

# エンピリカルソフトウェア工学の 現状と展望

奈良先端科学技術大学院大学  
文部科学省EASEプロジェクト  
松本健一

日科技連「第24回ソフトウェア品質シンポジウム」  
招待発表 2005/9/9

EASE  
EASE PROJECT

# エンピリカルソフトウェア工学(ESE)

日科技連「第24回ソフトウェア品質シンポジウム」  
招待発表 2005/9/9

EASE  
EASE PROJECT

## ソフトウェア開発の危機

ソフトウェア開発の高信頼化, 高生産化が求められている。

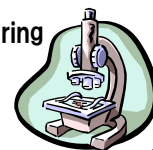
- ソフトウェアのバグは多大な社会的損失を引き起こす。
  - 2005年3月22日 300系新幹線でATCのプログラムミス. 計100編成で機能停止.
  - 2005年2月1日 乗用車8車種でエンジン制御プログラムに不具合. リコールに.
- 開発者一人当たりの開発量は増える一方である。
  - 第3世代携帯電話のプログラム規模はおよそ500万行. 1985年から銀行に導入された第3次オンラインシステムに匹敵する.
  - 1998年から2003年の5年間で, 情報サービス産業の売上高は45%増加したが, 就業者数は6%しか増えていない.



3

## エンピリカルソフトウェア工学(ESE)とは

- ソフトウェア工学に実証性の概念を前提とするアプローチ.
- Empirical=Experiment+Experience.
- 目的に応じた**測定データ(実証・実績データ)**に基づいてソフトウェアの生産性や信頼性の向上を目指す。
  - 科学的根拠に基づいてプロジェクトの改善を行うには必須.
- 国際的なジャーナル, 国際会議, 研究会なども軌道に乗ってきている。
  - International Journal of Empirical Software Engineering
  - International Symposium on Empirical Software Engineering
  - International Software Engineering Research Network (ISERN)



4

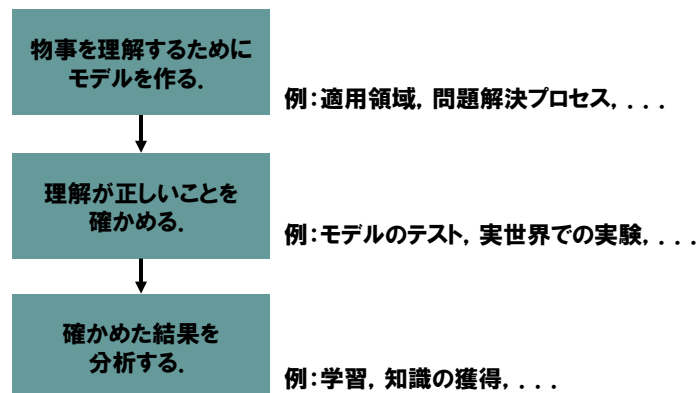
# International Symposium on Empirical Software Engineering

- 第1回:2002年, 日本(奈良)
- 第2回:2003年, イタリア(ローマ)
- 第3回:2004年, 米国(ロサンゼルス)
- 第4回:2005年, オーストラリア(ノーザヘッド)
  - 主なセッションテーマ
    - Project Management, Cost and Effort
    - Maintenance, Reuse
    - Metrics, Testing, Quality, Evaluation
    - Experiments and Experimental Support
    - Software Process
    - Requirements
    - Lessons Learned, Industrial Studies



## ESEの必要性(1) (by Prof. Victor R. Basili)

- 他の工学分野ではごく当たり前のアプローチである。



## 他分野における測定と高信頼化・高生産化

- 自動車業界では、開発現場での測定データに基づいて、現在でも、年間100万件の改善を実施し、年間1000億円の合理化を達成している。



- POSシステム(販売時点管理)

