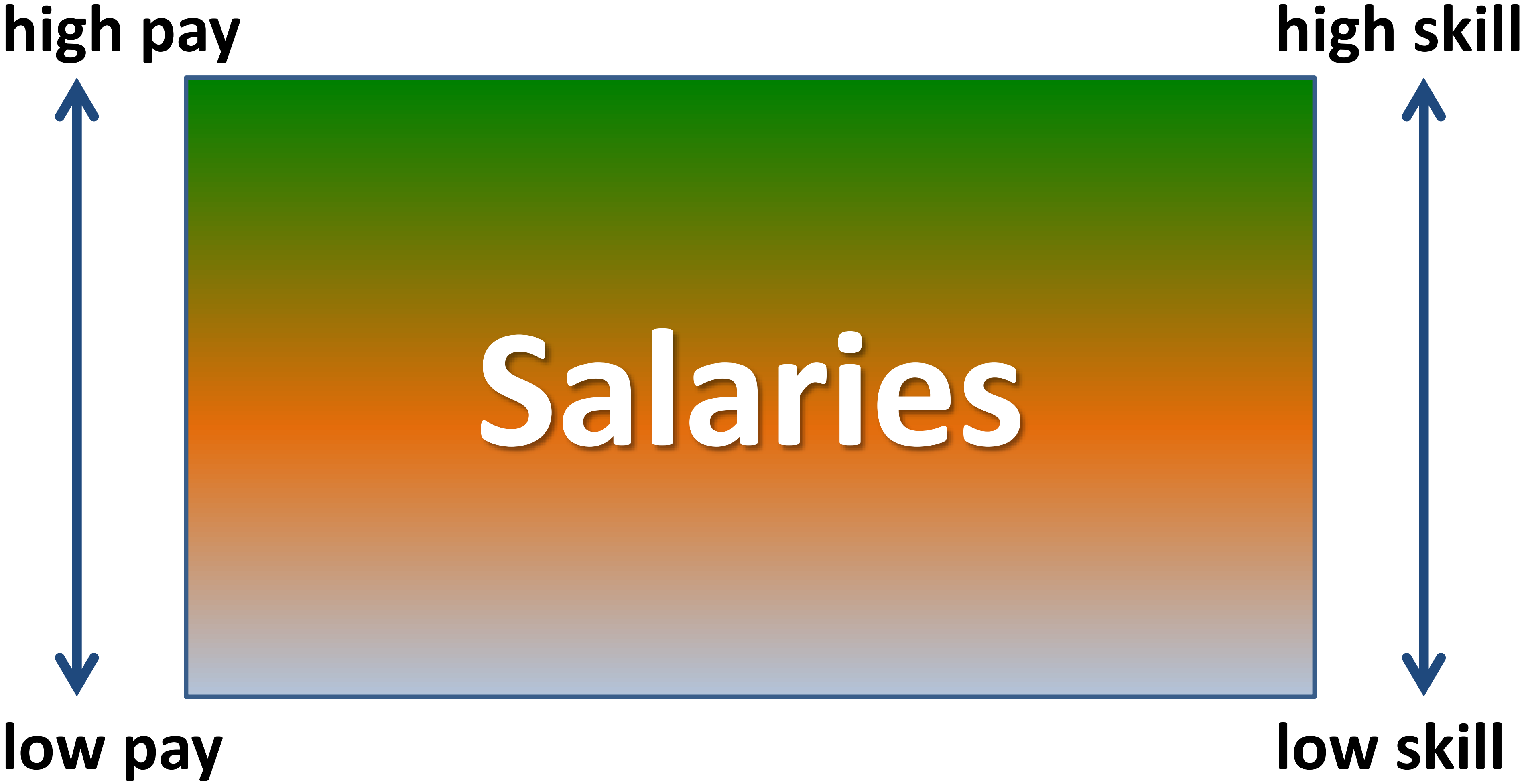
record shops
music download

pay phone
mobile phone

digital camera
smart phone

film photography
digital photography

The Salary Spectrum



high pay

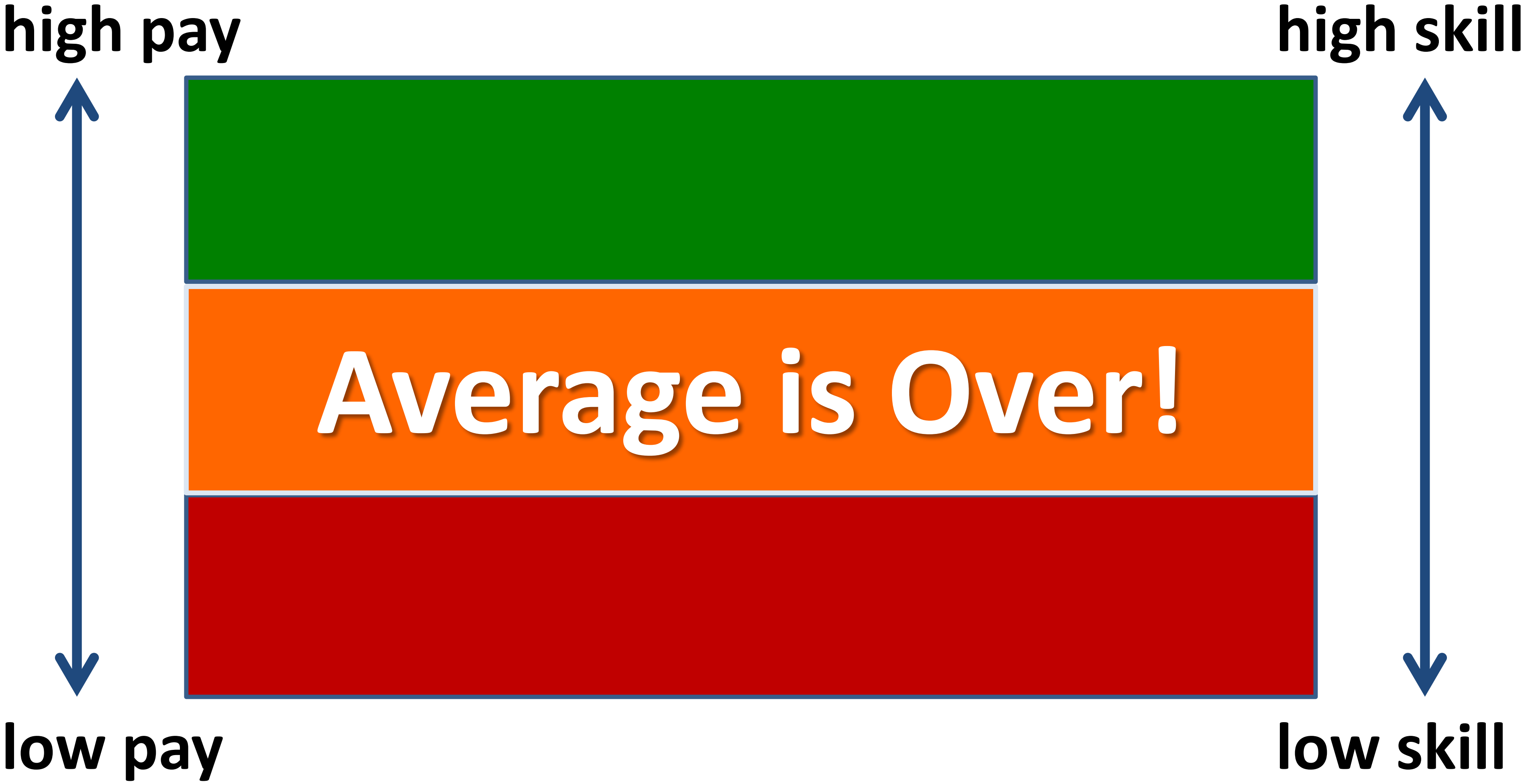
high skill

Salaries

low pay

low skill

“Average is over!”



high pay

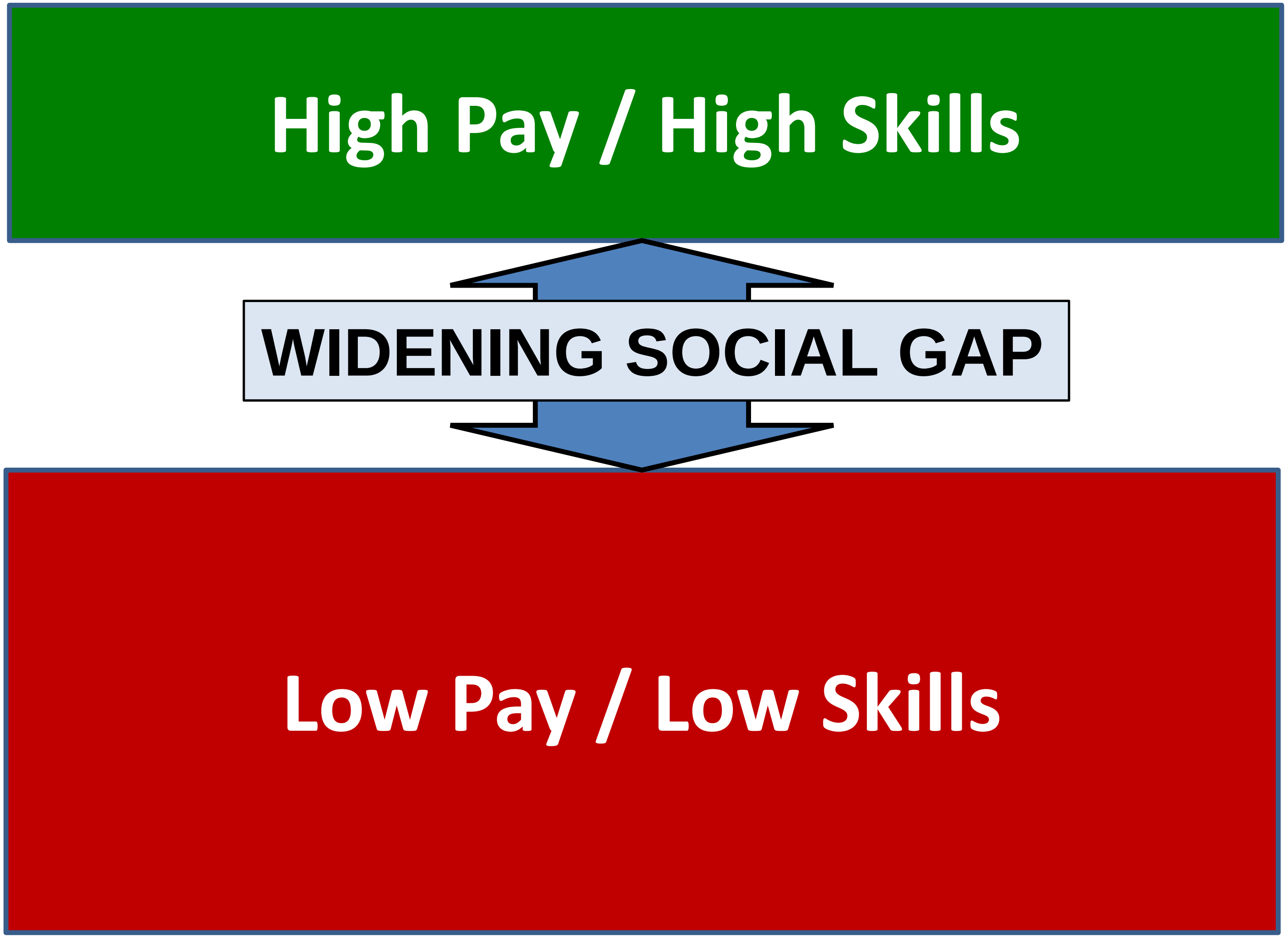
high skill

Average is Over!

low pay

low skill

“Average is over!”



High Pay / High Skills

WIDENING SOCIAL GAP

Low Pay / Low Skills

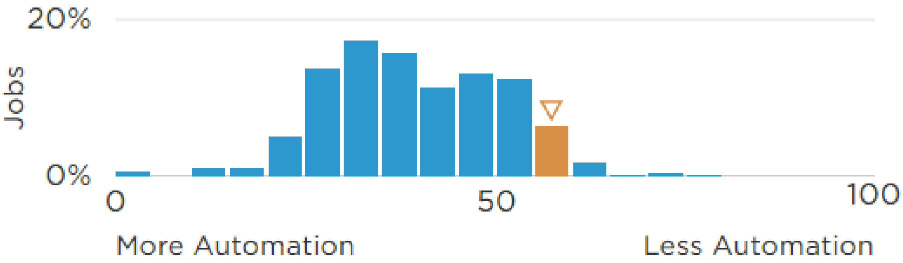
Will your job be done by a machine?

Chance of job being automated in next 20 years…	
High School Teacher	<1%
Airline Pilots	18%
Actors	37%
Judges	40%
Computer Programmers	48%
Miners	59%
Librarians	65%
Real Estate Agents	86%
Taxi Drivers	89%
Umpires & Referees	98%

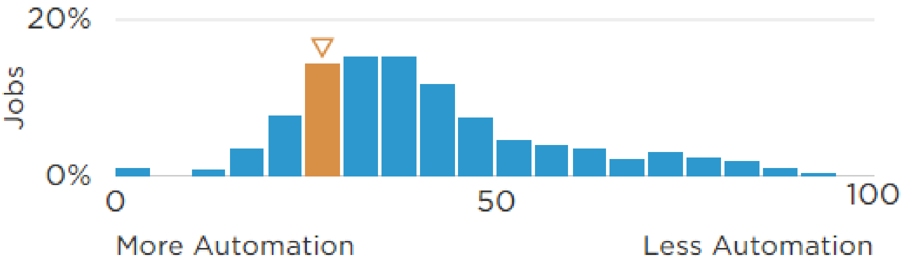
NPR Research, 2015

Software Applications Developers - 4.2%

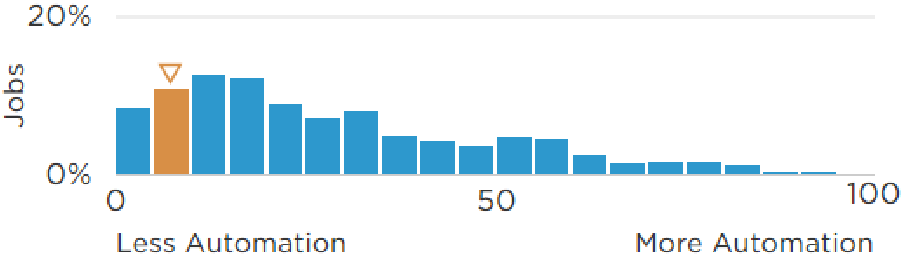
Need to come up with clever solutions?



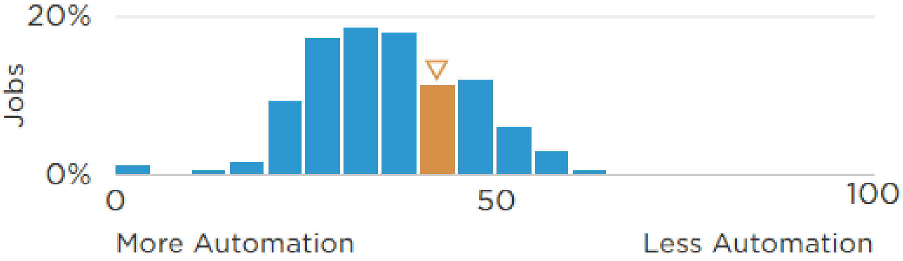
Required to personally help others?



Job needs you to squeeze into small spaces?



Job requires you to negotiate?



