

ソフトウェア開発データの分析に基づくエンジニアリング研究の推進 ～収集データの活用に向けた IPA/SEC の取組み～

Promotion of Software Engineering Research Based on Software Development Projects Data: Activities on Utilization of Collected Data in IPA/SEC

山下 博之¹ 秋田 君夫¹
Hiroyuki YAMASHITA Kimio AKITA

1. はじめに

独立行政法人情報処理推進機構ソフトウェア・エンジニアリング・センター(IPA/SEC)では、2004年の発足以来毎年、ソフトウェア開発プロジェクトの基本要素(規模、工数、工期や信頼性等)の実績データをプロジェクト情報と共に収集し、それらの分布や要素間の関係等の分析を行った上で、「ソフトウェア開発データ白書」として発行してきている。これらのデータを有効利用することにより、ソフトウェアの品質や信頼性向上につながる新たな分析手法の発見やそれに伴う開発手法へのフィードバック等が期待される。本稿では、IPA/SECに蓄積されているソフトウェア開発データの一層の活用に向けた取組みを紹介する。

2. ソフトウェア開発データ

2.1 IPA/SEC の取組み

「ソフトウェア開発データ白書」には、工数と工期、規模との関係、規模と発生不具合数との関係、生産性等のデータを収録している。2010-2011年版^[1]は、23社から提供を受けた2584件のプロジェクトデータを含む。プロジェクトマネージャは、表1に示すように、これら各種のデータを組み合わせたり自プロジェクトのデータを重ねたりすることにより、生産性や工期、品質、工程等の情報をさまざまな切り口で見ることができ、プロジェクト計画立案時や評価時の参考にすることができる。また、これらのデータの分析により、ソフトウェア開発の傾向を把握し、各種施策に反映することができる。

表1 ソフトウェア開発データの内容と用途

項番	区分	入力項目	狙い
1	生産性水準	規模(FP/SLOC)、工数、業種、言語、アーキテクチャ(システム構成)	工数当たりの開発規模の分布において、自社プロジェクトの位置を、規模の大きさ、業種、言語、アーキテクチャ等の層別に評価する。
2	工期水準	工数、工期	「工期は工数の三乗根に比例する」傾向を活用し、自社プロジェクトの工数と工期が適正な状態にあるかどうかを評価する。
3	品質水準	検出バグ数(結合、総合テスト分)不具合数(稼働後)規模(FP/SLOC)、言語	開発時の検出バグ数および稼働後の発生不具合密度の分布において、自社プロジェクトの位置を工程、言語等の層別に評価する。
4	工程比率	工程別の工数と工期	工程別の工数比の分布および工期比の分布より、標準的な工数比および工期比を参考に、自社プロジェクトの工程別工数、工期の配分をシミュレートする。

さらに、将来的には、対象のソフトウェア/システムの特長や開発組織・プロジェクトのプロファイル、開発方法の情報等と共に分析することにより、システム障害の要因究明や再発予防方法の発見が期待される^[2]。

2.2 取組みの位置づけ

IPA/SECの取組みは、図1に示す、ベンチマーキングの枠組みに関する国際標準(ISO/IEC 29155-1)^[3]に基づいている。同図中の各要素との対応は、次の通りである:

- ① データの収集・蓄積と機密室でのセキュアな管理
- ② データ白書の出版、大学との共同研究、等
- ③ プロジェクト診断ツール、データ活用ガイド、等
- ④ データ測定・提供支援ツール

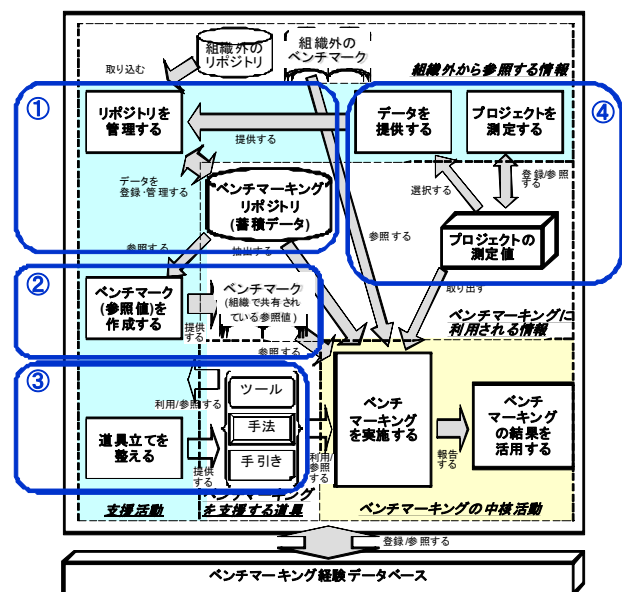


図1 ISO/IEC 29155-1のベンチマーキング枠組みモデル

3. ソフトウェア開発データの活用

3.1 ソフトウェア開発への活用

3.1.1 取組み

IPA/SECでは、入力した利用者のプロジェクトデータを、画面上に表示したデータ白書掲載のグラフ上に重ねてプロットできる「プロジェクト診断支援ツール」を公開し、他社データとのベンチマーキングを可能にしている。本ツ

¹ 独立行政法人情報処理推進機構ソフトウェア・エンジニアリング・センター、IPA/SEC

ルにより、ユーザ・ベンダ間での目標設定、プロジェクト状況の把握、プロジェクト特性や立ち位置の確認を、定量的かつ客観的に行うことができる。

さらに、「ソフトウェア開発データ白書」の具体的な利用方法・事例と、このような定量データ一般の具体的な活用方法・事例を説明した、「データ白書の見方と定量データ活用ポイント」^[4]を発行している。

