

統計指標に基づく
品質マネジメント実践集

～ベンチマーキングによる
信頼性向上のヒントと具体例～

2016年 7月

独立行政法人情報処理推進機構（IPA）

技術本部 ソフトウェア高信頼化センター（SEC）



目次

1.	はじめに	1
1.1.	定量的管理の勧め	3
1.2.	ベンチマーキングの基本的考え方	4
1.3.	本書の位置付け	7
1.4.	本書の特長	9
1.5.	公開ベンチマーク（白書等）を参考にしたベンチマーキングの勧め	10
1.6.	公開ベンチマーク（白書等）利用上の留意事項	14
2.	ベンチマーキングの主なシーン	15
2.1.	開発者用のベンチマーキングの主なシーン	15
2.2.	開発組織のマネジャー層用のベンチマーキングの主なシーン	17
2.3.	3章以降の見方について	19
3.	開発者用のベンチマーキング	20
3.1.	信頼性目標及び品質保証プロセスの妥当性評価	20
3.1.1.	ベンチマーキング方法の手順例	20
3.1.2.	白書等を用いた具体例	23
3.1.3.	IT企業での具体例1	27
3.1.4.	IT企業での具体例2	28
3.2.	開発総工数及び生産性目標の妥当性評価	29
3.2.1.	ベンチマーキング方法の手順例	29
3.2.2.	白書等を用いた具体例	30
3.2.3.	IT企業での具体例1	34
3.2.4.	IT企業での具体例2	38
3.2.5.	IT企業での具体例3	40
3.2.6.	IT企業での具体例4	42
3.3.	成果物量及び単位成果物量あたりの工数の妥当性評価	43
3.3.1.	ベンチマーキング方法の手順例	43
3.3.2.	白書等を用いた具体例	45
3.3.3.	IT企業での具体例1	49
3.4.	工期の妥当性評価	53
3.4.1.	ベンチマーキング方法の手順例	53
3.4.2.	白書等を用いた具体例	54
3.4.3.	IT企業での具体例1	58
3.5.	工程別の工数比率の妥当性評価	61
3.5.1.	ベンチマーキング方法の手順例	61
3.5.2.	白書等を用いた具体例	62
3.6.	工程別の工期比率の妥当性評価	64
3.6.1.	ベンチマーキング方法の手順例	64
3.6.2.	白書等を用いた具体例	65
4.	開発組織のマネジャー層用のベンチマーキング（プロジェクト・マネジメントの改善）	67
4.1.	品質保証プロセス関連標準類の見直し（品質マネジメント関連）	67

4.1.1.	ベンチマーキング方法の手順例	67
4.1.2.	白書等を用いた具体例	69
4.1.3.	IT 企業での具体例 1	77
4.1.4.	IT 企業での具体例 2	80
4.2.	テスト工程完了評価関連標準類の見直し（品質マネジメント関連）	84
4.2.1.	ベンチマーキング方法の手順例	84
4.2.2.	白書等を用いた具体例	85
4.3.	ユーザの関与（ステークホルダー・マネジメント関連）	88
4.3.1.	ベンチマーキング方法の手順例	88
4.3.2.	白書等を用いた具体例	89
4.4.	成果物スコープの明確化（スコープ・マネジメント関連）	91
4.4.1.	ベンチマーキング方法の手順例	91
4.4.2.	白書等を用いた具体例	92
5.	開発組織のマネジャー層用のベンチマーキング（組織の改善）	94
5.1.	組織体制の整備	94
5.1.1.	ベンチマーキング方法の手順例	94
5.1.2.	白書等を用いた具体例	95
5.2.	信頼性変動要因分析と信頼性向上に向けた重点強化領域の特定	97
5.2.1.	ベンチマーキング方法の手順例	97
5.2.2.	白書等を用いた具体例	98
5.2.3.	IT 企業での具体例 1	100
5.3.	生産性変動要因分析と見積り妥当性評価方法の見直し等	102
5.3.1.	ベンチマーキング方法の手順例	102
5.3.2.	白書等を用いた具体例	103
5.4.	業種等のドメイン別マネジメント	106
5.4.1.	ベンチマーキング方法の手順例	106
5.4.2.	白書等を用いた具体例	107
5.4.3.	IT 企業での具体例 1	108
5.5.	信頼性／生産性等の経年変化に基づく組織改善計画	109
5.5.1.	ベンチマーキング方法の手順例	109
5.5.2.	白書等を用いた具体例	110
5.5.3.	IT 企業での具体例 1	112
6.	関連ツールの紹介	113
6.1.	定量データに基づくプロジェクト診断支援ツール	113
6.2.	グラフデータ	119
<謝辞>	122
<参考文献>	122

1. はじめに

「統計指標に基づく品質マネジメント実践集」(以下、本書)は、ソフトウェアの定量的管理(特にベンチマーキング)を通じて信頼性向上を図るための、統計指標に基づく定量的な品質マネジメントの実践事例集である。

本書は、種々のベンチマーキングのシーンに沿って、信頼性向上に向けた品質マネジメントのノウハウと具体例を提示している。具体例については、IPA/SEC が提供している公開ベンチマークである「ソフトウェア開発データ白書」及び「ソフトウェア開発データが語るメッセージ」(以下、白書等)を用いた19件の具体例と、IT企業での13件の具体例を掲載している。各組織における定量的な品質マネジメントのヒントとして参考にして頂けると幸いである。

(注1) ベンチマーキングについて

「ベンチマーキング」は「定量的管理」の一種であり、本書では次のような特長を有する定量的管理を指す。

- ◇ 良い成績を収めているプロジェクト群と比較し、それらのやり方(開発プロセス、マネジメント・プロセス、組織の特性等)を参考にして、自組織の業務改善及び組織の改善を進める。
- ◇ ポストプロセス計測データ(完了プロジェクトの実績データ)を用いた定量的管理であり、インプロセス計測データ(開発プロジェクト実行中のデータ)を用いた定量的管理は含まない。

図 1-1 に、定量的管理の枠組み例と、本書でのベンチマーキングの位置付けを示す。ベンチマーキングに相当する部分を、次の図形で示している。

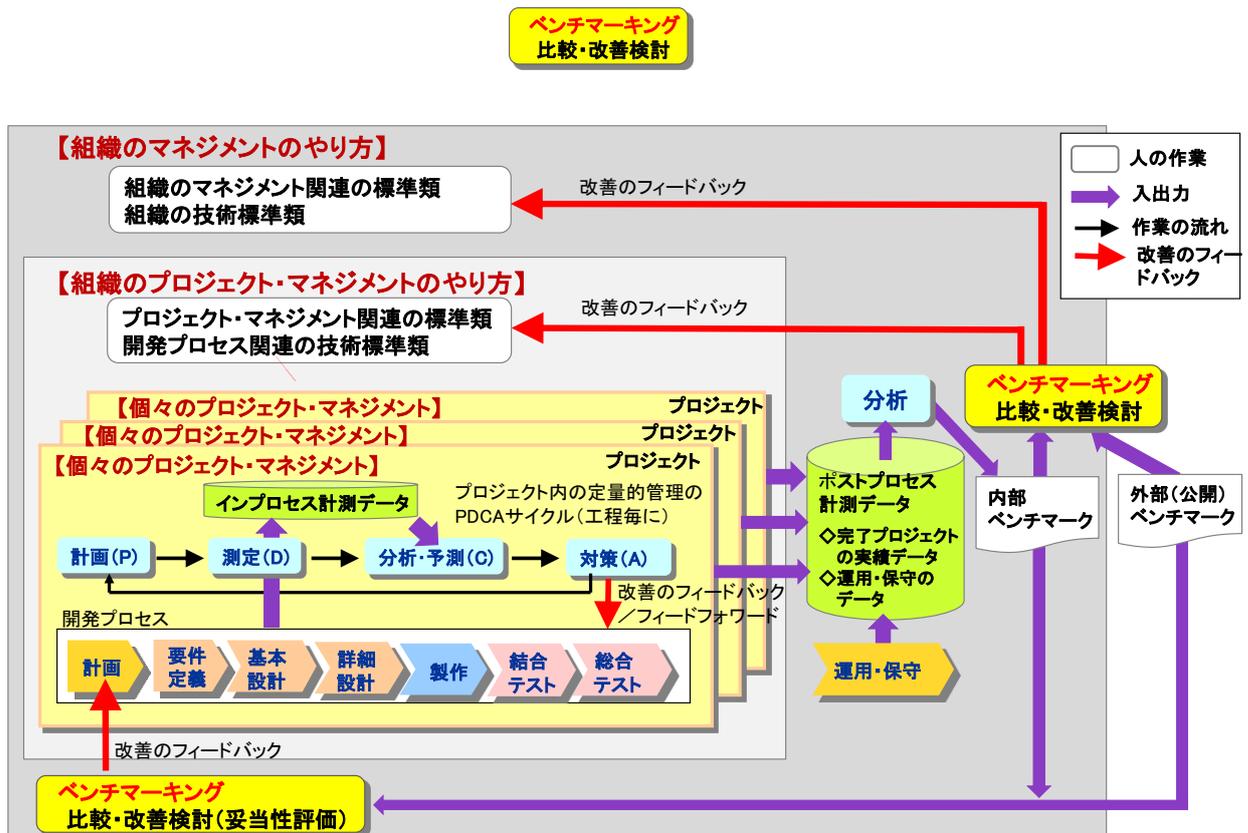


図 1-1 定量的管理の枠組み例と本書でのベンチマーキングの位置付け

(注2) 品質マネジメントに関して

本書のタイトルに「品質マネジメント」という用語を用いているが、次の点に留意されたい。

◇ 品質マネジメントの全スコープにわたっての記述は行っていない。

データで語れる範囲で、かつデータ収集できている範囲の次のようなマネジメントに限られる。

- ・プロジェクト計画、品質計画

- ・プロセス改善

品質保証関連プロセス（文書化、レビュー、テスト等）の改善、
信頼性変動要因分析による重点強化領域の特定等

- ・プロジェクト・マネジメントの改善（品質評価、ステークホルダー・マネジメント等）

- ・組織・体制の強化等

◇ 品質マネジメントのプロセスやアクティビティの詳しい解説は行っていない。

主に、データに基づいて品質マネジメントの指針や目安となる数値を提示したものである。

より良い品質マネジメントに向けてのヒントとして頂ければ幸いである。

<本書の見方について>

本書では「必要な箇所を拾い読み」することを想定しており、全体を通読しなくても必要な箇所を参照すれば良いように、できるだけ箇所ごとに完結するような構成になっている。その結果、同様な文章が複数箇所に出現することがあるが、ご容赦願いたい。

1.1. 定量的管理の勧め

ソフトウェア開発は勘と経験と度胸の世界すなわち KKD の世界であると言われることがある。KKD でのプロジェクト管理はうまくあたられば安い費用で効果を得ることもあるが、外れた場合のリスクは大きく最悪の場合はプロジェクトが破綻することになる。近年はソフトウェアに対する要求の多様化に伴って開発するソフトウェアは大規模かつ複雑になり、期待される品質も納期も年々厳しくなる傾向にある。このような状況に対応するためには、できる限り科学的な開発方法を採用し、KKD の領域を少なくする必要がある。

定量的管理は、科学的な開発方法の代表的なもののひとつであり、ソフトウェア開発のさまざまな側面で適用可能な方法である。もちろん定量的管理はソフトウェア工学における他のさまざまな手法と同様に、銀の弾丸いわゆる特効薬ではない。しかし、程度の差こそあれ確実にプロジェクト管理や組織の成熟度の向上、生産性や信頼性の向上に寄与する最も基本的な手法のひとつである。

ソフトウェア開発にはさまざまな側面があるが、定量化する必要がある、あるいは定量化することによって効果があると思われるのは、プロジェクト計画時のさまざまな見積もり、プロジェクト遂行中の QCD 管理（中間製品の品質管理、コスト管理、進捗管理など）、組織内のソフトウェア開発プロセスの改善などである。これらを定量的に行うことによって、より精度の高いプロジェクト管理が行えるようになり、よりレベルの高いソフトウェア開発プロセスを構築できるようになる。

ただし、定量的管理はあくまでも定性的管理の精度を高めるためのものであって、定量的管理が最初からあるわけではない。例えば、ある要因が生産性向上に効果があるとみなされたとき、その要因を定量的にどの程度変化させたらどの程度生産性向上に効果があるかがわかれば、プロジェクト管理の精度は向上し、効果的なプロセス改善を実施することができる。しかしどの要因を管理すべきかを判断するのは定性的なモデルによってであり、それがわからなければ定量的管理もできない。一方、定性的なモデルの構築は定量的な因果関係の分析結果に基づいて行われることが多く、その意味では定性的なモデルは定量的管理の結果とも言える。このように考えると、定性的モデルと定量的管理は相補的なものであることがわかる。

定量化は、まず対象を測定することによりその姿を正確に把握することから始めなければならない。どのようなメトリクス（測定量と測定方法）を用いて対象を測定すべきかは、目的によって異なる。目的に応じたメトリクスの選定方法については、既に GQM などさまざまな方法が提案されているのでそれらを参考にすればよい。測定によって得られた結果は単なる数値あるいは分類なので、そこから意味のある情報を得るためには様々な統計的処理が必要となる。ここでは統計の知識が必要となる。さらに、統計的処理を施した結果に対して正しい意味づけを行う必要がある。

このように記述すると定量化に対して怖気づいてしまうかもしれない。しかし、難しく考える必要はない。定量的管理は、効果が見えるようなところ、やれるところからやればよいのである。自社のデータ蓄積が十分でない場合は、まずは公開情報を利用する。分析方法も、散布図を眺め、平均と分散を計算し、箱ひげ図を描いてみて違いを可視化し、回帰分析を行って・・・と簡単などころから始めて一歩、一歩前進すればよい。少しでも目に見える効果が得られたならば、協力者も増え、効果は拡大していくものである。本書がそのために少しでも役立つようであれば幸いである。

1.2. ベンチマーキングの基本的考え方

(1) ベンチマーキングの定義

ベンチマーキングとは、IT プロジェクトの評価対象の特性を、プロジェクト同士で相互に、あるいは ベンチマークと対比する活動である。ここで、ベンチマークとは、特定の IT プロジェクトのパフォーマンスが、組織内外の IT プロジェクトと比較して どのレベルに位置するかを評価するため、比較対象として利用する組織内外の参照情報である (ISO/IEC29155-1「Information technology project performance benchmarking framework の Part 1:Concepts and definitions」の定義を仮訳)。

【補足】 ベンチマーキングについては、次のような解説もある。

ベンチマーキング(benchmarking)は、企業経営で用いられる分析手法である。

ベンチマーキングとは、「自己革新を目的とし、高い革新成果を達成している他社／他組織のやり方(ベスト・プラクティス)を学び、自己の革新を最高水準に高める方法を考え出すこと」である。

自社／自組織の製品・サービス及びプロセスを測定し、それを業界で成功を収めている企業(ベスト企業)／組織(ベスト組織)のもの(ベンチマーク)と比較しそのギャップを把握した上で、明らかになったギャップを埋めるためにベスト企業／ベスト組織のベストプラクティスを参考にして業務改善を進めることである。

本書では、次のような解釈を前提にする。

- ◇ベンチマークと対比する特性としては、単に製品(プロダクト)の信頼性実績や生産性実績だけでなく、それらの要因となっている可能性のある開発プロセス、マネジメント・プロセス、組織の特性等を含む。
- ◇ベンチマーキングとして行う活動には、良い成績を収めているプロジェクト群のやり方(開発プロセス、マネジメント・プロセス、組織の特性等)を参考にして、自組織のやり方の改善を図ることを含む。
- ◇ベンチマーキングは、定量的管理のうち、ポストプロセス計測データ(完了プロジェクトの実績データ)を用いたものを言う。インプロセス計測データ(開発プロジェクト実行中のデータ)を用いた定量的管理は含めない。

(2) ベンチマーキングの基本的概念

ISO/IEC29155-1「Information technology project performance benchmarking framework の Part 1:Concepts and definitions」の Figure 2—IT project performance benchmarking framework に基づいて作成したソフトウェア・ベンチマーキングの枠組みを図 1.2-1 に示す。

なお、IPA/SEC における「ソフトウェア開発データ白書」の取組みは、ソフトウェア・ベンチマーキングの国際標準 ISO/IEC29155 シリーズに準拠している。(IPA/SEC は、「ソフトウェア開発データ白書」の取組み内容を提案するなど、ソフトウェア・ベンチマーキングの国際標準化に貢献している)

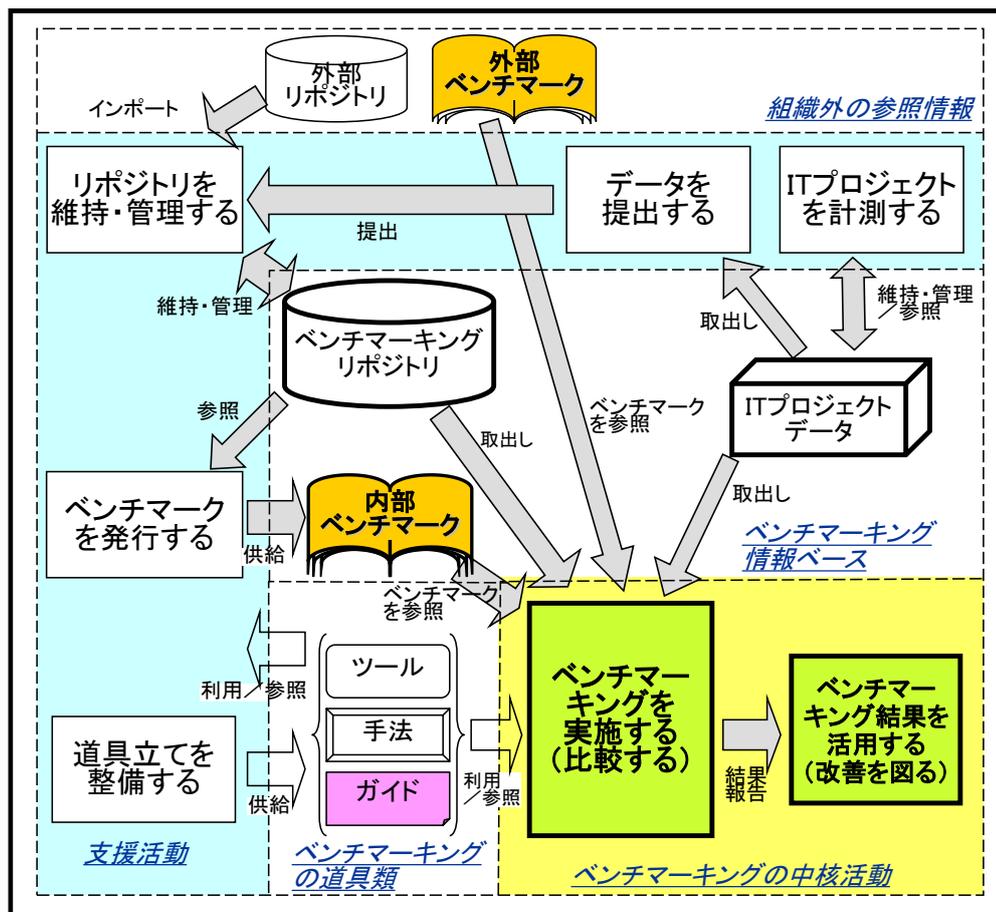


図 1.2-1 ISO/IEC29155-1 のソフトウェア・ベンチマーキングの枠組み

(3) 本書の狙い

本書は、上記の図 1.2-1 のガイド：

ガイド

の一つとして利用して頂くことを狙っている。

これまで企業が自組織／自プロジェクトの信頼性、生産性等に関するデータをどのようなベンチマーク項目と比較し、どのような改善策を講じるべきかといった定量的管理の実践面を支える適切な手引書は存在していなかった。国際標準 ISO/IEC29155 でも、自組織／自プロジェクトの信頼性、生産性等に関する定量的指標を、外部／内部ベンチマークと比較するところまでしか標準化対象となっていない。本書は、まだ国際標準化されていないがベンチマーキングの重要な部分である「ベンチマーキング結果を活用する(改善を図る)」部分も対象範囲に含めて手引きすることを狙っている。

(4) ベンチマーキングの要点

◇自組織のデータを用いたベンチマーキングが基本である。

組織内の IT プロジェクトデータから作成した内部ベンチマーク（又はベンチマーキング・リポジトリ）を利用してベンチマーキングを実施することが基本であるが、自組織のデータの収集やリポジトリ構築をまだ実施していない組織においても、実施できる部分から開始することが望ましい。

◇参考情報として、外部の公開ベンチマークを参照することも有意義である。

IPA/SEC が提供している公開ベンチマークとして、白書等がある。

◇収集データ及びメトリクスは、目的に応じて選定する必要がある。

目的を想定せずに多種多様なデータを収集するのは、現実的でなく意義が薄い。

また、既に明白になっていることに関してデータ収集するのは、ほとんど意味がない。

(参考) GQM 手法

目的指向のメトリクス選定に関する代表的な枠組みとして、GQM 手法がある。

計測モデルを、Goal 層 (概念レベル)、Question 層 (運用レベル) 及び Metric 層 (定量レベル) の3層で定義する手法である。

◇単にベンチマーキングを実施するだけでなく、ベンチマーキング結果を活用して改善を図ることが肝要である。

1.3. 本書の位置付け

(1) 本書のスコープ

本書は、ソフトウェア開発に関するベンチマーキング方法の手引書の形態を採っているが、図 1.2-1 の「ベンチマーキングの中核活動」を対象範囲とする。

◇ベンチマーキングの中核活動として、「ベンチマーキングを実施する（比較する）」だけでなく、「ベンチマーキング結果を活用する（改善を図る）」部分も対象範囲に含めている。