NPR Research, 2015

UK Situation – 2015 Figures (Deloitte)

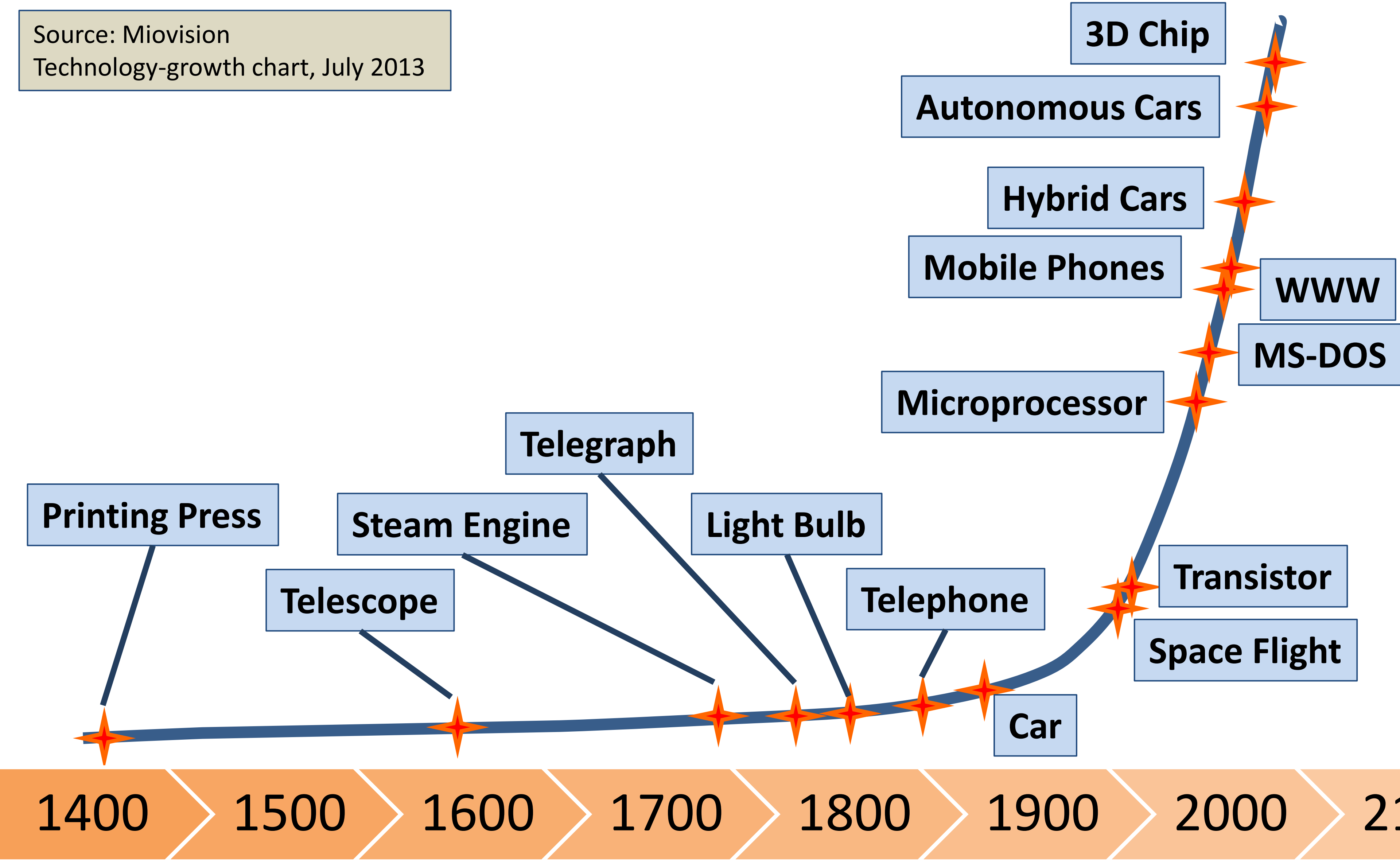
- "10 Million (35% of all) UK jobs are at risk of automation in the next 10 to 20 years"
- Last 15 years
  - Technology has already contributed to the loss of 800,000 lower-skill, higher-risk jobs
  - However, technology has already helped to create 3,500,000 new higher-skill, non-routine jobs
  - And, on average, each new job is paid approximately an extra ₩16 million adding more than ₩ 221 Trillion to the UK economy

US KPMG Study July 2016

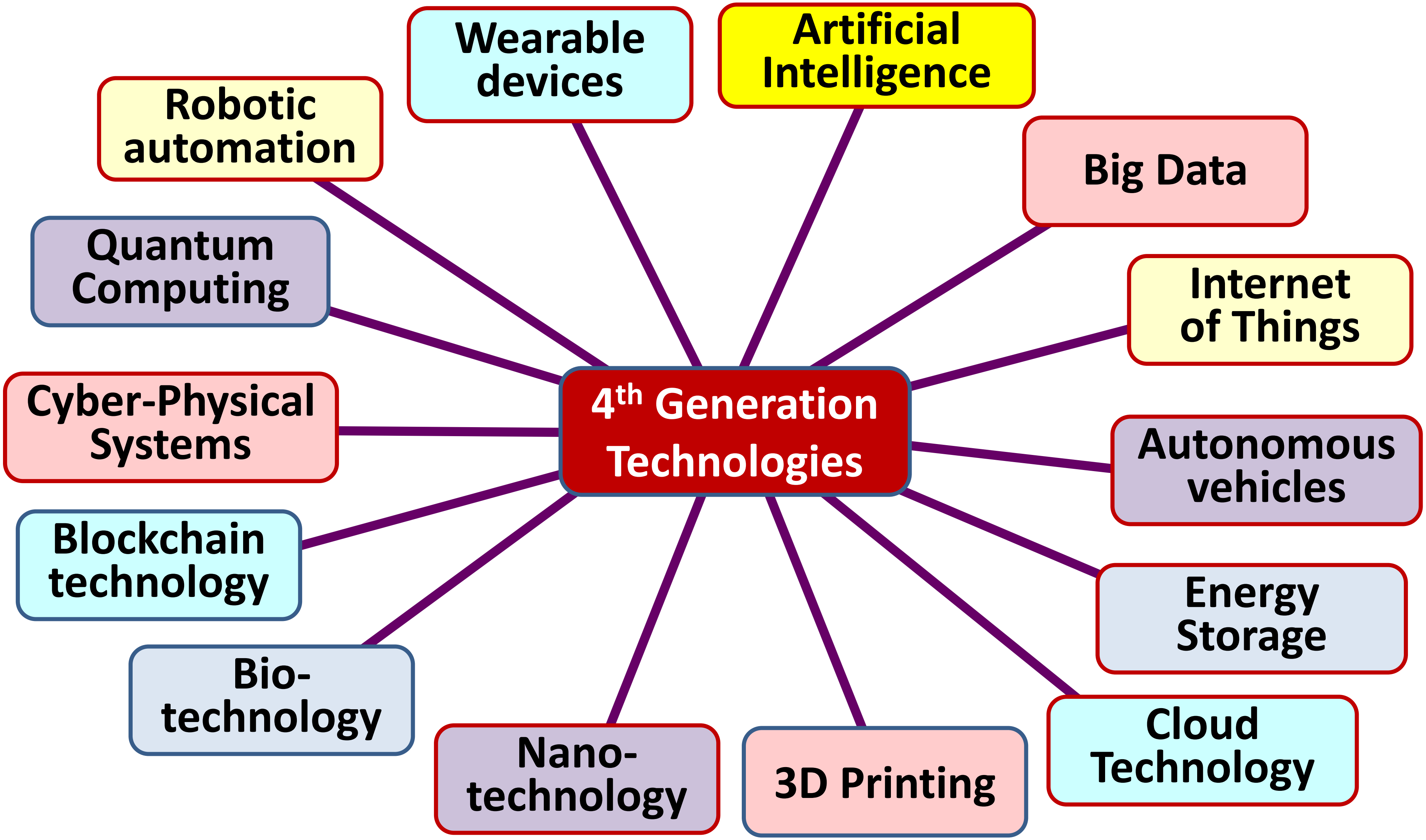
- **US Tech CEOs – next 3 years predict**
  - automation/machine learning will replace 5% of key workforce functions….BUT
  - only as they grow headcount
    - 95% expect an increase!

Accelerating Technology

Source: Miovision
Technology-growth chart, July 2013



4th Industrial Revolution – Not just robots



4th Generation Technologies

Robotic automation

Wearable devices

Artificial Intelligence

Big Data

Internet of Things

Quantum Computing

Cyber-Physical Systems

Autonomous vehicles

Blockchain technology

Energy Storage

Bio-technology

Nano-technology

3D Printing

Cloud Technology

User Expectations…

Mobility

Transparency

Choice

Time-to-market

Availability

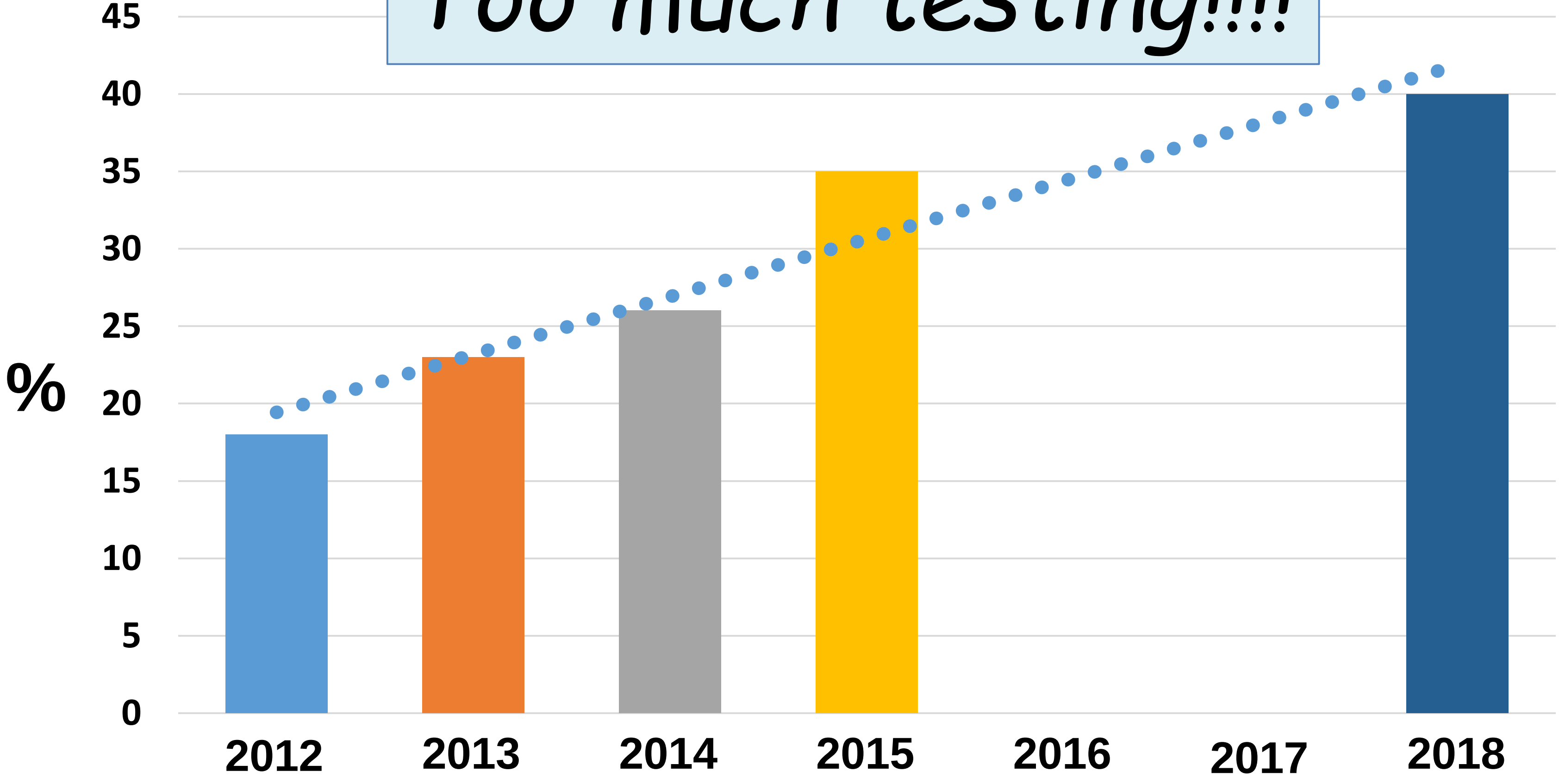
Trust

Customization

Quality

Smarter Software Testing

Testing as a Percentage of IT Budget

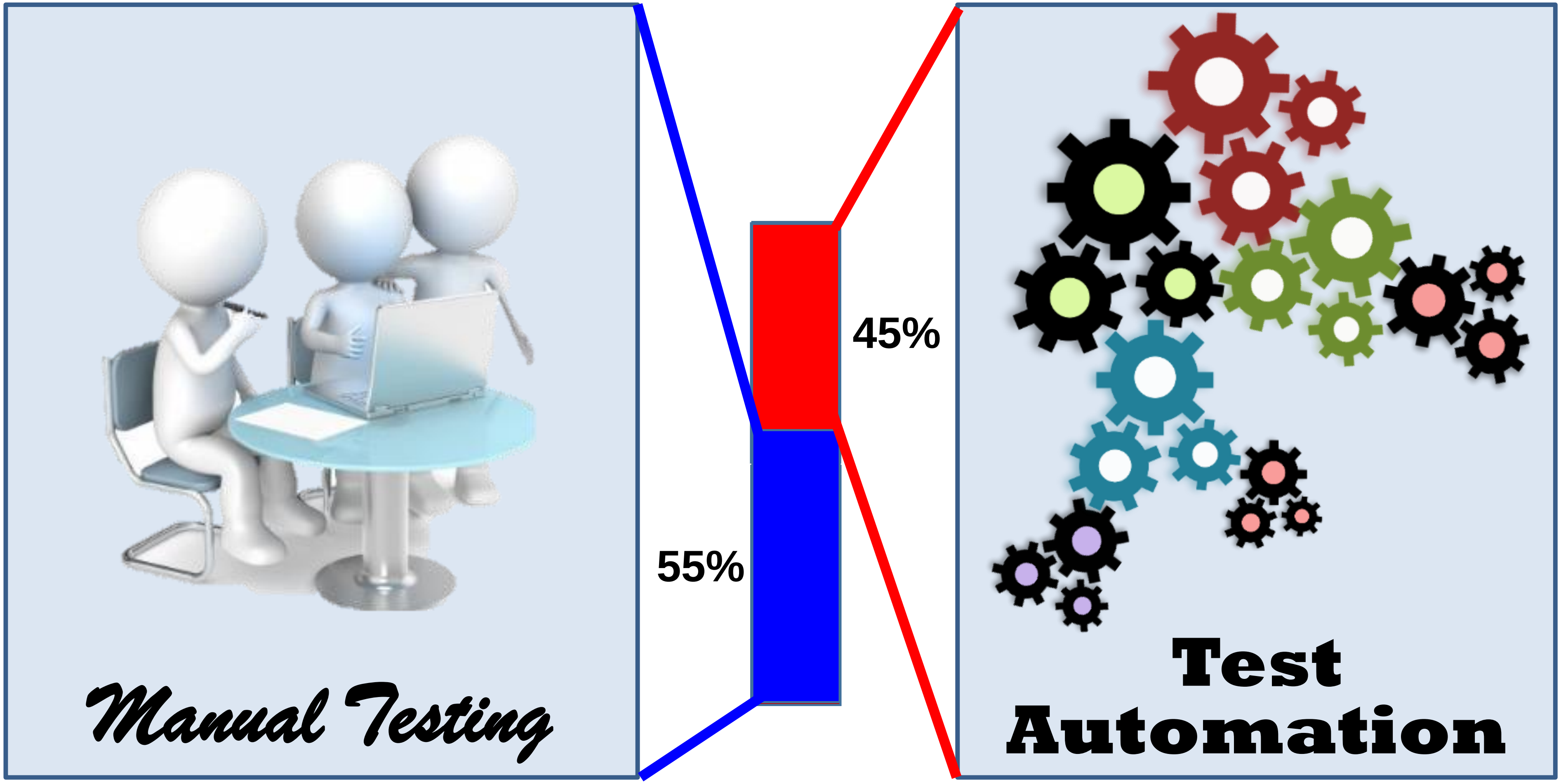


Too much testing!!!!

