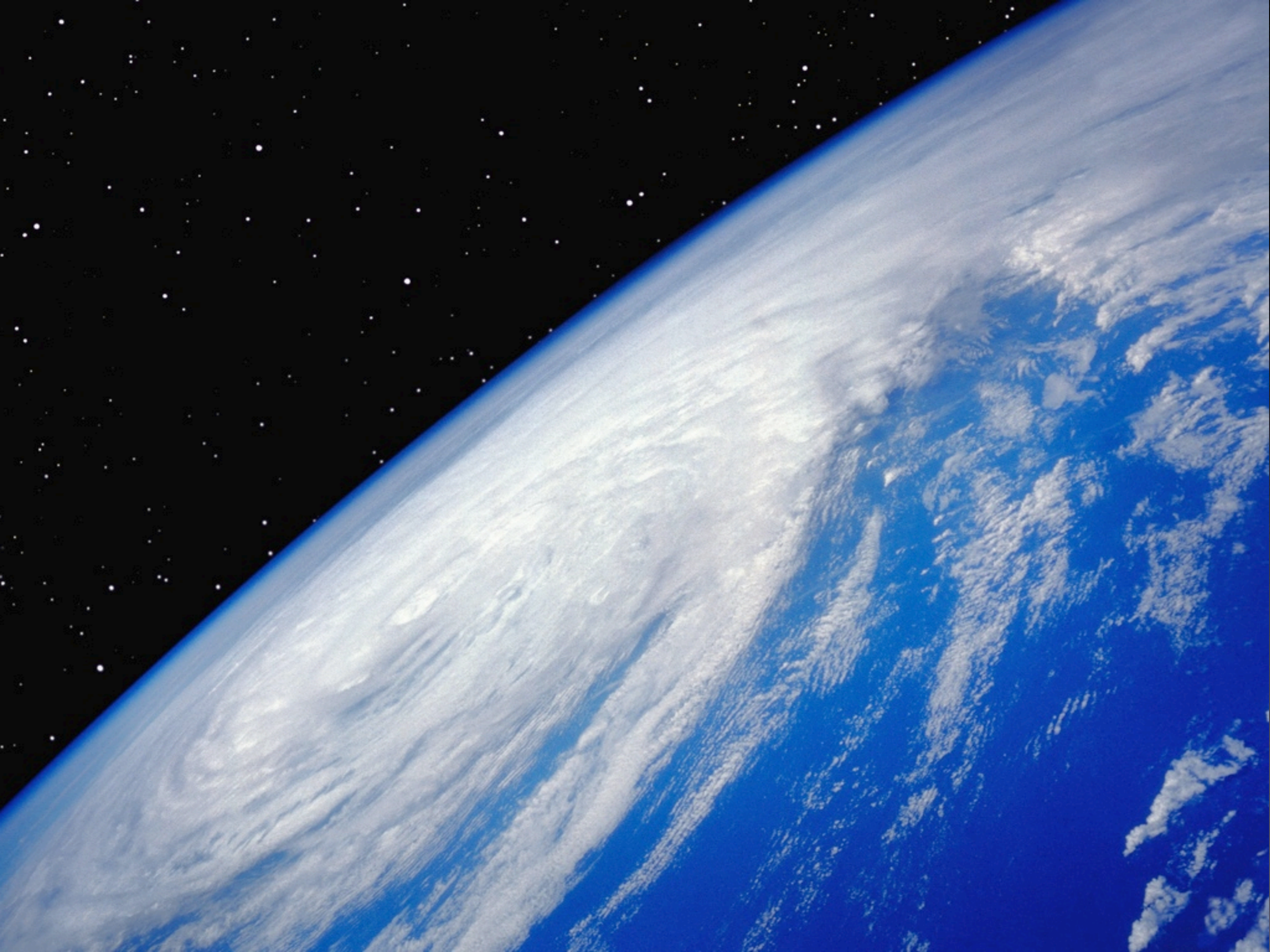


Learning from Software

Andreas Zeller

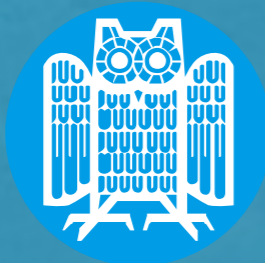
Saarland University



Saarbrücken



Saarbrücken



UNIVERSITÄT
DES
SAARLANDES



Max
Planck
Institute
for
Software Systems



max planck institut
informatik

Saarbrücken

- 28 tenured professors
- 55 professors at assistant level or higher
- 221 PhD students from all countries
- English as study and working language
- Largest computer science cluster in Europe



Saarbrücken



Saarbrücken





UNIVERSITÄT
DES
SAARLANDES

Excellence Initiative Proposal for the

Saarbrücken Graduate School of Computer Science

An integrated program from BSc to PhD



UNIVERSITÄT
DES
SAARLANDES

Excellence Initiative Proposal for a Cluster of Excellence

Multimodal Computing and Interaction

Robust, Efficient and Intelligent Processing of Text, Speech,
Visual Data and High Dimensional Representations

Two best-ranked
applications of
whole initiative

40 million €

- Multimodal Computing Excellence Cluster
- *20 junior research groups*
- assistant prof position
staff • funds • teaching
- www.mmci.uni-saarland.de



UNIVERSITÄT
DES
SAARLANDES

Excellence Initiative Proposal for a Cluster of Excellence

Multimodal Computing and Interaction

Robust, Efficient and Intelligent Processing of Text, Speech,
Visual Data and High Dimensional Representations



Apply today!





