


Annie Ying
IBM T.J. Watson Research Center



Building Wisdom from the Past

System Understanding and
Change Patterns

MSR 2004:
International Workshop on Mining Software Repositories

May 25, 2004

© 2004 IBM Corporation

A Chinese proverb...

It's a detriment to yourself
if you don't listen to the elders' sayings.

Does this apply to system understanding?

© 2004 IBM Corporation

2

Outline

- wisdom mined from repository data for system understanding
 - papers from MSR 2004
- a more detailed example
 - our work on mining change patterns on source code (joint with Murphy, Ng, Chu-Carroll)
- challenges

© 2004 IBM Corporation

3

Building maintenance strategies



- [Van Rysselberghe and Demeyer]
 - an approach to find FAC – “similar” changes that have frequently been changed
 - ideas on how to use FAC to build up maintenance strategies

© 2004 IBM Corporation

4

Analyzing usage data

UI re-engineering URL recommendation

- [El-Ramly and Stroulia]
 - an approach interaction patterns by sequential data mining usage data
 - applied the approach to UI re-engineering and URL recommendation

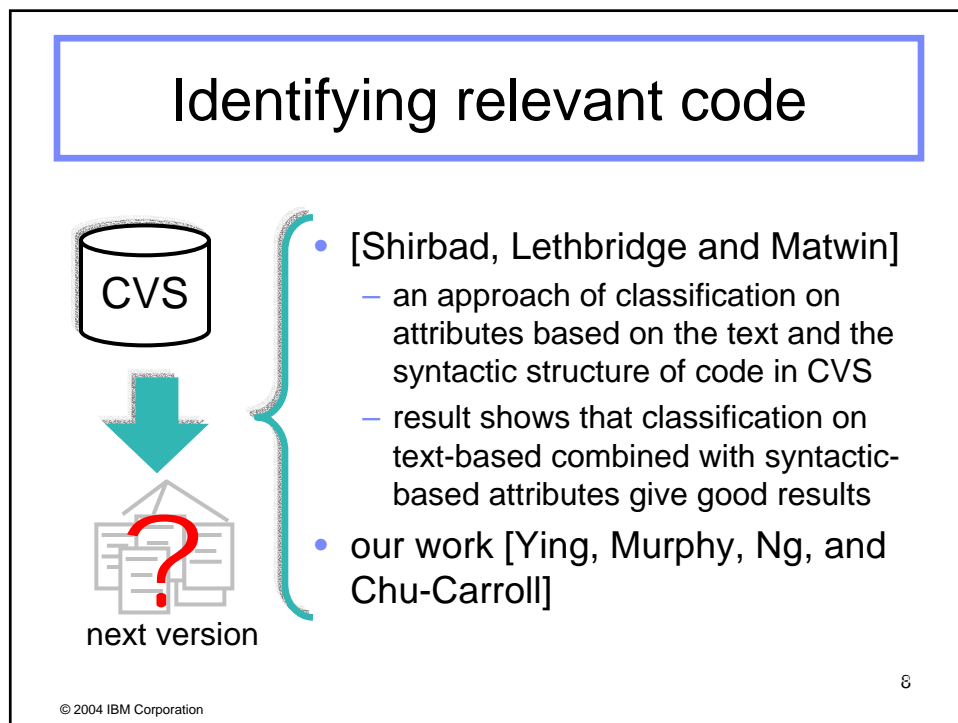
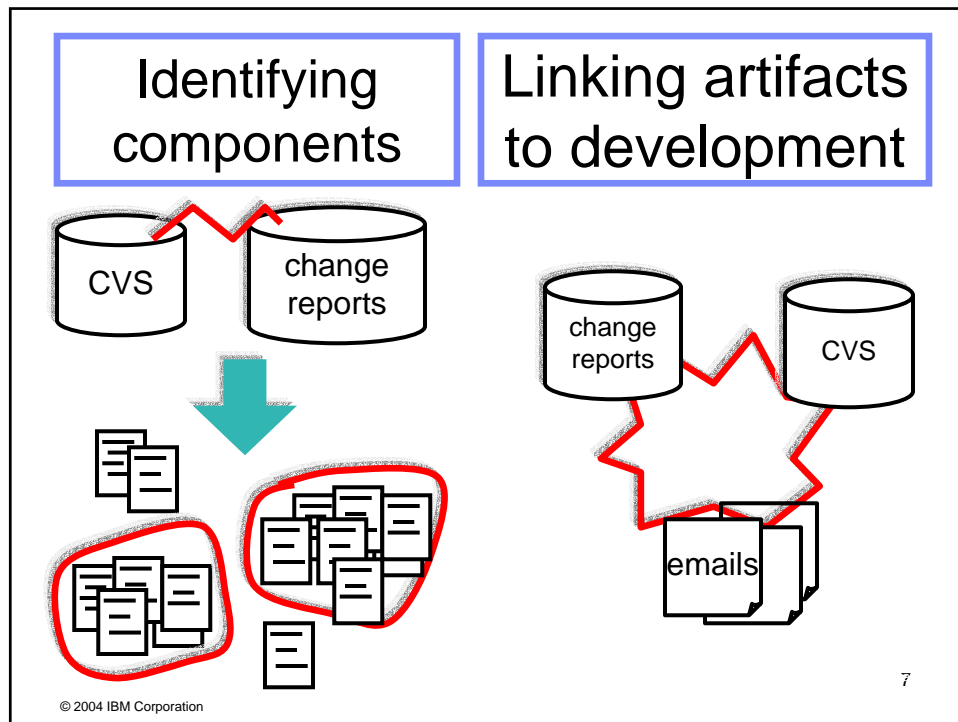
© 2004 IBM Corporation 5