◇一方、データ収集、ベンチマーク発行等の組織的支援活動については、本書の対象範囲に含めていない。

◇外部ベンチマークとしては、IPA/SEC が提供している白書等の公開ベンチマークの利用を想定する。

◇ベンチマーキングによって改善を進めることが肝要であり、本書は次のような改善をスコープに含めている。

- ・プロジェクト計画の改善
- ・プロジェクト・マネジメント関連標準類の改善
- ・組織のマネジメント関連標準類の改善

その概念図を図 1.3-1 に示す。

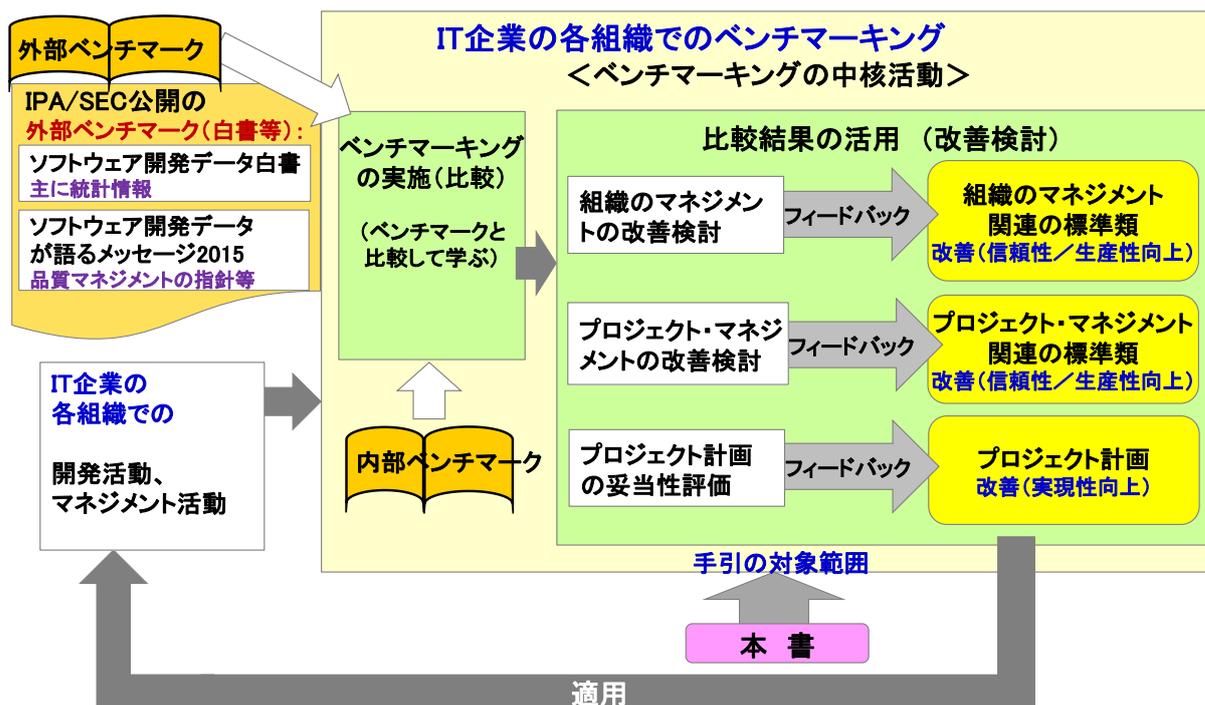


図 1.3-1 ベンチマーキングによる改善の概念

(2) 主な対象読者

白書等の読者を主な対象読者とする。

具体的には、開発者（プロジェクト・マネジャー及び担当者）と開発組織の管理層を主な対象読者として想定する（PMO 及び品質マネジメント推進部門が助言／支援する場合を含む）。

(3) 既存の定量的管理の手引書との違いについて

本書はソフトウェアのベンチマーキング方法の手引書の形態を採っており、定量的管理のうちポストプロセス計測データ（完了プロジェクトの実績データ）を用いた定量的管理を対象にしている。IPA/SEC 関連の定量的管理の手引書としては、既に「定量的品質予測のススメ」及び「続 定量的品質予測のススメ」がある。それらと本書との違いを表 1.3-1 に示す。

表 1.3-1 定量的品質予測のススメと本書との違い

比較項目	「定量的品質予測のススメ」及び 「続 定量的品質予測のススメ」	本書
データのスコープ	インプロセス計測データ (開発プロジェクト実行中のデータ)	ポストプロセス計測データ (データ提供企業から提供された完了プロジェクトの実績データ)
データの取扱い	非公開 (各社の機密データ)	データは非公開 (研究用には条件付きで開示) 統計情報は白書等で公開
マネジメントのスコープ	プロジェクト・マネジメント	プロジェクト・マネジメント及び組織の改善のためのマネジメント
主目的	対象プロジェクトを成功させるため、あるいはプロジェクト・マネジメントを改善するための定量的品質予測手法の手引書	上記マネジメントを改善するためのベンチマーキング方法の手引書
ベンチマーキングのための主な公開情報	特になし	IPA/SEC 白書データベースに裏付けられた白書等の公開情報 (下記) : ◇「ソフトウェア開発データ白書」 規模、工数、工期、信頼性、生産性及びそれらの変動要因に関する種々の統計情報 ◇「ソフトウェア開発データが語るメッセージ」 上記マネジメント改善に向けた参考用のメッセージ (マネジメントの指針及び目安となる数値)

1.4. 本書の特長

(1) 統計指標に基づく品質マネジメントの実践集

本書は、統計指標に基づいた品質マネジメントの実践集として利用できるものであり、信頼性向上等に向けた品質マネジメントの指針や方策の検討のためのヒントや目安を提供している。

- ① IT企業の品質マネジメント事例を含めて、豊富な具体例を掲載している。
 - ・IPA/SECの公開ベンチマーク（白書及びメッセージ2015）を利用した具体例：19件
 - ・IT企業の内部ベンチマークを利用した具体例：13件

- ② 信頼性向上のための品質マネジメントを主題に据えている。

本書は、統計指標に基づいて実証された品質マネジメント指針をヒントとして提示している（そのような文書は稀である）。また、品質マネジメントの指針や方策を検討するシーンについては、品質マネジメントにおける源流管理（注）の考え方を重視している。具体的には、目先のプロジェクトを対象にしたシーンだけでなく、プロジェクト・マネジメントのやり方の改善シーンや組織の改善のためのマネジメント・シーンを取り上げている。

（注）源流管理：仕事の仕組みや組織の改善を図るマネジメント。典型的には組織の標準類の改善につなげること。

- ③ 品質マネジメント指針に関しては、目安となる数値を併せて提示している。
例えば、信頼性向上のために設計レビューを強化するという方針についても、マネジメント・アクションに繋げるためにはどの程度レビュー強化すると良いのか、目安となる数値が重要である。

(2) IT業界初の体系的かつ具体的なベンチマーキング手引書

本書は、ソフトウェア・ベンチマーキングの手引書（ガイド）の形態を採っており、体系的かつ具体的にベンチマーキング方法を解説している。（ソフトウェアのベンチマーキング方法を具体的かつ体系的に解説した公開ガイドはまだ世の中に見当たらない。）

- ① 種々のベンチマーキングのシーン（15シーン）に沿って体系的に解説

＜ベンチマーキングのシーン例＞

- ・開発者用のベンチマーキング：信頼性目標及び品質保証プロセスの妥当性評価、工期の妥当性評価のシーン等
- ・開発組織のマネジャー層用のベンチマーキング：品質保証プロセス関連標準類の見直し、信頼性変動要因分析と信頼性向上に向けた重点強化領域の特定のシーン等

- ② IT企業の具体例を含めて、豊富な具体例を掲載している（上記(1)の①と同様）。

- ③ 改善を重視したベンチマーキング

ベンチマーキングの目的は、比較対象とのパフォーマンス・ギャップ（自組織の強み／弱み）を把握するだけでなく、そのギャップを効果的に埋めるために、パフォーマンスの良い組織のやり方を参考にしながら重点的に改善すべき対象領域を特定するなどによって改善を進めることにある。本書では、「パフォーマンスの良い組織のやり方を参考にして改善を進める」ことを重視しており、プロジェクト・マネジメントのやり方の改善シーンや組織の改善のためのマネジメント・シーンもスコープに含めている（上記(1)の②と同様）。

1.5. 公開ベンチマーク（白書等）を参考にしたベンチマーキングの勧め

以下のことから、IPA/SEC が提供している公開ベンチマーク（白書等）を、自組織のベンチマーキングの参考情報として利用することをお勧めする。

(1) IPA/SEC 白書データベースの特長

IPA/SEC では、ソフトウェア開発における定量的管理の普及促進を目的に、開発プロセスの標準化や見える化手法、定量的品質管理手法などの調査・検討を行っている。これらの活動の一環として、国内の多様で幅広い業種・業務から多数のソフトウェア開発に関するプロジェクトデータを収集し、ソフトウェア開発プロジェクト・データベース（以下、IPA/SEC 白書データベース）を構築・利用している。

IPA/SEC 白書データベースは、次の特長を有する世界的にも比類のない貴重なデータベースであり、種々の分析に活用できるものである。

◇データ件数が多い。

「ソフトウェア開発データ白書 2014-2015」作成時点で、3,541 件のプロジェクトデータを蓄積している。

◇偏りの少ない多様なプロジェクトデータから成っている。

データ提供企業は日本の代表的なソフトウェア開発ベンダ（現在約 30 社）であり、多様な業種、業務等に関するプロジェクトデータが蓄積されている。

◇生産性、信頼性等の要因分析が可能である。

プロジェクトデータとして、プロジェクトのプロファイルだけでなく、品質要求、開発プロセス、ユーザの協力度合い、組織の成熟度、開発環境等の広範なエリアの要因を分析するためのデータを備えている。

◇経年変化の分析が可能である。

データ収集開始してから 10 年以上が経過していることと、データ提供企業各社から毎年データ収集していることから、信頼性、生産性等の経年変化（推移）の分析が可能である。

(2) 公開ベンチマーク（白書等）の概要

IPA/SEC では、IT 業界の各組織におけるベンチマーキングの推進に向けて、参考用の公開ベンチマークとして白書等（「ソフトウェア開発データ白書」及び「ソフトウェア開発データが語るメッセージ」）を提供している。

① 「ソフトウェア開発データ白書」（以下、白書）

IPA/SEC 白書データベースのデータを分析し、生産性、信頼性等に関する統計情報を中心に「ソフトウェア開発データ白書」として取りまとめて書籍化し、2005 年から定期的に発行している。その最新版として「ソフトウェア開発データ白書 2014-2015」¹ を 2014 年 10 月 1 日に発行した。

また、「ソフトウェア開発データ白書 2016-2017」を 2016 年秋に発行予定である。

<主なベンチマーク情報>

以下の分析に関するデータ項目、メトリクス及び分析結果（統計情報）

◇規模、工数及び工期の関係

- ・規模と工数の関係
- ・工数と工期の関係

◇工程別情報

- ・工程別実績工数の比率
- ・工程別実績工期の比率

¹ 「ソフトウェア開発データ白書 2014-2015」の Web 公開サイト：<https://www.ipa.go.jp/sec/publish/tn12-002.html>

- ・レビュー実績工数及びレビュー指摘件数
- ・テスト工程別のテストケース密度と検出バグ密度

◇信頼性実績

規模別、開発言語別、業種別、アーキテクチャ別、開発フレームワークの利用別、検出バグ密度との関係等

◇生産性実績

規模別、開発言語別、業種別、アーキテクチャ別、プラットフォーム別、月あたりの要員数別、開発フレームワークの利用別、信頼性要求の高さとの関係、外部委託比率との関係等

② 「ソフトウェア開発データが語るメッセージ」

同じIPA/SEC 白書データベースに基づいて信頼性向上や生産性向上のために有用と考えられる新たな項目を中心に分析を行い、それらの分析結果とプロジェクトや組織のマネジメント指針とを纏めた「ソフトウェア開発データが語るメッセージ 2015」² を、2015年9月25日にIPAのWebサイトに公開するとともにプレスリリースを行った。

<主なベンチマーク情報>

以下の分析に関するデータ項目、メトリクス、分析結果（統計情報及び読み取れる傾向）及びメッセージ（信頼性／生産性向上に向けたマネジメントの指針及び目安となる数値）

◇エンジニアリング編

主に開発プロジェクトのプロジェクト・マネジャー及び担当者に向けて、開発プロセスの定量的コントロールの参考にして頂くことを目的として、レビュー、テスト等に関するいくつかの新規分析結果とメッセージを提示している。特に、品質保証プロセス（文書化、レビューおよびテスト）については、経済的に信頼性を確保するためにどの程度注力すれば良いかの指針と目安となる数値を提示している。

◇マネジメント編

主に開発組織の管理層の方々に向けて、ステークホルダー・マネジメント、スコープ・マネジメント等のプロジェクト・マネジメントや、変動要因分析結果に基づく重点強化領域の特定、組織体制の整備など、組織改善のためのマネジメント等の参考にして頂くことを目的として、いくつかの新規分析結果とメッセージ（マネジメント指針）を提示している。

◇トレンド編

主にIT業界や国等の政策立案担当者の方々に向けて、信頼性および生産性のマネジメントにおける目標や施策を検討する際の参考情報として利用して頂くことを期待して、信頼性及び生産性のトレンドに関する新規分析結果とメッセージを提示している。

² 「ソフトウェア開発データが語るメッセージ 2015」のWeb公開サイト：
<http://www.ipa.go.jp/sec/reports/20150925.html>

(3) 白書等を参考にしたベンチマーキングのメリット

一般に、ベンチマーキングを行う際には、組織内のソフトウェア開発プロジェクトの実績データを収集・蓄積して作成したベンチマーク（内部ベンチマーク）を利用することが望ましいが、ベンチマーキングの経験が浅い組織では、ベンチマーキングの導入や理解のための参考情報として公開ベンチマークを利用することができる。また、公開ベンチマークは、既にベンチマーキングを実践している組織でも、IT 業界内における自組織のパフォーマンスや強み/弱みを把握し、重点的に改善すべき対象領域を検討・特定するため等に利用することができる。

既にベンチマーキングを実践している組織においても、ベンチマーキングを更に進めるために、白書等の公開ベンチマークが以下のように有用な参考情報となると考えられる。白書等を参考にして自組織内でのベンチマーキングを一層推進することが望ましい。

① 国内 IT 業界における自組織のパフォーマンスや強み/弱みを把握するために、目安となるベンチマークとして白書等の統計情報を参考にすることができる。

白書等の統計情報は、国内 IT 業界におけるソフトウェアの信頼性実績、生産性実績、開発プロセスの実施実績等の一つの相場観を呈示している。

(注) 「1.6 留意事項」に示したように、「白書等の統計情報は、定量的管理されたプロジェクトの統計情報と見るのが妥当であって、世間相場観よりも良い値を示している可能性がある」ことに留意した上で、目安として参考にされたい。

自組織のソフトウェアの信頼性実績、生産性実績、開発プロセスの実施実績等の統計情報と白書等の統計情報とを比較し、自組織のパフォーマンスや強み/弱みを把握することが、重点的に改善すべき対象領域を検討・特定する契機となる。

(例) 白書等の統計情報と比較して、信頼性実績が低い（発生不具合密度（規模あたりの稼働後に発生した不具合数）が高い）ことが分かった。また、その要因となっている可能性がある品質保証プロセスに着目すると、設計文書化密度（規模あたりの設計書ページ数）及び設計レビュー工数密度（規模あたりの設計レビュー工数）が白書等の統計情報と比較して低いことが分かった。

これらのことから、設計書の充実と設計レビュー強化（設計レビュー工数の増強と設計レビュー効率向上）によって信頼性向上を図ることを重点課題として検討することとした。

② 自組織の信頼性/生産性向上に向けたプロジェクト・マネジメントや組織改善のマネジメントの見直しのために、白書等の分析結果（傾向）及びメッセージをヒントにすることができる（具体的なマネジメント・アクションの指針として参考にすることができる）。

(例 1) 白書等には、信頼性や生産性に影響を与える要因（変動要因）として、IPA/SEC 白書データベースから読み取れる変動要因とそれらの変動要因による変動の程度が具体的に示されている。必要に応じて自組織の変動要因が何かを探り直す時に、それらを参考にすることができる。

信頼性/生産性が向上する方向に変動要因を制御できると信頼性/生産性が向上すると期待できることから、組織改善のマネジメントにおいて自組織の変動要因を把握することが重点改善領域の検討・特定に役立つ。

(例 2) 白書等のメッセージには、経済的に信頼性を確保するための具体的なマネジメント・アクションの指針として、「レビューやテストにどの程度注力すると良いのか」が、IPA/SEC 白書データベースに基づいて定量的に示されている。自組織のプロジェクト・マネジメントに関する標準類（管理基準等）の見直しの際に、白書等のメッセージに示されたマネジメント指針をヒントとして参考にすることができる。また、白書等のメッセージに示された数値を目安として参考にすることができる。

- ③ 自組織のベンチマーキングの方法を見直す際に、目的に応じて、白書等のデータ項目、メトリクス及び分析方法を参考にすることができる。
- ④ 品質向上方針／方策を実施するためには、多くの場合、企業が持つ（人的・物的・資金的）リソースを動かす必要がある。従って、それらのリソースを動かす権限を持っているマネジャー層（あるいは経営層）の理解と協力を得ることが重要となる。そのためには、当たり前のようなメッセージであってもデータで実証して上申すること、またどの程度注力すると良いのかを数値で示すことが効果的なアプローチとなる。ベンチマーキングに限らず定量的管理の大きな狙いの一つは、マネジャー層・経営層に働きかけることにある。
- 白書等の公開ベンチマーク及び本書が、品質向上方針／方策の検討及び上申の後押しになれば幸甚である。

1.6. 公開ベンチマーク（白書等）利用上の留意事項

- (1) 白書等の統計情報を参考にする場合には、次の点に留意して「目安」として参考にされたい。
 - ① 白書等の分析結果は、基本的に種々の組織、業種、プロファイル等が混在したサンプルデータ集合から得られたものである。信頼性、生産性等の評価にあたっては、妥当と評価できる範囲が業種、品質要求レベルを始めとする種々の変動要因によって変動することを勘案する必要がある。一般に、システムはその重要性や停止時の影響の大きさによりクラス分けされる（クリティカリティクラス）。本書のメッセージに示される推奨値については、システムのクリティカリティクラスやその他の変動要因に応じて適宜補正した上で、「目安」として利用されたい。
 - ② 白書等で用いているサンプルデータ集合には、定量的管理が行われている組織のプロジェクトのデータが多く含まれていることから、白書等の統計情報は定量的管理されたプロジェクトの統計情報と見るのが妥当である。従って、白書等の生産性／信頼性に関する統計値は世間相場観よりも良い値を示している可能性がある。
 - ③ 今後、新しい開発プロジェクトのデータが追加されていくことに伴って、統計値や傾向が変化する可能性がある。特にサンプルデータ数がまだ少ない分析項目については、統計値や傾向が変化する可能性が高い。
- (2) 白書等の外部の公開ベンチマークは、次のように必要に応じて参考にされたい。
 - ・プロジェクトの実績データが蓄積されていない場合の代替情報として目安にする。
 - ・国内 IT 業界水準との差異を把握するために参照する。
- (3) 白書等に記載された回帰式を利用するにあたっては、白書の 3.4 節「回帰式利用上の注意事項」を参照のこと。
- (4) 分析項目によっては分析に必要なデータ項目群が異なること、データ項目ごとにデータ欠損しているプロジェクト群が異なることから、各分析項目の対象プロジェクト集合は、必ずしも同一でないことに注意されたい。また、開発規模のメトリクスとしてファンクションポイント（FP）を採用しているプロジェクト集合と、プログラムコード行数（SLOC）を採用しているプロジェクト集合とは異なる集合であることに注意されたい。

2. ベンチマーキングの主なシーン

本書では、ベンチマーキングの主なシーンとして、以下のシーンを取り上げる。

2.1. 開発者用のベンチマーキングの主なシーン

(1) プロジェクト計画の妥当性を評価するシーン

開発者（開発プロジェクトのプロジェクト・マネジャー及び担当者）が、プロジェクト計画の策定時に、プロジェクト計画の実現性を高めるために、以下のような観点からプロジェクト計画の妥当性を評価するシーン（開発組織のマネジャー層、PMO 及び品質マネジメント推進部門が、プロジェクト計画をレビューし助言／支援するシーンを含む）

- ① 信頼性目標及び品質保証プロセス（文書化、レビュー、テスト等）の妥当性
- ② 開発総工数及び生産性目標の妥当性
- ③ 成果物量及び単位成果物量あたりの工数の妥当性
- ④ 工期の妥当性
- ⑤ 工程別の工数比率の妥当性
- ⑥ 工程別の工期比率の妥当性等

<要点>

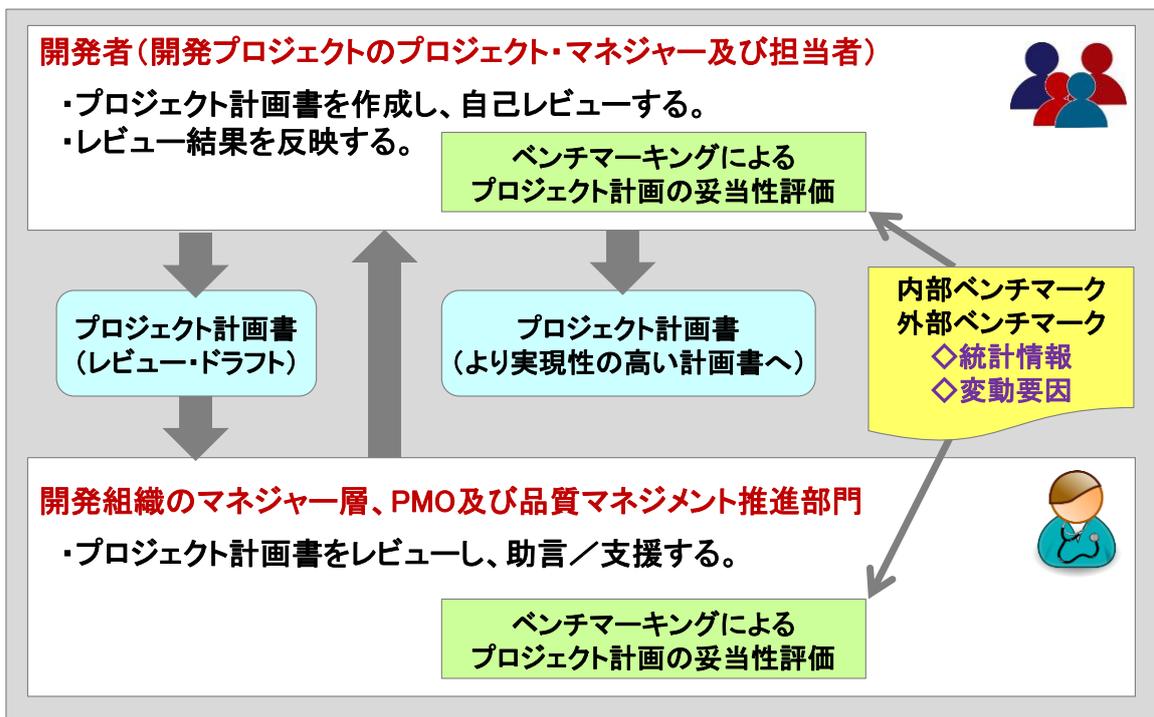


図 2.1-1 プロジェクト計画の妥当性を評価するシーン

- ◇ 内部ベンチマークの統計情報と比較して妥当性評価するのが基本である。
外部の公開ベンチマークは、次のように必要に応じて参考にする。
 - ・プロジェクトの実績データが蓄積されていない場合の代替情報として目安にする。
 - ・国内IT業界水準との差異を把握するために参照する。

- ◇ ベンチマーク中の統計情報の妥当と評価できる範囲（P25 から P75 の範囲等）と比較する。
ただし、妥当と評価できる範囲については、ベンチマーク中の変動要因情報を参照し、変動要因によって上方/下方修正した上で比較することが望ましい。
比較の結果、範囲外となる場合、範囲外となる合理的理由がなければ計画を見直す。

(注1) 自組織に妥当性評価の管理基準が設定されていれば、それに従う。

(注2) より高い信頼性目標/生産性目標を目指して改善を図るケースについては、2.2節「開発組織のマネジャー層用のベンチマーキングの主なシーン」の方を参照されたい。

2.2. 開発組織のマネジャー層用のベンチマーキングの主なシーン

主に開発組織のマネジャー層が、以下のような観点から、プロジェクト・マネジメントの改善や組織の改善のためのマネジメントを推進するシーン（PMO 及び品質マネジメント推進部門が、品質マネジメントを支援するシーンを含む）