UNIVERSITÄT
DES
SAARLANDES

Excellence Initiative Proposal for the

Saarbrücken Graduate School of Computer Science

An integrated program from BSc to PhD



UNIVERSITÄT
DES
SAARLANDES

Excellence Initiative Proposal for a Cluster of Excellence

Multimodal Computing and Interaction

Robust, Efficient and Intelligent Processing of Text, Speech,
Visual Data and High Dimensional Representations

Two best-ranked
applications of
whole initiative





UNIVERSITÄT
DES
SAARLANDES

Excellence Initiative Proposal for the

Saarbrücken Graduate School of Computer Science

An integrated program from BSc to PhD

Apply today!

- Graduate School from BSc to PhD
- 40 stipends @ 700–1400 €/month
- Research *what thou wilt*
- gradschool.cs.uni-sb.de

Saarbrücken

Program
Verification

Artificial
Intelligence

Information
Processing

Algorithms

Visual Computing

Software
Systems

Bioinformatics

Software
Engineering

Actually, I am not a software engineer. I am interested in programmers, at a somewhat much lower level.

Models

Specs

Code

Traces

Profiles

Tests

Learning from Software

Andreas Zeller

Saarland University

e-mail

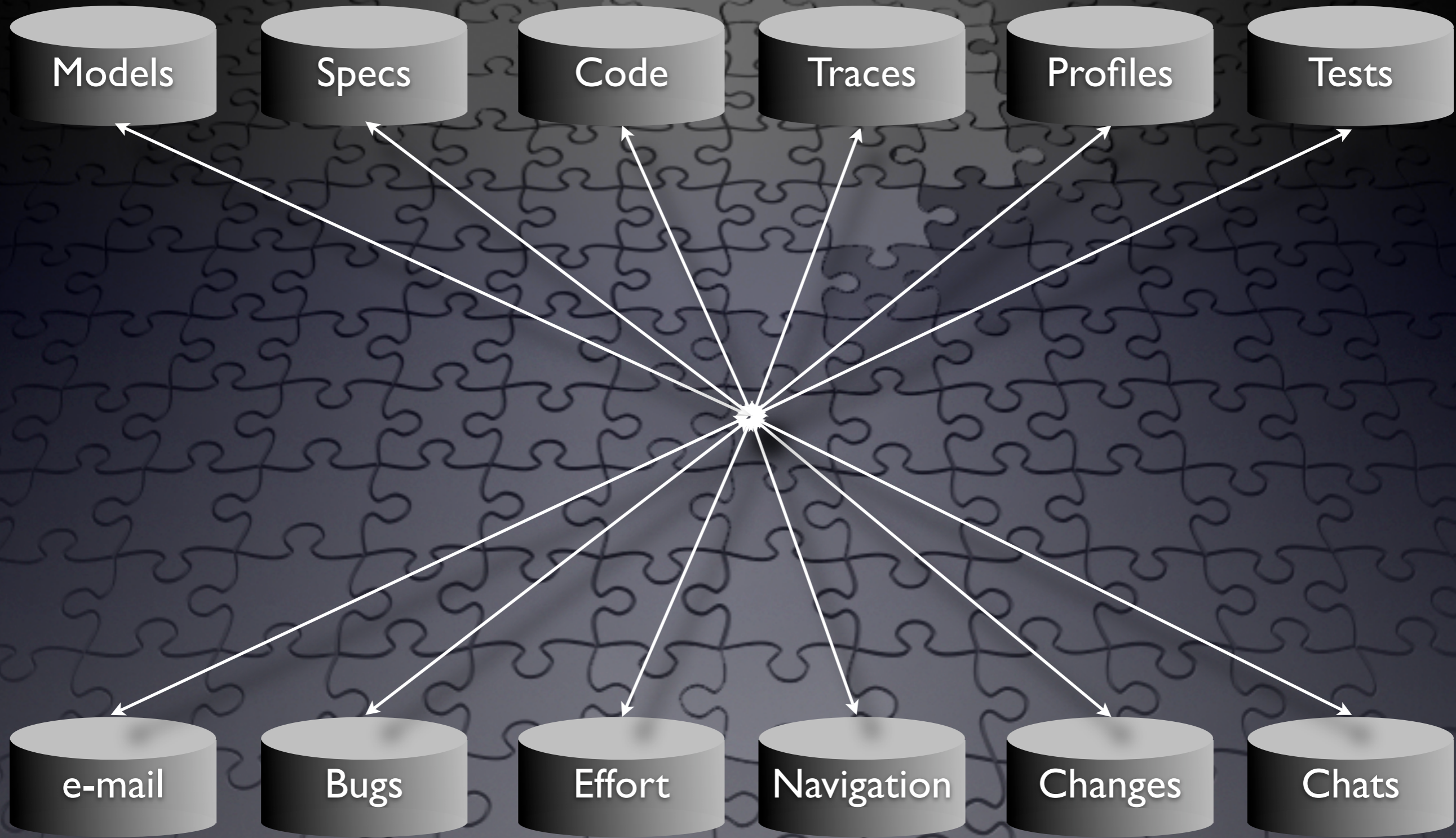
Bugs

Effort

Navigation

Changes

Chats



Problem Report for Keynote

Problem and system information:

Date/Time: 2006-03-07 23:35:25.516 +0100
OS Version: 10.4.5 (Build 8H14)
Report Version: 4

Command: Keynote
Path: /Applications/iWork '06/Keynote.app/Contents/MacOS/Keynote
Parent: WindowServer [79]

Version: 3.0.0 (423)
Build Version: 1
Project Name: iWork

Please describe what you were doing when the problem happened:

Your report will help Apple improve this software. Your personal information is not sent with this report. You will not be contacted in response to this report. For Apple product support, visit www.apple.com/support or contact your Apple dealer.