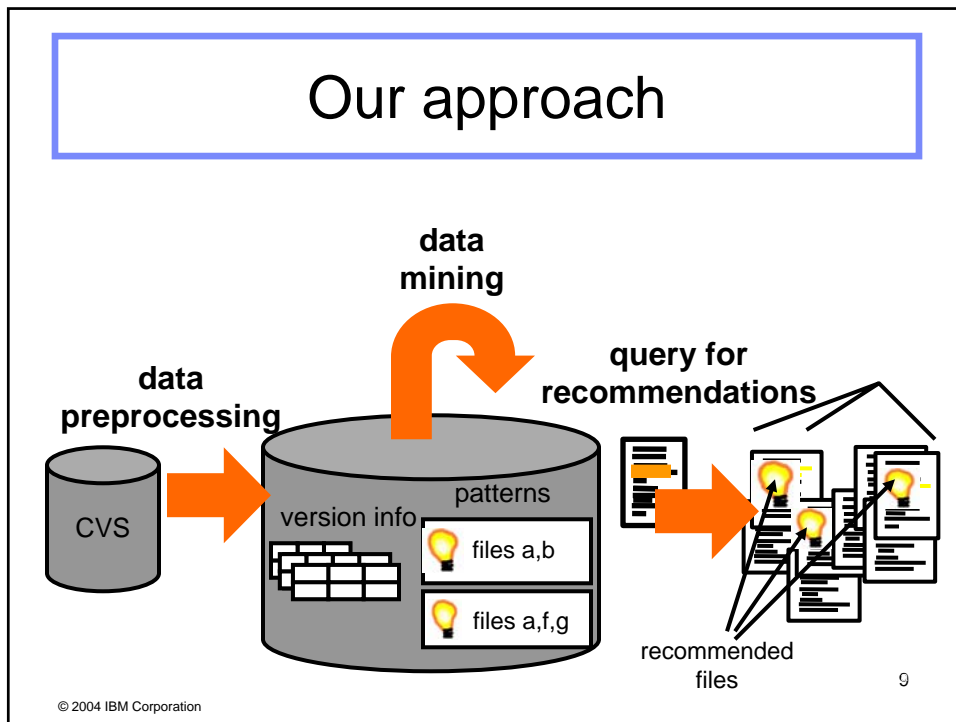
Visualizing Repository data

Tracing code evolution

[Godfrey, Dong, Kapsner, and Zou]

© 2004 IBM Corporation 6





An “interesting” scenario

mozilla.org

Bugzilla Bug 150339:
 huge font crashes X Windows web page with font size > 10,000 pixels