Is it good enough testing?

World Quality Report 2015-16

Automation is only a partial answer



Manual Testing

55%

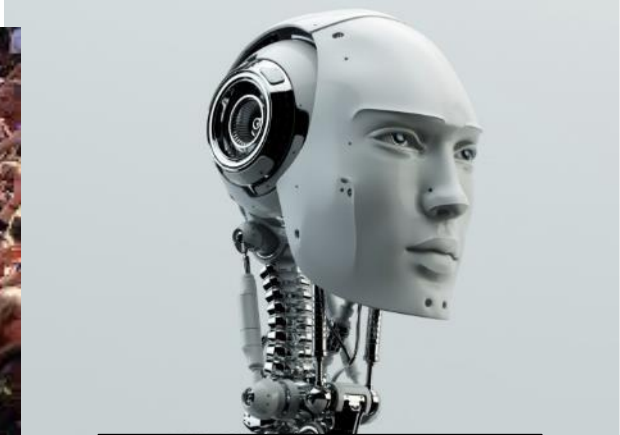
45%

Test Automation

World Quality Report 2015-16

Testing WITHOUT Test Design

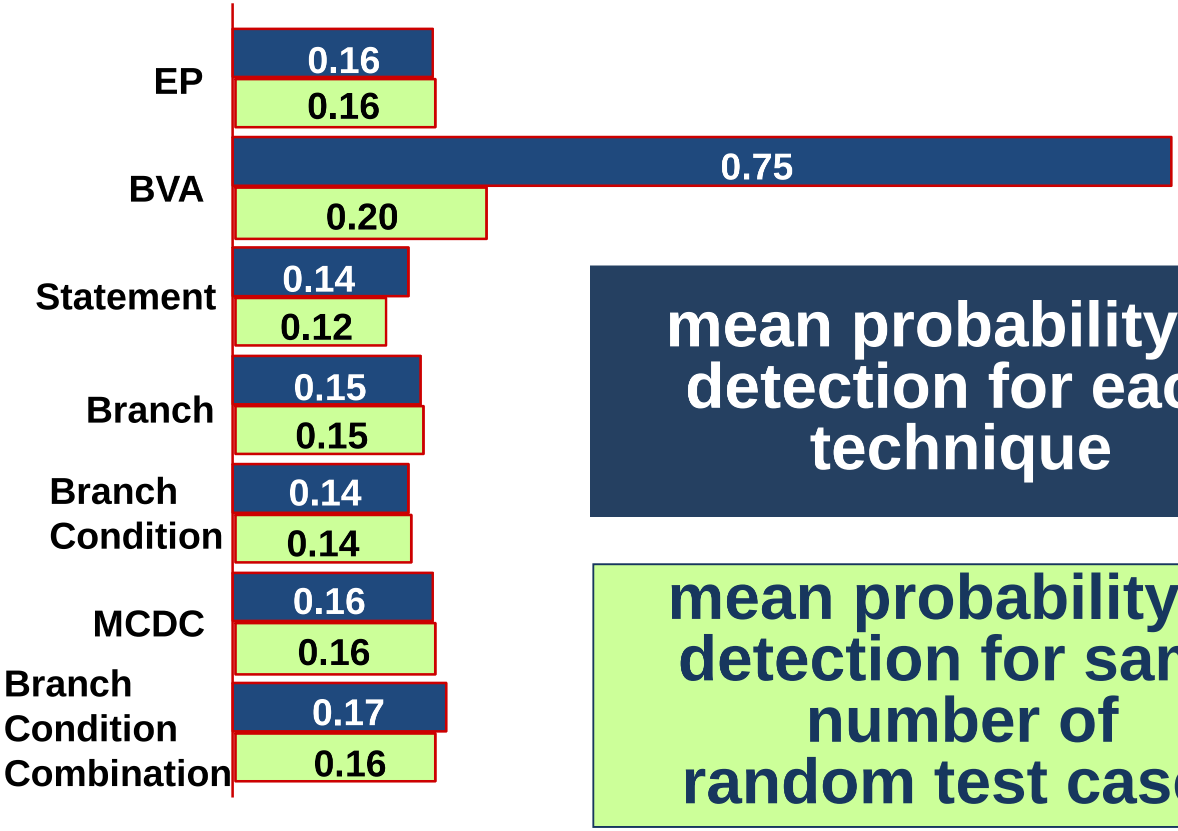
- **Tests generated using random test generation**
  - pure random and fuzz testing
- **Tests (generated and) run by the end users**
  - crowd testing and A/B testing
- **Tests generated by artificial intelligence**
  - regression tests and stress tests



Picture: http://careers2030.cs

Random Testing
&
Fuzz Testing

Experimental Evidence– Random vs Systematic Test Design

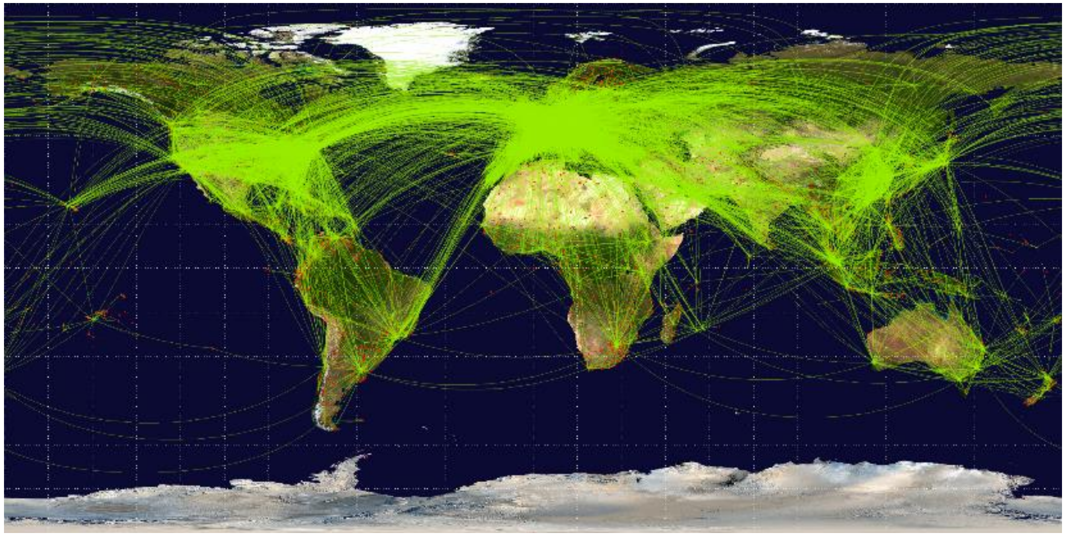


mean probability of detection for each technique

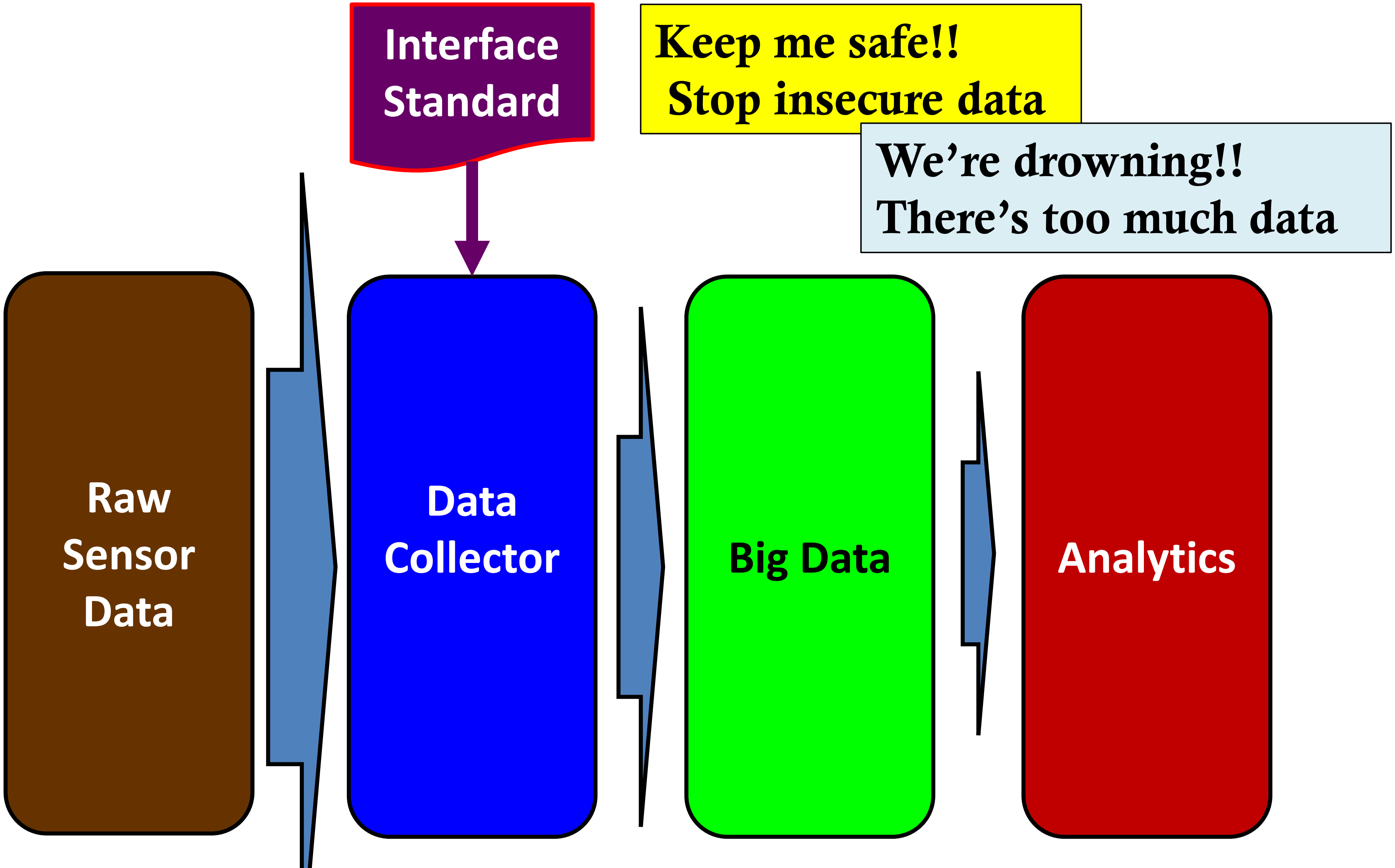
mean probability of detection for same number of random test cases

Big Data - Example Data Volumes

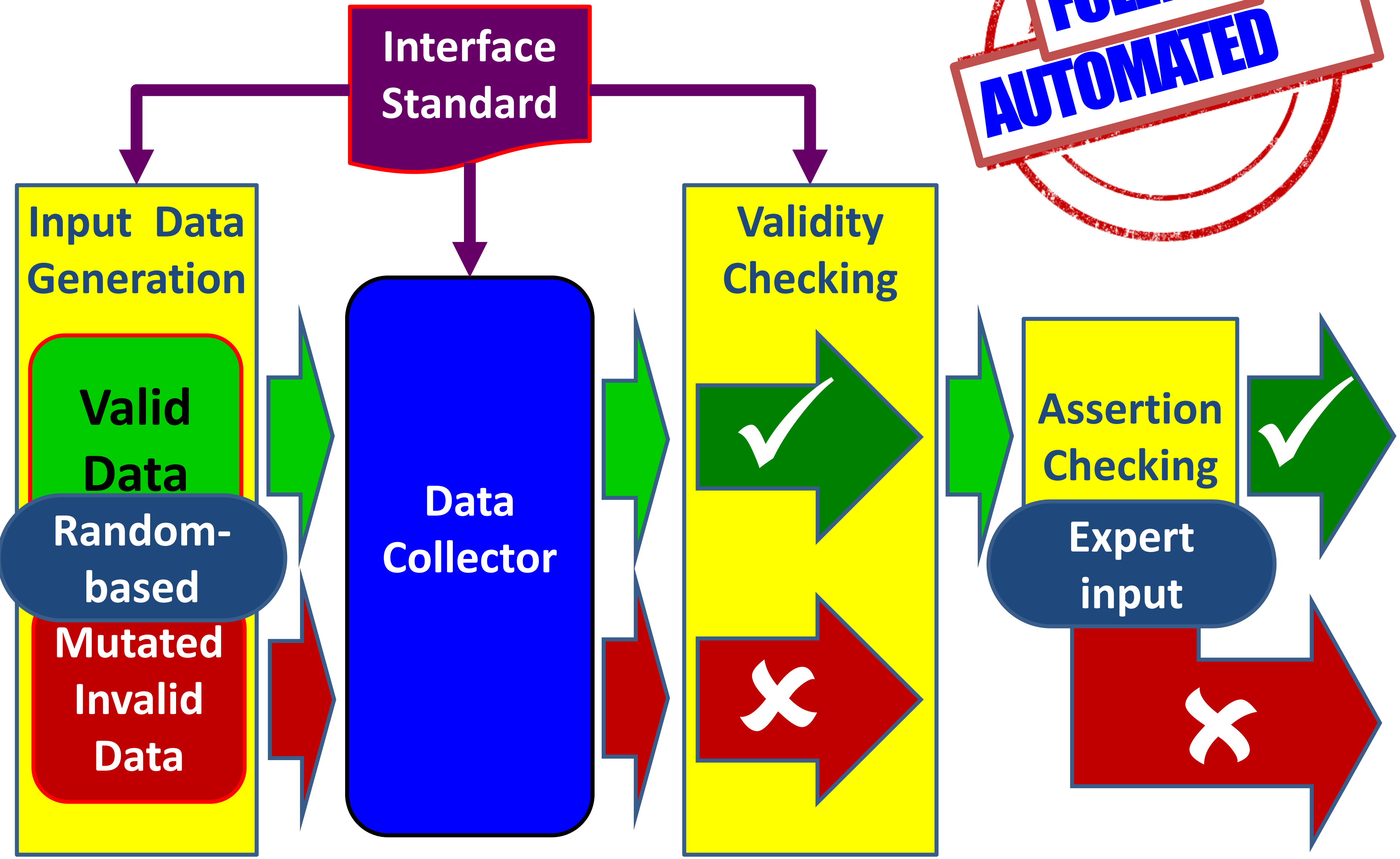
- **Airliner**
  - 3 GB/hour
- **CCTV**
  - 1 TB/month
- **Mobile Phone Operator**
  - 3 TB/day
- **Globally**
  - 2.5 Million TB/day



Data Validity from Internet of Things



Fuzz Testing



FULLY AUTOMATED

Interface Standard

Input Data Generation

Valid Data

Random-based

Mutated Invalid Data

Data Collector

Validity Checking

✓

✗

Assertion Checking

Expert input

✓

✗

Users as Testers - Crowd Testing & A/B Testing

What about the users as testers?

