

Empirical Software Engineering: No Silver Bullets

Mike Barker
Professor, Nara Institute of
Science and Technology
September 6, 2006



Today's Agenda

- **Introductions**
- Collecting Empirical Data
- Analyzing Empirical Data
- Reporting Empirical Data
- Summary and Conclusions



The EASE Project

NAIST OSAKA UNIVERSITY **ease**



What is the EASE project?

- **E**mpirical **A**pproach to **S**oftware **E**ngineering
- One of the leading projects of the Ministry of Education, Culture, Sports, Science and Technology (MEXT).
- 5 year project starting in 2003.
- Budget: \$2 million US / year.
- Project leader: Koji Torii, NAIST
Sub-leader: Katsuro Inoue, Osaka University
Kenichi Matsumoto, NAIST

<http://www.empirical.jp/english/>

ease Nara Institute of Science and Technology
EASE Project
Empirical Approach to Software Engineering



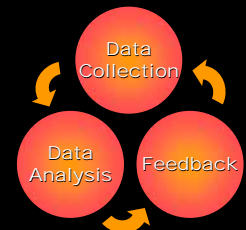
The purpose of the EASE project

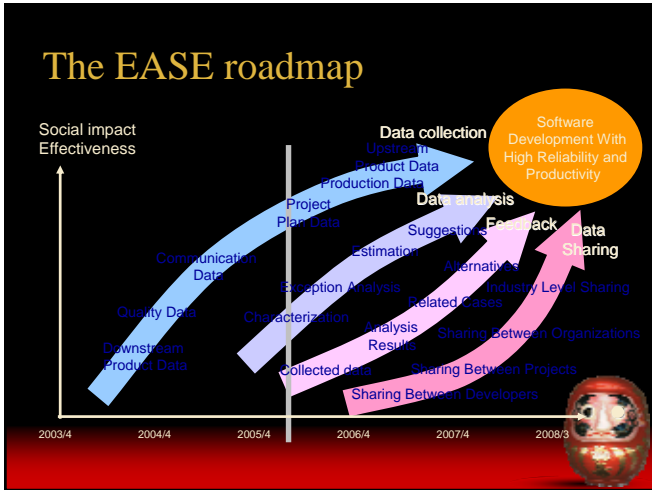
- Achievement of software development technology based on quantitative data
 - Construction of a quantitative data collection system
 - **Result 1: Making of EPM open source**
 - Construction of a system that supports development based on analyzed data
 - **Result 2: EPM application experience**
 - **Result 3: Coordinated cooperation with SEC**
- Spread and promotion of software development technology based on quantitative data to industry sites
 - **Result 4: Activation of the industrial world**



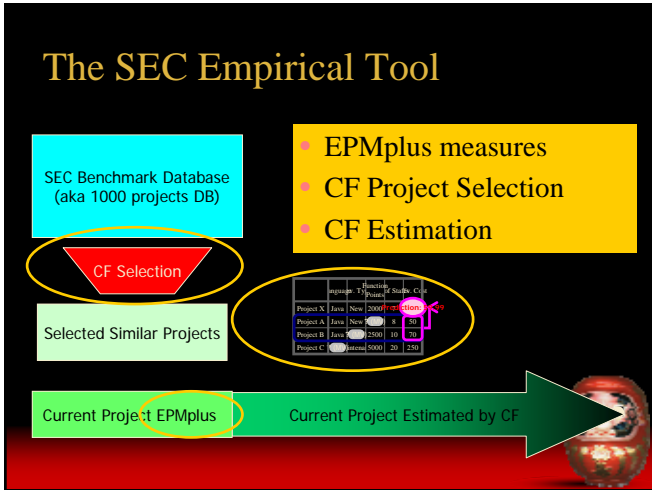
Empirical activities in EASE

- Data collection in real time, e.g.
 - configuration management history
 - issue tracking history
 - e-mail communication history
- Analysis with software tools, e.g.
 - metrics measurement
 - project categorization
 - collaborative filtering
 - software component retrieval
- Feedback to stakeholders for improvement, e.g.
 - observations and rules
 - experiences and instances in previous projects





- ### SEC and EASE
- Software Engineering Center (SEC) Japan is a project under METI
 - Strong cooperation
 - SEC (IPA) has issued an RFP to EASE for a commercial quality toolset using EPM and collaborative filtering. This three year project has the following stages:
 - Develop easy-to-use distribution kit of measurement tools
 - Practical usage in ten trial projects
 - Propose service business using measurement database
-



A Worksheet

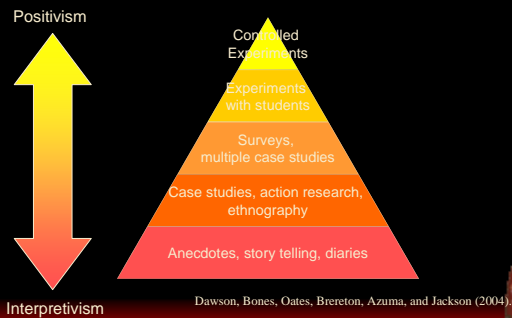
Goals	Measurements	Activities/Tasks	Analysis	Reporting

The GQM Template

What do we want to measure?	
Why do we want to measure it?	
What aspect are we interested in?	
Who wants to know?	
Where are we going to work?	