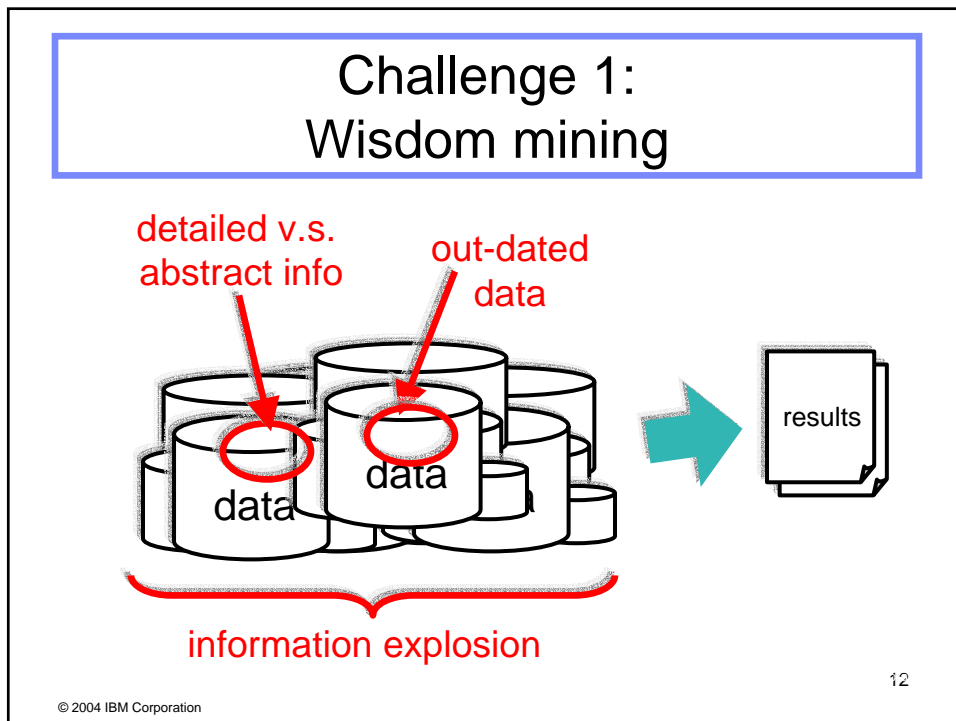
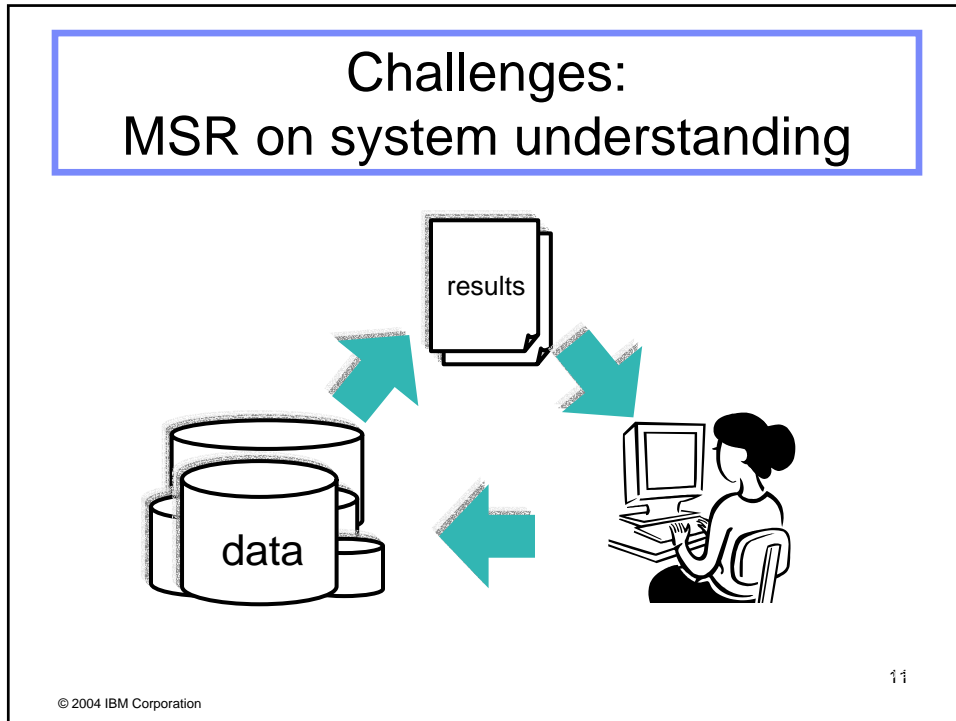
Step 2. enter into address bar: <http://www.adeliesolutions.com/Projects>

Actual Results: X windows crashes
 Created an attachment (id=87422) changed gtk/nsFontMetricsGTK.cpp

This patch limits the size of fonts to twice the display height.

The patch misses the Xlib gfx version should also change xlib/nsFontMetricsXlib.cpp

© 2004 IBM Corporation 10



Challenge 2: “Interesting” wisdom presentation

better UI integration →

```
public void testValue() {  
    Integer[] expected= new Integer[3];  
    for
```

“interesting”?

© 2004 IBM Corporation 13

Challenge 3: Privacy and security

user information exploited by hackers?

well-informed about data collection?

© 2004 IBM Corporation 14

Summary

- Many kinds of wisdom from MSR for system understanding, as seen in the MSR papers and previous work.
- e.g., our work uses source code change patterns and we evaluated on “interesting” source code recommendations.
- Challenges
 - determining the applicability and the level of abstraction from voluminous data
 - presenting “interesting” wisdom
 - privacy and security