Send to Apple...



Submit Report... Cancel

Bugs

Changes

Map bugs to
code locations

Bugs

Changes

Where do these bugs come from?

org.eclipse.core.internal

org.eclipse.core.runtime

org.eclipse.pde.internal

org.eclipse.jface.text

org.eclipse.jface.viewers

org.eclipse.jface.dialogs org.eclipse.jface.action

org.eclipse.jdt.compiler.parser
org.eclipse.jdt.compiler.parser.java

org.eclipse.jdt.ui

org.eclipse.jdt.core

org.eclipse.help.internal

org.eclipse.debug.internal

org.eclipse.ant.internal

org.eclipse.help.ui

Is it the Developers?

**Does experience
matter?**

**Bug density
correlates with
experience!**

Is it History?

I found lots of bugs here. Will there be more?

Yes! (But where did these come from?)

How about metrics?

Do *code metrics* correlate with bug density?

Sometimes!

Uh. Coverage?

Does *test coverage* correlate with bug density?

Yes –
the more coverage,
the more bugs!

Ah! Language features?

**Are gotos
harmful?**

No correlation!

Ok. Problem Domain?

Which tokens
do matter?

import • extends
• implements

Eclipse Imports

71% of all components importing compiler show a post-release defect

```
import org.eclipse.jdt.internal.compiler.lookup.*;
import org.eclipse.jdt.internal.compiler.*;
import org.eclipse.jdt.internal.compiler.ast.*;
import org.eclipse.jdt.internal.compiler.util.*;
...
import org.eclipse.pde.core.*;
import org.eclipse.jface.wizard.*;
import org.eclipse.ui.*;
```

14% of all components importing ui show a post-release defect

Eclipse Imports

Correlation with failure

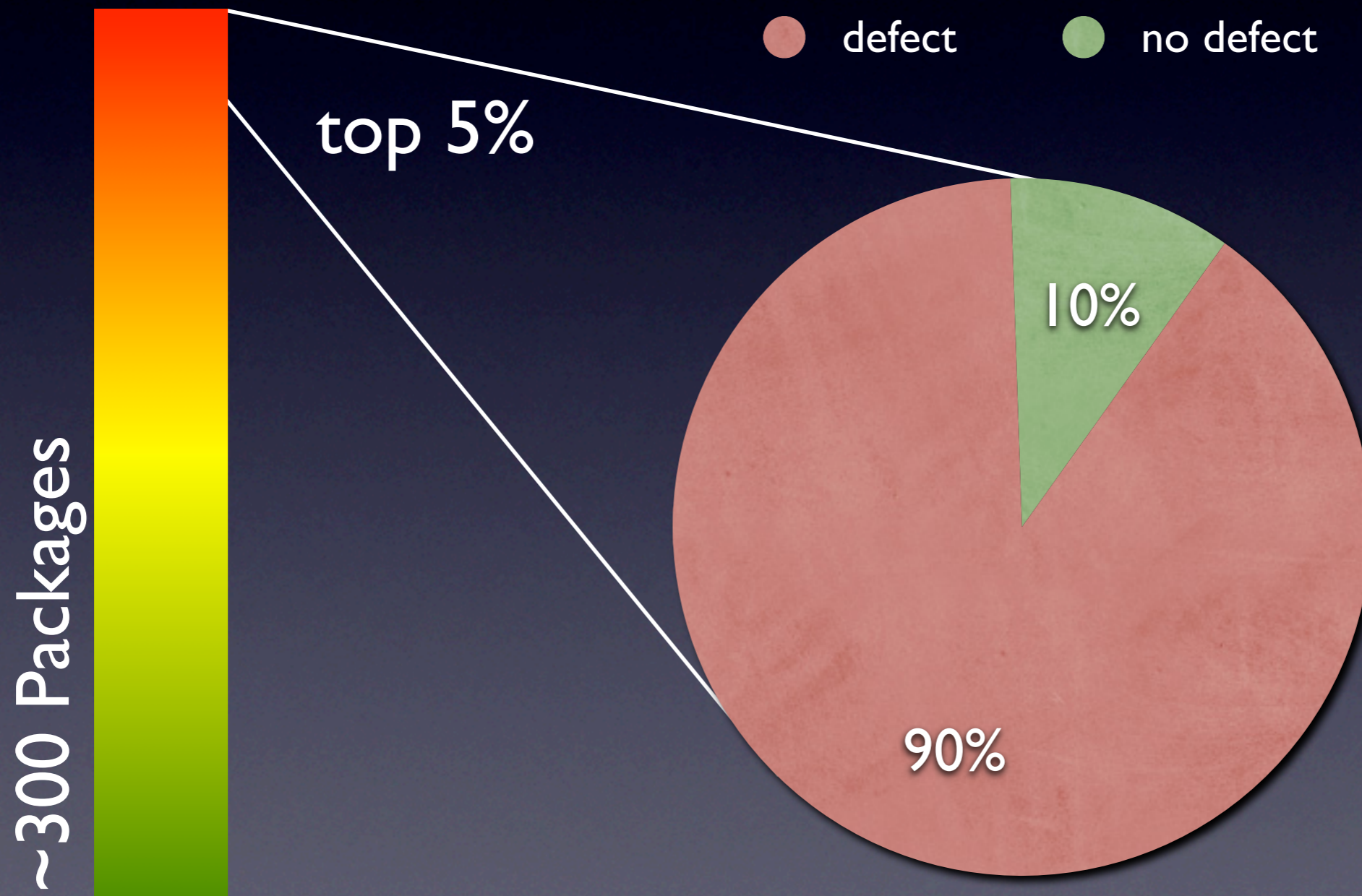
```
import org.eclipse.jdt.internal.compiler.lookup.*;  
import org.eclipse.jdt.internal.compiler.*;  
import org.eclipse.jdt.internal.compiler.ast.*;  
import org.eclipse.jdt.internal.compiler.util.*;
```