- ### Today's Agenda
- Introductions
 - Collecting Empirical Data
 - Analyzing Empirical Data
 - Reporting Empirical Data
 - Summary and Conclusions
-

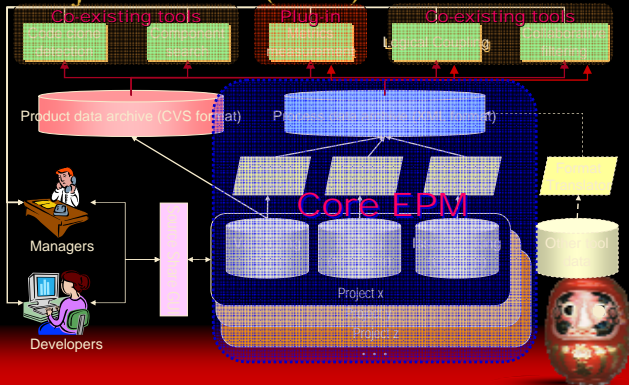
A Pyramid of Empirical Research Types



EPM: the Empirical Project Monitor

- An application supporting empirical software engineering
- EPM **automatically collects** development data accumulated in development tools through everyday development activities
 - Configuration management system: CVS
 - Issue tracking systems: GNATS
 - Mailing list managers: Mailman, Majordomo, FML

Implementation of the Empirical Project Monitor (EPM)



Automated data collection in EPM

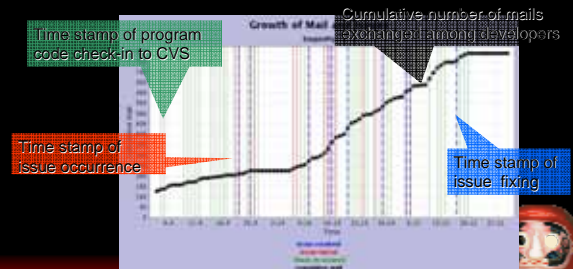
- Reduces the reporting burden on developers
 - without additional work for developers
- Reduces the project information delay
 - data available in real time
- Avoids mistakes and estimation errors
 - uses real (quantitative) data

EPM's GUI



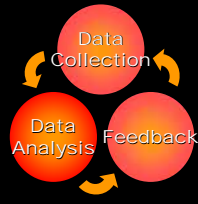
An example of output

- EPM can put data collected by CVS, Mailman, and GNATS together into one graph.

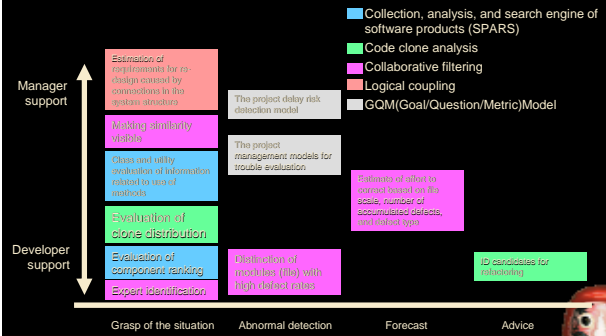


Today's Agenda

- Introductions
- Collecting Empirical Data
- **Analyzing Empirical Data**
- Reporting Empirical Data
- Summary and Conclusions



Analysis targets



Today's Agenda

- Introductions
- Collecting Empirical Data
- Analyzing Empirical Data
- **Reporting Empirical Data**
- Summary and Conclusions



Some Methods of Reporting

- Reports!
- Simulations
- Workshops and Training



Project Replayer An Investigation Tool to Revisit Processes of Past Project

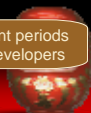
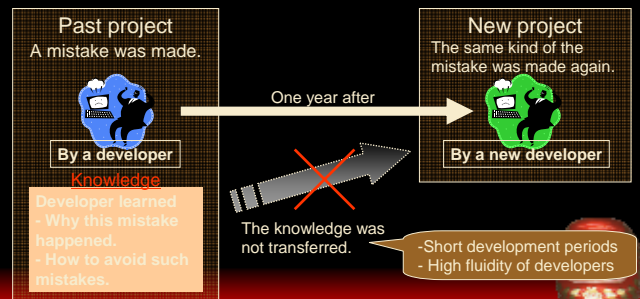
Keita Goto*, Noriko Hanakawa**, and Hajimu Iida*

