

Software Analytics Structural Quality Measurement and the State of IT Software

Dr. Bill Curtis

Chief Scientist, CAST & SVP, CAST Research Labs Director, Consortium for IT Software Quality (CISQ)



1 PSM 2011—Curtis — 13JUL11 — © CAST 2011

How Do We Get to Dependable Software

National Research Council Software for Dependable Systems



"As higher levels of assurance are demanded...testing cannot deliver the level of confidence required at a reasonable cost."

"The cost of preventing all failures will usually be prohibitively expensive, so a dependable system will not offer uniform levels of confidence across all functions."

"The correctness of the code is rarely the weakest link."



Jackson, D. (2009). Communications of the ACM, 52 (4)

Internal Quality Is Often Overlooked

Quality The degree to which a product meets its specified requirements

Problem—Customers struggle to state functional requirements. They do not understand non-functional requirements.



"...a failure to satisfy a non-functional requirement can be critical, even catastrophic...non-functional requirements are sometimes difficult to verify. We cannot write a test case to verify a system's reliability...The ability to associate code to non-functional properties can be a powerful weapon in a software engineer's arsenal."

What Is An Application





Multi-Tier, Multi-Language Applications





Revising Our Understanding of Defects

Study of defects across 1 open source and 2 large NASA applications

Observation	Percent of cases			
Fixes mapping to \geq 2 files	≈ 60%			
Fixes mapping to \geq 3 files	≈ 30-40%			
Fixes mapping to \geq 2 components	≈ 10-36%			
Fixes mapping to <u>></u> 2 subsystems	≈ 10-20%			
Narrow spread of faults	80% of faults in 20% of files			
Types of defects	Requirements33%Coding33%Data14%			

M. Hamill & K. Goseva-Popstojanova (2009). Common faults in software fault and failure data. IEEE TSE, 35 (4), 484-496.



Application Quality vs. Code Quality



Application Quality

Application quality <u>also</u> measures how well the individual components work together to make up the overall business system

Component Quality

Code quality is the measure of individual components for compliance with standards and best practices in the context of a specific language

Good component quality ≠ Good application quality



7 PSM 2011—Curtis — 13JUL11 — © CAST 2011

The 4th Wave in Software Engineering

4 vijieren en e	V V V V	Vhat: Vhen: Vhy:	Architecture, Quality characteristics, Reuse 2000→ Ensure software is constructed to standards that meet the lifetime demands placed on it
3 W Process	What: When: Why:	CMN 1990 Prov prof	A, ITIL, PMBOK, Agile 0-2000 vide a more disciplined environment for essional work incorporating best practices







Software Physiology





Comparing Ancient Languages





10 PSM 2011— Curtis — 13JUL11 — © CAST 2011

How the Data Were Collected

APPLICATION ENCE PLATFORM PARSERS ANGUAGE Z U C CAST ш

Oracle PL/SQL Sybase T-SQL **SQL Server T-SQL IBM SQL/PSM** C, C++, C# , Pro C COBOL CICS **Visual Basic VB.Net, ASP.Net** Java, J2EE **JSP** XML HTML Javascript **VBScript** PHP **PowerBuilder Oracle Forms PeopleSoft** SAP ABAP. **Netweaver** Tibco **Business Objects Universal Analyzer**

Security Performance Changeability







12 PSM 2011—Curtis — 13JUL11 — © CAST 2011

Uses of Structural Quality Measures



ACHIEVE INSIGHT. DELIVER EXCELLENCE



The Study Sample

	No. of	No. of
industry	Companies	Apps
Energy & Utilities	4	23
Financial Services	10	54
Insurance	12	25
I⊤ and Business Consulting	6	10
Manufacturing	11	25
Other	11	21
Public Administration	10	63
Software ISV	5	16
Telecommunications	5	51
Grand Total	74	288

14 PSM 2011—Curtis — 13JUL11 — © CAST 2011



Distribution of Application Size in KLOC



15

ACHIEVE INSIGHT. DELIVER EXCELLENCE

Industry Segments & Technologies

						Mixed	Oracle	
Industry	.NET	ABAP	C/C++	COBOL	Java EE	Technologies	4GL	Other
Energy & Utilities	2	1			17	2		1
Financial Services	1			18	6	26	1	2
Insurance	3		1	7	12	1		1
IT and Business Consulting		2		1	6	1		
Manufacturing	3	10	2		7		2	1
Other	3	1	3		10	2	1	1
Public Administration		3		1	22	2	35	
Software ISV	1	2			8	5		
Telecommunications	1	4	3		35	6	1	1
Grand Total	14	23	9	27	123	45	40	7



Top 4 Violations by Language

Java EE	Use of accessors to Private Fields Artifacts with High Fan-Out Unreferenced Methods Unreferenced Fields
.NET	Declaring Public Class Fields Artifacts with High Fan-Out Classes with a High Lack of Cohesion Artifacts with High Fan-In
С	Artifacts with High Fan-Out Large Functions - too many Lines of Code Functions with SQL statement including Subqueries Artifacts with High Cyclomatic Complexity
C++	Data Members that are not Private Artifacts with High Fan-Out Included files including other files Artifacts with High Cyclomatic Complexity
ABAP	Artifacts with a Complex SELECT Clause Artifacts with High Fan-Out Artifacts with High Cyclomatic Complexity Artifacts with High Essential Complexity



Distribution of Security Scores



18 PSM 2011— Curtis — 13JUL11 — © CAST 2011

Achieve Insight. Deliver Excellence.

Security Distributions by Language



ACHIEVE INSIGHT. DELIVER EXCELLENCE.

Distribution of Performance Scores



20 PSM 2011—Curtis — 13JUL11 — © CAST 2011

ACHIEVE INSIGHT, DELIVER EXCELLENCE.

Performance Distributions by Language



21 PSM 2011—Curtis — 13JUL11 — © CAST 2011

CHIEVE INSIGHT. DELIVER EXCELLENCE.

Distribution of Changeability Scores



22 PSM 2011— Curtis — 13JUL11 — © CAST 2011



Changeability Distribution by Industry



Does Size Matter?



24 PSM 2011— Curtis — 13JUL11 — © CAST 2011

Achieve Insight. Deliver Excellence

R² – TQI with Size for COBOL Apps



K Lines of Code



% Complex Components by Language



ACHIEVE INSIGHT, DELIVER EXCELLENCE

GO TO Considered Eternal

COBOL

Use of PERFORM ... THROUGH | THRU Missing WHEN OTHER when using EVALUATE Components with High Cyclomatic Complexity



27 PSM 2011— Curtis — 13JUL11 — © CAST 2011

CISQ— An Industry Response