(1) プロジェクト・マネジメントの改善を推進するシーン

① 品質マネジメント関連

品質保証プロセス関連標準類の見直し

（文書化、レビュー、テスト、上流工程の不具合摘出比率、工程の完了評価基準等）

② ユーザの関与（ステークホルダー・マネジメント関連）

ユーザ担当者の要求仕様関与度合いの向上等

③ 成果物スコープの明確化（スコープ・マネジメント関連）

要求仕様の明確化等

(2) 組織の改善のためのマネジメントを推進するシーン

体制の整備

信頼性変動要因分析と信頼性向上に向けた重点強化領域の特定

生産性変動要因分析と生産性向上に向けた重点強化領域の特定

業種等のドメイン別マネジメント

信頼性／生産性等の経年変化に基づく組織改善計画

<要点>

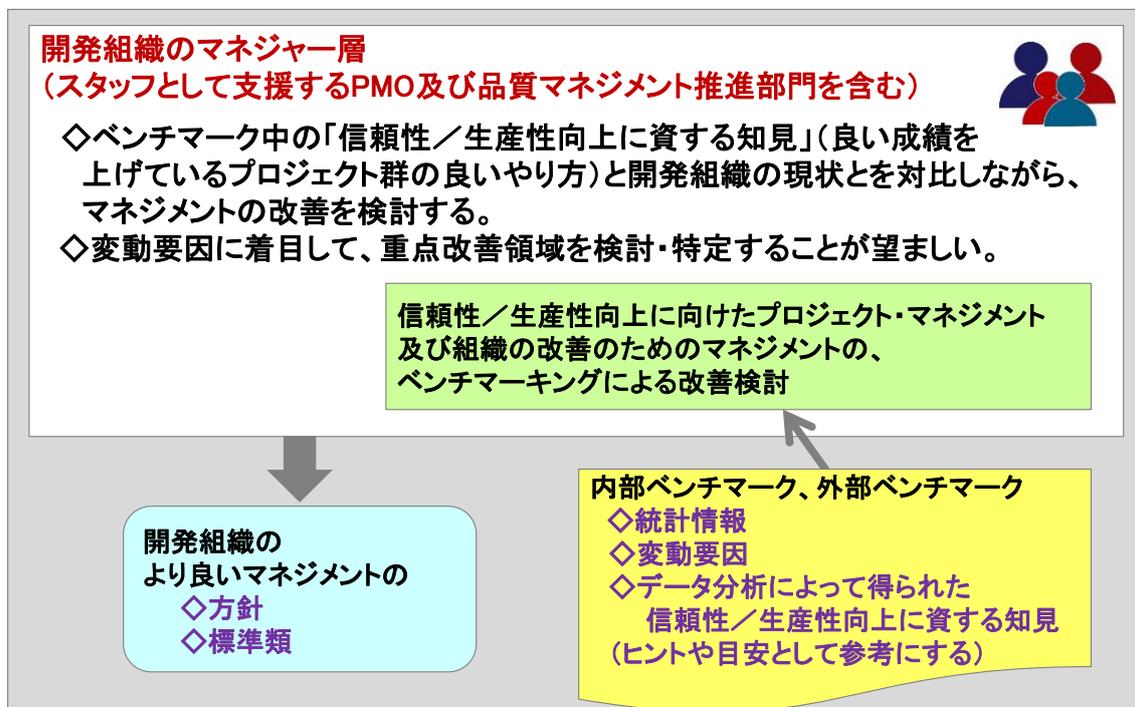


図 2.2-1 マネジメントの改善を推進するシーン

- ◇ 内部ベンチマークの「信頼性／生産性向上に資する知見」（良い成績を上げているプロジェクト群の良いやり方）と開発組織の現状とを対比しながら、マネジメントの改善を検討することが基本である。また、変動要因に着目して、重点改善領域を検討・特定することが望ましい。外部の公開ベンチマークは、次のように必要に応じて参考にする。
 - ・プロジェクトの実績データが蓄積されていない場合の代替情報として目安にする。
 - ・国内 IT 業界水準との差異を把握するために参照する。
- ◇ 改善検討結果は、開発組織のマネジメントの方針及び標準類に反映する。
- ◇ 「信頼性／生産性向上に資する知見」及び「開発組織の変動要因」を含む内部ベンチマークについては、PMO 及び品質マネジメント推進部門が作成し定期的に更新することを想定している。

2.3. 3章以降の見方について

3章から5章は、ベンチマーキングのシーンによって次のように構成されている。参考にされたいベンチマーキング・シーンに該当する節をご覧ください。

3章「開発者用のベンチマーキング」

2.1節「開発者用のベンチマーキングの主なシーン」の各シーン

4章「開発組織のマネジャー層用のベンチマーキング（プロジェクト・マネジメントの改善）」

2.2節「開発組織のマネジャー層用のベンチマーキングの主なシーン」のプロジェクト・マネジメントの改善関連の各シーン

5章「開発組織のマネジャー層用のベンチマーキング（組織の改善）」

2.2節「開発組織のマネジャー層用のベンチマーキングの主なシーン」の組織の改善関連の各シーン

3章から5章中の各節は、次の三つの項から構成されている。必要に応じて、適宜各項をご覧ください。

x.x 節	
x.x.1 項「ベンチマーキング方法の手順例」 ベンチマーキング方法の手順を要約的に例示している。 まず x.x.1 項を一読してから、x.x.2 項又は（／及び） x.x.3 項以降を読み進めることをお勧めする。	
x.x.2 項「白書等を用いた具体例」 1 項の手順に沿って、白書等をベンチマークとしたベンチマーキングの具体例を示している。 ベンチマーキングの経験が浅く内部ベンチマークを保有していない組織の方々には、x.x.3 項以降の前にこちらをご覧くださいようお勧めする。 既に内部ベンチマークを用いてベンチマーキングを実践されている組織の方々にも、必要に応じて参考のためにご覧ください。ベンチマーキングを更に進めるためのご参考になれば幸いです。	x.x.3 項以降「IT 企業での具体例」 1 項の手順に沿って、内部ベンチマークを用いた企業内のベンチマーキングの具体例（実施している事例）を示している。 （注） x.x.3 項以降が掲載されていない節もある。種々の着眼点がヒントになる事例なので、内部ベンチマークを保有していない組織の方々も、保有してベンチマーキングを実践されている組織の方々にも、是非ご覧ください。

3. 開発者用のベンチマーキング

3.1. 信頼性目標及び品質保証プロセスの妥当性評価

3.1.1. ベンチマーキング方法の手順例

ベンチマーキング方法の基本的な手順の例を以下に示す。

- (1) ベンチマーキングの目的を確認する。
ベンチマーキングの目的が、「プロジェクト計画におけるプロダクト目標（信頼性目標）及びプロセス目標（信頼性目標達成のための品質保証プロセスの目標）が、信頼性を確保するために妥当な水準に設定されているか否かを評価し、必要に応じて調整すること」であることを確認する。
- (2) 妥当と評価できる範囲を設定し、必要に応じて調整する。
信頼性に対応するベンチマーク中の統計情報を目安にして、信頼性目標の妥当と評価できる範囲を設定する。同様に、品質保証プロセスに対応するベンチマーク中の統計情報を目安にして、各品質保証プロセスの妥当と評価できる範囲を設定する。

<妥当と評価できる範囲の例>

同じ業種の統計情報の 25 パーセンタイル（P25）から 75 パーセンタイル（P75）までの範囲
散布図では信頼区間 50%の範囲に相当する。箱ひげ図では箱の範囲に相当する。

（備考 1）ベンチマークとしてどの統計情報を採用するかについては、評価対象プロジェクトとできるだけ同種の（例えば同じ業種の）ドメインの統計情報を採用することが望ましい。

（備考 2）ベンチマークとしては、次の点からも内部ベンチマークを利用することが望ましい。
白書等の公開ベンチマークの統計情報は種々の属性のプロジェクトが混在したサンプル集合から導出されたものである。一方、内部ベンチマークは比較的同種のプロジェクトのサンプル集合から導出できる可能性が高いと考えられる。従って、白書等の公開ベンチマークの統計情報と比較して、変動要因の個数が少なくばらつきが小さい（P25～P75 の範囲が小さいなどの）ことが期待される。

（備考 3）妥当と評価できる範囲の調整（妥当性評価に共通的な留意事項）

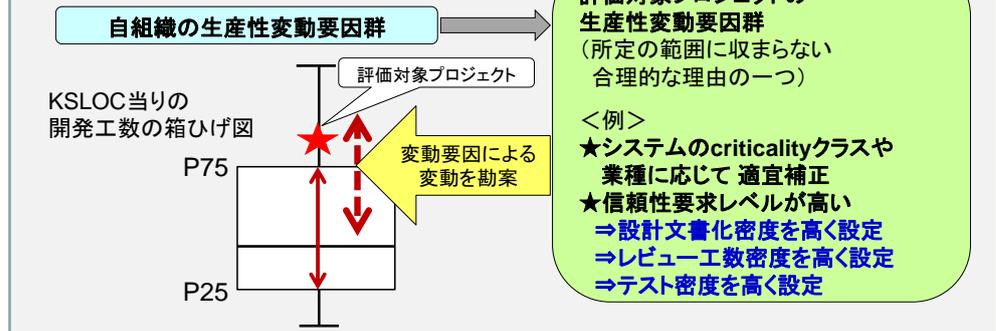
信頼性及び生産性は変動要因の影響によって変動する。妥当と評価できる範囲は、評価対象プロジェクトに該当する変動要因の影響を勘案して調整（上方修正／下方修正）することが望ましい。代表的な変動要因としては、システムのクリティカリティクラス（品質要求の厳しさを表すクラス）、品質要求レベル、業種等がある。

具体的な変動要因については、ベンチマーク中の変動要因に関する情報を参照されたい。「ソフトウェア開発データが語るメッセージ」では、「信頼性変動要因の分析」及び「生産性変動要因の分析」の記事がある。なお、「ソフトウェア開発データ白書 2016-2017」では、「信頼性変動要因の分析」及び「生産性変動要因の分析」の節が追加される見込みである。

ただし、ベンチマーク中に示された各変動要因が自組織にとって主要な変動要因に該当するか否かを判断する必要がある。なお、自組織において固定的になっている変動要因は、自組織にとっての変動要因には当たらない。

一定の範囲に収まっていないというだけで妥当でないと評価するのは早計である。どのようなプロジェクトであるかによっては、変動要因による変動を始めとして一定の範囲に収まらなくなる合理的な理由が存在する可能性がある。**評価対象プロジェクトに該当する変動要因によって生じる変動幅を勘案して、所定の妥当な範囲を上方修正/下方修正**しながら妥当性評価することが望ましい。その結果においても妥当な範囲外となり、かつ変動要因以外の合理的な理由がない場合には、計画や見積りを見直すことが望ましい。

<変動要因の活用イメージ例>



(備考4) 妥当と評価できる範囲に四分位数を採用している理由について

白書等で取り扱っているほとんどのデータの分布は、右裾広がり非対称性の強い分布になっており、正規分布とは程遠いものである。従って、白書等では代表値として平均値よりも中央値を重視している。また、妥当と評価できる範囲を平均値±標準偏差のように設定するのは適切でない。四分位数は、データ分布の形によらず分布の様子を把握するのに便利なので、妥当と評価できる範囲の代表的な例として、P25～P75の範囲を提示している。なお、P25は第1四分位数、中央値は第2四分位数、P75は第3四分位数である。P25～P75の範囲におよそ50%のデータが含まれることになる。

(注) プロダクト目標及びプロセス目標について組織の目標が設定されている場合には、それに従って妥当性評価すること。

<信頼性のメトリクス例>

◇発生不具合密度 (FP発生不具合密度又はSLOC発生不具合密度) : 稼働後一定期間に発生した不具合数÷開発規模 (FP規模又はSLOC規模)

<品質保証プロセスに関するメトリクス例>

◇レビュー関連

- ・設計レビュー工数比率 : 設計レビュー工数÷開発総工数
- ・設計レビュー工数密度 : 設計レビュー工数÷開発規模 (FP規模又はSLOC規模)
- ・設計レビュー指摘密度 : 設計レビュー指摘数÷開発規模 (FP規模又はSLOC規模)

◇テスト関連

- ・テスト密度 : テストケース数÷開発規模 (FP規模又はSLOC規模)
- ・テスト検出不具合密度 : テスト検出不具合数÷開発規模 (FP規模又はSLOC規模)
- ・テスト検出能率 : テスト検出不具合数÷テストケース数

◇文書化関連

・設計文書化密度：設計書ページ数÷開発規模（FP 規模又は SLOC 規模）

◇上流工程での不具合摘出比率：上流工程での不具合摘出数÷（上流工程での不具合摘出数+下流工程での不具合摘出数）

- (3) 評価対象プロジェクトのプロダクト目標及びプロセス目標が、それぞれの妥当と評価できる範囲内に収まっているか否かを判定する。
- (4) 妥当と評価できる範囲外となる場合、評価対象プロジェクトの特性を勘案しながら、範囲外となる合理的な理由がないか検討する。合理的な理由があれば妥当と評価し、合理的な理由がなければ目標の見直しを検討する。
- (注) 信頼性向上を目指して高い目標を設定したことによって妥当と評価できる範囲外となる場合については、4章「開発組織のマネジャー層用のベンチマーキング（プロジェクト・マネジメントの改善）」の「品質保証プロセス関連標準類の見直し」等を参考にして検討されたい。

<合理的な理由の検討（又は妥当と評価できる範囲の調整）のケース例>

◇評価対象プロジェクトの属性が生産性変動要因に該当しているケース

- ・ミッション・クリティカルであり、信頼性目標及び品質保証プロセス（文書化、レビュー、テスト等）の（実施度合いの）目標を高く設定しているケース
- ・業種が混在したサンプル集合に基づいて作成されたベンチマークを用いてベンチマーキングする場合で、評価対象プロジェクトの業種が金融・保険業に該当しているケース

3.1.2. 白書等を用いた具体例

(1) 目的（活用シーン）

開発者（開発プロジェクトのプロジェクト・マネジャー及び担当者）が、プロジェクト計画の策定時に、プロジェクト計画の実現性を高めるために、次のような観点からプロジェクト計画の妥当性を評価するシーン（開発組織のマネジャー層、PMO 及び品質マネジメント推進部門が、プロジェクト計画をレビューし助言／支援するシーンを含む）

◇プロジェクト計画におけるプロダクト目標（信頼性目標）及びプロセス目標（信頼性目標達成のための品質保証プロセスの目標）が、信頼性を確保するために妥当な水準に設定されているか否かを評価し、必要に応じて調整する。

(2) ベンチマーク

信頼性及び品質保証プロセスに関する白書等の統計情報を、目安として参考にする。

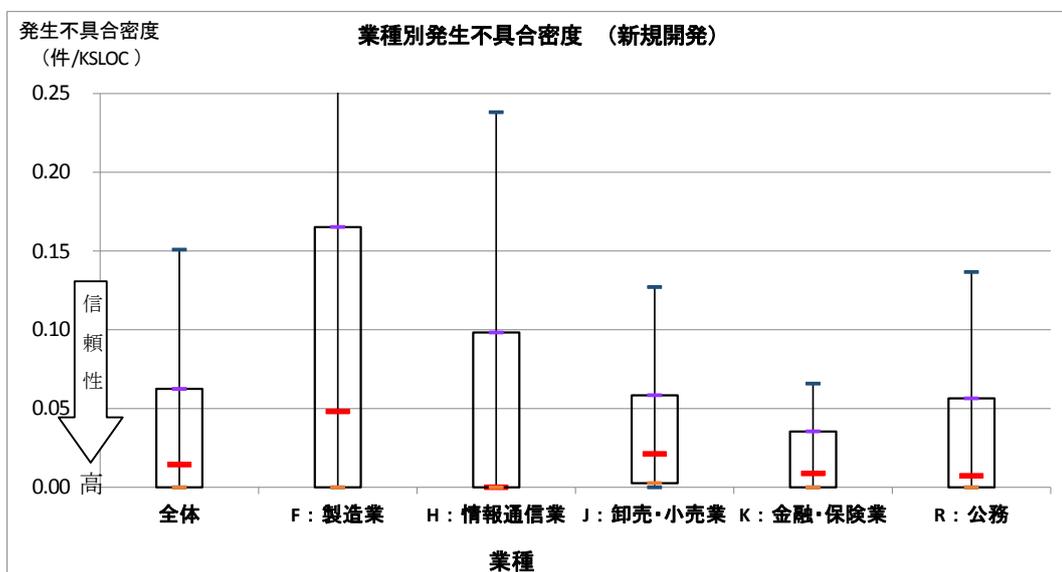
<参照箇所>

詳しくは、「ソフトウェア開発データが語るメッセージ」の「ドメイン別マネジメント」、あるいは「ソフトウェア開発データ白書」の「工程別の分析」を参照されたい。

① 信頼性（SLOC 発生不具合密度）の統計情報の例

SLOC 発生不具合密度：稼働後（最長 6 ヶ月以内）に発生した不具合数÷開発規模（KSLOC）

- 業種別 SLOC 発生不具合密度（新規開発、主開発言語グループ、開発 5 工程あり）

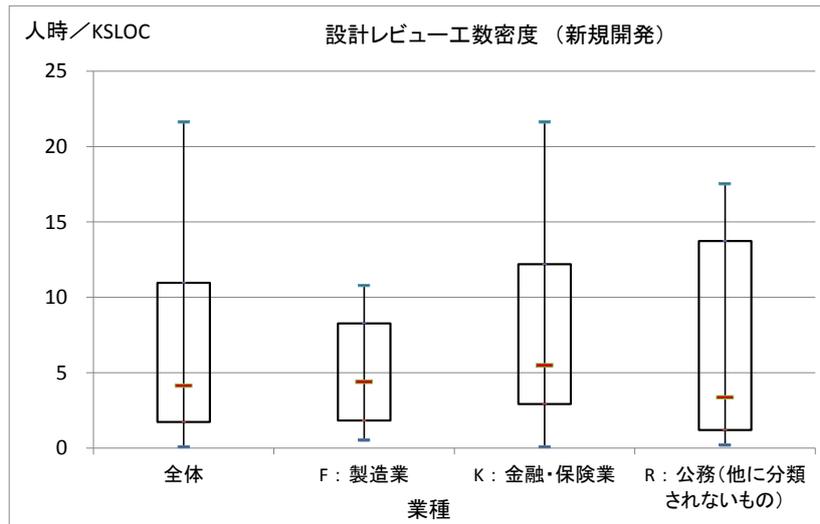


大業種	N	発生不具合密度（件/KSLOC）						
		最小	P25	中央	P75	最大	平均	標準偏差
全体	373	0.000	0.000	0.014	0.063	2.413	0.081	0.204
F: 製造業	45	0.000	0.000	0.048	0.165	1.013	0.123	0.200
H: 情報通信業	43	0.000	0.000	0.000	0.098	0.483	0.066	0.112
J: 卸売・小売業	34	0.000	0.003	0.021	0.058	2.413	0.127	0.416
K: 金融・保険業	128	0.000	0.000	0.009	0.035	1.117	0.044	0.134
R: 公務（他に分類されないもの）	51	0.000	0.000	0.007	0.056	1.269	0.114	0.265

② 品質保証プロセス関連の統計情報の例

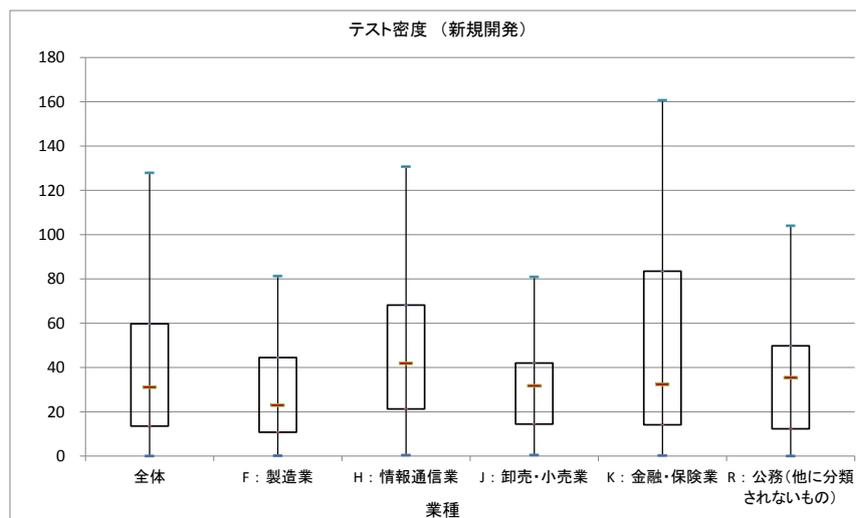
- ・設計レビュー工数密度：基本設計から製作までの設計レビュー工数÷開発規模（KSLOC）
- ・テスト密度：（結合テストケース数＋総合テスト（ベンダ確認）ケース数）÷開発規模（KSLOC）
- ・設計文書化密度：（基本設計書のページ数＋詳細設計書のページ数）÷開発規模（KSLOC）
- ・上流工程での不具合摘出比率：基本設計から製作までのレビュー指摘数÷（基本設計から製作までのレビュー指摘数＋結合テストから総合テスト（ベンダ確認）までの検出不具合数）

● 業種別設計レビュー工数密度（新規開発、主開発言語グループ、開発5工程あり）



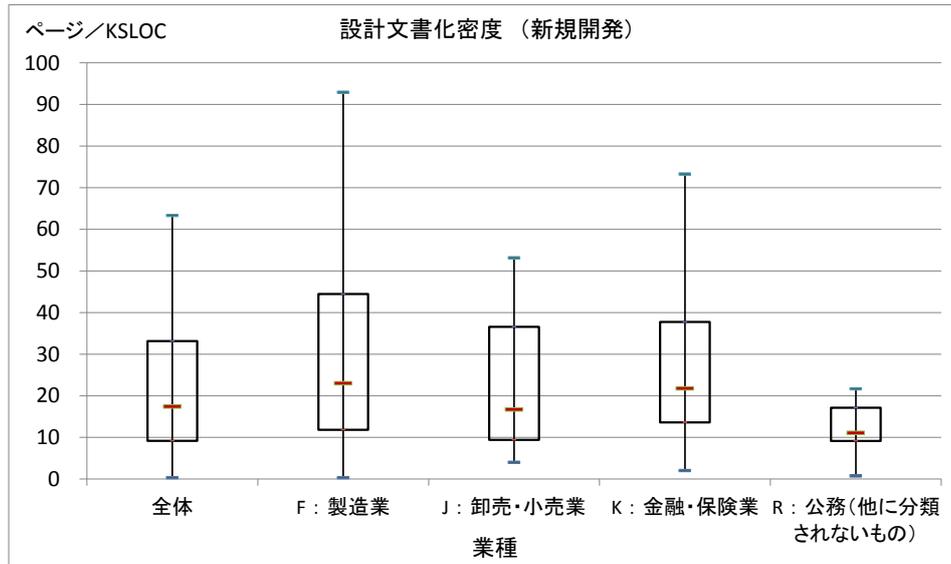
業種	N	設計レビュー工数密度 人時/KSLOC							
		最小	P25	中央	P75	最大	平均	標準偏差	
全体	104	0.08	1.73	4.13	10.96	1559.36	34.11	173.07	
F: 製造業	16	0.53	1.83	4.39	8.27	55.63	10.10	14.87	
K: 金融・保険業	39	0.08	2.92	5.49	12.20	1559.36	70.85	274.66	
R: 公務(他に分類されないもの)	23	0.20	1.20	3.36	13.74	17.54	6.74	6.39	
H: 情報通信業	5								データ不足
J: 卸売・小売業	5								データ不足