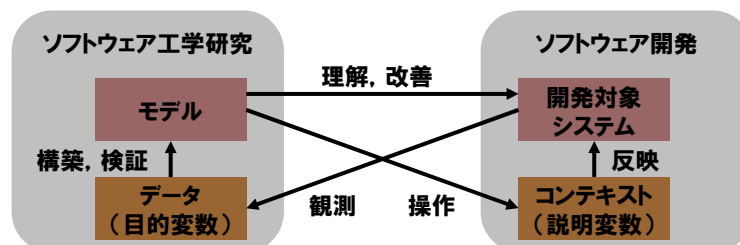
- スーパーマーケットやコンビニエンスストアの店舗で、商品を販売することに商品の販売情報を記録し、集計結果を在庫管理やマーケティング材料として用いるシステム。
- 緻密な在庫・受発注管理ができるようになるほか、複数の店舗の販売動向を比較したり、天候と売り上げを重ね合わせて傾向をつかむなど、他のデータと連携した分析・活用が可能となる。



7

## ESEの必要性(2) (by Prof. Victor R. Basili)

- 開発対象システム(とその開発プロセスや管理)に関する定量的/定性的データの科学的な扱いを可能にし、研究と開発の共生関係を実現する。



8

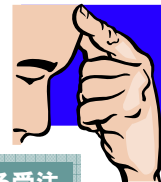
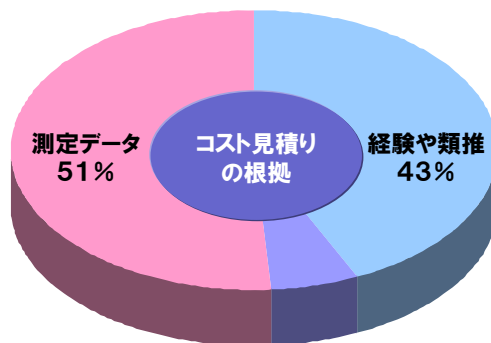
## ESEにおける3つの段階

- 観察(Empirical observations)
  - 実験や調査で確認された事実.
  - 現象を**表現**することができる.
- 法則(Laws)
  - 繰り返される観察.
  - 現象が起こるコンテキストを理解し, 現象を**予測**することができる.
- 理論(Theories)
  - 因果関係.
  - 現象を**説明**することができる.

A. Endres, D. Rombach, A Handbook of Software and Systems Engineering: Empirical Observations, Laws, and Theories, Addison Wesley (2003).  
吉鋪紀子訳, EASEプロジェクト監修, ソフトウェア工学・システム工学ハンドブック: エンピリカルアプローチによる法則とその理論, コンピュータ・エージ社(2005).

9

## 受注ソフトウェア開発における測定の実態



(社)情報サービス産業協会(JISA), 「情報サービス産業における受注ソフトウェア開発の技術課題に関わるアンケート調査」, 2004年.

10

# EASEプロジェクト

日科技連「第24回ソフトウェア品質シンポジウム」  
招待発表 2005/9/9

EASE  
EASE PROJECT

## 概要

- Empirical Approach to Software Engineering.
- 文部科学省リーディングプロジェクト.
  - e-Society基盤ソフトウェアの総合開発.
    - データ収集に基づくソフトウェア開発支援システム.
- 主要組織.
  - 奈良先端科学技術大学院大学, 大阪大学,
  - NTTソフトウェア, 日立製作所, 日立公共システム, SRA先端技術研究所.
- 平成15年度から5年計画で実施中.

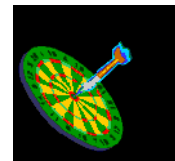
12

日科技連「第24回ソフトウェア品質シンポジウム」  
招待発表 2005/9/9

EASE  
EASE PROJECT

## 目標(とこれまでの成果)

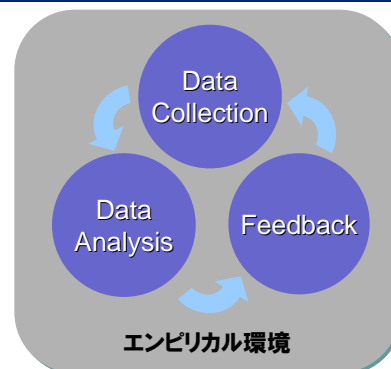
- 実証・実績データに基づくソフトウェア開発技術の実現
  - 実証・実績データ収集システムの構築
    - 成果1: EPMのオープンソース化
  - 分析したデータに基づいて開発作業を支援するシステムの構築
    - 成果2: EPM適用事例
    - 成果3: SECとの連携協力
- 実証・実績データに基づくソフトウェア開発技術の現場への普及, 促進
  - 成果4: 産業界の活性化