...

```
import org.eclipse.pde.core.*;  
import org.eclipse.jface.wizard.*;  
import org.eclipse.ui.*;
```

Correlation with success

Prediction

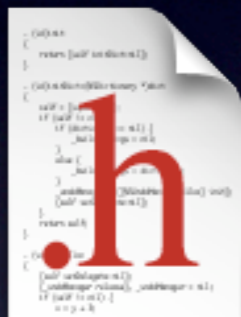


Firefox vulnerabilities

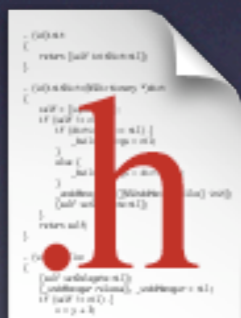




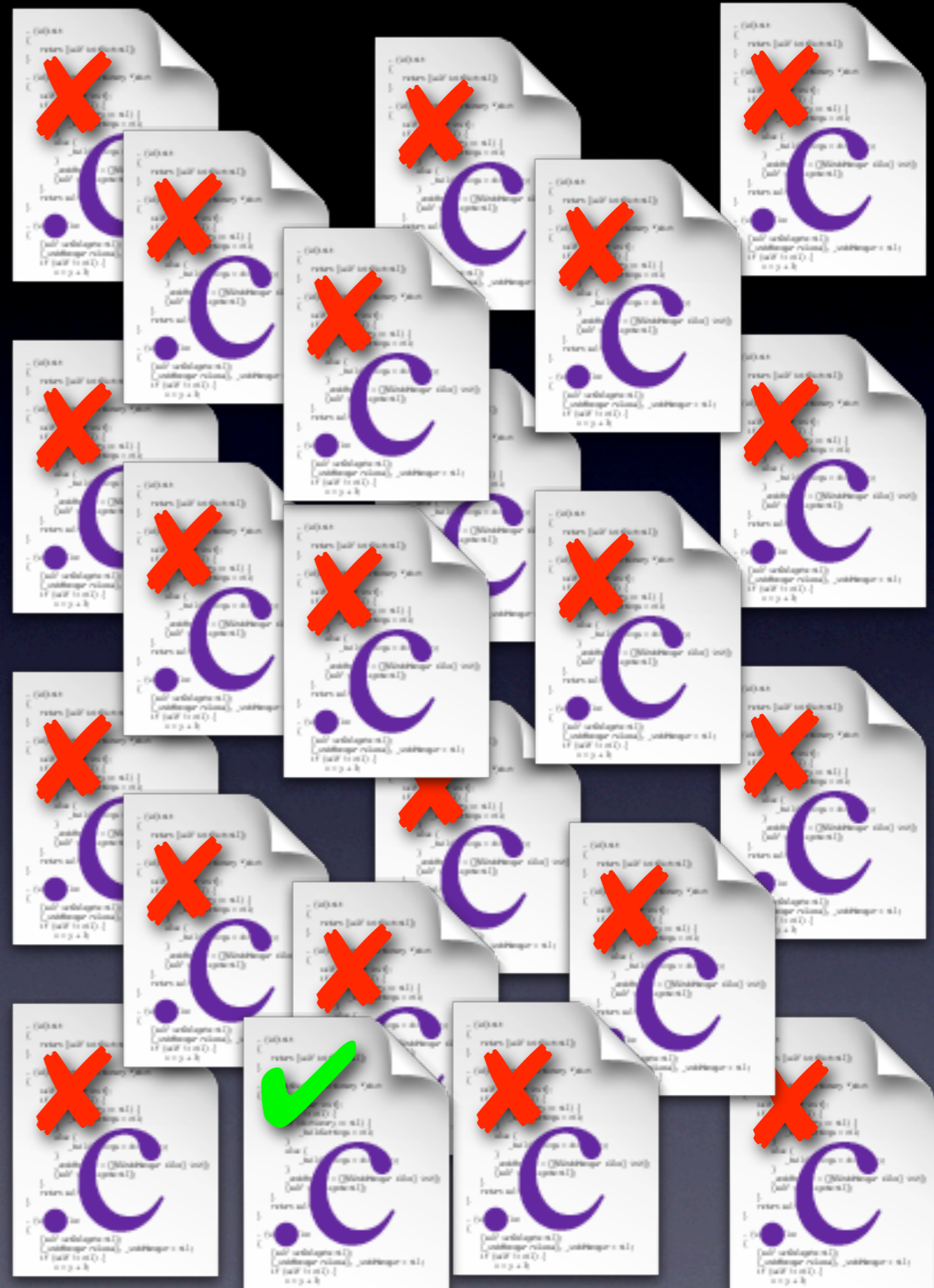
nsIContent.h



nsIContentUtils.h

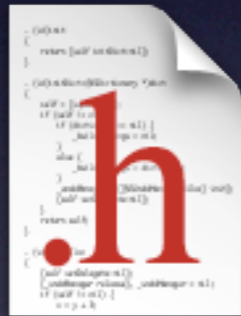


nsIScriptSecurityManager.h

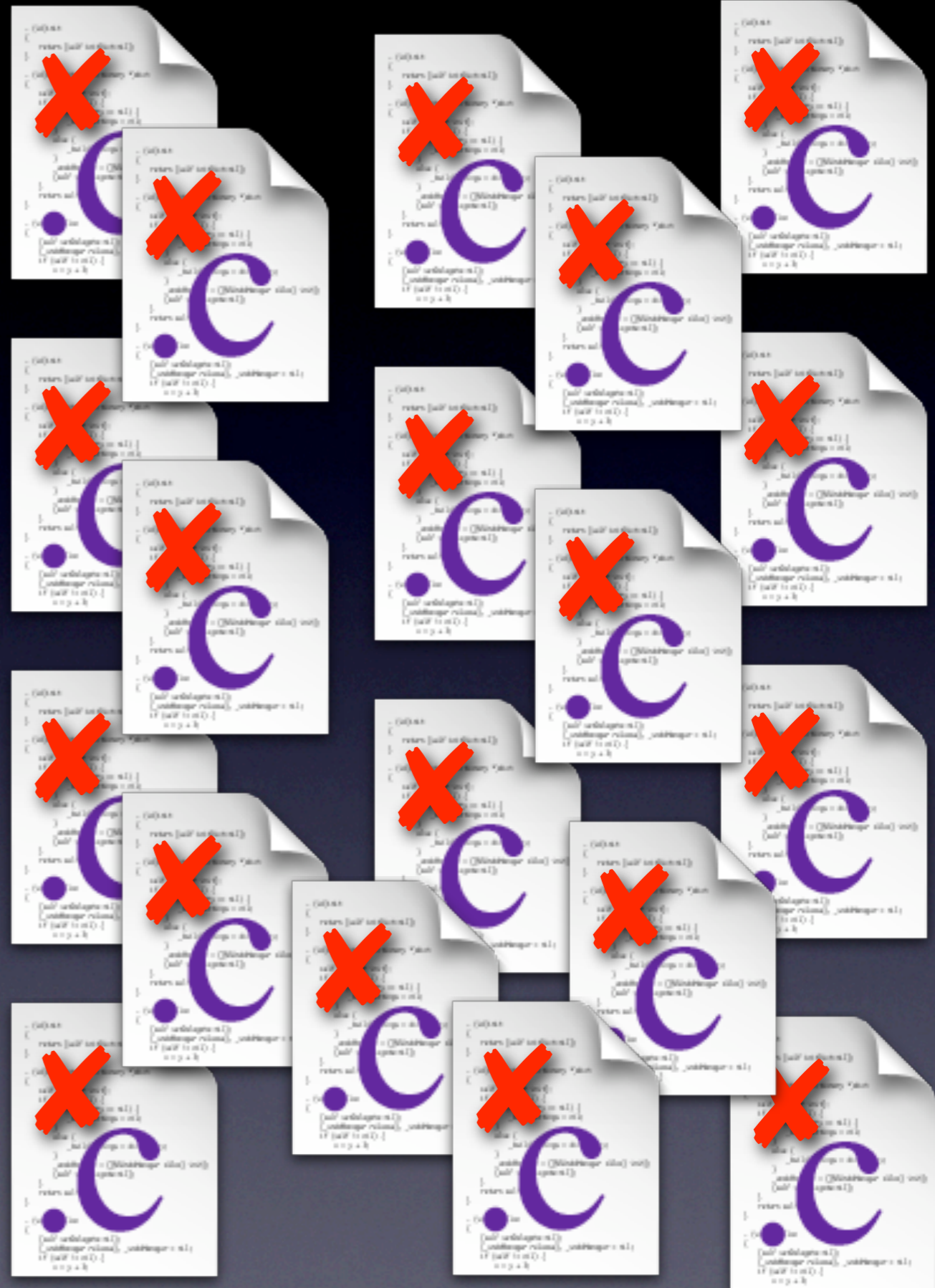




nsIPrivateDOMEvent.h



nsReadableUtils.h



Prediction	Component	Fact
1	nsDOMClassInfo	3
2	SGridLayout	95
3	xpcprivate	6
4	jspxml	2
5	nsGenericHTMLElement	8
6	jsgc	3
7	nsIEnvironment	12
8	jsfun	1
9	nsHTMLLabelElement	18
10	nsHttpTransaction	35

Software Archives

- contain full record of project history
- maintained via programming environments
- automatic maintenance and access
- freely accessible in open source projects