* Graduate School of Information Science, NAIST, Japan
** Faculty of Management Information, Hannan Univ., Japan



A Problem in Recent Software Development (2/2)

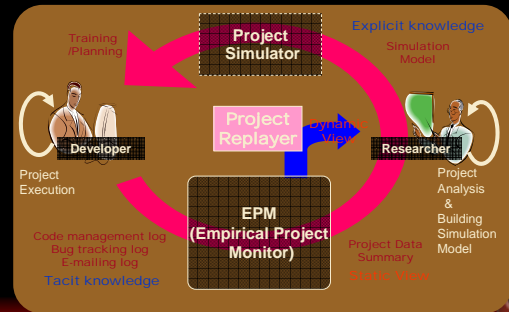
- The same mistakes are repeated in an organization.



Goal and Approach

- Goal
 - Establishing knowledge feedback cycle from past projects to new projects.
- Approach
 - We propose a Knowledge Feedback Cycle (KFC) framework.
 - The following 3 tools play very important roles in the KFC framework.
 - **Empirical Project Monitor (EPM)** : automatically collects software project data. (EPM was made by EASE project Japan)
 - **Project Replayer** : replays history of a past project for researchers to help extracting knowledge.
 - **Project Simulator**: reuses the extracted knowledge to provide estimation for new projects.

Knowledge Feedback Cycle (KFC) Overview



Views provided by Replayer

Event List View shows all events listed in time order.

Graph View shows graph of project progress data.

Developer View shows current status of each member in project

File View shows current status of each source file in project

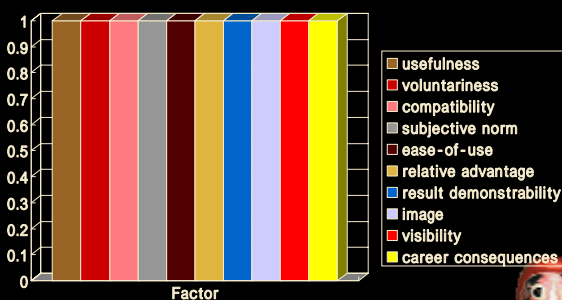
Time control slider

Acceptance of New Methods

1. **Usefulness:** will it help me?
2. **Voluntariness:** do I have a choice?
3. **Compatibility:** how well does it fit with values, needs, and past experiences?
4. **Subjective norm:** do important others think I should use this?
5. **Ease of use (or complexity):** how hard is it to use?
6. **Relative advantage:** how much better is it?
7. **Result demonstrability:** can it show real advantages?
8. **Image:** Does it make me look better?
9. **Visibility:** can others see what I'm doing?
10. **Career Consequences:** what is the long-term payoff for using this?

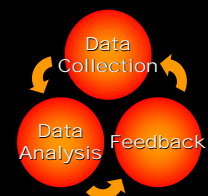
Riemenschneider, Hardgrave, and Davis (2002).

Acceptance of New Methods

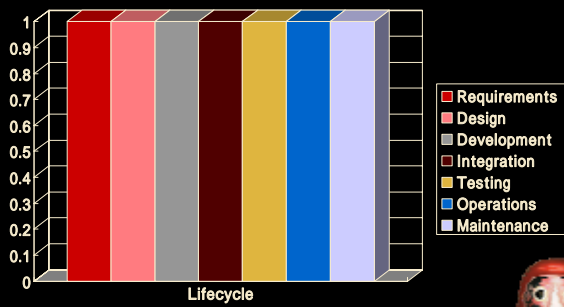


Today's Agenda

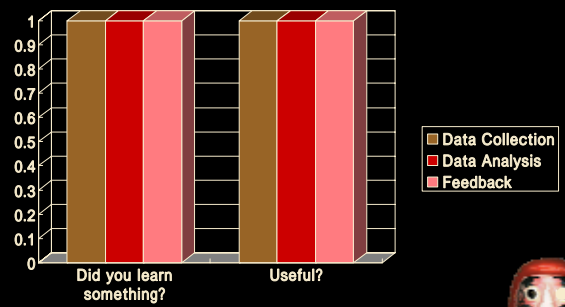
- Introductions
- Collecting Empirical Data
- Analyzing Empirical Data
- Reporting Empirical Data
- **Summary and Conclusions**



Which area most needs empirical methods?



Instant Vote!



Thank you all!