13

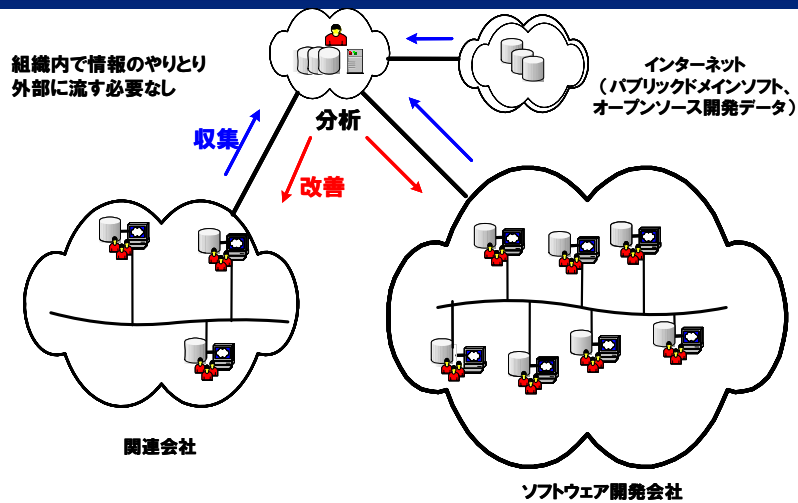
## 実証・実績データに基づく開発支援

- 自動データ収集
  - 構成管理履歴
  - 障害履歴
  - メール履歴
- データ分析
  - メトリクス
  - プロジェクト分類
  - 協調フィルタリング
  - ソフトウェア部品検索
- 生産性, 信頼性改善のためのフィードバック
  - 観察と規則化
  - 過去のプロジェクトの具体的な事例



14

# エンピリカル環境の概念



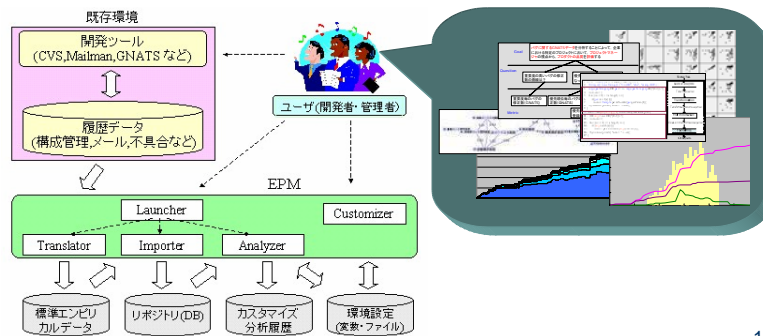
15

日科技連「第24回ソフトウェア品質シンポジウム」  
招待発表 2005/9/9

EASE  
EASE PROJECT

# EPM: Empirical Project Monitor

- エンピリカル環境の一部を実現したシステム。
- ソフトウェア開発プロジェクトデータを自動収集可能。
- オープンソースソフトウェアがフロントエンド。



16

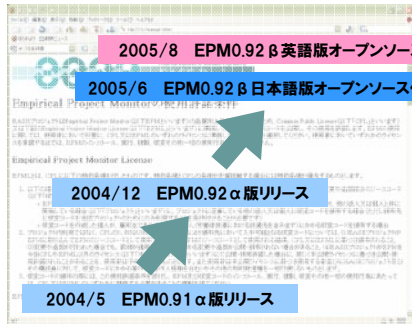
日科技連「第24回ソフトウェア品質シンポジウム」  
招待発表 2005/9/9