3.1.2 事例

ツールを用いた活用事例としては、図2に示すように、プロジェクト全体の工数と工期の分布図を用いて、自社プロジェクトの傾向を分析し、プロジェクト計画及びプロジェクト遂行中の参考値とする、工数と工期についての自社基準(目安)を策定する、といった活用がある。

別の活用事例としては、図3に示すように、工程別工数比率の箱ひげ図を用いて、PMによる開発計画策定局面において、プロジェクトで計画した工程別の工期、工数の配分の適正度を評価し、計画実現の達成度、リスク度を認識することができる。

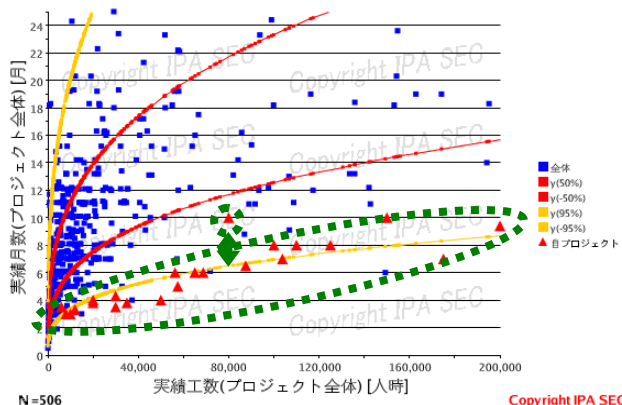


図2 工数・工期の妥当性確認に活用する一例

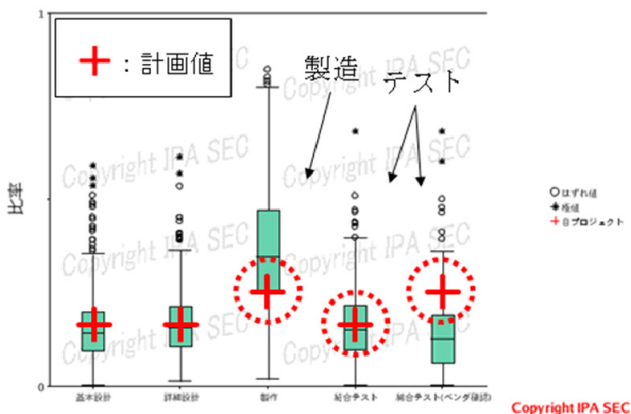


図3 工程別工数比率による工期配分妥当性確認の一例

3.2 新たな知見を目指した活用

3.2.1 取組み

IPA/SEC に収集・蓄積されているソフトウェア開発データについて、高度な分析手法の研究等、一層の有効活用を図るため、大学等の公的研究機関に対して一定の守秘条件

の下で開示する仕組みを整備している。この仕組みに基づいて、現在、数件の研究が進行中である。

3.2.2 事例

一つの研究は、「ヒューマン・ファクタが開発にどう影響するか」というテーマでの定量データの分析である。

別の研究は、規模と工数間の関係分析である。一次分析の結果、総工数はファンクションポイント(FP)規模の1.156乗に比例すること、開発言語によらずソースコード行数(SLOC)規模のべき乗に比例するべき乗の値は1より小さいこと、が分かり、現在、詳細解析中である。

4. データを活用した共同研究の仕組み

データの開示は、データ提供元と IPA/SEC との機密保持契約の範囲内で、研究機関と IPA/SEC との「共同研究」の枠組みで行う。ただし、研究の遂行においては、研究機関の意向を最大限尊重する。

原則として研究機関からの申し出をトリガに、データ利用条件の確認の後、研究企画書を IPA/SEC に提出して頂く。研究内容が IPA/SEC の事業と整合すれば、研究計画の調整を経て「共同研究」の合意に至る。その後、機密保持に関する誓約書の提出の後、データを引き渡す。

主要な利用条件は、次の通りである：

- (1) IPA/SEC 事業への展開
 - ・研究結果が SEC の定量的管理の普及に関連する事業に資すること
 - ・研究成果が SEC で利用できること
- (2) 成果の公表
 - ・成果は学会等公共性のあるところに公開され、多くの関係者が平等に成果にアクセス可能であること
- (3) データの匿名性確保
 - ・プロジェクトの匿名性が担保された状態で公表されること
 - ・研究成果公表前に SEC による確認を必ず受けること
- (4) データのセキュリティ
 - ・データがセキュリティの確保される適切な環境(研究室、利用機器、体制等)で管理されること

5. おわりに

IPA/SEC は、ソフトウェア開発データの分析に基づく研究が一層促進され、産業界にとって有益な成果が生み出されることを期待している。その結果として、企業から開発データを提供して頂くことが困難になってきている状況^[2]がいくらかでも打開されることを願っている。

大学等のみなさんには、是非、ご利用を検討頂きたい。

参考文献

[1] IPA ソフトウェア・エンジニアリング・センター著作・監修, “ソフトウェア開発データ白書 2010-2011”, IPA (2010).
 [2] 山下博之, “社会の“インテリジェンス”活用推進に向けた考察”, SEC journal, No.23 (2011).
 [3] ISO/IEC 29155-1:2011, Systems and software engineering -- Information technology project performance benchmarking framework -- Part 1: Concepts and definitions
 [4] “ソフトウェア開発データ白書の見方と定量データ活用ポイント”, http://sec.ipa.go.jp/reports/20110331_2.html (2011).