- **Their personas are perfect ☺**
  - – we don’t have to guess their habits or preferences
- **Their test environments are truly representative ☺**
- **Can we trust them? ☹**
  - – use Non-Disclosure Agreements (NDAs)
  - – don’t tell them!

Mobile

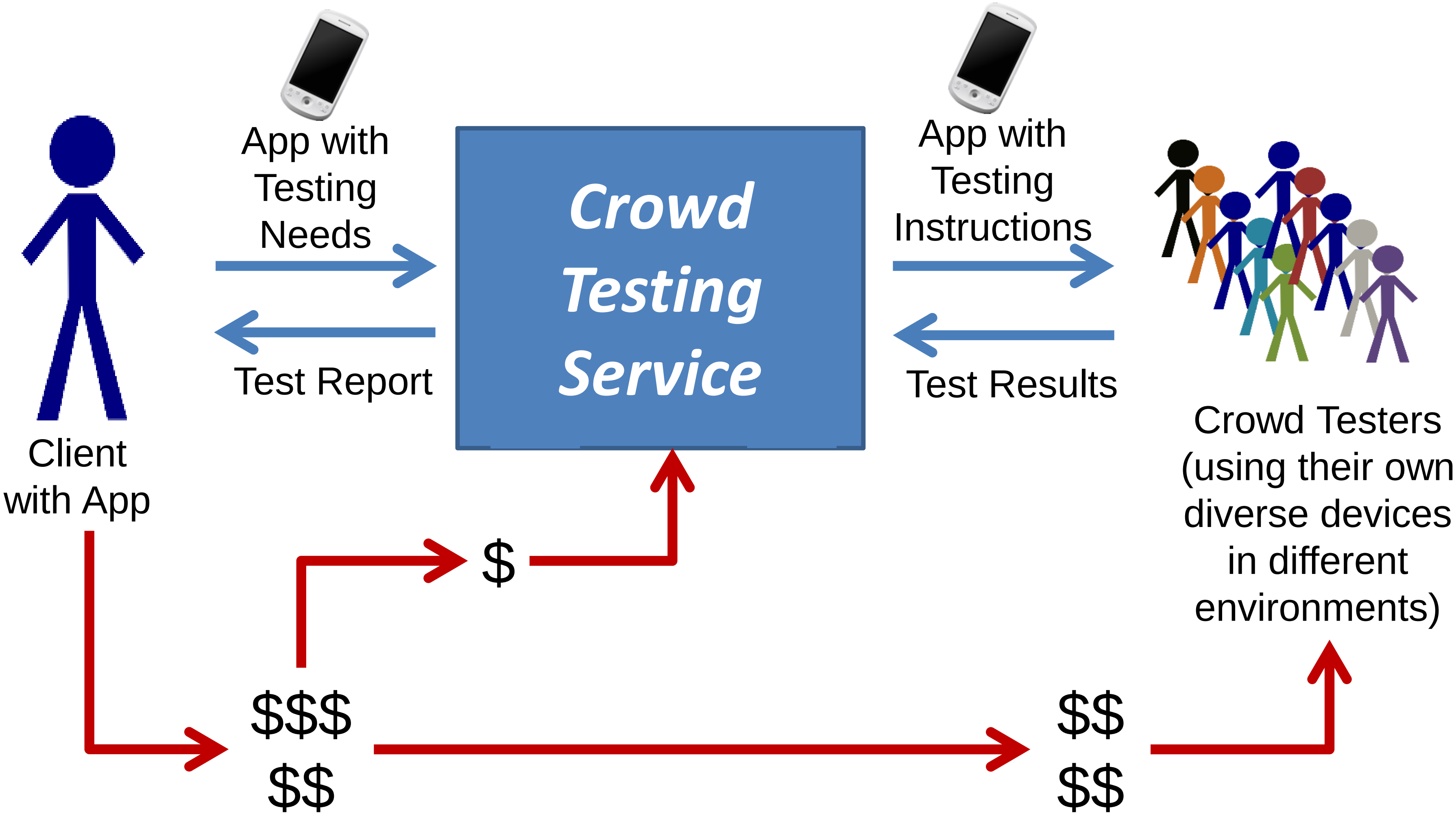
Search
Help
Games
Calendar
Maps
Music
Contacts
Messaging
Internet
Store
Chat
Camera
N
W
E
S

Android Handset Fragmentation

XT920
ST23
iris405+
C16
5 (Nexus 5)
Z10
XT925
K01
K007
GT-S7582
Moto X (XT1058)
ST26
SM-G7102
C1905
D6503
GT-I9105

24,000 distinct devices (Aug 2015)

Crowd Testing for Multiple Devices/Env’ts



App with Testing Needs

Test Report

Crowd Testing Service

App with Testing Instructions

Test Results

Client with App

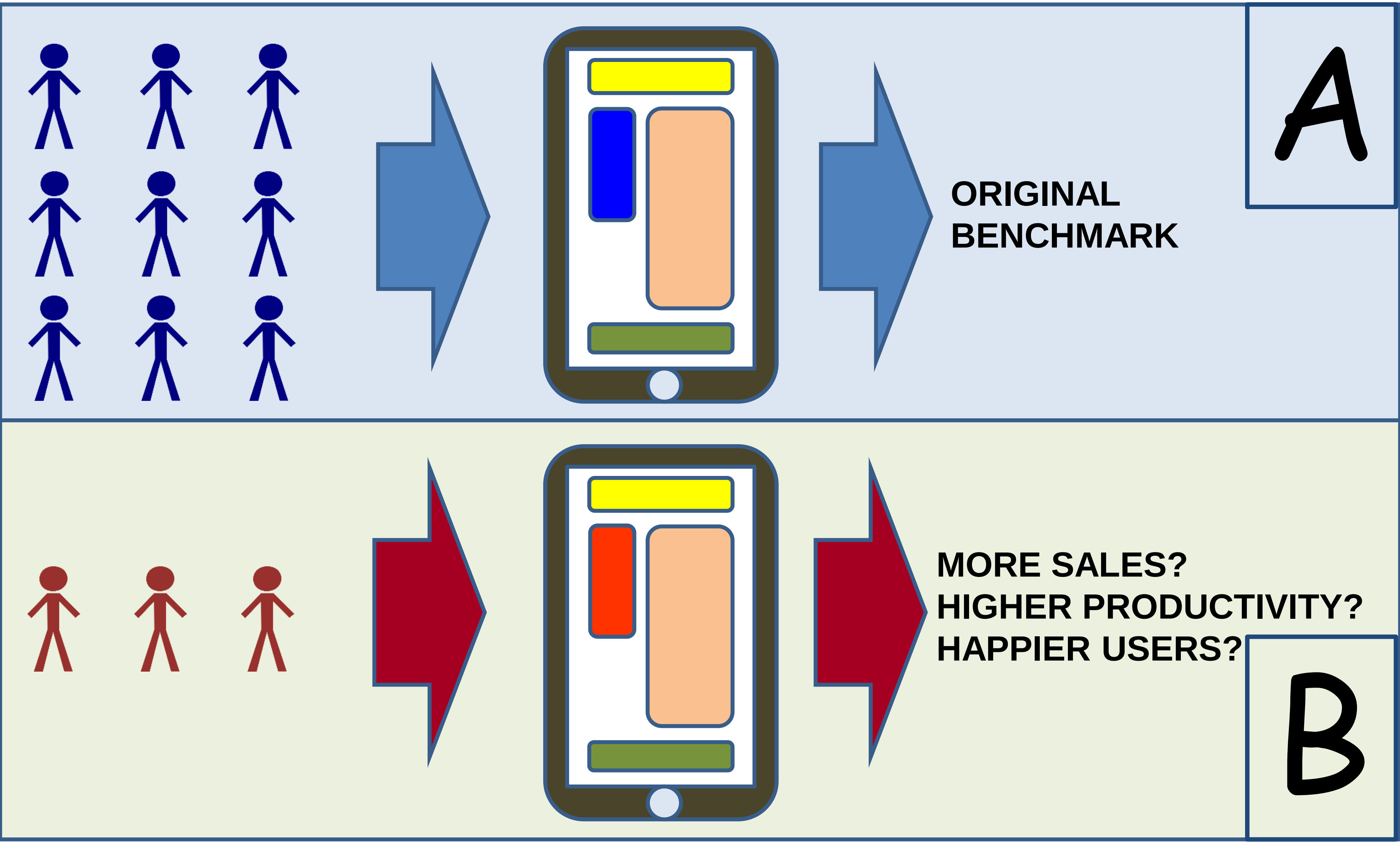
Crowd Testers (using their own diverse devices in different environments)

$

$$$
$$

$$
$$

A/B Testing

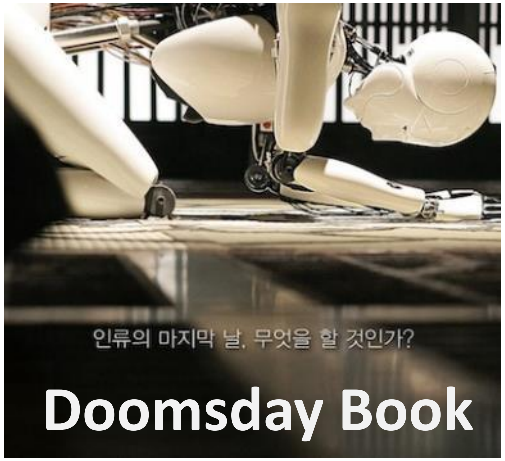


ORIGINAL
BENCHMARK

MORE SALES?
HIGHER PRODUCTIVITY?
HAPPIER USERS?

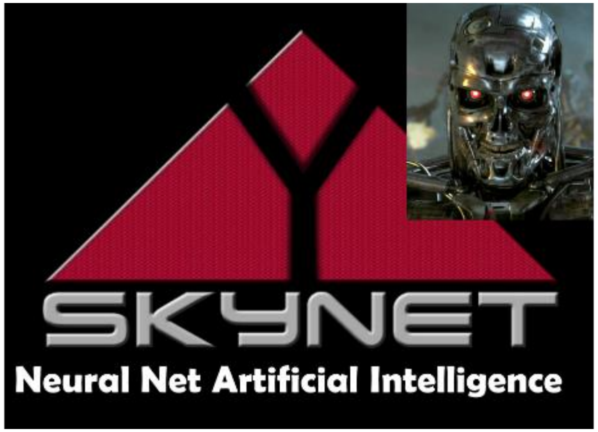
Testing with Artificial Intelligence

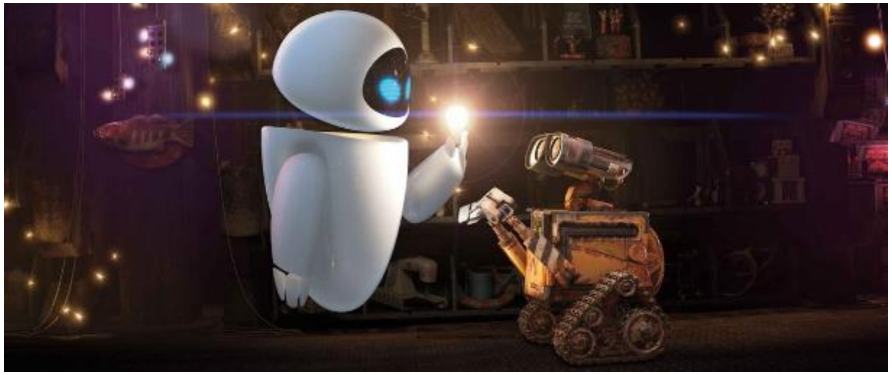
Artificial Intelligence (AI) in the Cinema



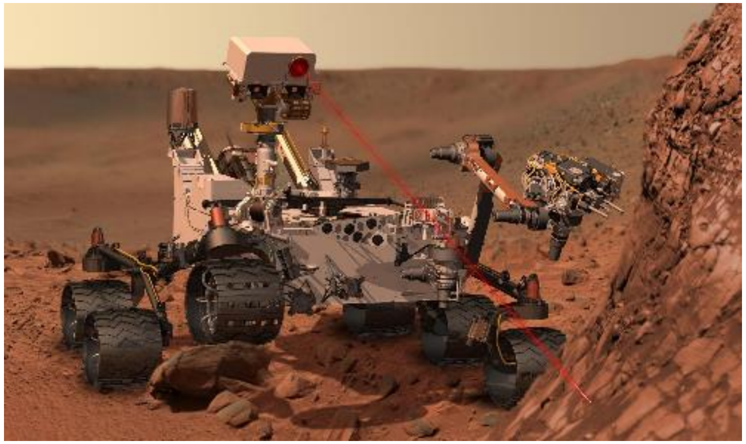
Doomsday Book

I’m sorry Dave…





Artificial Intelligence (AI) Works!

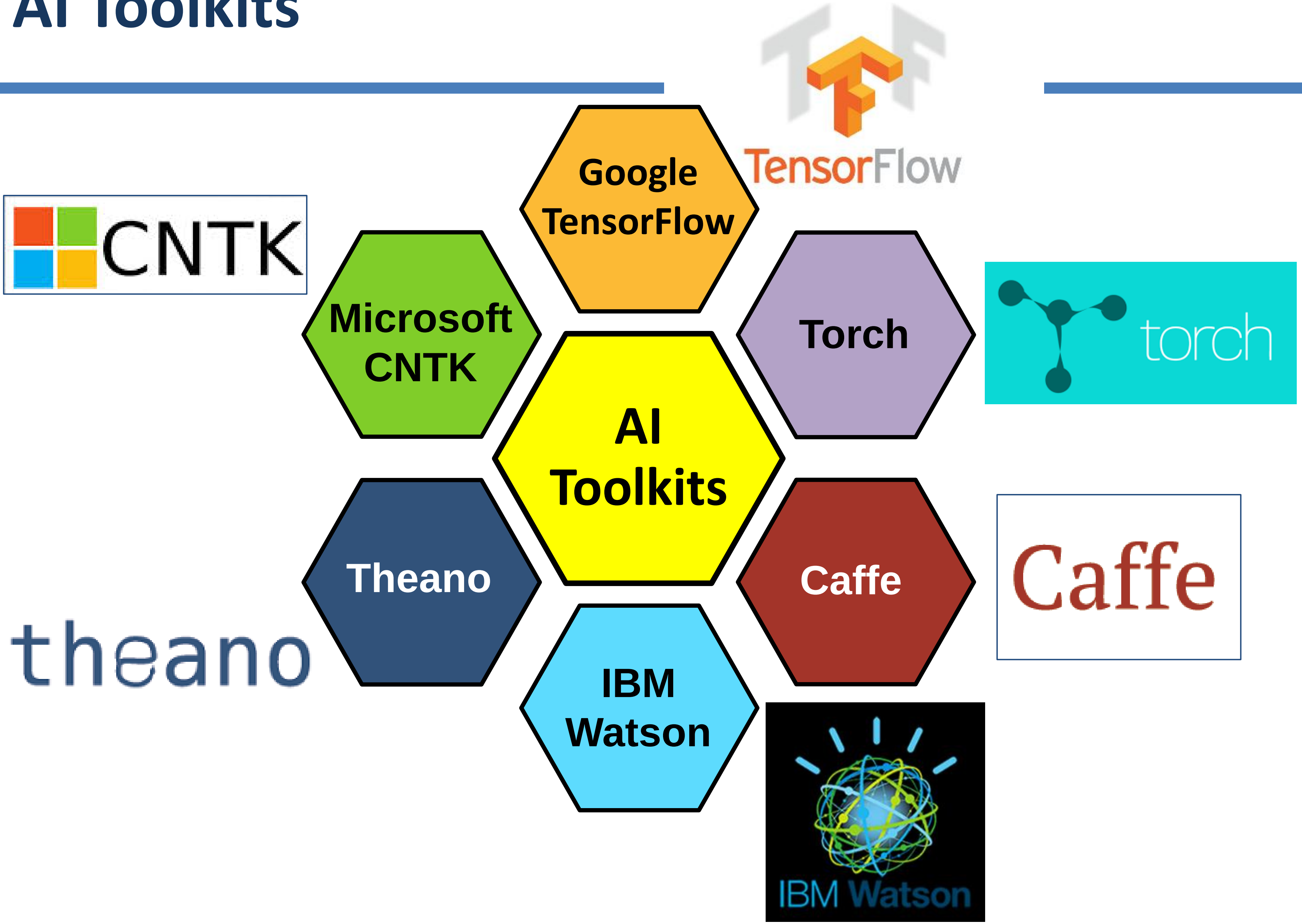


Artificial Intelligence Techniques

- Neural Networks
- Expert/Knowledge-based Systems
- Machine Learning
- Fuzzy and Probabilistic Logic
- Classification
- Search and Optimization