EASE  
EASE PROJECT

# 成果1: EPMのオープンソース化

<http://www.empirical.jp/research/epm.html>

- Common Public License (CPL) に特約条項を付した Empirical Project Monitor License (EPML) を作成した。
- 使用者は、CPLかEPMLのいずれかを選択する。



## CPL

米IBMが、IBM Public Licenseをベースに作成したライセンス規定。CPLのソースコードと独自開発のソースコードを組み合わせたソフトを作成して、そのソフトのオブジェクトコードを頒布する場合、CPLのソースコードの部分だけを公開すればよい。

## EPML特約条項

EPMを利用してソフトウェアを開発する共同研究プロジェクトなどで、プロジェクトメンバーにEPMへの変更や追加部分のソースコード(変更コード)を頒布する場合、また、変更コードを作成した個人が、雇用主である法人または個人にかかる変更コードを頒布する場合には、**変更コードを公開したり、頒布先に交付したりする必要はない。**

...

17

日科技連「第24回ソフトウェア品質シンポジウム」招待発表 2005/9/9

EASE  
EASE PROJECT

# 成果2: EPM適用事例

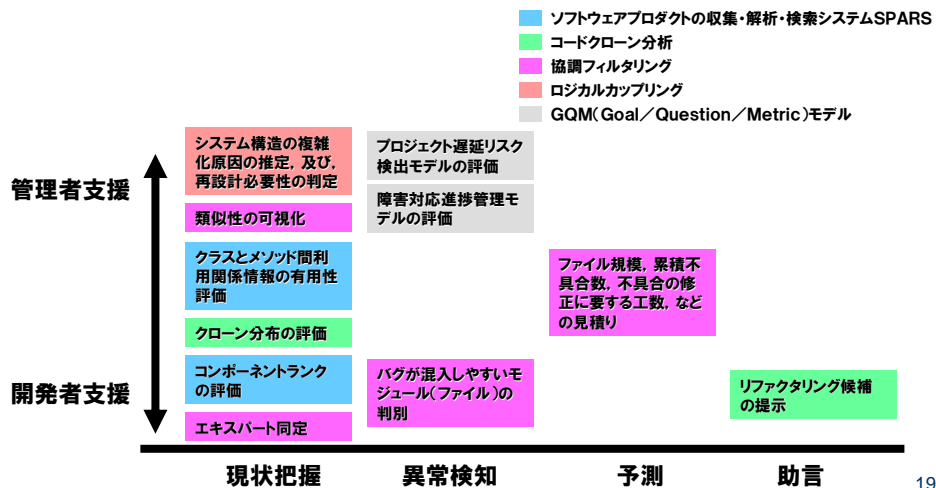
連携企業名	日立公共システムエンジニアリング	三菱スペース・ソフトウェア
適用対象ソフトウェア	自治体庁内業務向けパッケージソフトウェア	ある企業から委託されたソフトウェア
開発言語	Java	Java 他
開発期間	6ヶ月	10ヶ月
開発規模	130,000行	250,000行
EPM導入・運用コスト	25人日	11人日
開発者によるEPM主観評価	自動データ収集は有用。分析結果の提示方式は、ソフトウェアとその開発過程を客観的に見る有用な手段。	データ収集は、開発作業を妨げない。プロジェクトの推移を、収集データで客観的に把握可能。