● 業種別テスト密度（新規開発、主開発言語グループ、開発5工程あり）



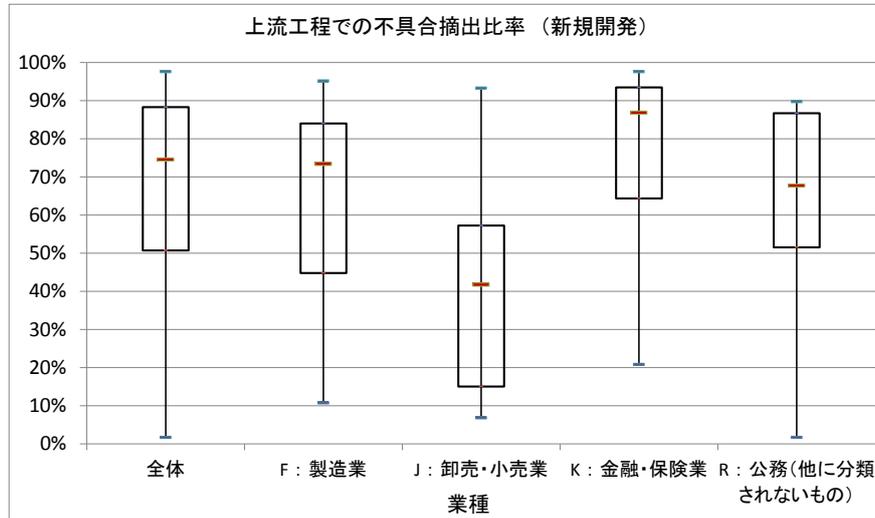
業種	N	テスト密度 (ケース/KSLOC)						
		最小	P25	中央	P75	最大	平均	標準偏差
全体	509	0.0	13.5	31.1	59.8	1703.3	65.4	146.1
F: 製造業	66	0.2	10.8	22.9	44.5	302.8	43.9	59.9
H: 情報通信業	87	0.4	21.2	41.9	68.2	970.7	64.5	111.5
J: 卸売・小売業	42	0.5	14.4	31.7	42.0	587.1	49.1	95.0
K: 金融・保険業	162	0.2	14.1	32.3	83.5	1703.3	94.1	225.2
R: 公務(他に分類されないもの)	58	0.0	12.3	35.4	49.8	698.5	59.4	101.2

● 業種別設計文書化密度（新規開発、主開発言語グループ、開発5工程あり）



業種	N	設計文書化密度（ページ/KSLOC）						
		最小	P25	中央	P75	最大	平均	標準偏差
全体	160	0.3	9.2	17.4	33.1	1527.9	39.2	130.4
F: 製造業	24	0.3	11.8	23.0	44.4	109.9	33.6	30.7
J: 卸売・小売業	14	4.0	9.4	16.7	36.6	97.3	30.5	31.5
K: 金融・保険業	68	2.0	13.6	21.7	37.7	1527.9	61.3	196.8
R: 公務（他に分類されないもの）	24	0.7	9.1	11.1	17.1	47.1	13.5	8.9
H: 情報通信業	3	4.5	6.2	7.9	17.7	27.4	13.3	12.4

● 業種別上流工程での不具合摘出比率（新規開発、主開発言語グループ、開発5工程あり）



業種	N	上流工程での不具合摘出比率（%）						
		最小	P25	中央	P75	最大	平均	標準偏差
全体	133	1.7%	50.7%	74.6%	88.3%	97.6%	65.9%	27.5%
F: 製造業	24	10.8%	44.8%	73.4%	84.0%	95.2%	63.2%	26.6%
J: 卸売・小売業	10	6.9%	15.0%	41.7%	57.2%	93.3%	40.9%	28.6%
K: 金融・保険業	41	1.8%	64.4%	86.8%	93.5%	97.6%	74.0%	27.7%
R: 公務（他に分類されないもの）	27	1.7%	51.5%	67.7%	86.7%	89.7%	62.1%	28.1%
H: 情報通信業	6	66.3%	71.3%	75.0%	83.1%	89.5%	76.9%	8.9%

(3) ベンチマーキング方法

- ① 自組織が金融・保険業の業種ドメインに該当するのであれば、妥当と評価できる範囲を、白書等の金融・保険業の各統計情報の 25 パーセンタイル (P25) から 75 パーセンタイル (P75) までの範囲とする。

<P25～P75 の範囲の例 (開発規模が SLOC で金融・保険業の新規開発の場合) >
(上記(2)ベンチマークから抜粋)

メトリクス	単位	P25	P75
SLOC 発生不具合密度	件/KSLOC	0.000	0.035
設計レビュー工数密度	人時/KSLOC	2.92	12.20
テスト密度	ケース/KSLOC	14.1	83.5
設計文書化密度	ページ/KSLOC	13.6	37.7
上流工程での不具合摘出比率	%	64.4	93.5

- ② 評価対象プロジェクトの信頼性及び品質保証プロセスに関する目標が、それぞれの P25～P75 の範囲内に収まっているか否かを判定する。ただし、品質保証プロセスに関する目標の妥当性評価については、評価対象プロジェクトの信頼性目標値の高低あるいは信頼性要求レベルに応じて、妥当と評価できる範囲を調整 (上方修正/下方修正) しながら評価することが望ましい。
- ③ 評価対象プロジェクトが P25～P75 の範囲外となる場合、評価対象プロジェクトの特性を勘案しながら、範囲外となる合理的な理由がないか検討する。合理的な理由があれば妥当と評価し、合理的な理由がなければ目標の見直しを検討する。

3.1.3. IT 企業での具体例 1

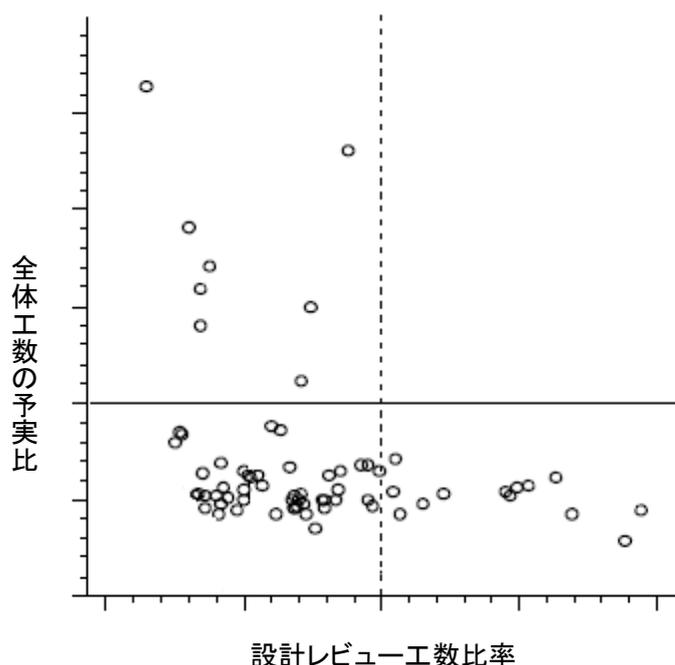
<設計レビュー工数比率の妥当性評価>

(1) 目的（活用シーン）

計画工数に対する実績工数の乖離度合い（全体工数の予実比）の観点から、プロジェクト計画における設計レビュー工数比率の妥当性を評価する。

(2) ベンチマーク

過去プロジェクトの実績に基づくベンチマークでは、プロジェクト全体の計画工数に対する実績工数の比率（全体工数の予実比）が一定以上に乖離するプロジェクトは、ある一定以下の設計レビュー工数比率となっている。



(3) ベンチマーキング方法

このベンチマークをプロジェクトにおける設計レビュー工数比率の目標を設定する際に利用し、ある一定以上の設計レビュー工数比率を目標として設定する。それ以下の設計レビュー工数比率とする場合は、プロジェクト特性に応じてその妥当性及びリスクを評価して目標を設定するようにする。

参考ポイント

設計レビュー工数比率の目標の妥当性評価については、稼働後（あるいは出荷後）の発生不具合密度との関係から妥当性評価する方法や、設計レビュー検出能率（設計レビュー指摘件数÷設計レビュー工数）との関係から妥当性評価する方法などがある。

この例では、計画工数に対する実績工数の乖離度合い（全体工数の予実比）との関係から設計レビュー工数比率の目標を妥当性評価する方法を採っており、妥当性評価のためのベンチマーキング方法の一つとして参考になると考えられる。

3.1.4. IT 企業での具体例 2

<海外ソフトハウスの品質評価基準の見直し>

(1) 目的（活用シーン）

海外ソフトハウスの納入品質を確保するため、社内の品質基準と海外ソフトハウスの実力値を比較し、プロジェクト進行中に品質基準の見直しを行いながら、品質評価及び是正処置を行う。

(2) 考え方

海外ソフトハウスの実力にばらつきがあり、社内の品質基準をクリアすることで品質確保が行えるかどうかの判断が難しい。海外ソフトハウスに委託開発する場合には、社内の品質基準を基に海外ソフトハウスの品質の作り込みレベルを判断し、品質基準の見直しを継続的に行いながら品質確保する。

(3) ベンチマーク

<検証に用いる内部ベンチマークの抽出方法>

過去のプロジェクトデータから閾値を設定した、社内のゾーンモデルを初期の品質基準として準備する。

<検証に用いるベンチマーク（メトリクス）>

- ① 欠陥密度（件数／KSLOC）
- ② テスト密度（項目／KSLOC）
- ③ テスト密度と欠陥密度の閾値の区域（ゾーン判定）

(4) ベンチマーキング方法

- ① 海外ソフトハウスの実施した第一週目のテスト結果（テスト密度と欠陥密度）をゾーンモデルに適用。良好と判断されたプログラムをサンプリングし、受入テストを実施。
- ② 不良と判断されたプログラムのテスト結果を基に、品質目標とゾーンモデルの閾値を見直し、品質基準を再設定。
見直した品質基準をクリアできないプログラムについて、追加試験を実施。
- ③ 第二週目以降のテスト結果についても、上記の①②を繰り返し、品質基準を見直しつつ品質確保を実施。

【備考】

海外ソフトハウスを例に具体例を記載したが、国内ソフトハウスにおいても実力が不明な場合には有効な事例である。ただし、ある程度規模（プログラム数、機能数など）が大きな発注の場合には有効であるが、小規模の発注の場合は、個別に品質分析を行った方が効率的に品質確保を行えると考えられる。

参考ポイント

組織の開発力／成熟度が不透明な企業に委託開発する場合には、プロジェクト計画時の品質目標及び品質保証プロセス目標の妥当性評価やテスト時の品質評価等において、社内の品質基準をそのまま適用することは適切でない可能性がある。

この例では、初期の品質基準には社内の品質基準を適用するものの、テストの途中結果に応じて品質基準を見直し（再設定）しながらベンチマーキングを繰り返すという方法を採用しており、組織の開発力／成熟度が不透明な企業に委託開発する場合において、現実的かつ実践的なベンチマーキング方法として参考になると考えられる。

3.2. 開発総工数及び生産性目標の妥当性評価

3.2.1. ベンチマーキング方法の手順例

ベンチマーキング方法の基本的な手順の例を以下に示す。

- (1) ベンチマーキングの目的を確認する。
ベンチマーキングの目的が、「工数見積りの妥当性を評価し（具体的にはプロジェクト計画における開発総工数及び生産性目標が、妥当な水準に設定されているか否かを評価し）、必要に応じて調整すること」であることを確認する。
- (2) 妥当と評価できる範囲を設定し、必要に応じて調整する。
開発総工数及び生産性に対応するベンチマーク中の統計情報を目安にして、それらの妥当と評価できる範囲を設定する。

<妥当と評価できる範囲の例>

25 パーセンタイル (P25) から 75 パーセンタイル (P75) までの範囲

散布図では信頼区間 50%の範囲に相当する。箱ひげ図では箱の範囲に相当する。

(備考) 3.1.1 「ベンチマーキング方法の手順例」の(備考1)～(備考4)と同様。3.1.1 項を参照されたい。

(注) 生産性目標について組織の目標が設定されている場合には、それに従って妥当性評価すること。

<生産性のメトリクス例>

◇FP 生産性：単位工数（人時）あたりの開発規模（FP 規模）

◇SLOC 生産性：単位工数（人時）あたりの開発規模（SLOC 規模）

- (3) 開発総工数及び生産性の目標が、妥当と評価できる範囲内に収まっているか否かを判定する。
- (4) 妥当と評価できる範囲外となる場合、評価対象プロジェクトの特性を勘案しながら、範囲外となる合理的な理由がないかどうか検討する。合理的な理由があれば妥当と評価し、合理的な理由がなければ目標の見直しを検討する。

<合理的な理由の検討（又は妥当と評価できる範囲の調整）のケース例>

◇評価対象プロジェクトの属性が生産性変動要因に該当しているケース

- ・ミッション・クリティカルであり、品質保証プロセス（文書化、レビュー、テスト等）の（実施度合いの）目標を高く設定しているケース
- ・業種が混在したサンプル集合に基づいて作成されたベンチマークを用いてベンチマーキングする場合で、評価対象プロジェクトの業種が金融・保険業に該当しているケース

◇開発規模が比較的小さい、次のようなプロジェクト

生産性メトリクスのせいで生産性が不当に低く評価されるケース

- ・流用母体（OSS、パッケージ等）をテストし信頼性を確保するケース
- ・母体（前版部分）の信頼性向上を図るケース

◇当プロジェクトで対外システムとの接続テストを強化するケース

◇対外接続システムの数が多いケース

3.2.2. 白書等を用いた具体例

(1) 目的 (活用シーン)

開発者 (開発プロジェクトのプロジェクト・マネジャー及び担当者) が、プロジェクト計画の策定時に、プロジェクト計画の実現性を高めるために、次のような観点からプロジェクト計画の妥当性を評価するシーン (開発組織のマネジャー層、PMO 及び品質マネジメント推進部門が、プロジェクト計画をレビューし助言/支援するシーンを含む)

◇工数見積りの妥当性を評価し (具体的にはプロジェクト計画における開発総工数及び生産性目標が、妥当な水準に設定されているか否かを評価し)、必要に応じて調整する。

(2) ベンチマーク

開発総工数及び生産性に関する白書等の統計情報を、目安として参考にする。

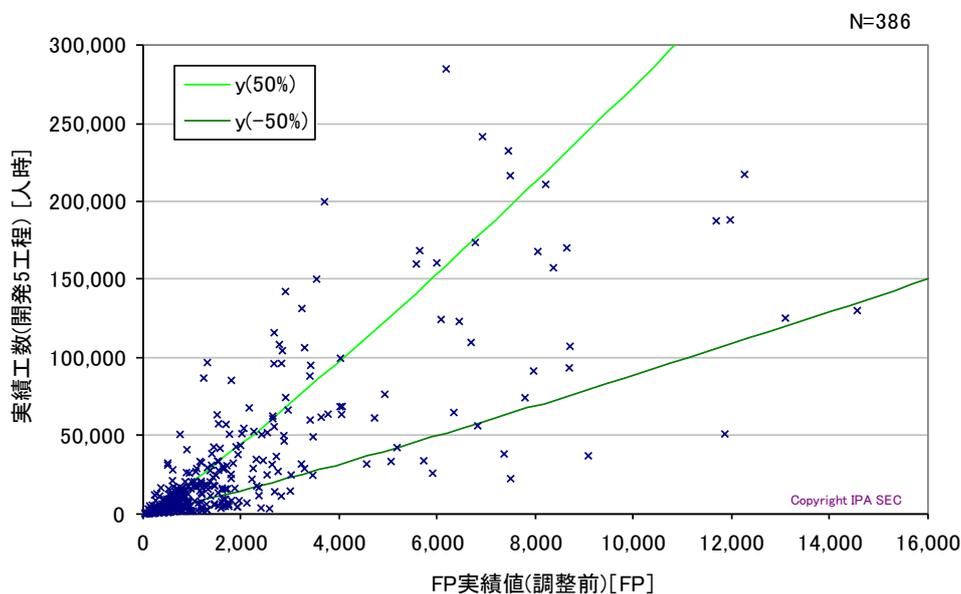
<参照箇所>

詳しくは、「ソフトウェア開発データ白書」の「FP 規模と工数」、「FP 生産性」、「SLOC 規模と工数」及び「SLOC 生産性」に関する記事を参照されたい。

なお、「ソフトウェア開発データ白書 2016-2017」では、「業種編」の形で、業種別の統計情報が追加される見込みである。

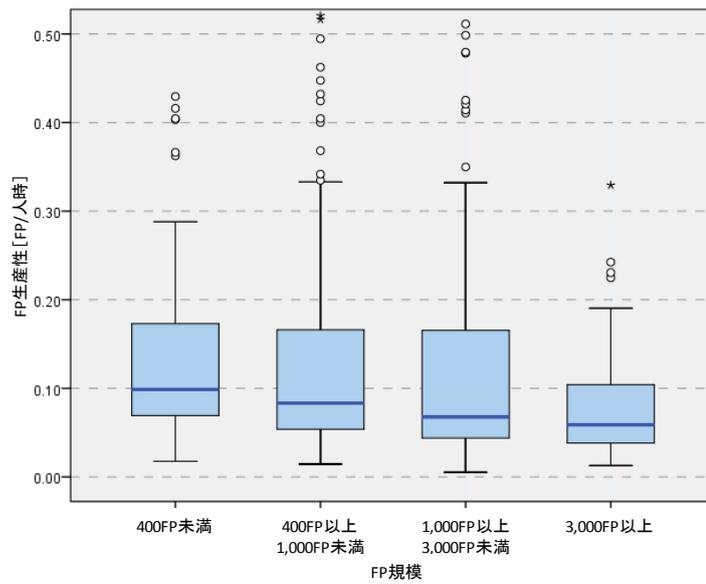
① FP 規模と工数の統計情報の例

- FP 規模と工数 (新規開発、IFPUG グループ) (信頼区間 50%付き)



② FP 生産性の統計情報の例

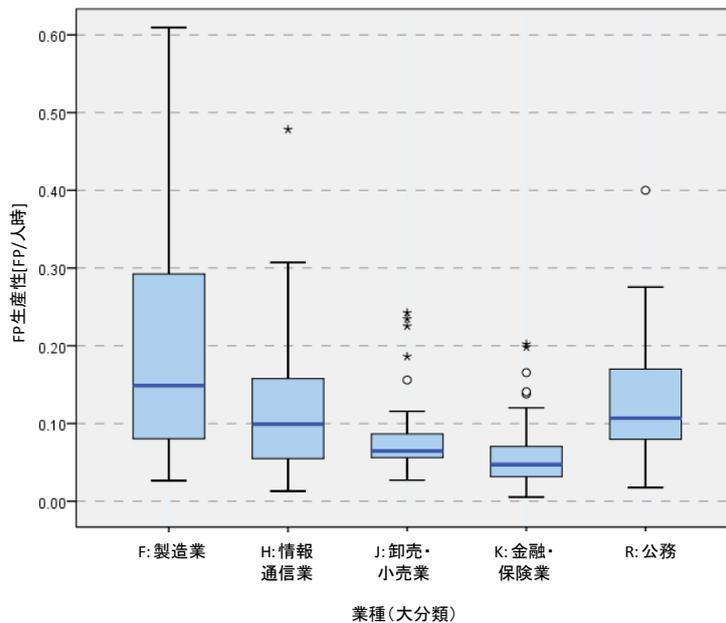
- FP 規模別 FP 生産性 (新規開発、IFPUG グループ)



[FP/人時、FP/160人時]

FP規模	単位	N	最小	P25	中央	P75	最大	平均	標準偏差
全体	FP/ 人時	386	0.005	0.050	0.079	0.154	0.739	0.124	0.118
400FP未満		84	0.018	0.070	0.099	0.172	0.429	0.129	0.098
400FP以上 1,000FP未満		124	0.014	0.054	0.083	0.164	0.521	0.134	0.120
1,000FP以上 3,000FP未満		121	0.005	0.044	0.068	0.165	0.739	0.131	0.143
3,000FP以上		57	0.013	0.038	0.059	0.104	0.329	0.082	0.064
全体	FP/ 160人時	386	0.84	7.97	12.65	24.69	118.21	19.90	18.91
400FP未満		84	2.82	11.13	15.82	27.55	68.71	20.66	15.61
400FP以上 1,000FP未満		124	2.30	8.67	13.29	26.28	83.31	21.50	19.19
1,000FP以上 3,000FP未満		121	0.84	7.03	10.81	26.47	118.21	20.96	22.94
3,000FP以上		57	2.06	6.12	9.37	16.66	52.69	13.07	10.22

● 業種別 FP 生産性 (新規開発、IFPUG グループ)

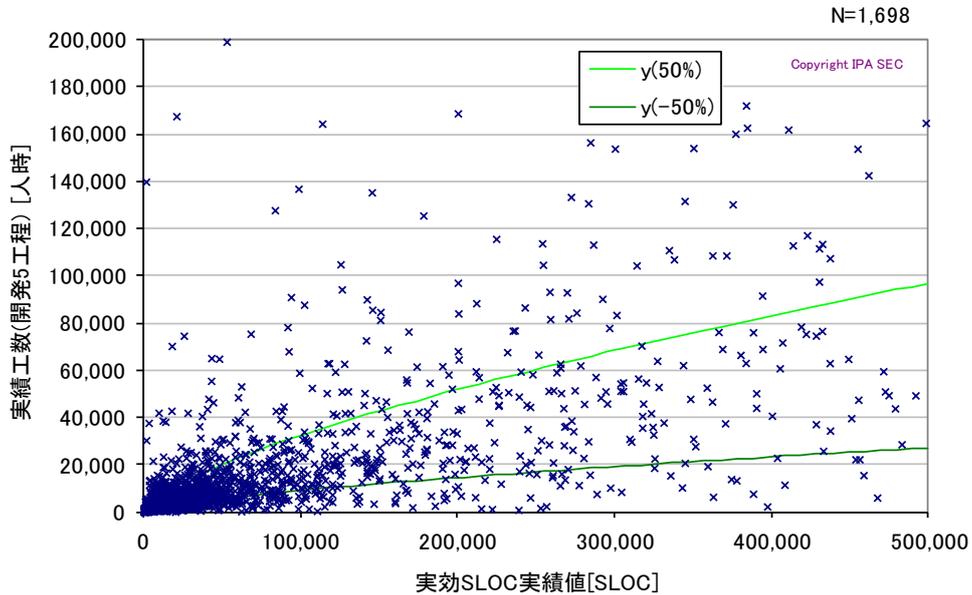


[FP/人時]

業種 (大分類)	N	最小	P25	中央	P75	最大	平均	標準偏差
F: 製造業	84	0.026	0.081	0.148	0.292	0.609	0.196	0.143
H: 情報通信業	29	0.013	0.055	0.099	0.158	0.478	0.120	0.100
J: 卸売・小売業	38	0.027	0.056	0.064	0.086	0.739	0.100	0.119
K: 金融・保険業	97	0.005	0.031	0.047	0.070	0.202	0.057	0.038
R: 公務 (他に分類されないもの)	25	0.018	0.080	0.107	0.170	0.400	0.129	0.089