- Much of today's effective AI uses a variety of overlapping techniques
  - and exploits the availability of processing power & storage

AI Toolkits



AI Toolkits

Google TensorFlow

Microsoft CNTK

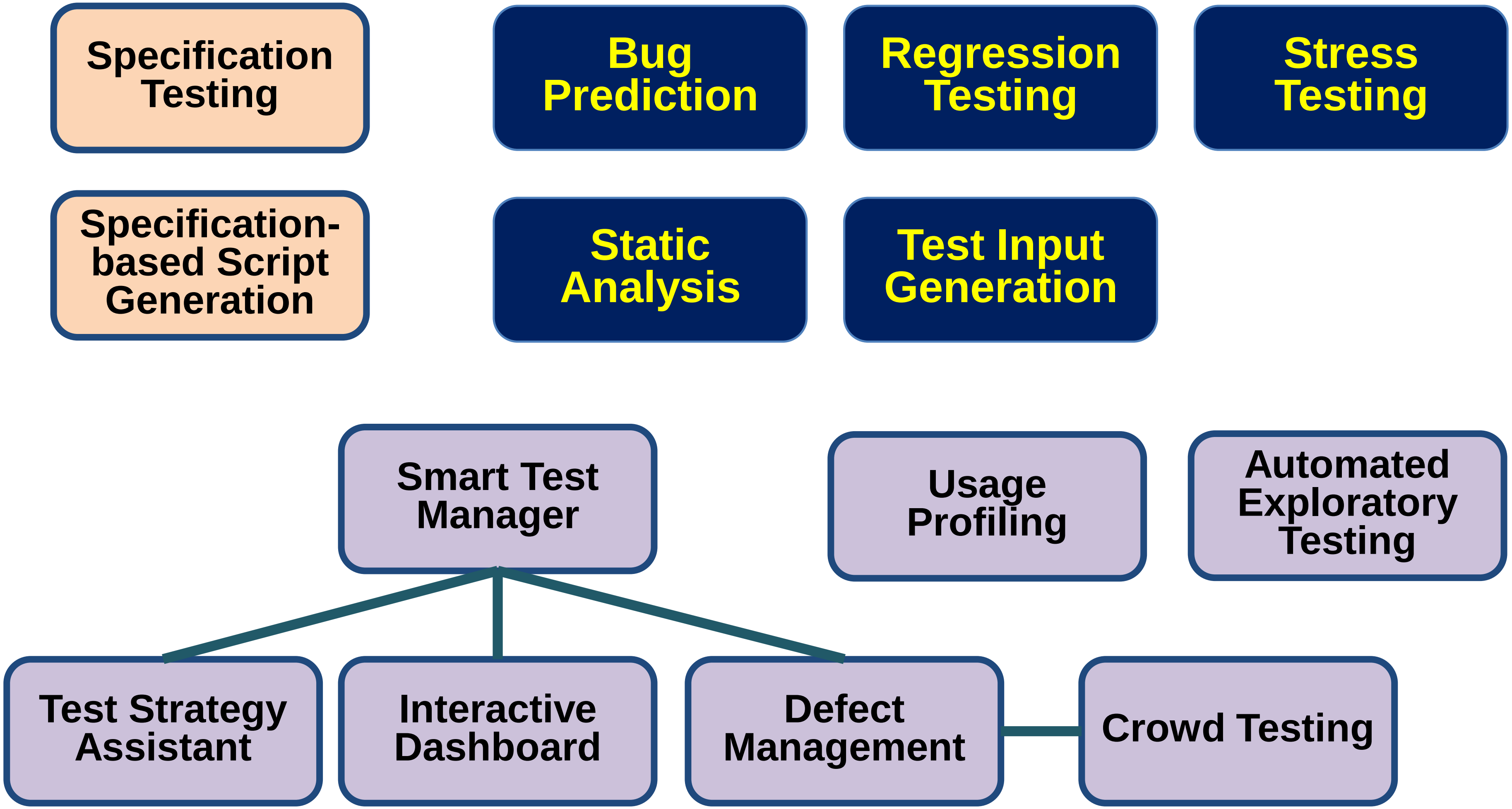
Torch

Theano

Caffe

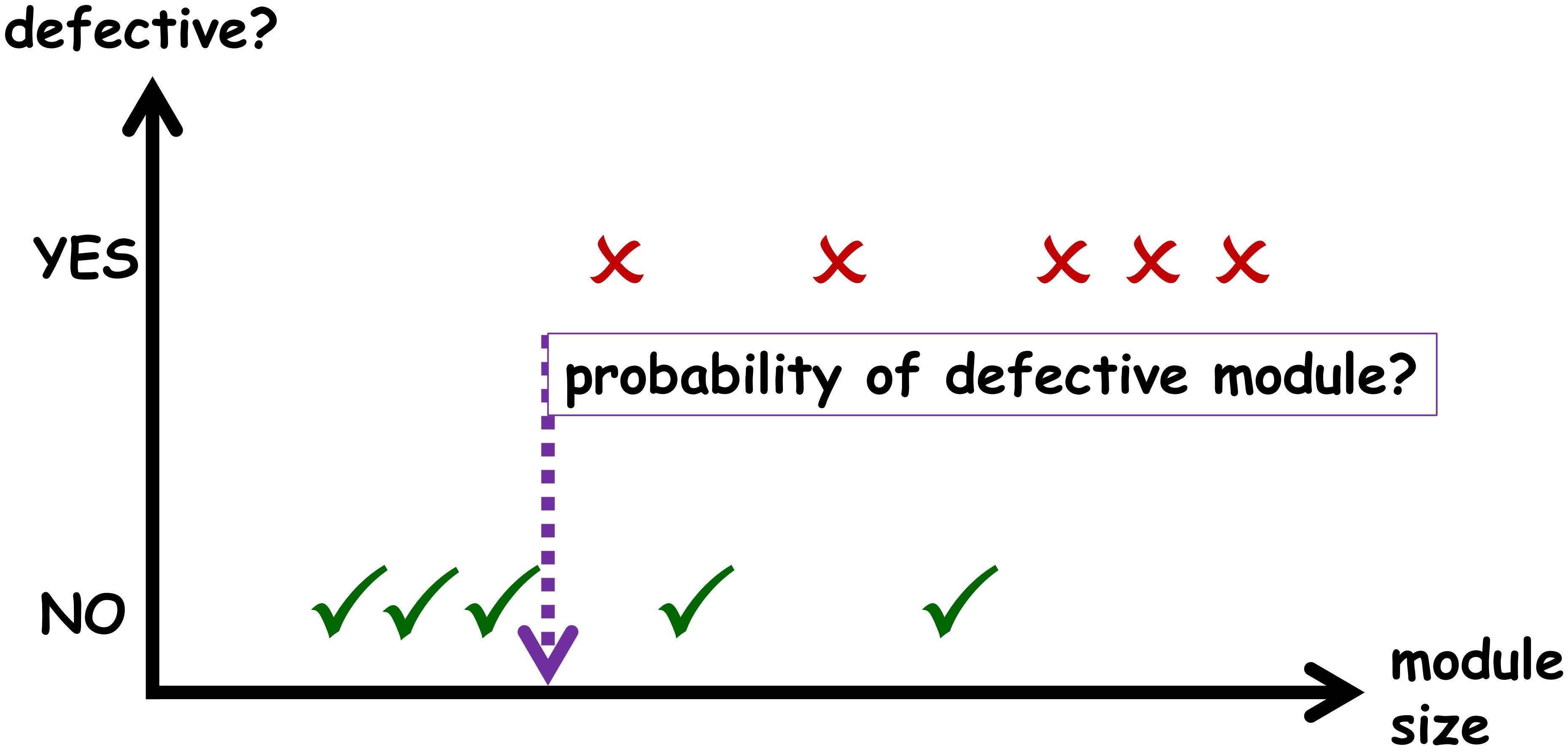
IBM Watson

AI - Smart Testing Opportunities

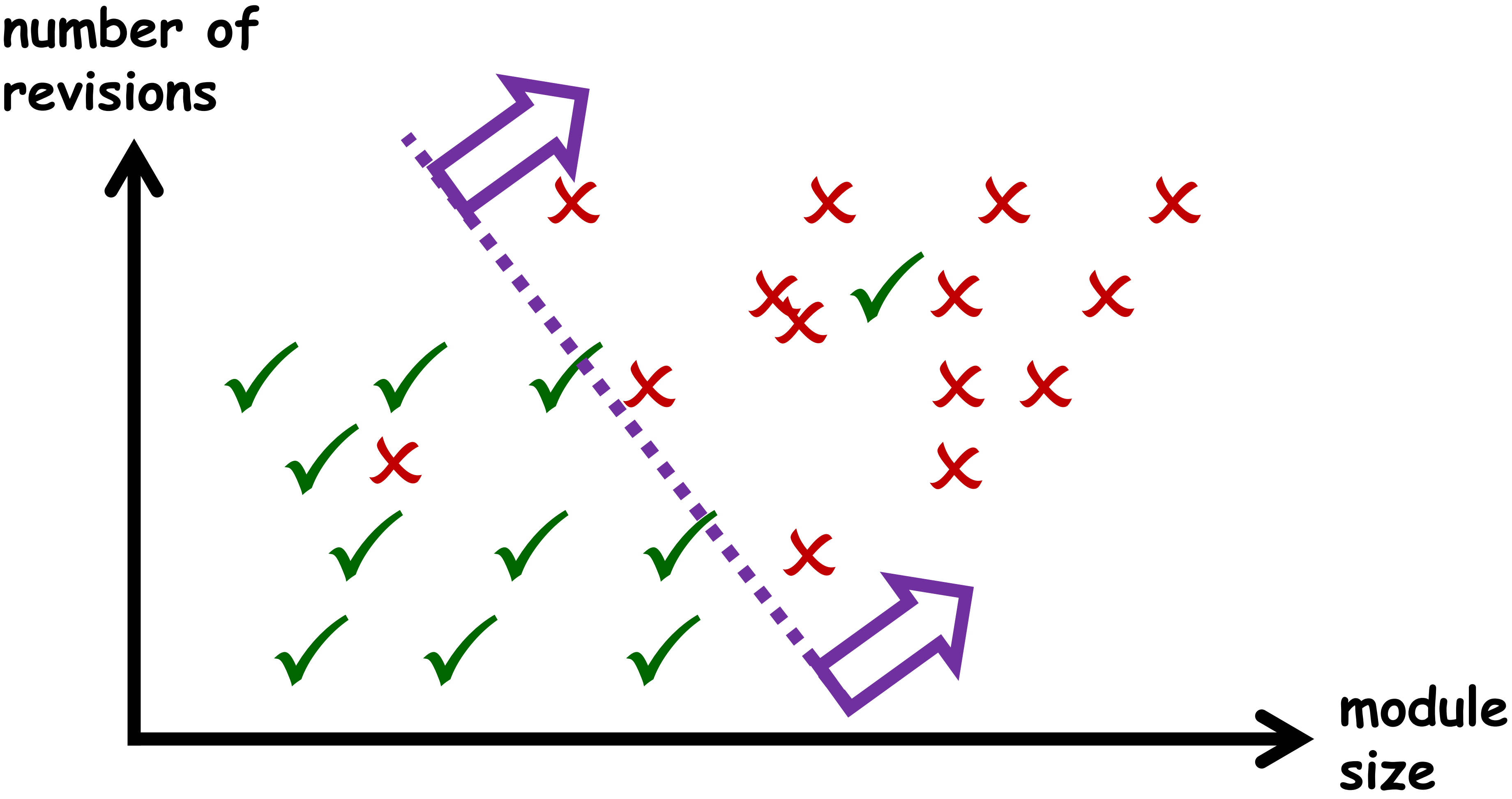


Bug Prediction

Bug Prediction – a Single Attribute



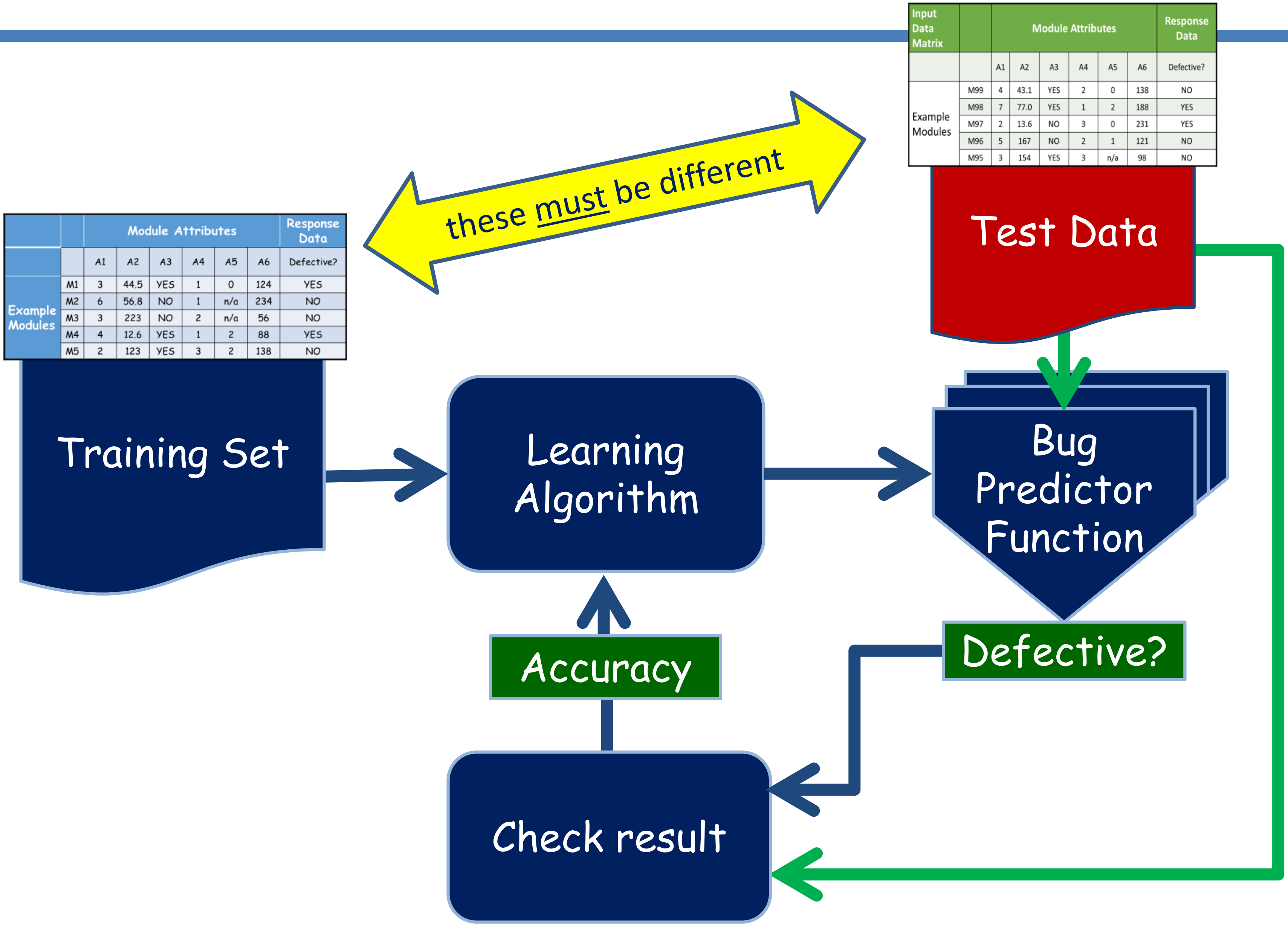
Bug Prediction – Two Attributes



Bug Prediction – Multiple Attributes

Input Data Matrix		Module Attributes						Response Data
		A1	A2	A3	A4	A5	A6	Defective?
Example Modules	M1	3	44.5	YES	1	0	124	YES
	M2	6	56.8	NO	1	n/a	234	NO
	M3	3	223	NO	2	n/a	56	NO
	M4	4	12.6	YES	1	2	88	YES
	M5	2	123	YES	3	2	138	NO