18

日科技連「第24回ソフトウェア品質シンポジウム」招待発表 2005/9/9

EASE  
EASE PROJECT

## 成果2:EASEデータ分析技術



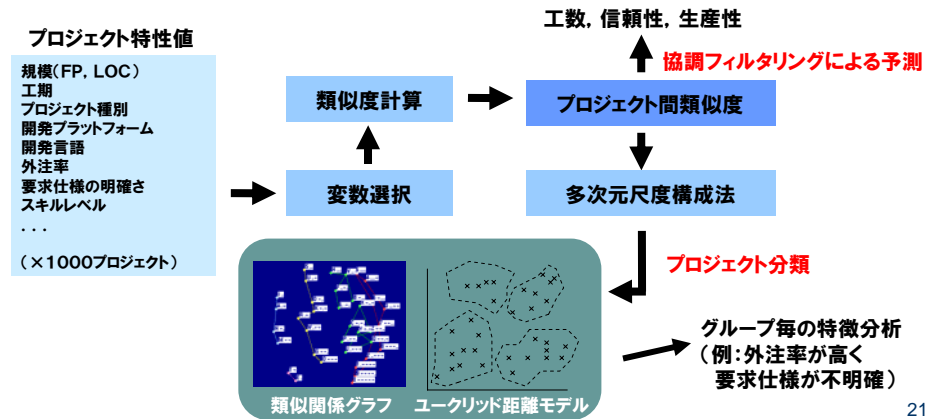
## 成果2:Lessons Learned

- **開発コストに比べ、データ収集コストはごくわずか。**
  - 従来は開発コストの10~20%にも達した。
- **データ分析コストの大幅減少。**
  - 適用事例が増えれば、より組織的な「分析作業プロセスの再利用」が期待される。
- **分析への具体的なニーズの明確化。**
  - 例:「プログラムモジュール別分析や異常値管理が必要である。」
- **収集データと分析結果のみによる現状把握や異常検知。**
  - 対象プロジェクトの知識や経験のないポスドクでも、ソフトウェア構造上の特異点や開発作業の変調を指摘することができた。