Models

Specs

Code

Traces

Profiles

Tests

miles



Mining and Learning from Software

e-mail

Bugs

Effort

Navigation

Changes

Chats

Very often



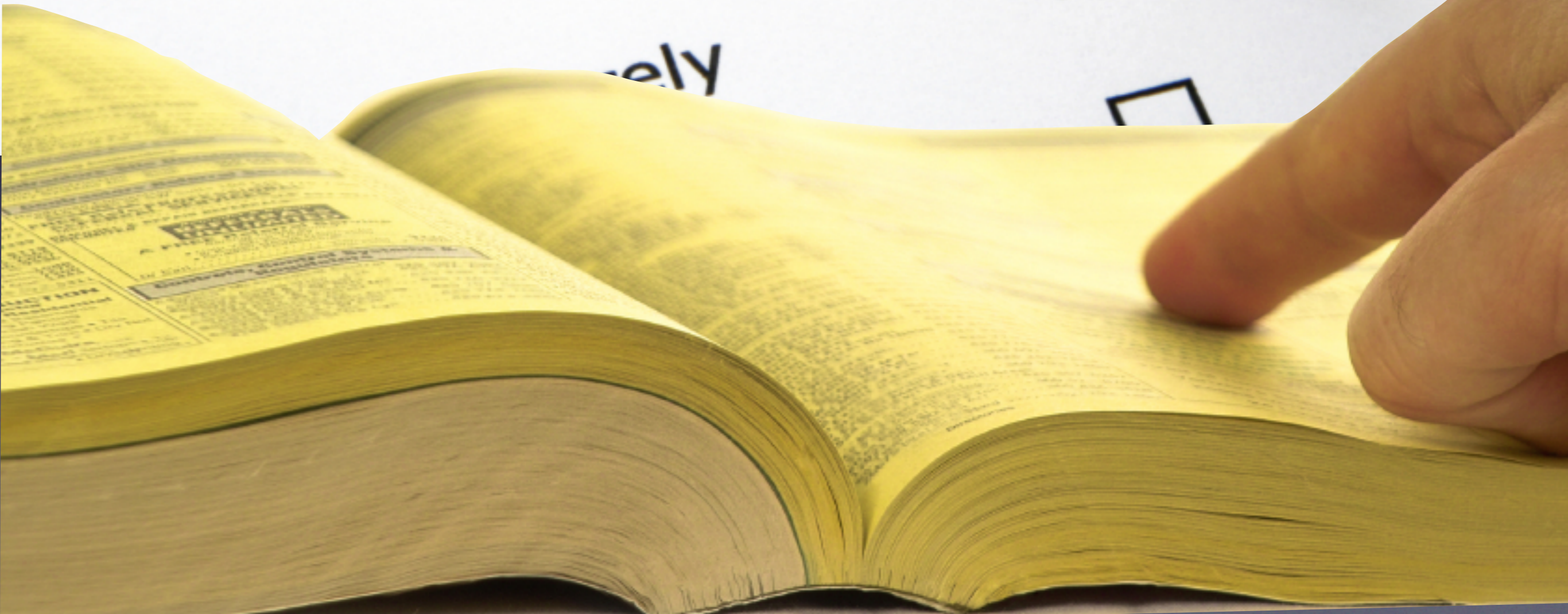
Often



Sometimes



ely





Jazz.net

The screenshot shows the IBM Jazz web application interface in a Mozilla Firefox browser. The page title is 'Jazz - Mozilla Firefox'. The address bar shows the URL: `http://jazz/jazz/web#perspectiveWorkItemAction=viewWorkItem&id=14463`. The page header includes the Jazz logo, navigation links (Home, Plans, Work Items, Reports), and user information: 'You are browsing as Guest | Login'. The main content area displays a defect report for 'Defect #14463' with the title 'NPE when clicking on Create Work Item'. The defect is categorized as 'Verified' and 'Fixed'. The 'Details' section shows the defect type as 'Defect', severity as 'Critical', created on 'Dec 14, 2006' by 'Michael Schneider', category as 'Eclipse UI', and tags as 'M9candidate'. The 'Description' section contains a stack trace for a 'java.lang.NullPointerException' in the 'ProjectAreaPicker' class. The 'Discussion' section shows a comment from Michael Schneider dated 'Dec 14, 2006 4:57 AM' with the same stack trace. The footer includes the IBM logo, 'Contact | Terms of Use', and 'Rational software'.

**IBM Jazz Faculty Award
for
Mining Jazz data to assess
development processes
25,000\$**



Eclipse Bugs

org.eclipse.core.internal

org.eclipse.core.runtime

org.eclipse.pde.internal

org.eclipse.jface.text

org.eclipse.jface.viewers

org.eclipse.jface.dialogs org.eclipse.jface.action

org.eclipse.jdt.internal.compiler.parser
AbstractCommentParser.java
pre: 33
post: 41