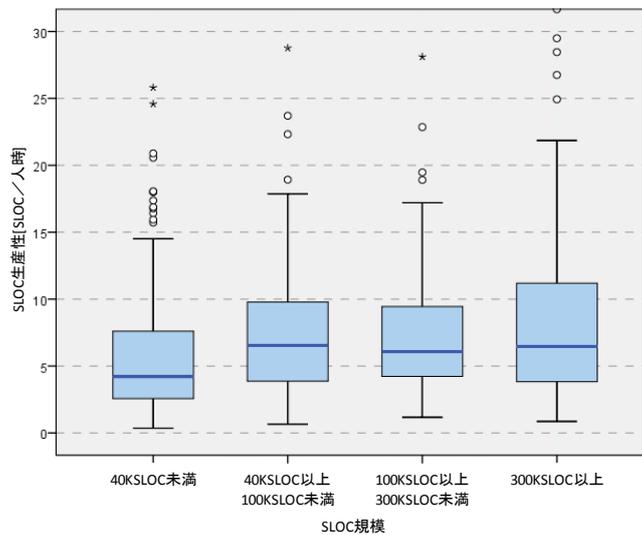
③ SLOC 規模と工数の統計情報の例

- SLOC 規模と工数（全開発種別、主開発言語混在）（信頼区間 50%付き）
拡大図（SLOC 規模 ≤ 500,000 & 工数 ≤ 200,000）



④ SLOC 生産性の統計情報の例

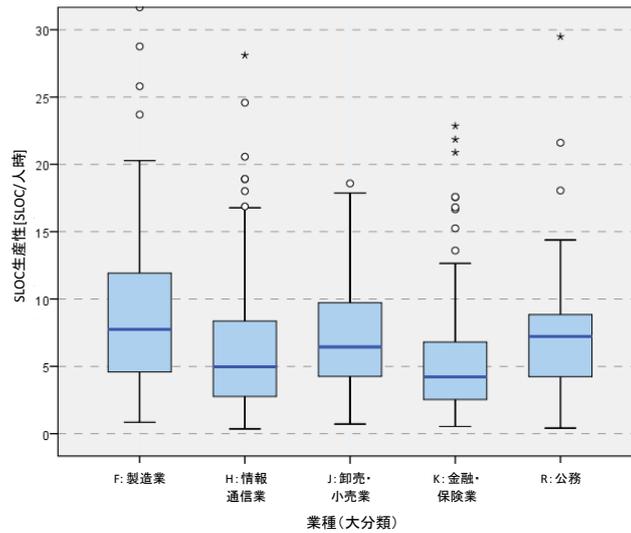
- SLOC 規模別 SLOC 生産性（新規開発、主開発言語グループ）



[SLOC/人時、KSLOC/160人時]

SLOC規模	単位	N	最小	P25	中央	P75	最大	平均	標準偏差
全体	SLOC/ 人時	630	0.4	3.4	5.6	9.0	82.6	7.8	8.3
40KSLOC未満		218	0.4	2.6	4.2	7.6	44.7	5.8	5.3
40KSLOC以上100KSLOC未満		139	0.7	3.9	6.5	9.8	71.7	8.6	8.9
100KSLOC以上300KSLOC未満		142	1.2	4.2	6.1	9.4	41.6	8.0	6.7
300KSLOC以上		131	0.9	3.8	6.5	11.2	82.6	10.1	12.1
全体	KSLOC/ 160人時	630	0.06	0.54	0.90	1.43	13.21	1.25	1.34
40KSLOC未満		218	0.06	0.41	0.67	1.22	7.15	0.93	0.84
40KSLOC以上100KSLOC未満		139	0.10	0.62	1.05	1.57	11.47	1.38	1.42
100KSLOC以上300KSLOC未満		142	0.19	0.68	0.97	1.50	6.65	1.28	1.07
300KSLOC以上		131	0.14	0.61	1.03	1.79	13.21	1.62	1.93

● 業種別 SLOC 生産性（新規開発、主開発言語グループ）



[SLOC/人時]								
業種（大分類）	N	最小	P25	中央	P75	最大	平均	標準偏差
F：製造業	91	0.8	4.6	7.7	11.9	76.1	10.6	10.6
H：情報通信業	99	0.4	2.8	5.0	8.4	33.5	6.7	6.0
J：卸売・小売業	55	0.7	4.3	6.4	9.7	82.6	9.0	11.7
K：金融・保険業	200	0.5	2.5	4.2	6.8	71.7	6.2	8.1
R：公務（他に分類されないもの）	64	0.4	4.3	7.2	8.8	29.5	7.4	4.7

(3) ベンチマーキング方法

- ① 妥当と評価できる範囲を、白書等の各統計情報の 25 パーセンタイル (P25) から 75 パーセンタイル (P75) までの範囲とする。

業種別の統計情報がある場合には、自組織が情報通信業の業種ドメインに該当するのであれば、妥当と評価できる範囲を情報通信業の各統計情報の 25 パーセンタイル (P25) から 75 パーセンタイル (P75) までの範囲とする。

<P25～P75 の範囲の例（開発規模が SLOC で金融・保険業の新規開発の場合）>
 (上記(2)ベンチマークから抜粋)

メトリクス	単位	P25	P75
FP 生産性	FP/人時	0.031	0.070
SLOC 生産性	SLOC/人時	2.5	6.8

- ② 評価対象プロジェクトの開発総工数及び生産性に関する目標が、それぞれの P25～P75 の範囲内に収まっているか否かを判定する。ただし、評価対象プロジェクトに該当する生産性変動要因を勘案して、妥当と評価できる範囲を調整（上方修正/下方修正）しながら評価することが望ましい。生産性変動要因については、白書等の知見をヒントとして参考にされたい。
- ③ 評価対象プロジェクトが P25～P75 の範囲外となる場合、評価対象プロジェクトの特性を勘案しながら、範囲外となる合理的な理由がないかどうか検討する。合理的な理由があれば妥当と評価し、合理的な理由がなければ目標の見直しを検討する。

3.2.3. IT 企業での具体例 1

<信頼性要求水準に着目した工数予測モデルの構築>

プロジェクトデータリポジトリのデータから工数予測モデルを構築するにあたり、信頼性要求水準に着目することにより、予測精度を向上することができた事例である。

(1) 目的（活用シーン）

プロジェクト計画における工数計画値が妥当な水準に設定されているか否かを、信頼性要求水準で層別した工数予測モデルを用いた予測工数と比較することによって評価し、必要に応じて調整する。

(2) 考え方

予測精度向上に向けて、信頼性要求水準で層別した工数予測モデルを構築するに至った考え方及び経緯を以下に示す。

① 工数予測モデルの構築と活用

まず、当社で構築している、工数予測モデルについて説明する。このモデルは、当社のプロジェクトデータリポジトリ中のプロジェクトデータを層別、「ソフトウェア開発データ白書 2014-2015」（以下データ白書）での分析と同様に、開発規模(SLOC)と開発工数(人時)を対数スケールに変換した上で回帰分析を行い、その回帰式：

$$\log(\text{開発工数}) = A \cdot \log(\text{開発規模}) + B$$

を 50% 予測区間³と共に線形スケールの開発規模-開発工数の散布図上に表示したものである。なお、線形スケールでの回帰式は以下の回帰曲線となる。

$$\text{開発工数} = \alpha \cdot \text{開発規模}^\beta \quad (\alpha = \exp(B), \beta = A)$$

次ページの図 3.2.3-1 に 2014 年度データで作成したモデルを示す。このモデルは、プロジェクトの開発工数計画値の妥当性判断の際に利用することを想定している。このモデルを用いて、見積規模に対する開発工数の 50% 予測区間内に開発工数計画値が入っていれば、妥当性が高いという目安にする。

このモデルの相関係数 R は 0.81 であり、データ白書の回帰分析の結果と比較しても遜色はない。しかしながら、実際のプロジェクトデータの中には、50% 予測区間の上限からも大きく乖離したグループ(図 3.2.3-1 の赤丸部分)が見られるという課題があった。

³データ白書では、「信頼区間」の語を用いているが、回帰式の信頼区間(“Confidence interval”)と区別のため、「予測区間」(“Prediction interval” または “Predictive confidence interval”)の語を用いる。

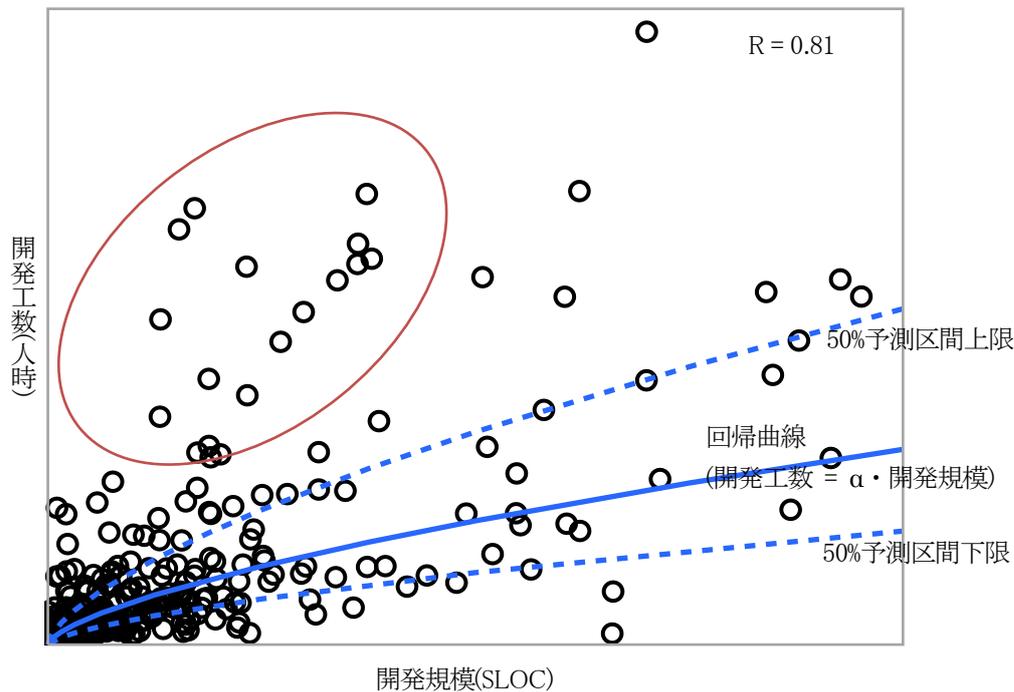


図 3.2.3-1 工数予測モデル

② 信頼性要求水準の予測モデルへの取り込み

①で述べた課題を解決するため、信頼性要求水準に着目した。信頼性要求水準としては、データ白書の「重要インフラ情報システムのシステムプロファイル」に準拠した尺度、「Type I」（社会的影響がほとんどない）～「Type IV」（人命に影響、甚大な経済損失）を使用、うち、「Type I」と「Type II」を「通常信頼性要求水準」、「Type III」と「Type IV」を「高度信頼性要求水準」と2つの層に分けて分析した。

図 3.2.3-2 に信頼性要求水準で層別した開発規模-開発工数の散布図を示す。これを見ると、図 3.2.3-1 の工数予測モデルの 50% 予測区間上限からも大きく乖離したグループのほとんどが「高度信頼性要求水準」のプロジェクトであることが分かる。そこで、信頼性要求水準を変数に加えた重回帰分析を行い、図 3.2.3-1 と同様にその回帰式：

$$\begin{aligned} \log(\text{開発工数}) &= A' \cdot \log(\text{開発規模}) + B'_{\text{通常}} && \text{通常信頼性要求水準} \\ \log(\text{開発工数}) &= A' \cdot \log(\text{開発規模}) + B'_{\text{高度}} && \text{高度信頼性要求水準} \end{aligned}$$

※A'は信頼性要求水準によらない共通定数、B'通常、B'高度は定数

を 50% 予測区間と共に線形スケールに変換して図 3.2.3-2 に表示した。

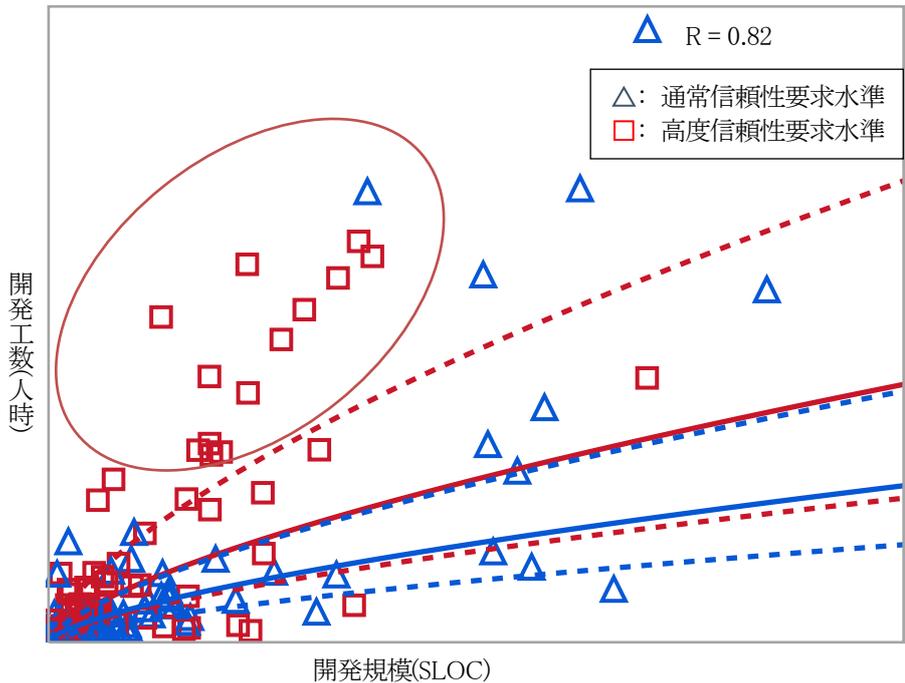


図 3.2.3-2 信頼性要求水準を考慮した工数予測モデル

なおこの式は以下の 1 つの重回帰式を見やすく変形したものである。
 以下のモデルの相関係数 R は 0.82 である。

$$\log(\text{開発工数}) = A' \cdot \log(\text{開発規模}) + (B'_{\text{高度}} - B'_{\text{通常}}) \text{信頼性要求水準} + B'_{\text{通常}}$$

信頼性要求水準は、通常信頼性要求水準の時は 1、高度信頼性要求水準の時は 0 の値をとる
 ダミー変数

線形スケールでの回帰式は以下の通りとなる。

$$\begin{aligned} \text{開発工数} &= \alpha'_{\text{通常}} \cdot \text{開発規模}^{\beta'} & (\alpha'_{\text{通常}} = \exp(B'_{\text{通常}}), \beta' = A') & \text{通常信頼性要求水準} \\ \text{開発工数} &= \alpha'_{\text{高度}} \cdot \text{開発規模}^{\beta'} & (\alpha'_{\text{高度}} = \exp(B'_{\text{高度}}), \beta' = A') & \text{高度信頼性要求水準} \end{aligned}$$

※指数 β' は信頼性要求水準によらず共通

この信頼度要求水準を加えた工数予測モデルを見ると、高度信頼性要求水準のプロジェクトは通常信頼性要求水準と比較して、同じ開発規模であっても多く開発工数が掛かっていることが分かる。

(3) ベンチマーク

信頼性要求水準で層別した次の工数予測モデルの 50% 予測区間を、ベンチマークとして採用する。

$$\begin{aligned} \text{開発工数} &= \alpha'_{\text{通常}} \cdot \text{開発規模}^{\beta'} & (\alpha'_{\text{通常}} = \exp(B'_{\text{通常}}), \beta' = A') & \text{通常信頼性要求水準} \\ \text{開発工数} &= \alpha'_{\text{高度}} \cdot \text{開発規模}^{\beta'} & (\alpha'_{\text{高度}} = \exp(B'_{\text{高度}}), \beta' = A') & \text{高度信頼性要求水準} \end{aligned}$$

※指数 β' は信頼性要求水準によらず共通

(4) ベンチマーキング方法

この信頼性要求水準を加えた予測モデルを用いて、開発するシステムの信頼性要求水準に応じた見積規模に対する開発工数の50%予測区間内に開発工数計画値が入っていれば、妥当性が高いという目安にする。

50%予測区間内に開発工数計画値が入っていない場合、評価対象プロジェクトの特性を勘案しながら、範囲外となる合理的な理由がないかどうか検討する。合理的な理由があれば妥当と評価し、合理的な理由がなければ目標の見直しを検討する。

参考ポイント

妥当性評価の精度を高めるには、業種別、アプリケーション種別、品質要求レベル等によって層別したベンチマークを用意して、評価対象プロジェクトが該当するカテゴリのベンチマークを用いて妥当性評価することが望ましい。

この例は、品質要求レベル（信頼性要求水準）によって層別したベンチマークを採用した、妥当性評価のベンチマーキング方法の事例として参考になると考えられる。

3.2.4. IT 企業での具体例 2

<開発規模別の工数見積りの妥当性評価>

(1) 目的（活用シーン）

プロジェクト計画における工数計画値が妥当な水準に設定されているか否かを、開発規模で層別した生産性のベンチマークを用いて評価し、必要に応じて調整する。

(2) 考え方

工数見積りに使用する生産性は、プロジェクト特性などで層別されていることが少なく、各組織で1つの指標値が設定されることが多い。そこで生産性の変動要因の一つである「開発規模」に着目して、開発規模別の生産性を用いた工数見積りの妥当性評価の例を示す。

SI 開発においては開発規模により変動する作業（間接工数など）と変動が少ない作業（どの開発規模でも必要となる作業）があり、小規模開発では変動が少ない作業の工数が影響して生産性が低くなる傾向になるため、開発規模が小さい時の工数見積りには特に注意が必要である。

規模は、小規模、中規模、大規模（、超大規模）などの3~4つの分類で層別する。各会社、組織で開発規模の定義（分類）がある場合は、その定義に従う。定義が存在しない場合は、規模の均等割り、分析に使用する標本数の均等割りなどの方法を検討する。

(3) ベンチマーク

生産性の変動要因の一つである規模に着目し、規模別に層別した指標を設定する。

まず、自組織の生産性の基本統計量を導出する。また、自組織の開発規模別の生産性を分析し、小規模開発の生産性が低く、ある規模での生産性が高く（ピークに）なる傾向が見られることを確認する。

[メトリクス]

生産性：開発規模 ÷（開発 5 工程の工数）（SLOC/人月）

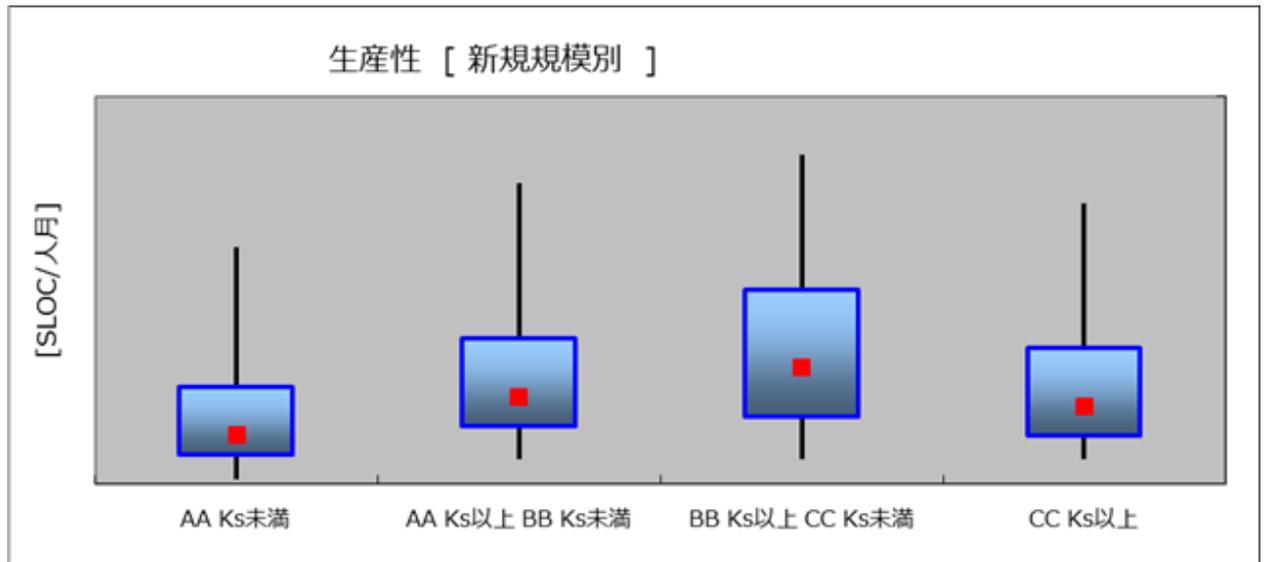
[プロジェクトデータの検索条件等]

- ・ 開発 5 工程が揃っているプロジェクト
- ・ 開発プロジェクトの種別が新規開発
- ・ 開発規模（SLOC 規模） > 0
- ・ 開発 5 工程の総工数 > 0

[分析結果（ベンチマーク）]

◇新規開発の場合の箱ひげ図（規模は AA<BB<CC）

開発規模が AA（KSLOC）以下の生産性は他規模の生産性に比べて低い傾向にあり、開発規模が BB（KSLOC）以上 CC（KSLOC）未満の生産性が一番高い傾向にある。



(4) ベンチマーキング方法

評価対象プロジェクトの開発規模と生産性を、自組織の規模別生産性における該当する規模カテゴリの中央値と比較することによって妥当性評価する。

中央値以上の場合には、開発規模以外の生産性変動要因が影響していないかどうか確認しながら、中央値以上となる合理的な理由がないか検討する。合理的な理由があれば妥当と評価し、合理的な理由がなければ目標の見直しを検討する。

本例のベンチマーキング方法を適用するにあたっては、3.2.5 項 「プロジェクト特性に着目した工数見積りの妥当性評価」の方法を併用することをお勧めする。

参考ポイント

妥当性評価の精度を高めるには、自組織の主要な生産性変動要因によって層別したベンチマークを用意し、評価対象プロジェクトが該当するカテゴリのベンチマークを用いて妥当性評価することが望ましい。(生産性変動要因の例としては、業種別、アプリケーション種別、品質要求レベル等が挙げられる。)

この例は、生産性の変動要因の一つである規模に着目し、開発規模によって層別したベンチマークを採用した、妥当性評価のベンチマーキング方法の事例として参考になると考えられる。

ただし、開発規模が生産性変動要因としてどのように作用しているかについては、必ずしも一定の傾向が見られるとは限らない。(小規模開発の生産性が低い傾向は共通的に見られる傾向であるが、大規模開発の生産性が低いかどうか、またどの程度大規模であれば生産性が低くなるかについては定かとは言えない。)従って、当例と同様なベンチマーキング方法を採用するにあたっては、開発規模と生産性との関係を自組織において把握した上で実施する必要がある。

3.2.5. IT 企業での具体例 3

<プロジェクト特性に着目した工数見積りの妥当性評価>

(1) 目的（活用シーン）

プロジェクト計画における工数計画値が妥当な水準に設定されているか否かを、プロジェクト特性でグループ分けした生産性のベンチマークを用いて評価し、必要に応じて調整する。

(2) 考え方

生産性は、プロジェクト特性で大きく変化するため、見積り時、評価時にはこの特性を考慮することが必要となる。そこで見積もった生産性の妥当性評価にプロジェクト特性を考慮した活用例を示す。