Supervised Learning Process



		Module Attributes						Response Data
		A1	A2	A3	A4	A5	A6	Defective?
Example Modules	M1	3	44.5	YES	1	0	124	YES
	M2	6	56.8	NO	1	n/a	234	NO
	M3	3	223	NO	2	n/a	56	NO
	M4	4	12.6	YES	1	2	88	YES
	M5	2	123	YES	3	2	138	NO

Input Data Matrix		Module Attributes						Response Data
		A1	A2	A3	A4	A5	A6	Defective?
Example Modules	M99	4	43.1	YES	2	0	138	NO
	M98	7	77.0	YES	1	2	188	YES
	M97	2	13.6	NO	3	0	231	YES
	M96	5	167	NO	2	1	121	NO
	M95	3	154	YES	3	n/a	98	NO

Bug Prediction Metrics

- **Source code metrics**
  - Lines of code
  - Number of comments
  - Cyclomatic complexity
  - Module dependencies
- **Process metrics**
  - Revisions made to module
  - Times refactored
  - Times fixed / when fixed
  - Lines of code changed (code churn)
  - Module age
- **People and organizational metrics**
  - Number of authors
  - Developer experience

Bug Prediction Results

- **“87% detection rate achieved average with 26% false alarms”**
  - [Tosun, 2010]
- **“73%-95% of faults can be predicted in just 10% of files” (across 7 projects)**
  - [Kim, 2007]
- **Best predictors are:**
  - People and Organizational measures (84%)
  - Module change (80%)
  - Fixed recently (and connected modules)
  - Reused module (more error-prone than new modules)
  - Module age

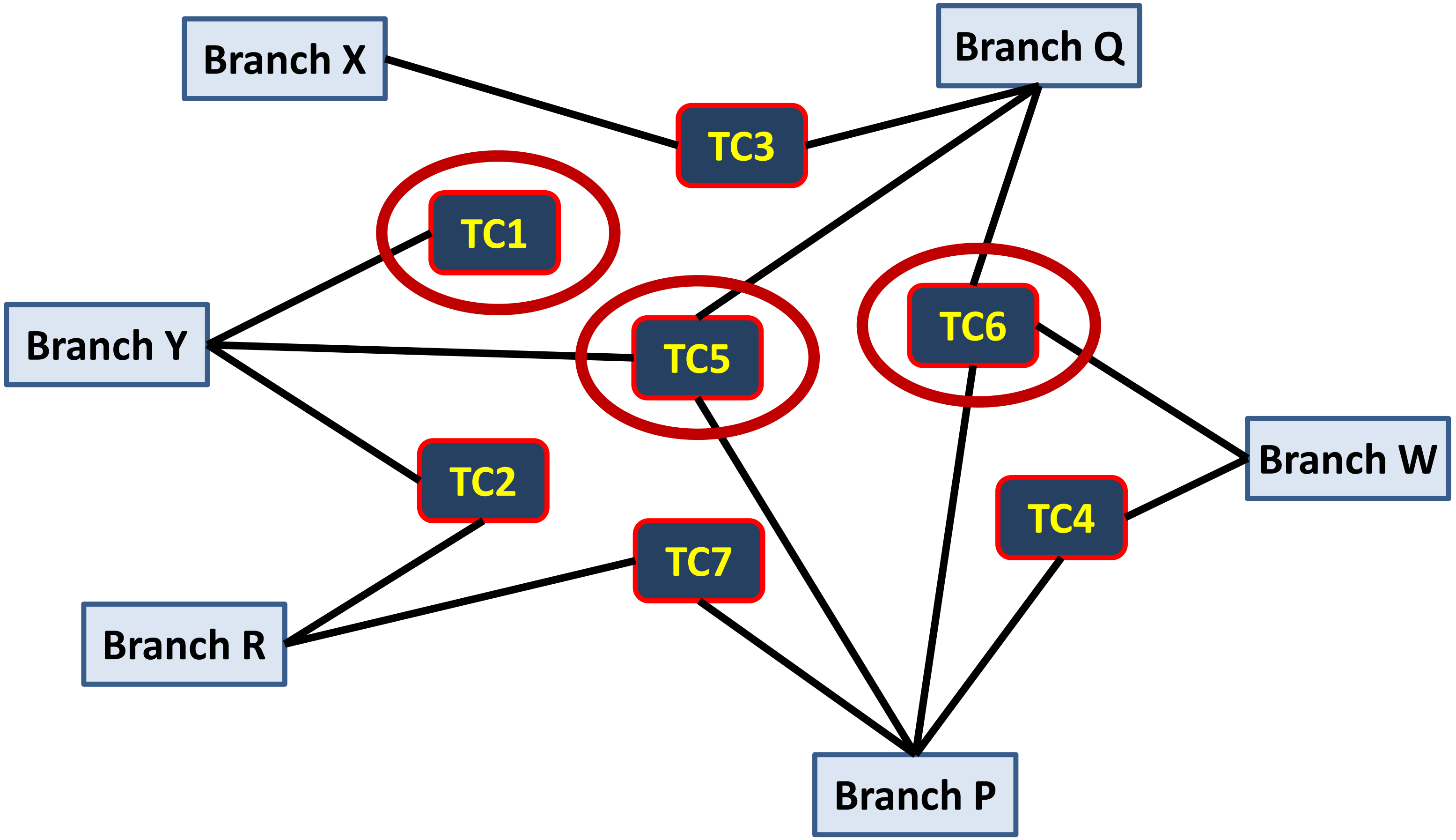
Static Analysis

Static Analysis Tool - Facebook – Infer

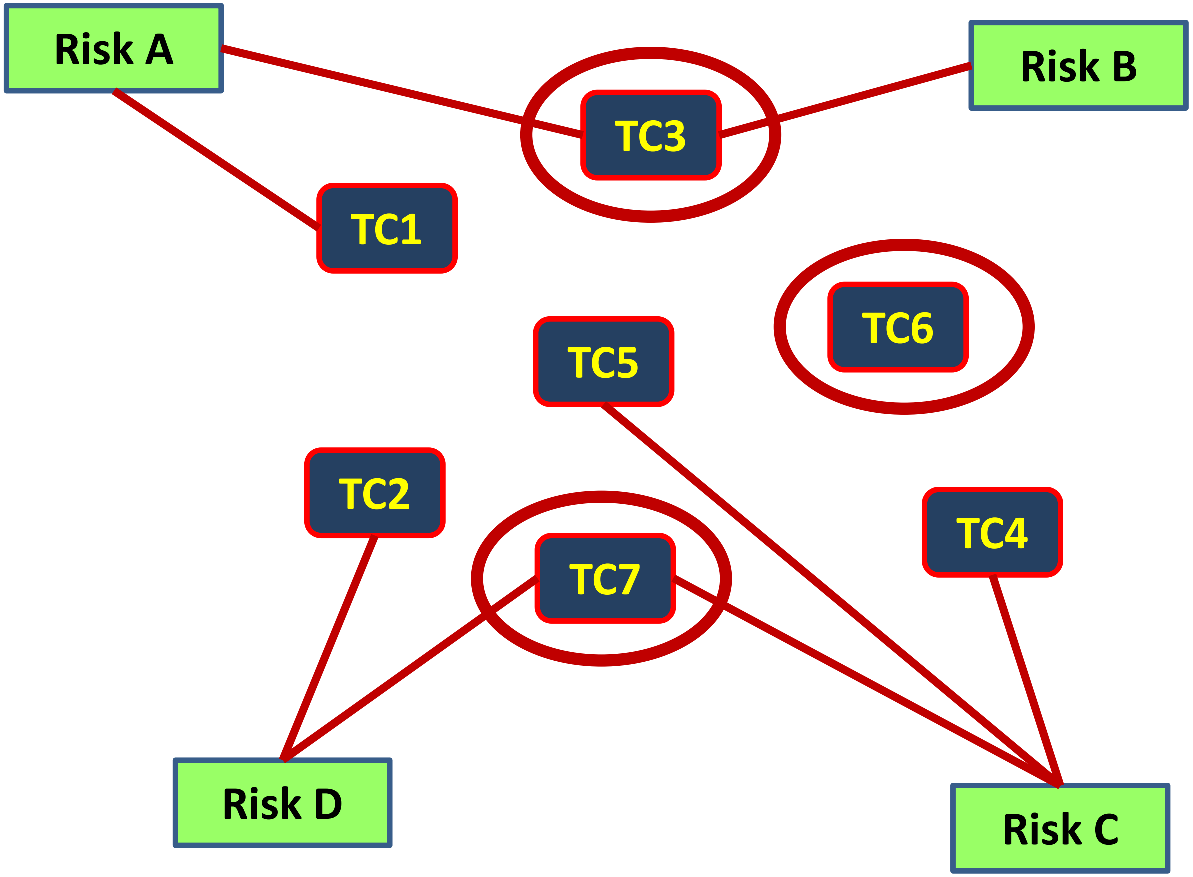
- **Open source**
- **Analyses C, Objective-C and Java**
  - on Android and iOS
- **Fast – can do millions of LOC in a few minutes**
  - ideal for continuous integration
- **Facebook claims that approximately 80% of raised issues are fixed (so are true faults)**
- **Also used by Instagram, Uber, Spotify, etc.**