## 成果3:SECとの連携協力(1)

### ● 1000プロジェクト定量データの分析



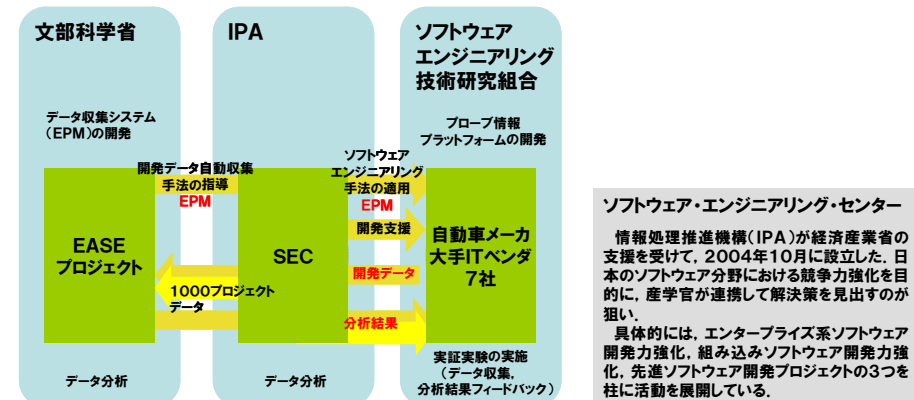
21

日科技連「第24回ソフトウェア品質シンポジウム」  
 招待発表 2005/9/9

EASE  
 EASE PROJECT

## 成果3:SECとの連携協力(2)

### ● 先進ソフト開発プロジェクトへのEPMの適用と分析



22

日科技連「第24回ソフトウェア品質シンポジウム」  
 招待発表 2005/9/9

EASE  
 EASE PROJECT

## 成果3:GNATS標準データ

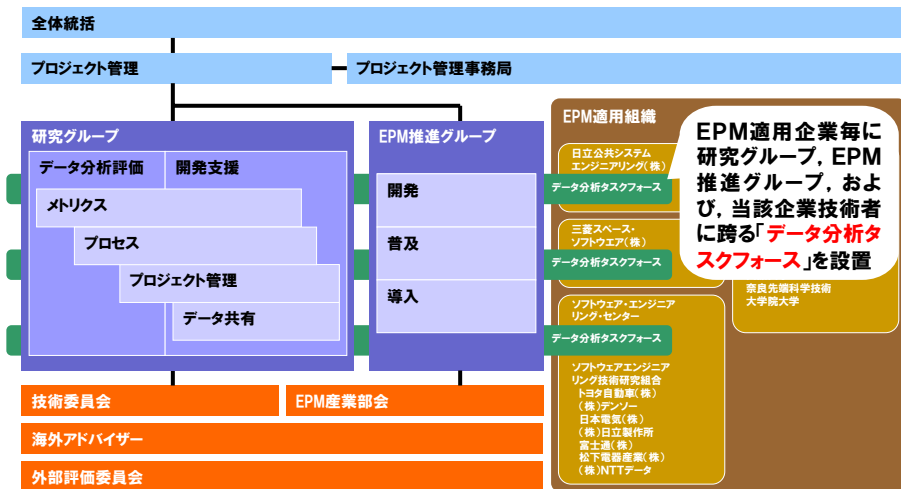
- 収集必須データ
  - 対応状況, 起票分類, 責任者, 発見日, テスト対象, 発見工程, 発見バージョン, 問題内容, 再現度, 重要度, 優先度, 問題原因[詳細], 混入工程, 問題を本来発見すべき工程, 修正工数, 修正内容, 完了日
  
- 備考
  - 「原因(Fault)の分析」を目的とする。
  - 可能な限り, 開発作業を阻害しない運用とする。
    - 問題票の再起票や分割などは行わない。

23

日科技連「第24回ソフトウェア品質シンポジウム」  
招待発表 2005/9/9

EASE  
EASE PROJECT

## プロジェクト体制(平成17年度)



24

日科技連「第24回ソフトウェア品質シンポジウム」  
招待発表 2005/9/9

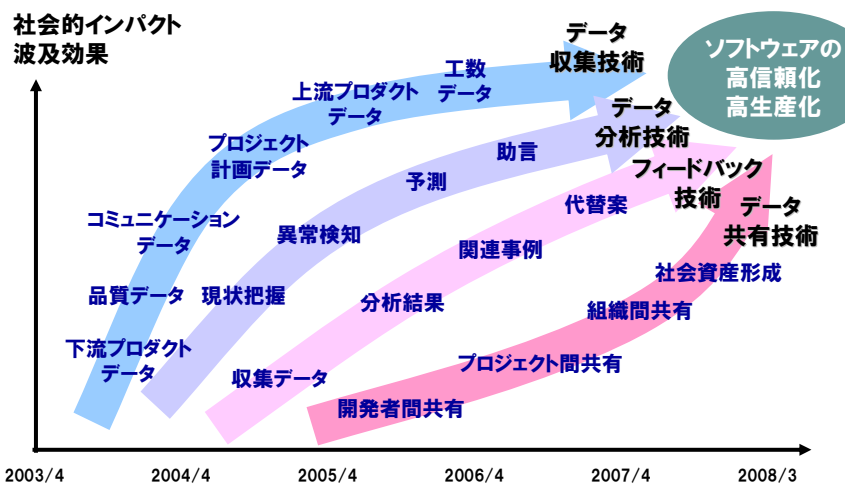
EASE  
EASE PROJECT

# 技術的課題

日科技連「第24回ソフトウェア品質シンポジウム」  
招待発表 2005/9/9

EASE  
EASE PROJECT

## 4つの基盤技術(EASEロードマップ)



日科技連「第24回ソフトウェア品質シンポジウム」  
招待発表 2005/9/9

EASE  
EASE PROJECT

## 自動データ収集

- データ収集コストをできるだけ小さくする。
  - 収集コストがプロジェクトの利益率を上回るようでは普及しない。
- ソフトウェア開発者・管理者の負荷を出来るだけ小さくする。
  - 負荷が大きいと、「失敗するプロジェクトはごく一部であり、自身のプロジェクトは失敗しない」と考えている開発者・管理者の協力は得られない。
- 出来るだけ共通化する。
  - フロントエンドにオープンソースソフトウェアを採用する。
  - データフォーマットを統一する。
- 「プロジェクト計画データ」、「上流プロダクトデータ」、「工数データ」へと収集対象を広げていく。