X 軸を規模、Y 軸を生産性の散布図を作成し、規模とプロジェクト特性で散布図をエリア分けする。規模は、小規模、中規模、大規模（、超大規模）などの3~4つの分類で層別する。各会社、組織で開発規模の定義がある場合は、その定義に従う。その他の方法としては規模の均等割り、分析に使用する標本数の均等割りなどの方法がある。

生産性は、生産性の四分位（P25、P75 等）等を用いて3~4つに分割する。それぞれのエリアでプロジェクトの特性を洗い出し、同じ特性が同じグループになるように四分位の区切りを調整する。

(3) ベンチマーク

自組織の生産性の基本統計量を導出する。また、自組織の X 軸を開発規模、Y 軸を生産性とした散布図を作成する。次に、生産性の変動要因の一つであるプロジェクト特性（定性情報）に着目し、特性毎にグループ化し、それぞれのグループのエリア（生産性範囲）を設定する。

[メトリクス]

生産性：開発規模 ÷（開発 5 工程の工数）（SLOC／人月）

[プロジェクトデータの検索条件等]

- ・ 開発 5 工程が揃っているプロジェクト
- ・ 開発プロジェクトの種別が新規開発
- ・ 開発規模（SLOC 規模） > 0
- ・ 開発 5 工程の総工数 > 0

[分析結果（ベンチマーク）]

◇新規開発の場合の散布図

各グループの特性例は、以下の通りである。

【A グループ】

- ・ 業種／業務ノウハウを保有する類似システム開発経験者による再利用。業務に特化したパッケージ適用プロジェクトが多く、開発規模が小さい傾向にある。

【B グループ】

- ・ 類似処理のパターン化／共通化による、プロジェクト内再利用。1画面当りの処理量が少ない傾向にある。

【C グループ】

- ・ 開発規模が小さく、システム特性から再利用が行いにくい。業務、処理内容が複雑な傾向にある。

【D グループ】

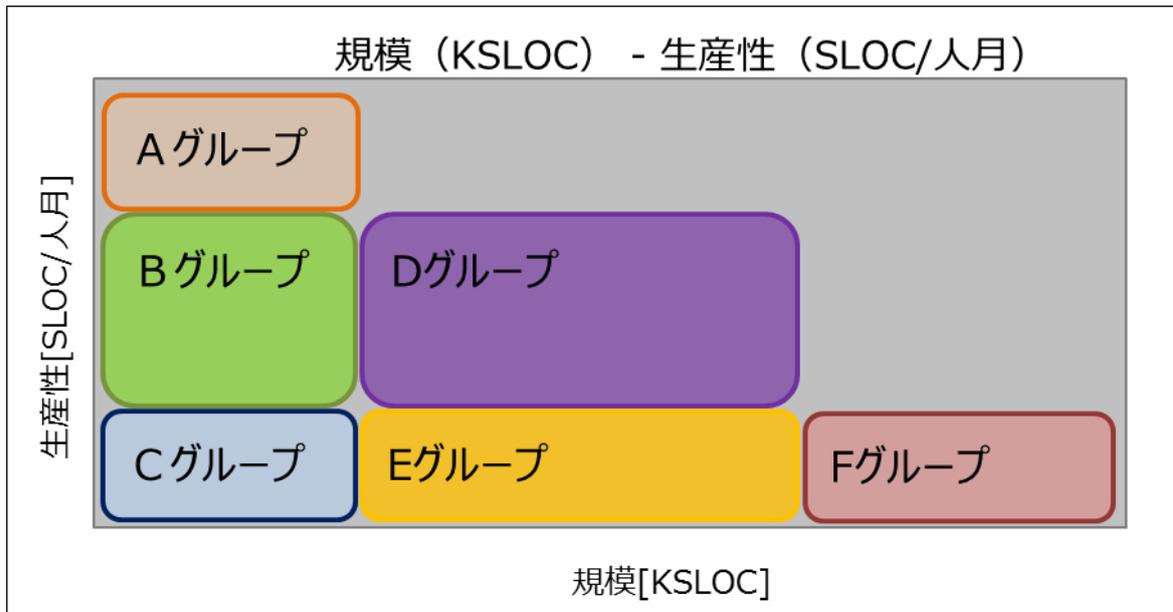
- ・業種／業務ノウハウを保有する類似システム開発経験者による再利用。パッケージ適用プロジェクトが多い傾向にある。

【Eグループ】

- ・業務、処理内容が複雑な傾向にある。1画面当りの処理量が多い傾向にある。

【Fグループ】

- ・超大規模かつ複雑なシステムで、プロジェクトの開発期間が長い。



(4) ベンチマーキング方法

評価対象プロジェクトの開発規模と生産性を散布図上にプロットして、どのエリアにプロットされるかを確認する。プロットされたエリアに在るグループのプロジェクト特性を確認し、評価対象プロジェクトと同様であるかどうかを確認することにより見積値の妥当性評価を行う。

評価対象プロジェクトの特性とグループの特性とが異なる場合には、その要因を洗い出し理由を明確にした上で、見積り結果を見直すことを検討する。

参考ポイント

妥当性評価の精度を高めるには、自組織の主要な生産性変動要因によって層別したベンチマークを用意し、評価対象プロジェクトが該当するカテゴリのベンチマークを用いて妥当性評価することが望ましい。(生産性変動要因の例としては、業種別、アプリケーション種別、品質要求レベル等が挙げられる。) この例は、生産性の変動要因の一つであるプロジェクト特性(定性情報)に着目し、プロジェクト特性によって層別したベンチマークを採用した、妥当性評価のベンチマーキング方法の事例として参考になると考えられる。

3.2.6. IT 企業での具体例 4

<マネジャー層による見積り結果の妥当性の検証>

(1) 目的（活用シーン）

自組織での実績／経験がないプロジェクト特性（開発言語など）の開発における見積り精度の向上を目的に、マネジャー層による見積り結果の妥当性の検証を行う。

(2) 考え方

プロジェクトの実績データ（他組織の実績も含む）から導出した内部ベンチマークに基づき、マネジャー層が見積り結果の妥当性を検証する。

(3) ベンチマーク

<検証に用いる内部ベンチマークの抽出方法>

- ① 検証対象のプロジェクト特性を考慮し、プロジェクト完了実績 DB から実績データを抽出する。
プロジェクト特性には、データ白書の PJ プロフィールに類似する項目を用意しているが、主に使用される特性としては、「業種（社内の事業分野）」「規模」「開発言語」があげられる。
- ② 抽出した実績データから、特殊事情で遂行するプロジェクト、失敗などにより QCD が参考にならないプロジェクトなど、特異なプロジェクトを除外して精査する。
- ③ 精査した実績データを元に、内部ベンチマークを導出する。内部ベンチマークを表す方法としては、白書と同様に「箱ひげ図」及び「散布図」を用いる。

<検証に用いるベンチマーク>

次のメトリクスの統計情報をベンチマークとして採用する。

- ・規模に対する工数（人時／KSLOC）
- ・規模に対する工期（日／KSLOC）
- ・ライン生産効率（SLOC／人時）
- ・ステップ別工数比率（%）など

(4) ベンチマーキング方法（妥当性の検証方法）

- ① マネジャー層は、見積り結果を内部ベンチマークと比較して差異があるかどうかを確認する。
- ② 差異がある場合、見積り担当者が合理的な見積り根拠（差異の理由）を説明できるかどうかをインタビュー形式で確認し、見積り結果の妥当性を検証する。合理的な理由を説明できれば妥当と評価し、合理的な理由を説明できなければ目標の見直しを検討する。

<インタビューによる確認項目の例>

- ・ 見積りの過程
見積り手法において分析的な過程（見積り対象のブレイクダウン）が存在するか？
- ・ リスク評価の織り込み
リスク評価を反映し、コンティンジェンシー予備が適切に見積られているか？

参考ポイント

この例は、見積りとベンチマークとの差異について、マネジャー層がインタビューすることによって差異に合理的理由があるかどうかを確認するというベンチマーキング方法の事例として参考になると考えられる。

3.3. 成果物量及び単位成果物量あたりの工数の妥当性評価

3.3.1. ベンチマーキング方法の手順例

ベンチマーキング方法の基本的な手順の例を以下に示す。

(1) ベンチマーキングの目的を確認する。

ベンチマーキングの目的が、「成果物量に着目して工程ごとに工数見積りの妥当性を評価し（具体的にはプロジェクト計画における各工程の成果物量及び単位成果物量あたりの工数が、妥当な水準に設定されているか否かを評価し）、必要に応じて調整すること」であることを確認する。

(備考) 成果物量に着目して工程ごとに工数見積りの妥当性を評価する方法のメリット

[メリット1] 成果物量には評価対象プロジェクトのスコープ、品質要求、制約条件等がある程度反映されるので、評価精度が高まる。（一方、開発規模から総開発工数の妥当性を評価する方法では、これらを反映できない。）

[例1] 評価対象プロジェクトでは品質要求が高いため、設計書ページ数及びテストケース数を多く設定する。

[例2] 評価対象プロジェクトでは関連システムとの接続テストを強化する必要があるため、総合テスト工程のテストケース数を多く設定する。

[メリット2] 工程ごとの成果物量および工数を吟味することになるので、開発計画の実現可能性を高めることに繋がる。

(2) 妥当と評価できる範囲を設定し、必要に応じて調整する。

ベンチマーク中の次の統計情報を目安にして、次の統計量について妥当と評価できる範囲を設定する。

- ・各工程の開発規模あたりの成果物量（各工程の成果物量÷開発規模）
- ・各成果物の単位成果物量あたりの工数（各工程の工数÷各工程の成果物量）

(注) これらの内部ベンチマークを整備するにあたっては、前提として次の確認が必要である。

- ・自組織の開発作業標準において、各工程の主要成果物が規定されており、各成果物の構成内容や記述要領が標準化されていること。
- ・開発規模と各工程の成果物量との間、及び各工程の成果物量と各工程の工数との間に、強い正相関が認められること。

<妥当と評価できる範囲の例>

25パーセンタイル (P25) から 75パーセンタイル (P75) までの範囲

散布図では信頼区間 50%の範囲に相当する。箱ひげ図では箱の範囲に相当する。

(備考) 3.1.1 「ベンチマーキング方法の手順例」の(備考1)～(備考4)と同様。3.1.1 項を参照されたい。

(注) 組織の目標が設定されている場合には、それに従って妥当性評価すること。

(3) 各工程の開発規模あたりの成果物量及び各成果物の単位成果物量あたりの工数が、それぞれの妥当と評価できる範囲内に収まっているか否かを判定する。

(4) 妥当と評価できる範囲外となる場合、評価対象プロジェクトの特性を勘案しながら、範囲外となる合理的な理由がないか検討する。合理的な理由があれば妥当と評価し、合理的な理由がなければ目標の見直しを検討する。

<合理的な理由の検討（又は妥当と評価できる範囲の調整）のケース例>

◇評価対象プロジェクトの属性が生産性変動要因に該当しているケース

- ・ミッション・クリティカルであり、品質保証プロセス（文書化、レビュー、テスト等）の（実施度合いの）目標を高く設定しているケース
- ・業種が混在したサンプル集合に基づいて作成されたベンチマークを用いてベンチマーキングする場合で、評価対象プロジェクトの業種が金融・保険業に該当しているケース

◇開発規模が比較的小さい、次のようなプロジェクト

生産性メトリクスのおかげで生産性が不当に低く評価されるケース

- ・流用母体（OSS、パッケージ等）をテストし信頼性を確保するケース
- ・母体（前版部分）の信頼性向上を図るケース

◇当プロジェクトで対外システムとの接続テストを強化するケース

◇対外接続システムの数が多いケース

3.3.2. 白書等を用いた具体例

(1) 目的（活用シーン）

開発者（開発プロジェクトのプロジェクト・マネジャー及び担当者）が、プロジェクト計画の策定時に、プロジェクト計画の実現性を高めるために、次のような観点からプロジェクト計画の妥当性を評価するシーン（開発組織のマネジャー層、PMO 及び品質マネジメント推進部門が、プロジェクト計画をレビューし助言／支援するシーンを含む）

◇成果物量に着目して工程ごとに工数見積りの妥当性を評価し（具体的にはプロジェクト計画における各工程の成果物量及び単位成果物量あたりの工数が、妥当な水準に設定されているか否かを評価し）、必要に応じて調整する。

(2) ベンチマーク

工程別の成果物量と工数に関する白書等の統計情報を、目安として参考にする。
白書等では、各工程の成果物量として次のものを採用している。

開発工程	成果物量（実績）
要件定義	要件定義書ページ数
基本設計	基本設計書ページ数
詳細設計	詳細設計書ページ数
製作	コード行数（SLOC）
結合テスト	結合テストケース数
総合テスト（ベンダ確認）	総合テスト（ベンダ確認）ケース数

新規開発の場合、開発規模（FP 規模又は SLOC 規模）と各工程の成果物量との間、及び各工程の成果物量と各工程の工数との間には、強い（又は中程度の）正相関が見られる。また、分析対象プロジェクトを特定の業種、例えば金融・保険業のプロジェクトに絞ると、より強い相関が見られる。

（備考）これらの相関は、「ソフトウェア開発データ白書」のデータにおいて、業種を始め様々なプロファイルのプロジェクトデータが混在した中でも見られる傾向であることから、各組織において同種のソフトウェア開発を行うドメインの中では、より強い相関が見られると考えられる。

<参照箇所>

詳しくは、「ソフトウェア開発データが語るメッセージ」の「成果物量のコントロールと見積り妥当性評価」に関する記事を参照されたい。

なお、「ソフトウェア開発データ白書 2016-2017」では、「工程別の成果物量と工数」として同様の統計情報が追加される見込みである。

① 各工程の、開発規模あたりの成果物量及び成果物量あたりの工数の中央値の例

●開発規模あたりの成果物量及び成果物量あたりの工数の中央値の一覧（新規開発、主開発言語グループ）

開発工程	要件定義	基本設計	詳細設計	製作	結合テスト	総合テスト (ベンダ確認)
データ数	74	137	131	535	324	328
開発規模当りの成果物量の中央値	KSLOC当りの要件定義書ページ数 (ページ/KSLOC)	KSLOC当りの基本設計書ページ数 (ページ/KSLOC)	KSLOC当りの詳細設計書ページ数 (ページ/KSLOC)		KSLOC当りの結合テストケース数 (ケース/KSLOC)	KSLOC当りの総合テストケース数 (ケース/KSLOC)
	1.22	6.39	12.10		32.20	9.00
成果物量当りの工数の中央値	要件定義書ページ当りの要件定義工数 (人時/ページ)	基本設計書ページ当りの基本設計工数 (人時/ページ)	詳細設計書ページ当りの詳細設計工数 (人時/ページ)	KSLOC当りの製作工数 (人時/KSLOC)	結合テストケース当りの結合テスト工数 (人時/ケース)	総合テストケース当りの総合テスト工数 (人時/ケース)
	10.90	4.57	2.49	52.20	1.04	2.39

●開発規模あたりの成果物量及び成果物量あたりの工数の中央値の一覧（新規開発、IFPUGグループ）

開発工程	要件定義	基本設計	詳細設計	製作	結合テスト	総合テスト (ベンダ確認)
データ数	86	127	93	184	192	229
開発規模当りの成果物量の中央値	FP当りの要件定義書ページ数 (ページ/FP)	FP当りの基本設計書ページ数 (ページ/FP)	FP当りの詳細設計書ページ数 (ページ/FP)	FP当りのKSLOC実績値 (KSLOC/FP)	FP当りの結合テストケース数 (ケース/FP)	FP当りの総合テストケース数 (ケース/FP)
	0.118	0.558	1.050	0.075	1.830	0.663
成果物量当りの工数の中央値	要件定義書ページ当りの要件定義工数 (人時/ページ)	基本設計書ページ当りの基本設計工数 (人時/ページ)	詳細設計書ページ当りの詳細設計工数 (人時/ページ)	KSLOC当りの製作工数 (人時/KSLOC)	結合テストケース当りの結合テスト工数 (人時/ケース)	総合テストケース当りの総合テスト工数 (人時/ケース)
	9.43	4.46	2.37	71.30	1.15	3.00

② 各工程の、開発規模あたりの成果物量及び成果物量あたりの工数の基本統計量の例

●KSLOCあたりの要件定義書ページ数（新規開発、主開発言語グループ）

							(ページ/KSLOC)
N	最小	P25	中央	P75	最大	平均	標準偏差
74	0.02	0.57	1.22	2.78	21.41	2.37	3.18

●KSLOCあたりの基本設計書ページ数（新規開発、主開発言語グループ）

							(ページ/KSLOC)
N	最小	P25	中央	P75	最大	平均	標準偏差
137	0.28	2.98	6.39	12.24	384.36	12.27	33.86

●KSLOCあたりの詳細設計書ページ数（新規開発、主開発言語グループ）

							(ページ/KSLOC)
N	最小	P25	中央	P75	最大	平均	標準偏差
131	0.4	6.3	12.1	23.5	1143.6	32.5	112.9

●KSLOCあたりの結合テストケース数（新規開発、主開発言語グループ）

							(ケース/KSLOC)
N	最小	P25	中央	P75	最大	平均	標準偏差
324	0.1	13.6	32.2	54.0	2115.3	60.7	162.9

●KSLOCあたりの総合テスト（ベンダ確認）ケース数（新規開発、主開発言語グループ）

							(ケース/KSLOC)
N	最小	P25	中央	P75	最大	平均	標準偏差
328	0.02	2.81	9.00	22.26	986.62	30.86	82.60

●要件定義書ページあたりの要件定義工数（新規開発、主開発言語グループ）

							(人時/ページ)
N	最小	P25	中央	P75	最大	平均	標準偏差
74	1.1	4.9	10.9	21.3	588.0	29.5	76.1

●基本設計書ページあたりの基本設計工数（新規開発、主開発言語グループ）

							(人時/ページ)
N	最小	P25	中央	P75	最大	平均	標準偏差
137	0.37	2.09	4.57	10.33	110.15	9.26	13.97

●詳細設計書ページあたりの詳細設計工数（新規開発、主開発言語グループ）

							(人時/ページ)
N	最小	P25	中央	P75	最大	平均	標準偏差
131	0.12	1.37	2.49	4.17	82.08	4.40	8.01

●KSLOCあたりの製作工数（新規開発、主開発言語グループ）

							(人時/KSLOC)
N	最小	P25	中央	P75	最大	平均	標準偏差
535	0.7	26.3	52.2	86.7	918.4	74.2	88.3

●結合テストケースあたりの結合テスト工数（新規開発、主開発言語グループ）

							(人時/ケース)
N	最小	P25	中央	P75	最大	平均	標準偏差
324	0.01	0.47	1.04	2.48	106.72	3.11	8.98

●総合テスト（ベンダ確認）ケースあたりの総合テスト（ベンダ確認）工数（新規開発、主開発言語グループ）

							(人時/ケース)
N	最小	P25	中央	P75	最大	平均	標準偏差
328	0.00	0.76	2.39	7.15	286.32	8.24	20.91

(3) ベンチマーキング方法

- ① 妥当と評価できる範囲を、白書等の各統計情報の 25 パーセンタイル (P25) から 75 パーセンタイル (P75) までの範囲とする。

<開発規模あたりの成果物量の P25～P75 の範囲の例 (開発規模が SLOC で新規開発の場合) >
(上記(2)ベンチマークから抜粋)

メトリクス	単位	P25	P75
要件定義書ページ数	ページ/KSLOC	0.57	2.78
基本設計書ページ数	ページ/KSLOC	2.98	12.24
詳細設計書ページ数	ページ/KSLOC	6.3	23.5
結合テストケース数	ケース/KSLOC	13.6	54.0
総合テスト (ベンダ確認) ケース数	ケース/KSLOC	2.81	22.26

<成果物量あたりの工数の P25～P75 の範囲の例 (開発規模が SLOC で新規開発の場合) >
(上記(2)ベンチマークから抜粋)

メトリクス	単位	P25	P75
要件定義工数	人時/ページ	4.9	21.3
基本設計工数	人時/ページ	2.09	10.33
詳細設計工数	人時/ページ	1.37	4.17
製作工数	人時/KSLOC	26.3	86.7
結合テスト工数	人時/ケース	0.47	2.48
総合テスト (ベンダ確認) 工数	人時/ケース	0.76	7.15

- ② 評価対象プロジェクトの開発規模あたりの成果物量及び成果物量あたりの工数に関する設定が、それぞれの P25～P75 の範囲内に収まっているか否かを判定する。ただし、評価対象プロジェクトに該当する生産性変動要因を勘案して、妥当と評価できる範囲を調整 (上方修正/下方修正) しながら評価することが望ましい。生産性変動要因については、白書等の知見をヒントとして参考にされたい。
- ③ 評価対象プロジェクトが P25～P75 の範囲外となる場合、評価対象プロジェクトの特性を勘案しながら、範囲外となる合理的な理由がないかどうか検討する。合理的な理由があれば妥当と評価し、合理的な理由がなければ目標の見直しを検討する。

3.3.3. IT 企業での具体例 1

<WBS 見積りの改善>

(1) 目的 (活用シーン)

機能規模 (FP) の測定結果を利用した WBS 見積りによって、見積りの妥当性を確認する。
また、プロジェクト遂行時のコントロールにも役立てる。

(2) 改善までの背景

<これまでの見積手順>

見積手法 1 : 案件対象のシステムの FP 値を測定し、自社内の生産性指標を利用し工数を算出。

$$\text{工数1} = \text{規模 (FP)} * \text{生産性 (工数/FP)}$$

見積手法 2 : プロジェクトの WBS (機能名に紐づいた作業と成果物) から個々の作業に対する工数を見積り、積み上げて全体工数を算出。

$$\text{工数2} = \text{作業1の工数} + \dots + \text{作業Nの工数}$$

工数 1 と工数 2 を比較検証し、工数 1 の生産性や工数 2 のある作業の工数を見直すことで、その差分を小さくするなどの調整を行い、最終的見積結果を導き出す。

<課題>

- ・複数の見積手法を利用して見積精度の向上を図ったが、単純な工数値の比較で終わっており、その妥当性確認が不十分であった。
- ・2つの見積手法は独立した内容になってしまっており、また、見積手法 1 の元となっている規模と工数のモデルはばらつきが大きいことから、あまり利用されなくなっていた。
- ・そのため、さらに FP を測定することも少なくなっていた。
- ・見積手法 2 のみで見積った工数に対し、プロジェクト遂行時に作業および工数の変動があっても、その内訳として機能規模の追加・変更か、内部的作業生産性の見誤りかを捉えることが難しかった。

(3) 改善の考え方

見積手法 2 において、作業に対する工数を類推的に見積るのではなく、見積手法 1 によって算出した FP 値を利用して、まず成果物量を見積り、見積った成果物量から作業の工数を導き出すよう見積手法を改善する。

<見積手法の考え方>

【見積対象の機能の粒度】

- ・ WBS を策定する際には、設計や製造の粒度の機能を対象とする。 (成果物対象機能)

例) ○○業務情報登録機能

- ・ FP を測定する際には IFPUG に準拠した機能の対象、粒度を考慮する。(FP 測定対象機能)

注意点) 上記の「成果物対象機能」と「FP 測定対象機能」の粒度が異なる場合がある。

往々にして成果物対象機能の粒度の方が大きい。(1 : n となる。)

例) 上記の○○業務情報登録機能が実は以下の FP 測定対象機能で構築されている。

- | | |
|-----------|-----------|
| a. 検索結果表示 | b. 情報新規登録 |
| c. 情報表示 | d. 情報変更 |

【見積手順】（設計工程を例としている）

- ① 対象案件の開発システムにおいて、上記「成果物対象機能」を洗い出す。

例) ○○業務情報登録機能

- ② 成果物対象機能を考慮した WBS を作成する。

例) ○○業務情報登録機能の基本設計-画面設計書作成 など

内訳として、上記 a.~d.画面の画面設計書が成果物としてあり、それらを作成する作業が作業名として洗い出される。

- ③ WBS の作業毎に成果物対象機能に対して、FP 測定対象機能を紐付ける。

例) ○○業務情報登録機能は、上記 a.~d.で構築されるという整理

- ④ FP 測定対象機能から FP 値を測定し、それらを利用して成果物対象機能の FP 値を見積る。

例) 上記 a.~d.の機能に対して FP 値を測定する。

それらを合計したものが「○○業務情報登録機能」の（合計）FP 値

- ⑤ 成果物対象機能の FP 値から、単位 FP 値に対する成果物量のデータを利用して成果物量を見積る。

例) 100FP に対して基本設計書 20 ページ作成という実績に基づく指標があった場合、

ある成果物対象機能の FP 値が 20FP として測定されたなら、成果物量は 4 ページとなる。

$$\begin{aligned} \text{成果物量} &= \text{測定 FP 値} * \text{単位 FP あたりのページ数} \\ &= 20\text{FP} * (20 \text{ ページ}/100\text{FP}) = 20 * 0.2 = 4 \text{ ページ} \end{aligned}$$

- ⑥ 見積った成果物量から生産性をかけて作業工数を算出する。

例) 1 日（人日）で基本設計書 2 ページ作成という実績に基づく指標があった場合、ある成果物量 4 ページに対する成果物作成工数は、

$$\begin{aligned} \text{作成工数} &= \text{見積対象成果物量} * \text{単位成果物量あたりの工数} \\ &= 4 \text{ ページ} * (1 \text{ 人日}/2 \text{ ページ}) = 4 * 0.5 = 2 \text{ 人日} \end{aligned}$$

(4) ベンチマーク

- ・単位 FP 値あたりの成果物量の統計情報
- ・単位成果物量あたりの工数の統計情報

(5) ベンチマーキング方法

上記(3)の見積手法において上記(4)ベンチマークの統計情報を活用することによって、工数見積りの精度又は工数見積りの妥当性評価の精度を高める。

評価対象プロジェクトが妥当と評価できない場合、その合理的な理由がないか検討する。合理的な理由があれば妥当と評価し、合理的な理由がなければ目標の見直しを検討する。

<適用イメージ>

設計作業（ドキュメント作成を伴う作業）の例：

【改善前】

No.	標準作業			業務機能別 作業	工数 (人日)
1	基本 設計	機能 設計	画面 設計	○○業務機能 画面設計	2. 2
2					
3					
4					
...		
...			帳票 設計

【改善後】

No.	標準作業			業務機能別 作業	機能	FP 値	FP 値 (合計)	成果物量 (ページ)	工数 (人日)
1	基本 設計	機能 設計	画面 設計	○○業務機能 画面設計	検索結果 表示	6	2 2 (個別 の測定 からの 合計)	4. 4 (計算 から)	2. 2 (計算 から)
2					情報 新規登録	9			
3					情報 表示	4			
4					情報 変更	1 3			
...		
...			帳票 設計
...	