Regression Testing

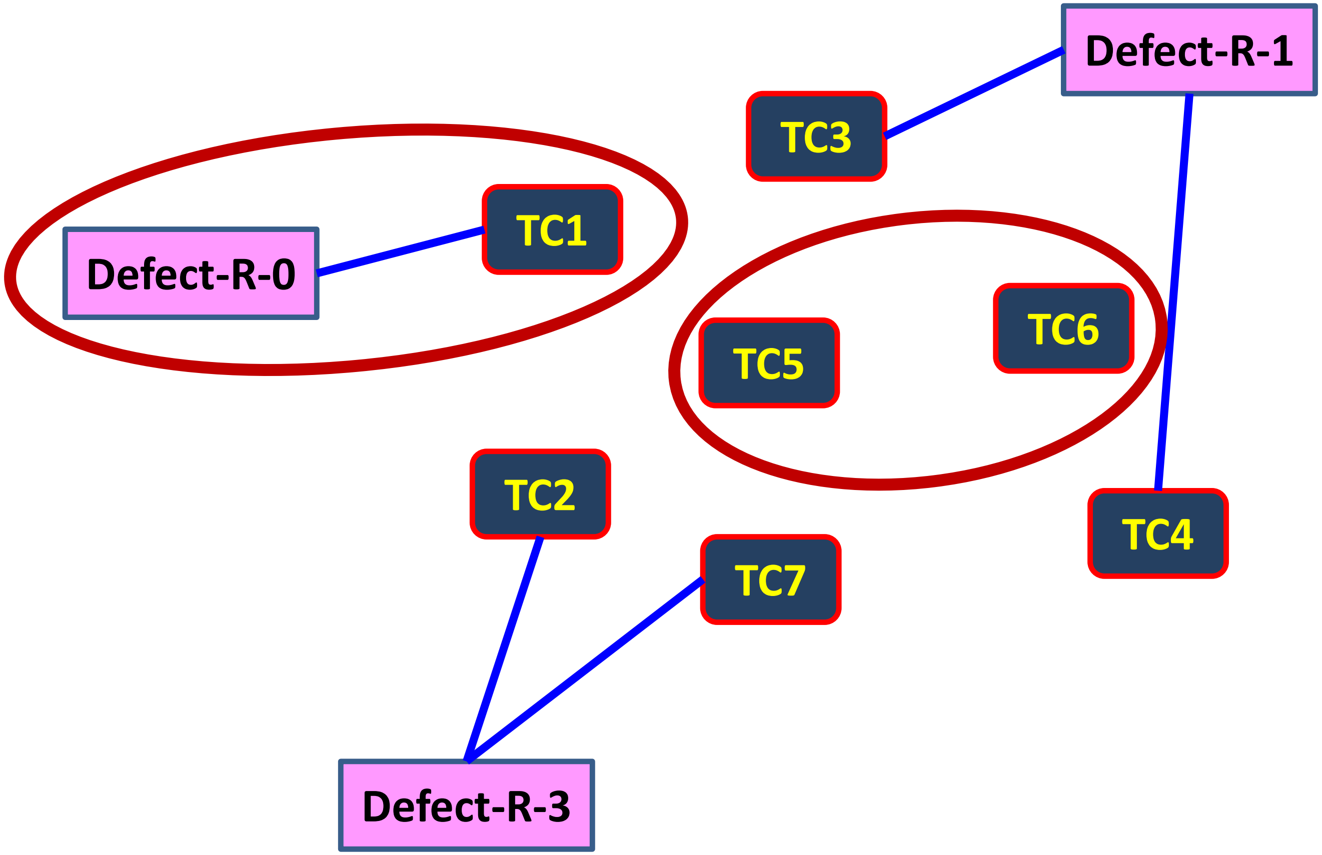
Regression Test Optimization



Regression Test Optimization



Regression Test Optimization



Regression Test Selection

Intelligent Agent

Branch Q

Risk A

Defect-R-1

Branch X

TC3

TC1

Defect-R-0

Risk B

TC6

TC5

Branch Y

Branch W

TC2

TC4

TC7

Branch R

Risk C

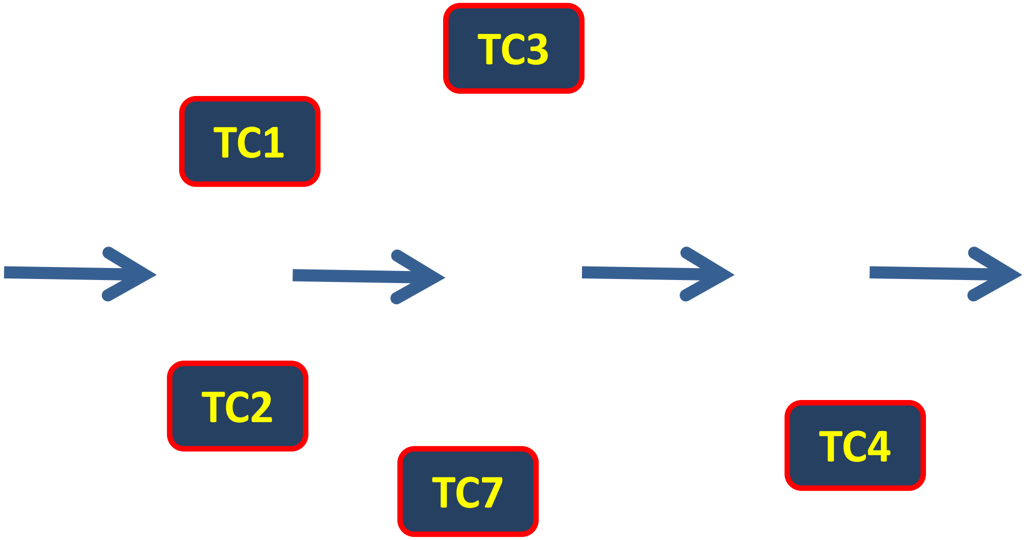
Defect-R-3

Branch P

Risk D

Regression Test Prioritization

Intelligent Agent



TC3

TC1

TC2

TC7

TC4

Regression Test Optimization Criteria

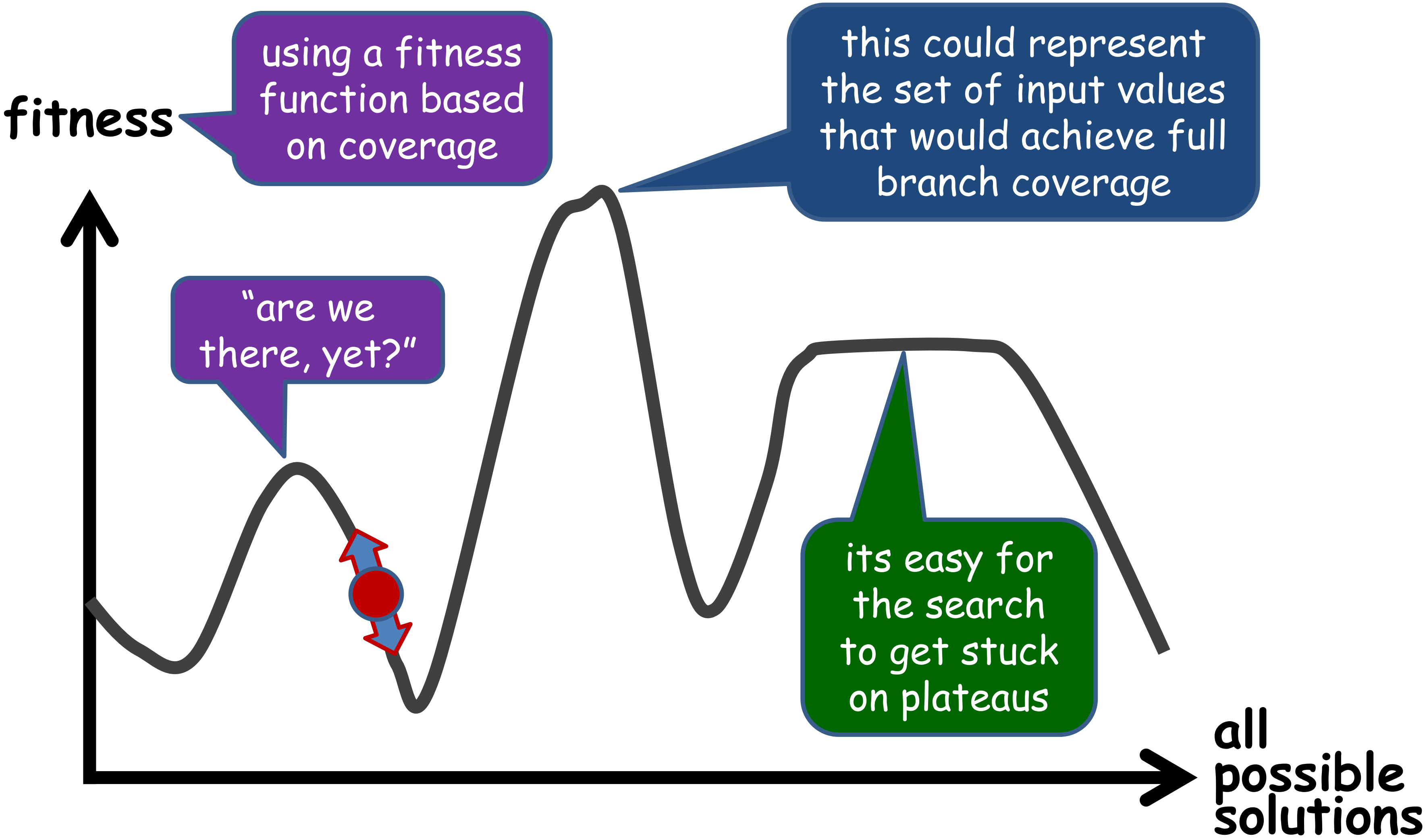
- Tests that found defects previously
- Tests that reduce execution time
- Reduce the number of tests needed
- Tests that achieve full coverage
- Test that exercise recently changed code
- Tests that address high risks
- etc.

Regression Test Optimization Results

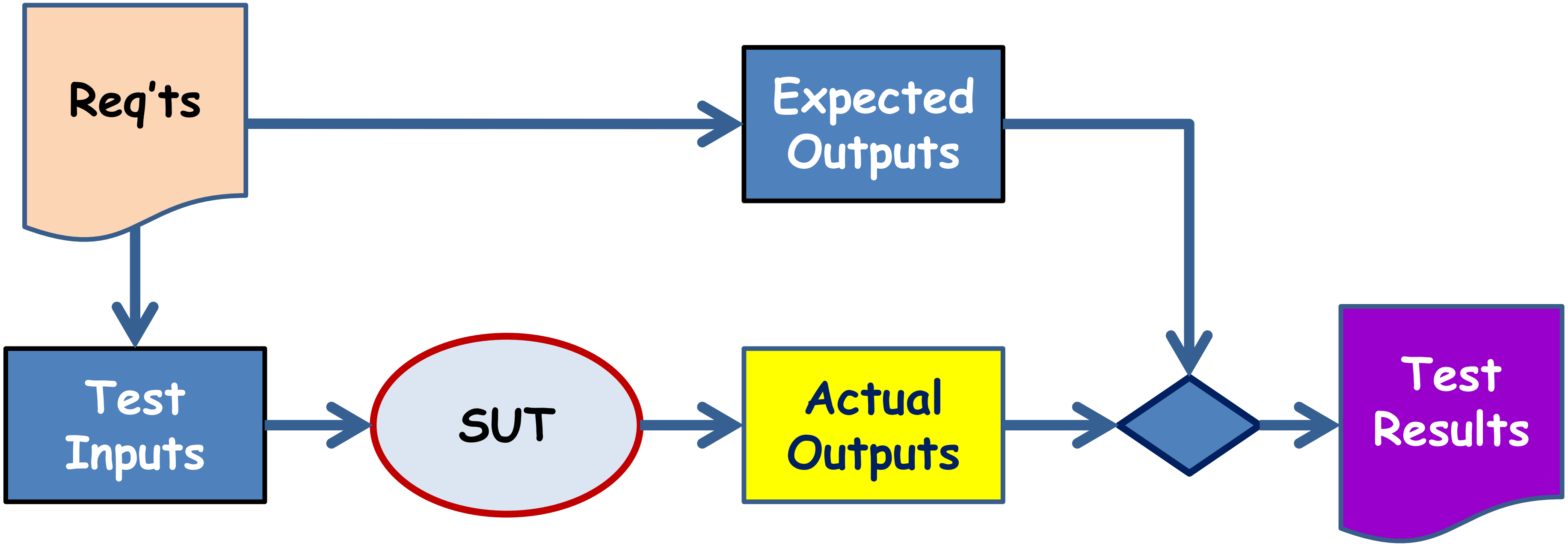
- **The algorithm reduces the test suite data by approximately 50%**
  - [Rai, 2014]
- **The techniques are 40-50% more effective in uncovering the first failure of the changed program**
  - [Jiang, 2009]
- **Average reduction in test suite size of 94% while achieving requirements-based coverage**
  - implemented in:
    - a continuous integration env’t with 30 seconds run time
    - implemented at Cisco, Norway
  - [Gotlieb, 2016]

Automated Test Input Generation

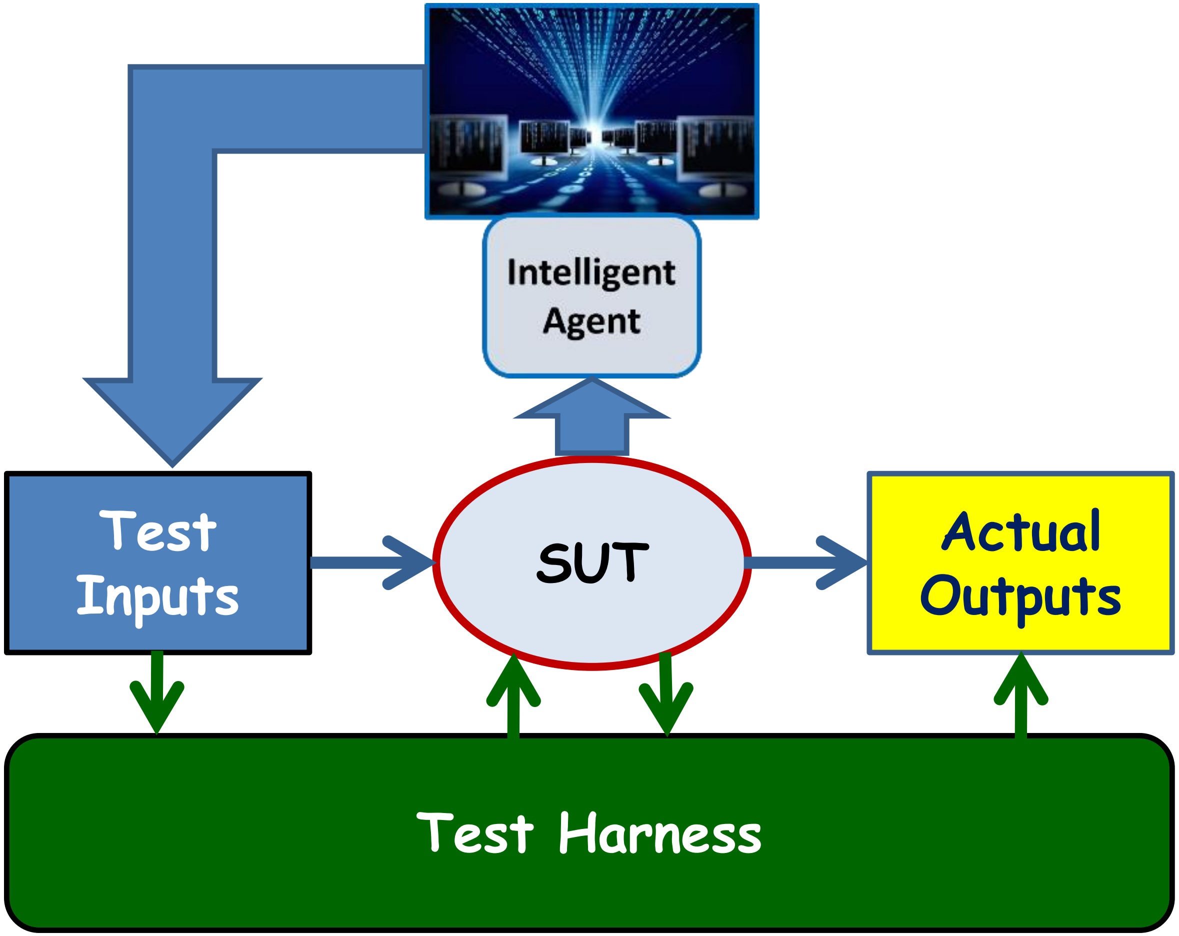
Example use of AI Search-Based Testing – Searching using a 'Hill Climb' Algorithm


fitness
using a fitness function based on coverage
this could represent the set of input values that would achieve full branch coverage
"are we there, yet?"
its easy for the search to get stuck on plateaus
all possible solutions