27

## 探索的データ分析

- 企業の多様なニーズに応える。
- プロジェクト管理, 特に, 問題解決プロセスに関わる暗黙知の表出を手助けする。
  - 仮説やモデルに基づく分析が可能となるまでにはもう少し時間が必要である。
- まずは「現状把握」や「異常検知」あたりから。
- 2つのアプローチ
  - Goal-driven Approach
    - 測定情報モデルやGoal/Question/Metricsモデルなどに基づいて、分析目的と計測データの対応付けが重要となる。
  - Data-driven Approach
    - 大量の蓄積データに対するデータマイニング技術やデータ可視化技術などが重要となる。

28

## リアルタイムフィードバック

- データ収集からフィードバックまでの遅延をできるだけ小さくする。
  - プロジェクト終了後では、その効果や利用範囲は限定される。
- 当事者である開発者や管理者に直接提示する。
  - 「異常検知」時などは、開発者や管理者からのリクエストがなくても提示する。
- 収集データや分析結果だけでなく、関連事例や代替案(問題解決案)などもあわせて提示する。

29

## 匿名性のあるデータ共有

- 個人情報の保護や匿名性を確保しつつ、組織横断的にデータを集約し、分析することのできるシステムの実現を目指す。
  - 収集データを(無理やり)一箇所に集積したりはしない。
  - データを収集した組織がそれぞれ保持することを許す。

30

## Point Of Development (開発時点管理)

- POSシステム(販売時点管理)の概念(の一部)をソフトウェア開発に導入する。
  - 開発作業(のある基本単位)を実施することに作業情報を記録し、その集計結果を、プロジェクト管理や組織運営に利用する。
  - 緻密な開発管理ができるようになるだけでなく、プロジェクトや組織の間で開発状況を比較したり、組織や社会が持つ他のデータと連携した分析・活用が可能となる。



31

## まとめにかえて

- エンピリカルソフトウェア工学がもたらすもの
  - 安心・安全な社会の実現
    - 故障しない基幹システムの実現
    - ソフトウェアトレーサビリティの確立
    - バグをなくすことで、コンピュータウィルスの90%をシャットアウト
  - より高度な人材
    - ソフトウェアアナリスト
    - 組織横断的開発管理者
- 参考文献等
  - EASEプロジェクトホームページ <http://www.empirical.jp/>
  - 松本 健一:“エンピリカルソフトウェア工学の現状と展望:SELが遺した13の教訓”, SEC Journal, Vol.1, No.2, pp.6-13 (平17-4).

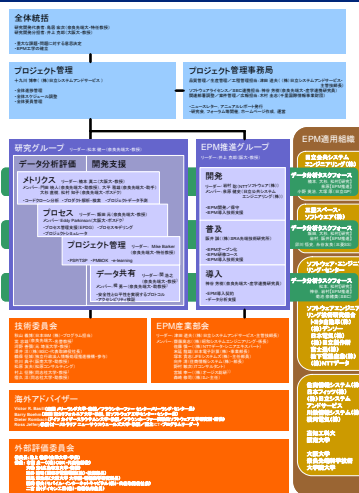
32

# EASEファクトデータ(2005年6月時点)

日科技連「第24回ソフトウェア品質シンポジウム」  
招待発表 2005/9/9



## プロジェクト体制図(平成17年度:詳細)



- **メンバー数**
  - 大学 24名(学生を除く)
  - 企業 27名(企業内協力者を除く)
- **15年度からの変更点**
  - プロジェクト管理機能の強化。
  - EPM推進グループに導入担当を追加。
  - 産学にまたがる「データ分析タスクフォース」の設置。
  - 企業技術者からなる「EPM産業部会」の設置。
  - 「外部評価委員会」の設置。

日科技連「第24回ソフトウェア品質シンポジウム」  
招待発表 2005/9/9



## EPM適用企業数(共同研究契約締結企業数)

- EPM適用完了, データ分析中: 2社
  - 日立公共システムエンジニアリング, 三菱スペース・ソフトウェア
- EPM適用中: 7社
  - トヨタ自動車, デンソー, 日本電気, 日立製作所, 富士通, 松下電器産業, NTTデータ
- EPM適用準備中: 4社
  - 住商情報システム, 日立システムアンドサービス, 川鉄情報システム, 横河電気