(6) 期待できる改善効果

改善後の対応を行っても、見積全体の対応としては、まだ以下が必要である。

- ・DB 設計などの機能と作業・成果物が直接紐づかない開発作業や、プロジェクト管理などの付帯作業は、これまでと同様の WBS 見積りを行う。
- ・FP 測定の ILF・EIF を含んだ全体の FP 値から、規模—工数モデルを利用した見積りは、これまでと同様に行う。

あくまでも WBS 見積りの内訳を詳細化したということであるが、これにより見積根拠がさらに明確になっただけでなく、以下の改善効果の期待がある。

- ・機能規模を捉える事で規模—工数モデルを利用した見積りも自ずと行うことになり、複数見積りによる妥当性確認の普及・展開が広がる。
- ・WBS 見積りの妥当性確認を、作業量（成果物量）、作業生産性の面から確認することができる。
- ・見積りの調整の際、顧客（発注者）との調整が必要な機能規模（要件・仕様）に関わる事と、自社内プロジェクトの内部調整が必要な成果物量、生産性との切り分けた対応が可能になる。
- ・ベースラインとして機能規模、成果物量、生産性などが定量的に明確になるので、プロジェクト遂行中に

においても、何が変動したかを捉える事ができ、次のアクションに繋げる事ができる。

機能を捉える粒度の標準化、有効なデータ収集、重み付けを考慮した見積指標の見直し、層別によるばらつきの極小化など課題もある。

しかし、「機能規模からは工数イメージは直感的に湧かない」という FP のイメージを払拭し、成果物を通して作業工数に紐付ける改善を通し、FP 利用の納得感醸成に繋がると考えている。

参考ポイント

この例は、WBS による積上げ見積りにおいて、内部ベンチマーク（単位 FP 値あたりの成果物量の統計情報及び単位成果物量あたりの工数の統計情報）を活用することによって、成果物量を見積り、見積った成果物量から作業工数を導き出すよう見積手法又は見積り妥当性評価方法を改善するものである。

この例は、機能規模を FP で定量化する場合の見積り手法又は見積り妥当性評価方法の改良例として参考になるとともに、成果物量に着目した統計情報を活用したベンチマーキング方法の事例として参考になると考えられる。

3.4. 工期の妥当性評価

3.4.1. ベンチマーキング方法の手順例

ベンチマーキング方法の基本的な手順の例を以下に示す。

- (1) ベンチマーキングの目的を確認する。
ベンチマーキングの目的が、「プロジェクト計画における工期（月数）が、妥当な範囲に収まっているか否かを評価し、必要に応じて調整すること」であることを確認する。
- (2) 妥当と評価できる範囲を設定し、必要に応じて調整する。
工数と工期（月数）の関係に関するベンチマーク中の統計情報（工数と工期（月数）の散布図と信頼区間）を目安にして、工期（月数）の妥当と評価できる範囲を設定する。

<妥当と評価できる範囲の例>

信頼区間 50%の範囲の工期（月数）

- (3) 工期（月数）が、妥当と評価できる範囲内に収まっているか否かを判定する。
- (4) 妥当と評価できる範囲外となる場合、評価対象プロジェクトの特性を勘案しながら、範囲外となる合理的な理由がないか検討する。合理的な理由があれば妥当と評価し、合理的な理由がなければ工期（月数）の見直しを検討する。
特に、信頼区間 95%の下限值を下回る場合は、工期（月数）を見直す。（信頼区間 95%の下限值を下回ったプロジェクトは今まで僅かしかなかったことから、実現可能性が極めて低いと考えられる。従って、工期短縮限界の一つの目安となる。）

<合理的な理由の検討（又は妥当と評価できる範囲の調整）のケース例>

◇ 対外接続システムの数が多いケース

3.4.2. 白書等を用いた具体例

<工数と工期の関係に関する統計情報を参考にして工期の妥当性を評価する例>

(1) 目的 (活用シーン)

開発者 (開発プロジェクトのプロジェクト・マネジャー及び担当者) が、プロジェクト計画の策定時に、プロジェクト計画の実現性を高めるために、次のような観点からプロジェクト計画の妥当性を評価するシーン (開発組織のマネジャー層、PMO 及び品質マネジメント推進部門が、プロジェクト計画をレビューし助言/支援するシーンを含む)

◇プロジェクト計画における工期 (月数) が、妥当な範囲に収まっているか否かを評価し、必要に応じて調整する。

(2) ベンチマーク

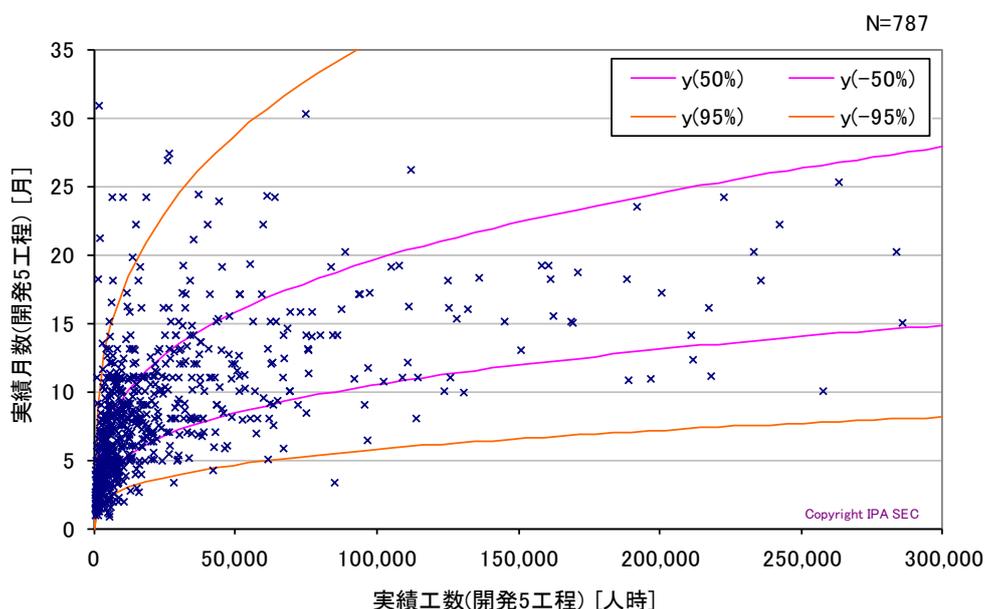
工数と工期に関する白書等の統計情報を、目安として参考にする。

<参照箇所>

詳しくは、「ソフトウェア開発データ白書」の「工数と工期」に関する記事を参照されたい。

① 開発5工程の工数と工期 (月数) の関係の例

- 開発5工程の工数と工期 (新規開発) (信頼区間 50%、95%付き)



(3) ベンチマーキング方法

- ① 工数と工期 (月数) の関係に関するベンチマーク中の統計情報 (工数と工期 (月数) の散布図と信頼区間) を目安にして、工期 (月数) の妥当と評価できる範囲を 信頼区間 50%の範囲の工期 (月数) に設定する。

<信頼区間 50%の範囲の例（新規開発の場合）>

（上記(2)ベンチマークから抜粋）

実績工数（開発 5 工程）	-50%	50%
50,000 人時	約 7.0 ヶ月	約 16.0 ヶ月
100,000 人時	約 10.5 ヶ月	約 19.5 ヶ月

- ② 評価対象プロジェクトの工期（月数）が、信頼区間 50%の範囲内に収まっているか否かを判定する。
- ③ 評価対象プロジェクトが信頼区間 50%の範囲外となる場合、評価対象プロジェクトの特性を勘案しながら、範囲外となる合理的な理由がないかどうか検討する。合理的な理由があれば妥当と評価し、合理的な理由がなければ目標の見直しを検討する。
- 特に、信頼区間 95%の下限値を下回る場合は、工期（月数）を見直す。（信頼区間 95%の下限値を下回ったプロジェクトは今まで僅かしかなかったことから、信頼区間 95%の下限値を下回る工期を計画することは実現可能性が極めて低いと考えられる。従って、信頼区間 95%の下限値が工期短縮限界の一つの目安となる。）

<工期の厳しさと信頼性実績の関係を参考にして工期の妥当性を評価する例>

(1) 目的（活用シーン）

開発者（開発プロジェクトのプロジェクト・マネジャー及び担当者）が、プロジェクト計画の策定時に、プロジェクト計画の実現性を高めるために、次のような観点からプロジェクト計画の妥当性を評価するシーン（開発組織のマネジャー層、PMO 及び品質マネジメント推進部門が、プロジェクト計画をレビューし助言／支援するシーンを含む）

◇プロジェクト計画における工期（月数）が、妥当な範囲に収まっているか否か（特に信頼性を確保できないほど厳しい工期になっていないかどうか）を評価し、必要に応じて調整する。

(2) ベンチマーク

白書のデータによると、工期が厳しいと信頼性が低下する傾向が見られる。

【工期の厳しさの評価方法】

計画工期の厳しさを、計画工期が短く（標準工期に対する計画工期の比率が低く）、かつ各工程の工期が重複している程度で評価する。

具体的には、計画工期の厳しさ = (標準工期に対する計画工期比率) ÷ (計画工期の重複度)。

1 より小さい領域で値が小さいほど、計画工期が厳しいと評価する。

ここで、標準工期は、開発 5 工程全体の実績工数と実績工期（月数）との回帰式を用いて、開発 5 工程全体の計画工数から算出した工期（月数）。標準的工期に対する計画工期比率は、1 より小さいものほど短納期であることを示す。

工期の重複度は、各工程の計画工期（計画終了日－計画開始日で求めた日数）の合計値÷開発 5 工程全体の計画工期（日数）計画工期の重複度は、1 より大きいものほど工程の重複度が高いことを示す。

【主なプロジェクトデータの検索条件】

◇新規開発、主開発言語グループ（SLOC 系）

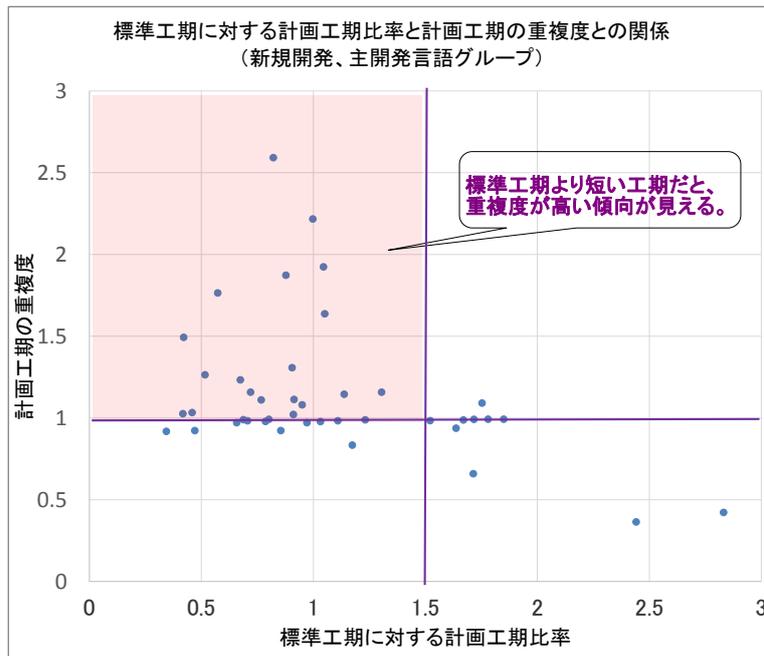
◇開発 5 工程が揃っているもの

◇SLOC 発生不具合密度が計算できるもの

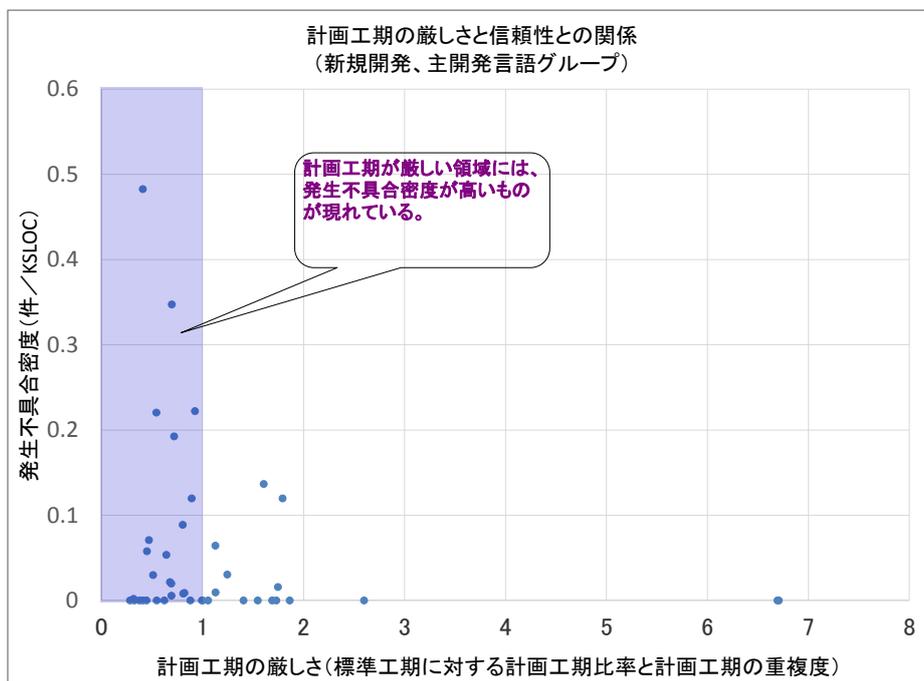
◇開発 5 工程の計画開始日及び計画終了日がすべて記入されているもの

【分析結果（ベンチマーク）】

① 標準工期に対する計画工期比率と計画工期の重複度との関係を見ると、標準工期より短い工期だと工期の重複度が高くなる傾向が見られる。



- ② 計画工期が厳しい領域（工期の厳しさが1未満の領域）には、発生不具合密度が高いものが現れている。工期の厳しさが1未満の場合、信頼性を確保できないリスクが高まると言える。



(3) ベンチマーキング方法

- ① 評価対象プロジェクトの計画工期及び工程計画から、工期の厳しさを算出する。
- ② 評価対象プロジェクトの工期の厳しさが1未満の場合、計画工期及び工程計画の実現性を再検討し、必要に応じて見直す。

3.4.3. IT 企業での具体例 1

<プロジェクト・マネジャーによる開発工期の妥当性の検証>

(1) 目的 (活用シーン)

見積り終了時に開発工期の実現性の確認を行う。

(2) 考え方

無理な工期がプロジェクトの失敗につながるという考えのもと、プロジェクトの実績データから導出した内部ベンチマークに基づいて、開発工期の実現性を確認する。

(3) ベンチマーク

<内部ベンチマークの設定方法>

当社では、無理な工期計画がプロジェクトの失敗に繋がるのではないかと仮説のもと、社内の過去プロジェクトの実績をもとに以下のことを検証し、内部ベンチマークを設定した。

開発工数と工期の関係を算出

プロジェクト毎に工期および工数の実績値を収集し、個々プロジェクトの全体工期と全体工数の関係を、ある公開ベンチマークで示されている下記の関係式 (工数の3乗根と工期の関係) を用いて係数A部分を算出。

$$(\text{全体工期}) = A \times (\text{全体工数})^B, \quad A=2.5, \quad B=1/3$$

・算出例 (全体工数=100人月、全体工期=9ヶ月の場合)

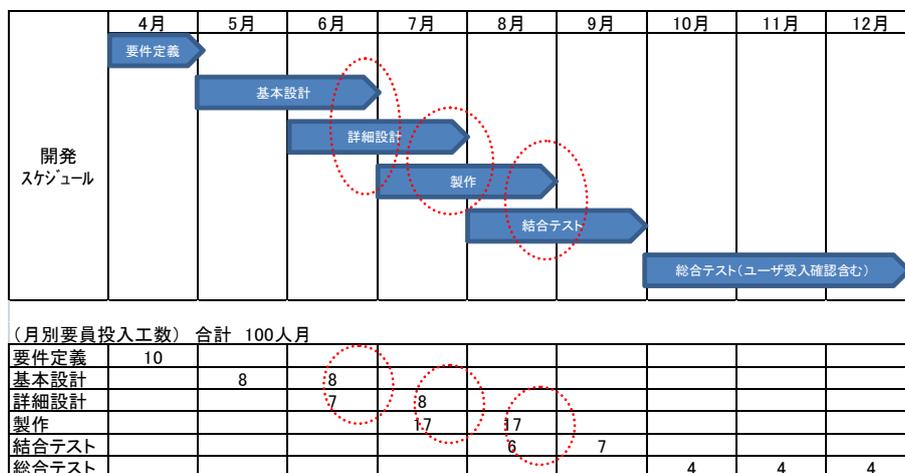
$$100 = A \times 9^{1/3}$$

$$A \approx 1.94$$

開発フェーズの重複度を算出

$$(\text{重複度}) = (\text{開発フェーズ重複期間中の工数合計}) \div \text{開発工数}$$

・開発フェーズ重複期間中の工数計算例

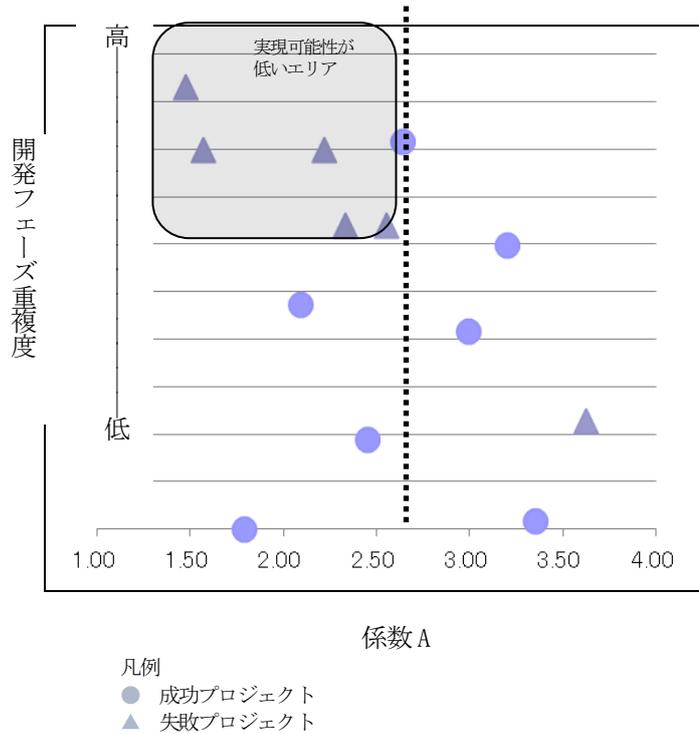


上図の破線で囲んだ部分の要員投入工数を重複期間の工数とする。

$$\begin{aligned}(\text{重複度}) &= (8+7+8+17+17+6) / 100 \\ &= 0.63\end{aligned}$$

<ベンチマーク>

上記結果を成功プロジェクトと失敗プロジェクトで散布図上に表示。



工数から見た標準的な工期より短い工期で、かつ開発フェーズ重複度の高い領域をプロジェクトの実現性が低い領域と設定する。(上図の網掛け部分)

- ・ 開発工期が短く、かつ開発フェーズ重複度が高い場合には下記の状況に陥ることが想定される。
 - ✓ 短期開発であるがゆえに一度遅延したら軌道修正が難しい。
 - ✓ 開発フェーズが並行であるがゆえに前フェーズでの遅延・不備の影響を大きく受けリカバリできない。

(4) ベンチマーキング方法 (実現性の確認方法)

プロジェクト・マネジャーは、見積り終了時に内部ベンチマークと比較し、開発工期面での実現性を確認し、必要に応じて再見積りを実施する。

参考ポイント

短納期開発ニーズの高まりは時代の趨勢であり、開発者としてもできるだけそれに応えたいところである。一方、短すぎる工期はプロジェクト崩れや生産性・信頼性の低下に繋がること、定性的には共通認識になっていると考えられる。