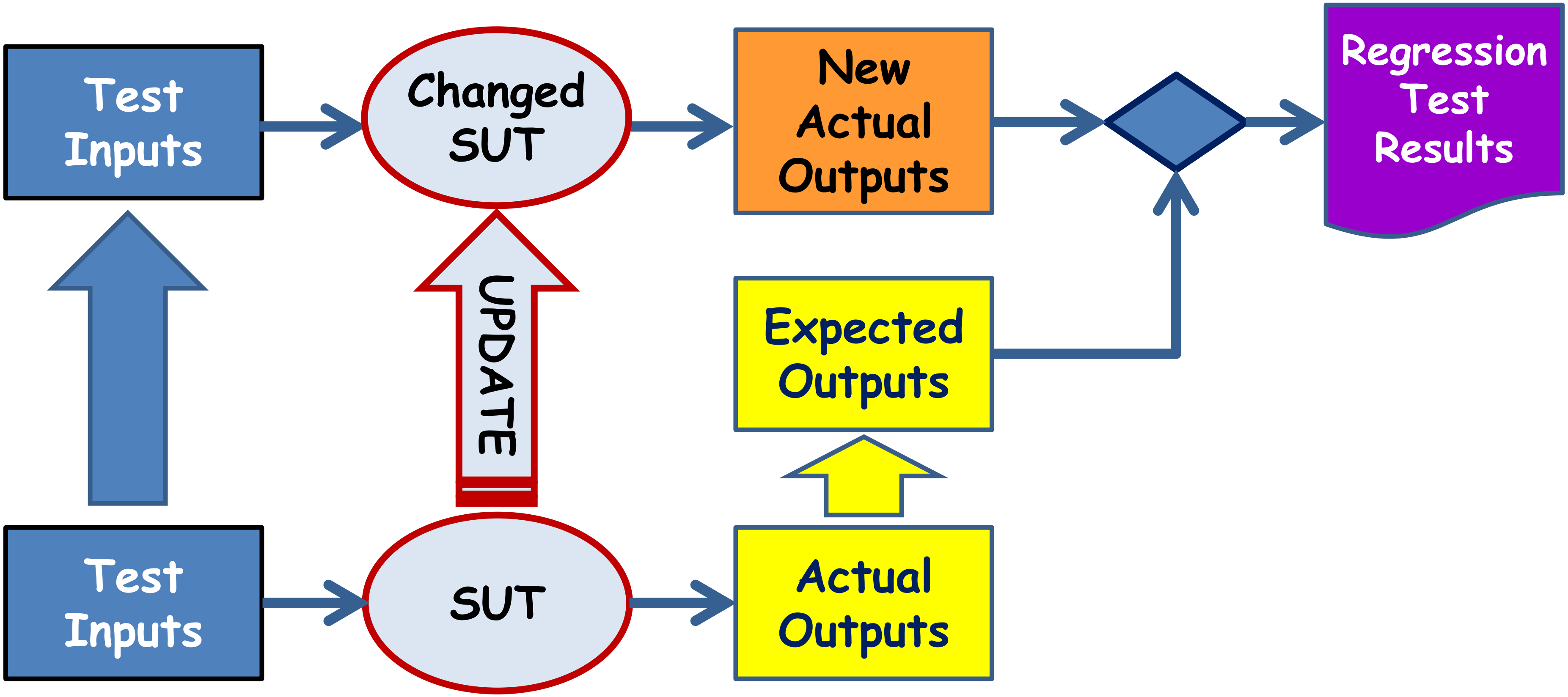
Manual Test Process



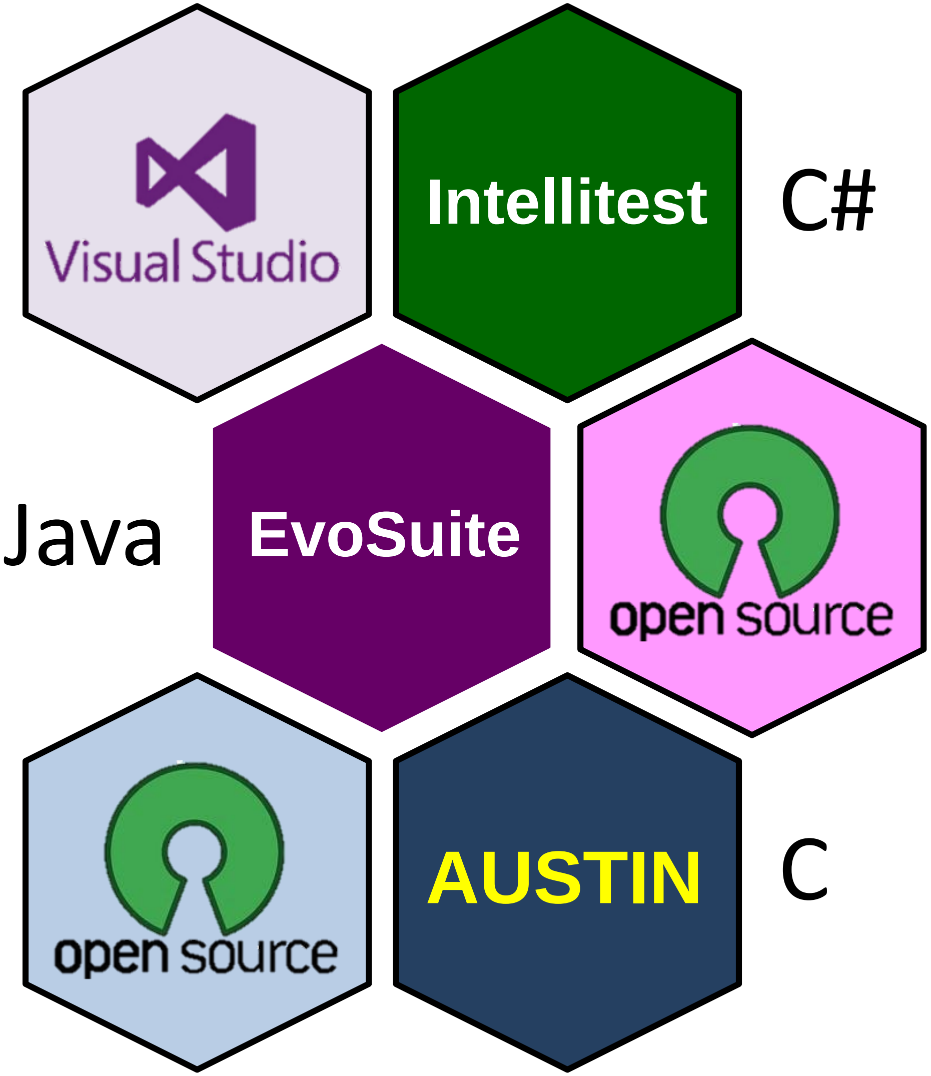
Automated Test Input Generation



Automated Test Input Generation



Example Tools



Intellitest

C#

Java

EvoSuite

AUSTIN

C

Automated Test Input Generation - Summary

- **Empirical studies have shown:**
  - tool support can lead to improvements in code coverage of up to 300%
  - that there is no measurable improvement in the number of bugs actually found by developer/testers – even though more branches are covered
- **But, a set of automatically-generated regression tests providing full coverage is an excellent starting point when you change or refactor the code**

- **Danger!!!**
  - testers rely on the tool → little or no black box testing
  - testers use the tool to meet safety-related test standards

Automated Stress Testing

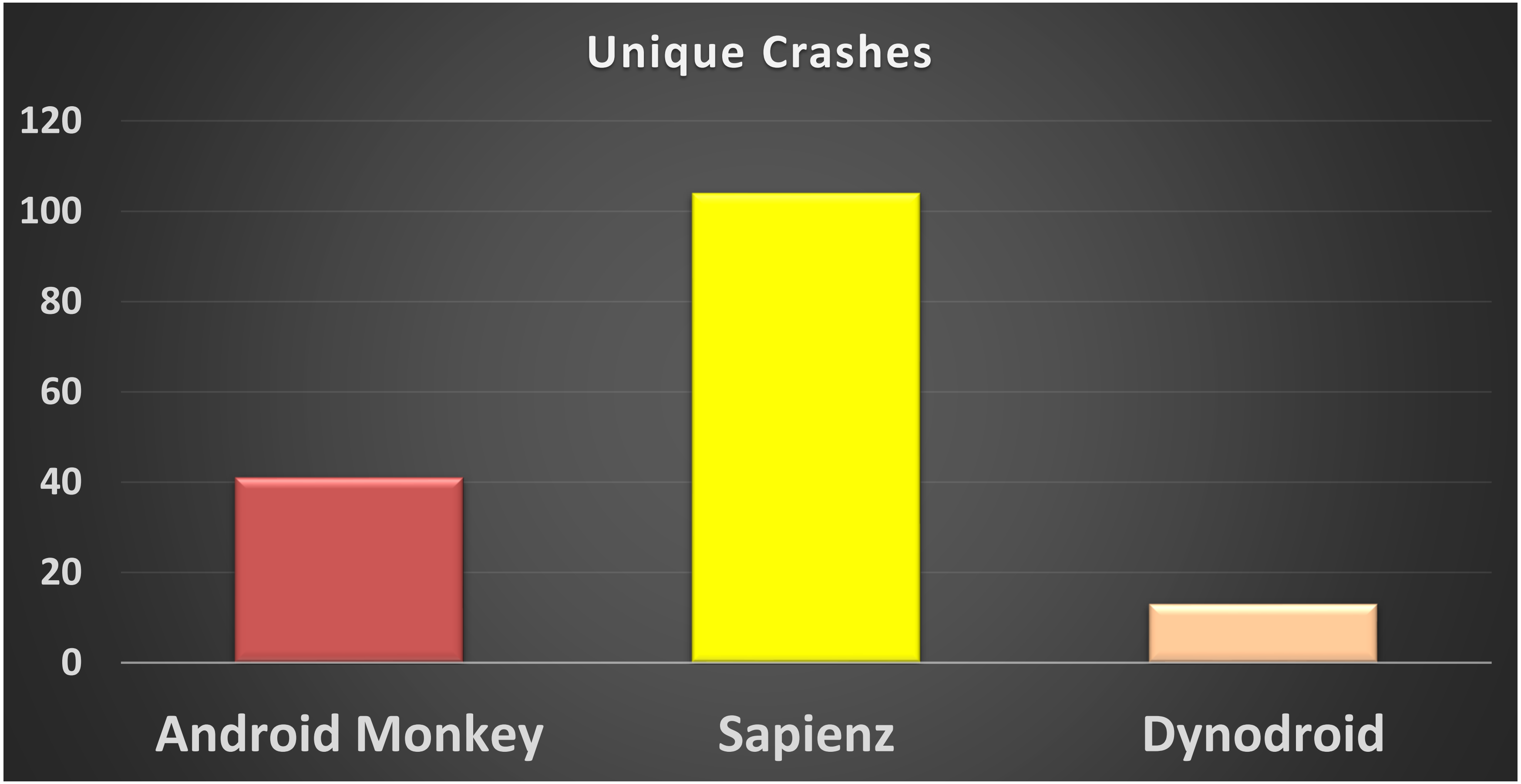
Automated Stress Testing Tools

- **Generate pseudo-random streams of user events such as clicks, touches, or gestures, as well as a number of system-level events**
  - they pretend they are a ‘stupid’ tester
- **Aim to cause an ANR (‘Application Not Responding’) or for the app to simply crash**
  - so test result is easy to observe
- **Require little tester input**
  - except to check-out the reported failures

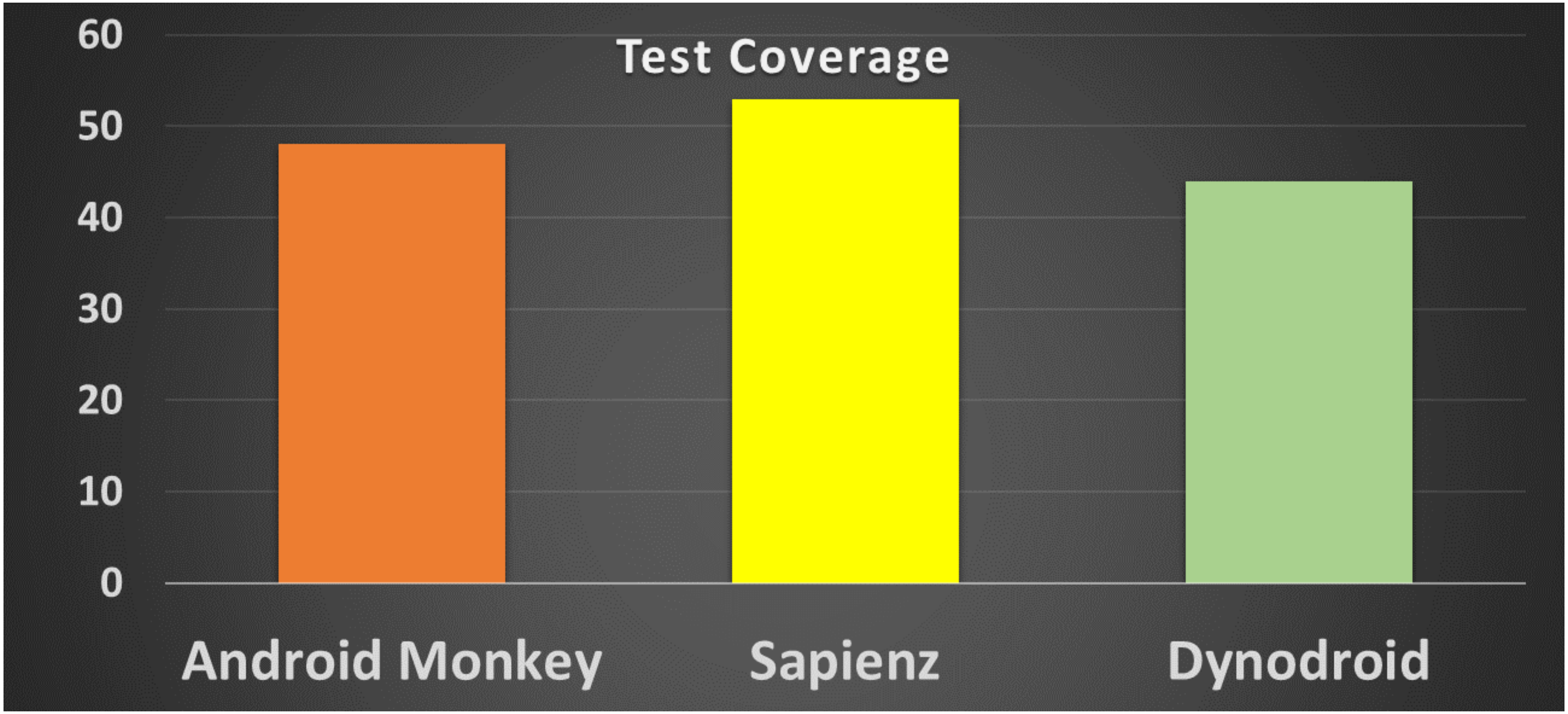
Example - Android Stress Testing Tools

- **Google Monkey**
  - built into the Android development platform - free
  - fuzz testing tool – random inputs
- **Sapienz**
  - open source
  - search-based testing tool
  - when applied to the top 1,000 Google Play apps, Sapienz found 558 unique, previously-unknown faults
- **Dynodroid**
  - open source
  - allows interleaving of human and tool
  - when applied to the top 1,000 Google Play apps, Dynodroid found 6 unique, previously unknown faults

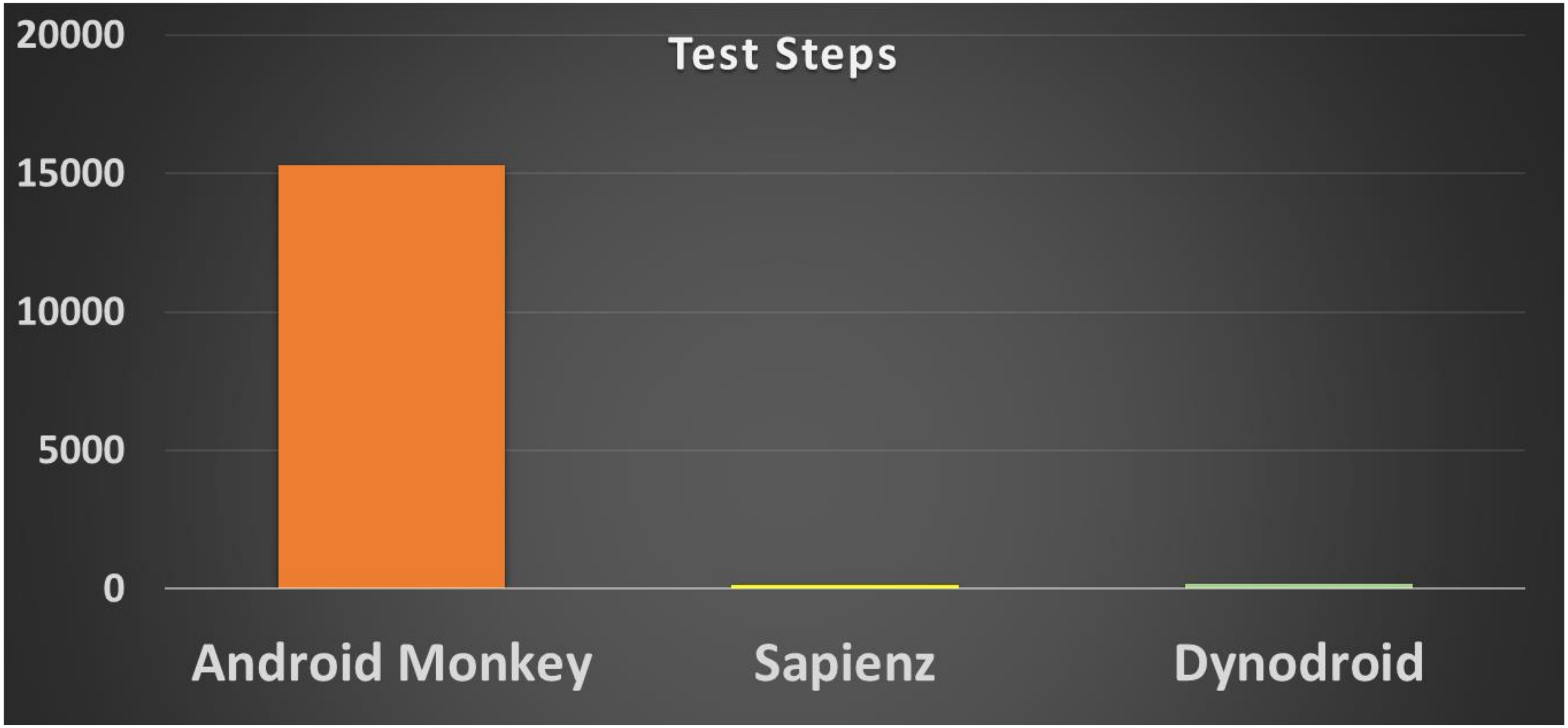
Defect Detection Effectiveness



Test Coverage



Fault Revealing Steps



Conclusions

- **The 4th Industrial Revolution**
  - Robots
  - Digital Disruption
  - Technology & User Expectations
- **Smarter Testing**
  - Random Testing & Fuzz Testing
  - Crowd Testing & A/B Testing
  - Testing using Artificial Intelligence (AI)

Thank you for listening ☺

Any Questions?