org.eclipse.jdt.ui

org.eclipse.jdt.core

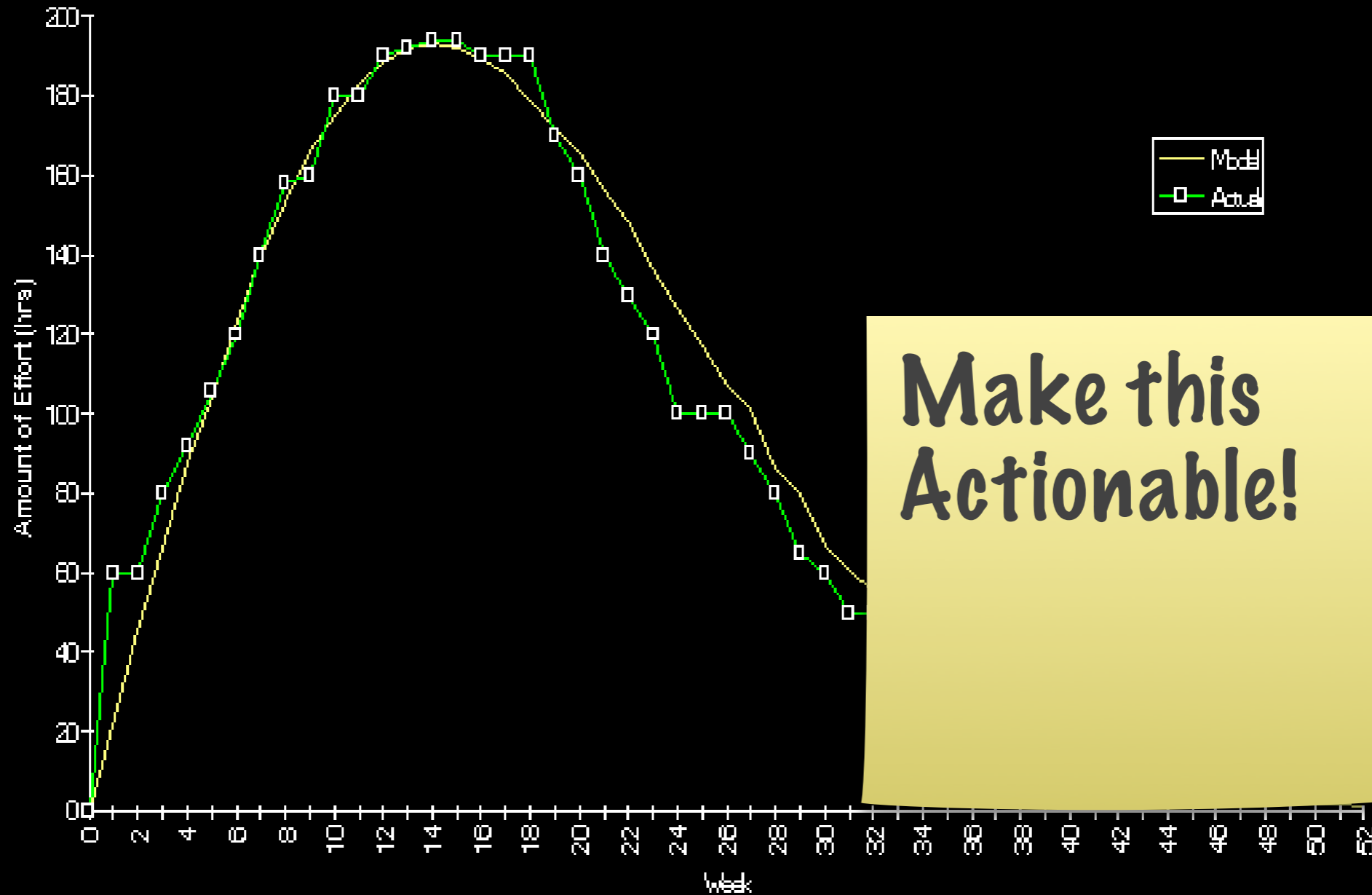
org.eclipse.help.internal

org.eclipse.debug.internal

org.eclipse.ant.internal

org.eclipse.help.ui

Studies



Rosenberg, L. and Hyatt, L. "Developing An Effective Metrics Program"
European Space Agency Software Assurance Symposium, Netherlands, March, 1996



Lynn Bridge City Hall Station

Fulton

Theatre Alley

Church St

P



240

yards

0.30 mi

33 mph

10:13 a

10:15 a

Broadway

tomtom



Assistance

The screenshot shows the Eclipse IDE interface. The main editor displays the `StandardSourcePathProvider.java` file with the following code snippet:

```
    } else {  
        // recover persisted source path  
        entries = recoverRuntimePath(configuration, IJavaLaunchConfigurationConstants.ATTR_SOURCE_PATH);  
    }  
    return entries;  
}
```

The Package Explorer on the left shows the project structure, including `StandardSourcePathProvider.java`, `StandardSourcePathProvider`, `VMRunnerConfiguration.java 1.3`, `VMStandin.java 1.12 (ASCII -ko)`, and `package.html 1.3 (ASCII -ko)`.

The Risky Locations view at the bottom left shows a table of risky locations:

Element	Risk
resolveClasspath	0.8888
run	0.8182
initializeProviders	0.7777
abort	0.7272
dispose	0.7000
init	0.6000
translate	0.5555
cleanup	0.5000
launch	0.4545

The Bugzilla search results at the bottom right show a table of bugs:

Version	Date	Author	Description
1.12	4/7/04 9:01 PM	darin	Bug 29890 - Debug Platform Source Lookup Facilites
1.10	1/29/04 9:38 PM	darin	Bug 34297 - allow a launch configuration classpath to be "default plus"
1.8	10/23/03 3:48 AM	darin	Fallback fix for bug 44877
1.7	10/22/03 9:18 PM	lboulier	Bug 44877 - Wrong JDK source lookup if Compile-JDK <> Debug-JDK

The bug details for Bug 34297 are shown below the table:

allow a launch configuration classpath to be "default plus"

Bug 34297 - allow a launch configuration classpath to be "default plus"

I often want to launch a program using its default classpath plus one or two additional jars or directories. The first time this is easy - just turn on the "use default classpath" checkbox, then turn it off and add the additional

At the bottom of the IDE, the status bar shows "Writable", "Smart Insert", and "65 : 1".