35

## 学術的成果(1)

- 研究論文 93編
  - 内訳: 学術論文誌掲載論文 29編, 査読つき国際会議発表 64編
    - ポストクと企業出向者による共著論文掲載
      - 大平雅雄, 横森勳士, 阪井誠, 岩村聡, 小野英治, 新海平, 横川智教, “ソフトウェア開発プロジェクトのリアルタイム管理を目的とした支援システム”, 電子情報通信学会論文誌D-I, Vol.J88-D-I, No.2, pp.228-239, Feb. 2005.
    - ソフトウェア工学分野で最も権威のある学術論文誌への論文掲載
      - K. Inoue, R. Yokomori, T. Yamamoto, M. Matsushita, S. Kusumoto: “Ranking Significance of Software Components Based on Use Relations,” *IEEE Transactions on Software Engineering*, Vol.31, No.3, pp.213-225, Mar. 2005.



36

## 学術的成果(2)

- 研究論文(つづき)

- 産学連携論文特集号への論文掲載

- 角田雅照, 大杉直樹, 門田暁人, 松本健一, “協調フィルタリングを用いたソフトウェア開発工数予測方法”, 情報処理学会論文誌, Vol.46, No.5, pp.1155-1164, May 2005.

- 社会人学生論文特集号への論文掲載

- 田中康, 飯田元, 松本健一, “成果物間の関連に着目した開発プロセスのモデルPReP”, 情報処理学会論文誌, Vol.46, No.5, pp.1233-1245, May 2005.

- 知的所有権 3件

- 受賞 8件

- 研究発表 25件

- 講演 19件 (国際シンポジウムなどでの招待講演 7件)



37

## メディアへの情報発信

- 雑誌掲載 9件

- “特別対談:ソフトウェア・エンジニアリングが日本のIT産業を変える”, コンピュートピア, 2005年6月号.

- 新聞報道 5件

- “ソフト開発:定量化し工程管理”, 日刊工業新聞, 平成15年11月14日朝刊.

- ウェブ記事への掲載 2件

- 有賀貞一のやぶにらみシステム論, Vol.131, “文部科学省「EASE」vs.経産省「SEC」”, 月間ソリューションIT, 1月号.

[http://www3.ric.co.jp/sol/contents/sol\\_0401/yabu04\\_01.html](http://www3.ric.co.jp/sol/contents/sol_0401/yabu04_01.html).

- 学術著書での引用 1件

- 学術書(和訳本)監修 1件



38

## 産業界への情報発信

- **ウェブサイト開設**
  - 訪問者数のべ80,865名, ページ閲覧数のべ266,238ページ.
- **EASE国際フォーラム開催**
  - 2003年11月, 会場:東京国際交流館, 参加者約410名.
- **エンピリカルソフトウェア工学研究会開催**
  - 年5回, 会場:東京・キャンパス・イノベーション・センター, 参加者各回約50名.
- **ツール出展**
  - 関西オープンソース2004, 来場者数約2000名.
- **技術指導**
  - 会場:大阪・千里 エンピリカルソフトウェア工学ラボ, 計8回.



39

日科技連「第24回ソフトウェア品質シンポジウム」  
招待発表 2005/9/9

EASE  
EASE PROJECT

## 海外の類似研究との比較

- EASEプロジェクトで提案するような, 実用的な自動データ収集/分析ツールの開発, データや分析結果のリアルタイムフィードバックや共有の検討は行われていない.
  - ドイツフランフォーファーIESE (Institute for Experimental Software Engineering)
  - 米国 COCOMO-II
  - 米国CeBASE (NSF Center for Empirically Based Software Engineering)
  - 豪州 NICTA (National Information and Communication Technology Australia)
- EASEプロジェクトでは, これら類似プロジェクトの研究者との共同研究を積極的に行っている.



40

日科技連「第24回ソフトウェア品質シンポジウム」  
招待発表 2005/9/9

EASE  
EASE PROJECT