Future environments will

- mine patterns from program + process
- apply rules to make predictions
- provide assistance in all development decisions
- adapt advice to project history

Wikis
Joy of Use
Participation
Usability
Recommendation
Social Software
Collaboration
Perpetual Beta
Simplicity
Trust
Economy
Remixability
DataDriven
The Long Tail

Empirical SE 2.0

Challenges





Models



Specs



Code



Traces



Profiles



Tests

Program Data

Process Data



e-mail



Bugs



Effort



Navigation



Changes



Chats

Models

Specs

Code

Deductive Reasoning

Traces

Profiles

Tests

Inductive Reasoning

e-mail

Bugs

Effort

Navigation

Changes

Chats

Models



Specs



Code



Traces



Profiles



Tests



e-mail



Bugs



Effort



Navigation



Changes



Chats

Models

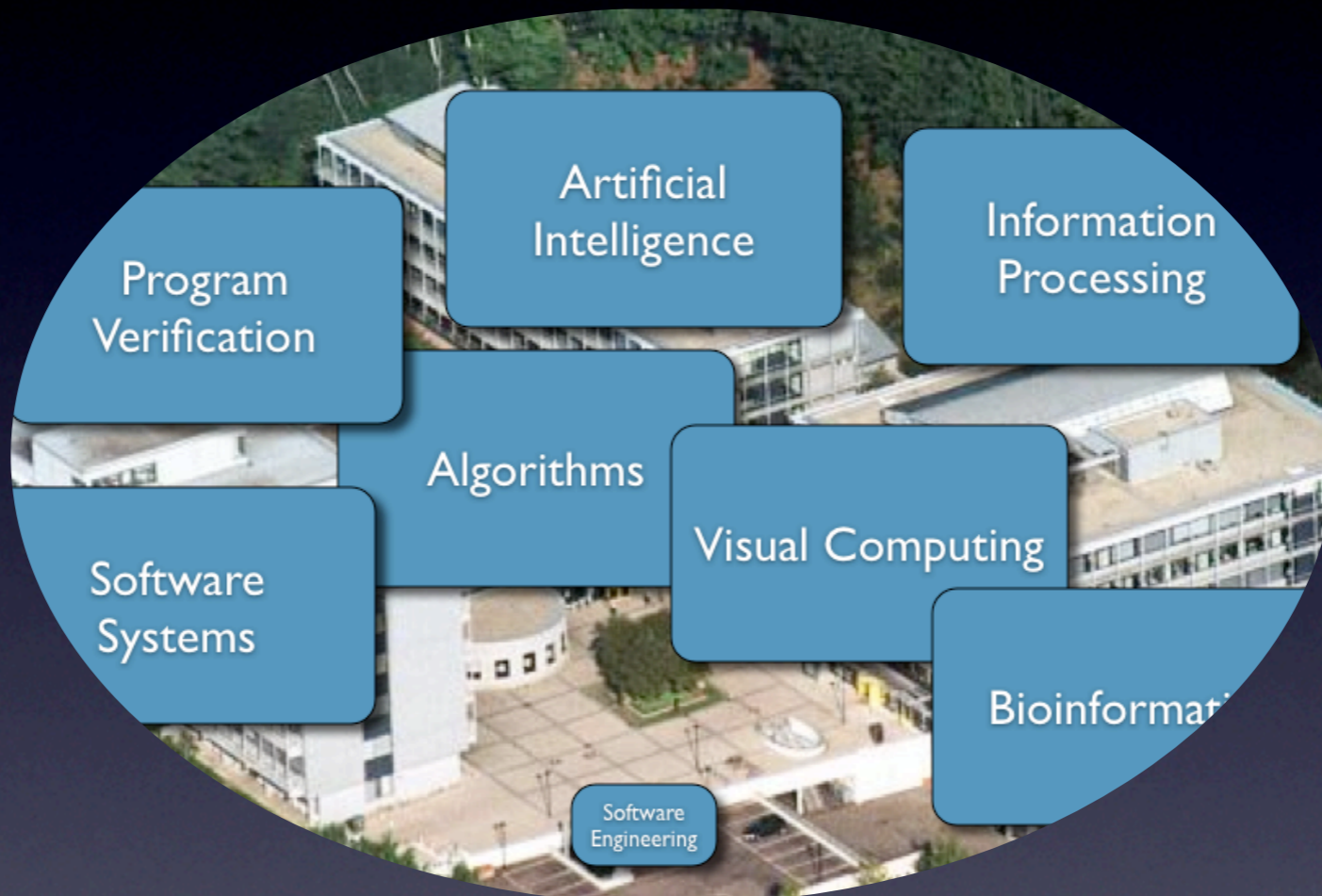
Specs

Code

Traces

Profiles

Tests



e-mail

Bugs

Effort

Navigation

Changes

Chats

Models

Specs

Code

Traces

Profiles

Tests



e-mail

Bugs

Effort

Navigation

Changes

Chats