ここで、短すぎる（あるいは無理な）工期かどうかの判断には、開発規模（又は工数）だけでなく、機能要求・非機能要求の難易度や開発能力（体制、スキル、経験等）を始めとして種々の要因が関わってくるので、それらを総合的に勘案することが望ましいと考えられる。しかしながら、最も基本的な要因である工数に対する工期の側面においても、ある限界を超えるとプロジェクトが失敗する、あるいは信頼性が低下するということを定量的に示すベンチマークは今まで見当たらなかった。

この例は、最も基本的な要因である工数に対する工期の側面から、短すぎて実現可能性が低い工期かどうかの目安となる定量的なベンチマークを用意できた例である。具体的には、工数と工期の関係式による標準的な工期に対して一定値以下の長さであるかどうかだけでなく、開発工程の重複度が一定値以上に高いかどうかを併せて評価することによって、工期の実現可能性を定量的に評価できるベンチマーキングとして成功している。

この例は、今までなかった、工期の妥当性を定量的に評価するためのベンチマーキング方法の事例として参考になると考えられる。

短すぎる（あるいは無理な）工期を計画することによって開発プロジェクトが失敗に陥ることは、ユーザ企業、開発ベンダの双方にとって不幸なことである。ユーザ企業の発注者の方々にも参考にして頂きたいベンチマーキング例である。

なお、3.4.2 項「白書等を用いた具体例」の〈工数と工期の関係に関する統計情報を参考にして工期の妥当性を評価する例〉も併せて参考にされたい。

3.5. 工程別の工数比率の妥当性評価

3.5.1. ベンチマーキング方法の手順例

ベンチマーキング方法の基本的な手順の例を以下に示す。

- (1) ベンチマーキングの目的を確認する。
ベンチマーキングの目的が、「プロジェクト計画における工程別の工数の比率が、妥当な範囲に収まっているか否かを評価し、必要に応じて調整すること」であることを確認する。
- (2) 妥当と評価できる範囲を設定し、必要に応じて調整する。
工程別の実績工数の比率に関するベンチマーク中の統計情報を目安にして、各工程の工数比率の妥当と評価できる範囲を設定する。

<妥当と評価できる範囲の例>

25 パーセンタイル (P25) から 75 パーセンタイル (P75) までの範囲

- (3) 各工程の工数比率が、それぞれの妥当と評価できる範囲内に収まっているか否かを判定する。
- (4) 妥当と評価できる範囲外となる場合、評価対象プロジェクトの特性を勘案しながら、範囲外となる合理的な理由がないか検討する。合理的な理由があれば妥当と評価し、合理的な理由がなければ工程別の工数比率の見直しを検討する。

<合理的な理由の検討（又は妥当と評価できる範囲の調整）のケース例>

◇開発規模が比較的小さい、次のようなプロジェクト

- ・流用母体（OSS、パッケージ等）をテストし信頼性を確保するケース
- ・母体（前版部分）の信頼性向上を図るケース

◇当プロジェクトで対外システムとの接続テストを強化するケース

◇対外接続システムの数が多いケース

3.5.2. 白書等を用いた具体例

(1) 目的（活用シーン）

開発者（開発プロジェクトのプロジェクト・マネジャー及び担当者）が、プロジェクト計画の策定時に、プロジェクト計画の実現性を高めるために、次のような観点からプロジェクト計画の妥当性を評価するシーン（開発組織のマネジャー層、PMO 及び品質マネジメント推進部門が、プロジェクト計画をレビューし助言／支援するシーンを含む）

◇プロジェクト計画における工程別の工数の比率が、妥当な範囲に収まっているか否かを評価し、必要に応じて調整する。

(2) ベンチマーク

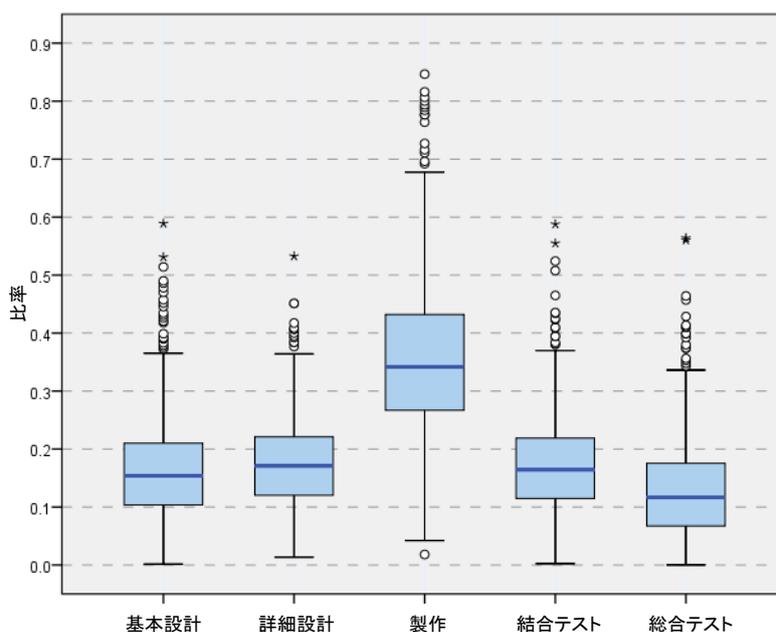
工程別の実績工数の比率に関する白書等の統計情報を、目安として参考にする。

<参照箇所>

詳しくは、「ソフトウェア開発データ白書」の「工程別工数」に関する記事を参照されたい。

① 工程別の実績工数の比率の例

● 工程別の実績工数の比率（新規開発）



工程	N	最小	P25	中央	P75	最大	平均	標準偏差
基本設計	679	0.001	0.104	0.154	0.210	0.589	0.167	0.092
詳細設計	679	0.014	0.120	0.171	0.221	0.533	0.175	0.077
製作	679	0.018	0.267	0.342	0.432	0.847	0.356	0.135
結合テスト	679	0.002	0.115	0.164	0.219	0.588	0.173	0.087
総合テスト	679	0.000	0.067	0.117	0.175	0.564	0.130	0.085

(3) ベンチマーキング方法

- ① 工程別の実績工数の比率に関するベンチマーク中の統計情報を目安にして、各工程の工数比率の妥当と評価できる範囲を各工程の工数比率の 25 パーセンタイル (P25) から 75 パーセンタイル (P75) までの範囲とする。

<P25～P75 の範囲の例 (新規開発の場合) >
(上記(2)ベンチマークから抜粋)

工程	単位	P25	P75
基本設計	比率	0.104	0.210
詳細設計	比率	0.120	0.221
製作	比率	0.267	0.432
結合テスト	比率	0.115	0.219
総合テスト (ベンダ確認)	比率	0.067	0.175

- ② 評価対象プロジェクトの各工程の工数比率が、P25～P75 の範囲内に収まっているか否かを判定する。
- ③ 評価対象プロジェクトが P25～P75 の範囲外となる場合、評価対象プロジェクトの特性を勘案しながら、範囲外となる合理的な理由がないかどうか検討する。合理的な理由があれば妥当と評価し、合理的な理由がなければ目標の見直しを検討する。

3.6. 工程別の工期比率の妥当性評価

3.6.1. ベンチマーキング方法の手順例

ベンチマーキング方法の基本的な手順の例を以下に示す。

- (1) ベンチマーキングの目的を確認する。
ベンチマーキングの目的が、「プロジェクト計画における工程別の工期（月数）の比率が、妥当な範囲に収まっているか否かを評価し、必要に応じて調整すること」であることを確認する。
- (2) 妥当と評価できる範囲を設定し、必要に応じて調整する。
工程別の実績工期（月数）の比率に関するベンチマーク中の統計情報を目安にして、各工程の工期（月数）比率の妥当と評価できる範囲を設定する。

<妥当と評価できる範囲の例>

25 パーセンタイル（P25）から 75 パーセンタイル（P75）までの範囲

- (3) 各工程の工期（月数）比率が、それぞれの妥当と評価できる範囲内に収まっているか否かを判定する。
- (4) 妥当と評価できる範囲外となる場合、評価対象プロジェクトの特性を勘案しながら、範囲外となる合理的な理由がないか検討する。合理的な理由があれば妥当と評価し、合理的な理由がなければ工程別の工期（月数）比率の見直しを検討する。

<合理的な理由の検討（又は妥当と評価できる範囲の調整）のケース例>

◇開発規模が比較的小さい、次のようなプロジェクト

- ・流用母体（OSS、パッケージ等）をテストし信頼性を確保するケース
- ・母体（前版部分）の信頼性向上を図るケース

◇当プロジェクトで対外システムとの接続テストを強化するケース

◇対外接続システムの数が多いケース

3.6.2. 白書等を用いた具体例

(1) 目的（活用シーン）

開発者（開発プロジェクトのプロジェクト・マネジャー及び担当者）が、プロジェクト計画の策定時に、プロジェクト計画の実現性を高めるために、次のような観点からプロジェクト計画の妥当性を評価するシーン（開発組織のマネジャー層、PMO 及び品質マネジメント推進部門が、プロジェクト計画をレビューし助言／支援するシーンを含む）

◇プロジェクト計画における工程別の工期（月数）の比率が、妥当な範囲に収まっているか否かを評価し、必要に応じて調整する。

(2) ベンチマーク

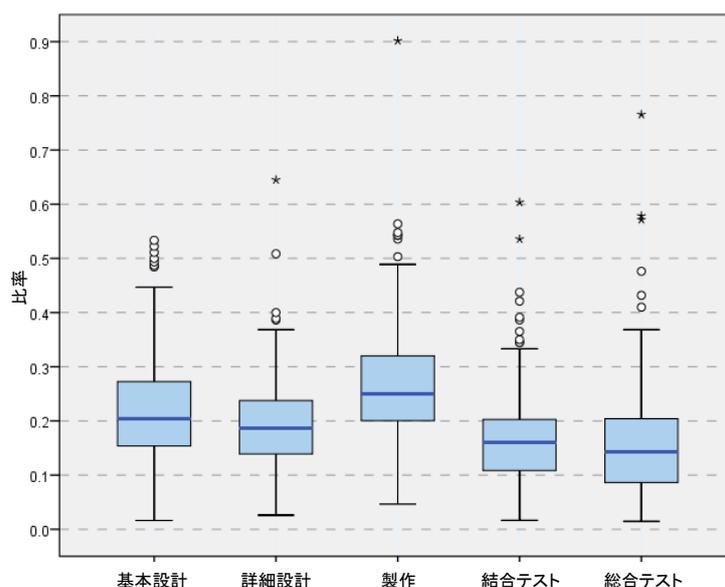
工程別の実績月数の比率に関する白書等の統計情報を、目安として参考にする。

<参照箇所>

詳しくは、「ソフトウェア開発データ白書」の「工程別工期」に関する記事を参照されたい。

工程別の実績月数の比率の例

● 工程別の実績月数の比率（新規開発）



[比率]								
工程	N	最小	P25	中央	P75	最大	平均	標準偏差
基本設計	239	0.016	0.154	0.205	0.273	0.533	0.218	0.095
詳細設計	239	0.026	0.139	0.186	0.238	0.645	0.192	0.083
製作	239	0.047	0.201	0.250	0.320	0.902	0.266	0.108
結合テスト	239	0.016	0.108	0.160	0.203	0.604	0.165	0.084
総合テスト	239	0.014	0.086	0.143	0.204	0.765	0.159	0.099

(3) ベンチマーキング方法

- ① 工程別の実績月数の比率に関するベンチマーク中の統計情報を目安にして、各工程の月数比率の妥当と評価できる範囲を各工程の月数比率の 25 パーセントイル (P25) から 75 パーセントイル (P75) までの範囲とする。

<P25～P75 の範囲の例 (新規開発の場合) >

(上記(2)ベンチマークから抜粋)

工程	単位	P25	P75
基本設計	比率	0.154	0.273
詳細設計	比率	0.139	0.238
製作	比率	0.201	0.320
結合テスト	比率	0.108	0.203
総合テスト (ベンダ確認)	比率	0.086	0.204

- ② 評価対象プロジェクトの各工程の月数比率が、P25～P75 の範囲内に収まっているか否かを判定する。
- ③ 評価対象プロジェクトが P25～P75 の範囲外となる場合、評価対象プロジェクトの特性を勘案しながら、範囲外となる合理的な理由がないかどうか検討する。合理的な理由があれば妥当と評価し、合理的な理由がなければ目標の見直しを検討する。

4. 開発組織のマネジャー層用のベンチマーキング（プロジェクト・マネジメントの改善）

4.1. 品質保証プロセス関連標準類の見直し（品質マネジメント関連）

4.1.1. ベンチマーキング方法の手順例

ベンチマーキング方法の基本的な手順の例を以下に示す。

(1) ベンチマーキングの目的を確認する。

ベンチマーキングの目的が、「ベンチマーク中の「信頼性向上に資する知見」（良い信頼性実績を上げているプロジェクト群の品質保証プロセスにおける良いやり方）と自組織の現状とを対比しながら、信頼性向上に向けてプロジェクト・マネジメント（特に品質マネジメント）の改善を検討する。検討結果を踏まえて、品質マネジメント関連の標準類を見直すこと」であることを確認する。

(2) ベンチマーク中の「信頼性向上に資する知見」（良い信頼性実績を上げているプロジェクト群の品質保証プロセスにおける良いやり方）をヒントとして参考にし、自組織の品質保証プロセスの現状と対比しながら、信頼性向上に向けた自組織のマネジメント方針（品質保証プロセスの重点強化領域）を検討する。

<方針検討の例>

- ・設計レビューとテストのどちらを重点領域にすると良いか
- ・上流工程での不具合摘出比率をどの程度に設定すると良いか

上流工程での不具合摘出比率：上流工程での不具合摘出数÷（上流工程での不具合摘出数＋下流工程での不具合摘出数）

(3) 上記の方針に沿って、品質保証プロセスの現状を眺みながら、できるだけ経済的に信頼性向上を図る方策を検討する。

(備考) プロダクト目標（信頼性／生産性の目標）及びプロセス目標の設定について

目標値は、長年にわたる定量的管理の実践の中で、各組織のベースラインとして確立し更新していくのが基本である。システムのクリティカリティやシステムリスクに応じて、あるいは業種等のドメインに応じて目標設定するのが良い。また、組織の成熟度に応じて段階的に目標設定するのが良いと考えられる。

<設計レビュー強化策検討の例>

◇設計レビュー工数（設計レビュー工数比率又は設計レビュー工数密度）と信頼性（発生不具合密度）との関係から、設計レビュー工数に関する要件を検討する。

発生不具合密度：稼働後一定期間の発生不具合数÷開発規模（FP 規模又は SLOC 規模）

設計レビュー工数比率：設計レビュー工数÷開発総工数

設計レビュー工数密度：設計レビュー工数÷開発規模（FP 規模又は SLOC 規模）

◇設計レビュー工数密度と設計レビュー検出能率との関係から、経済的に（効率的に）レビュー強化するための目安を検討する。

設計レビュー検出能率：設計レビュー指摘件数÷設計レビュー工数
◇設計レビューのパフォーマンスを向上させる方策を検討する。
設計文書化密度：設計書ページ数÷開発規模（FP 規模又は SLOC 規模）
設計書作成標準やレビュー・チェックリストの改良等

(4) 上記の検討結果に基づいて、品質マネジメント関連の標準類を見直す。

<見直す標準化事項の例>

◇レビュー関連

レビュー工数比率の管理基準値

レビュー工数密度の管理基準値等

設計文書化密度の管理基準値

設計書作成標準やレビュー・チェックリストの改良等

◇上流工程での不具合摘出比率の管理基準値

4.1.2. 白書等を用いた具体例

<信頼性向上に向けて自組織のマネジメント方針を検討する例>

(1) 目的（活用シーン）

主に開発組織のマネジャー層が、次のような観点から、プロジェクト・マネジメントの改善を推進するシーン（PMO 及び品質マネジメント推進部門が、マネジメントを支援するシーンを含む）

◇ベンチマーク中の「信頼性向上に資する知見」（良い信頼性実績を上げているプロジェクト群の品質保証プロセスにおける良いやり方）と自組織の現状とを対比しながら、信頼性向上に向けてプロジェクト・マネジメント（特に品質マネジメント）の改善を検討する。検討結果を踏まえて、品質保証プロセスについて重点強化方針を検討する。また、品質マネジメント関連の標準類を見直す。

(例)

自組織のソフトウェアの SLOC 発生不具合密度の実績は、「ソフトウェア開発データ白書」の SLOC 発生不具合密度の統計情報と比較して、相対的に高い（信頼性が低い）位置に分布している。関連事業のニーズから言っても、ソフトウェアの信頼性向上が課題と認識している。良い信頼性実績を上げているプロジェクト群の品質保証プロセスに学んで、品質保証プロセスに関して重点強化方針を検討したい。

(2) ベンチマーク

公開ベンチマーク（「ソフトウェア開発データが語るメッセージ」）を参照すると、以下のことが示されているのでヒントとして参考にする。

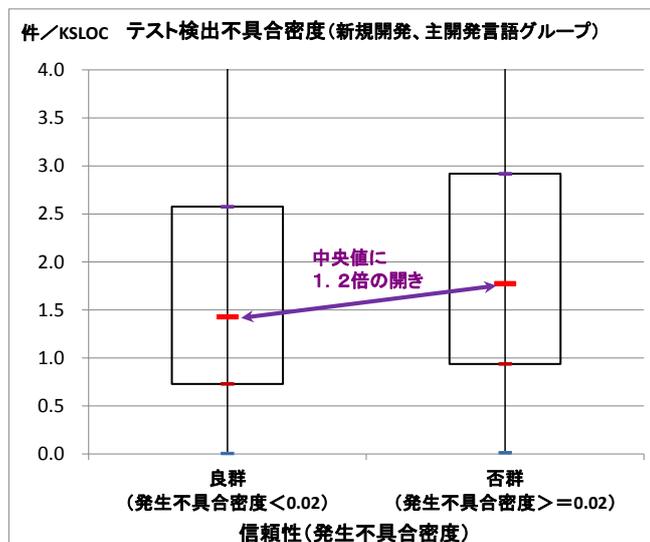
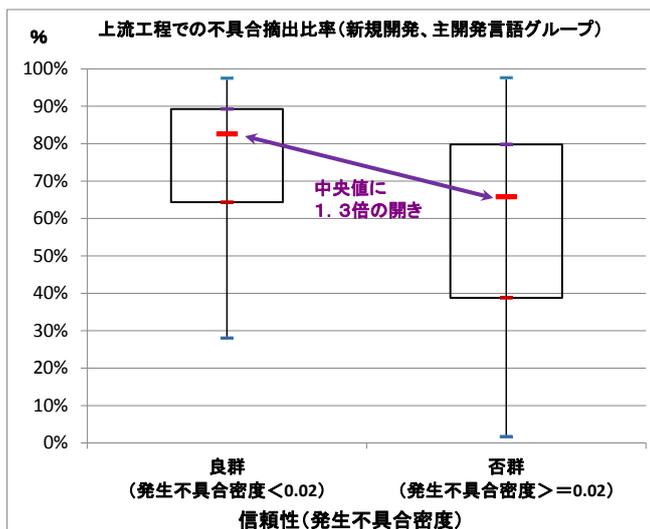
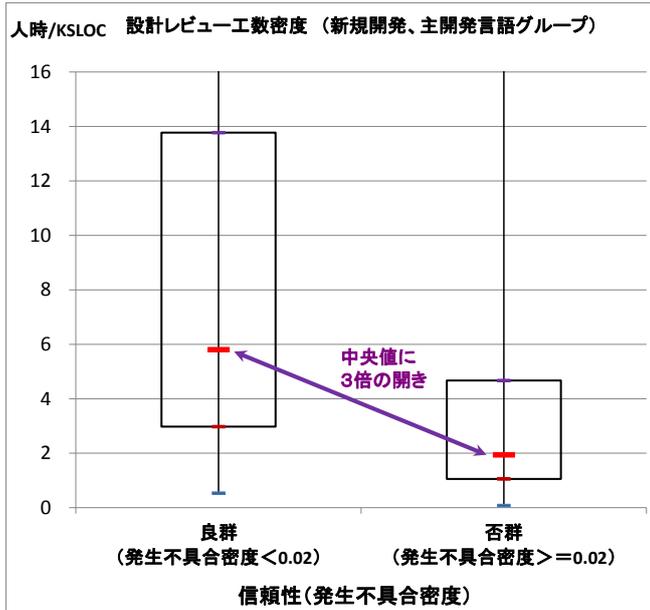
発生不具合密度が 0.02 件/KLSLOC 未満のプロジェクトを良群、0.02/KLSLOC 以上のプロジェクトを否群に分けて分析すると、良群に次の傾向が見られる。

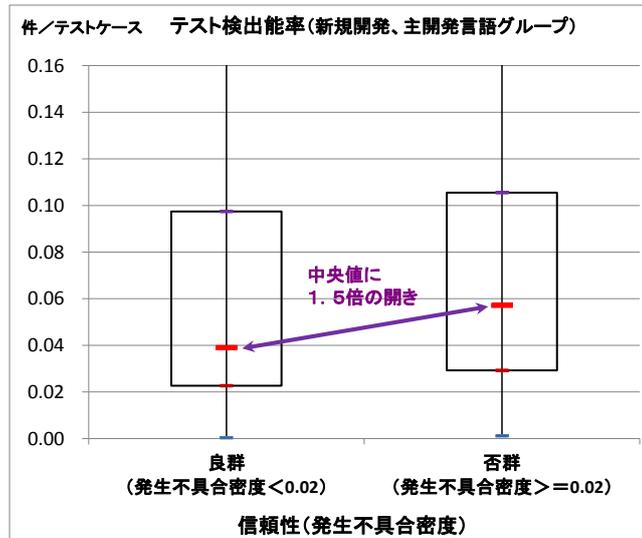
- ・設計レビュー工数密度が高い。
- ・上流工程での不具合摘出比率が高い。
- ・テスト検出不具合密度（テスト検出不具合数÷開発規模）が低い。
- ・テスト検出能率（テスト検出不具合数÷テストケース数）が低い。

これらのことは、要約的に次のことを示唆している。

◇上流工程での設計レビューを強化することによって、信頼性が向上する。

◇テストで信頼性を高める（作込み品質の低さを挽回する）のは困難である。（相対的に、テストで見つかる不具合が少ないプロジェクトの信頼性実績は良く、テストで見つかる不具合が多いプロジェクトの信頼性実績は良くない。）





<参照箇所>

詳しくは、「ソフトウェア開発データが語るメッセージ」の「信頼性変動要因の分析」に関する記事を参照されたい。

(3) ベンチマーキング方法

① 自組織の品質保証プロセスの現状を上記(2)の傾向に照らしてみると、次のように見える。

- ・設計レビュー工数密度が相対的に低い（中央値が3人時/KSLOC程度であり否群の分布に近い）。
- ・テスト検出能率が相対的に高い（中央値が0.07件/テストケース程度であり否群の分布に近い）。

また、テストで検出した不具合や稼働後の不具合の要因の傾向からも、レビューが足りないことを実感している。

② 上記のことから、「設計レビュー強化」を自組織のマネジメントの重点方針とする。

設計レビュー工数を増やすとともに、設計レビューのパフォーマンス向上を図りたい。

設計レビューのパフォーマンス向上に向けては、設計書の書き方、レビュー手法、レビュー・チェックリストの改良などを幅広く検討する。

<経済性を勘案した設計レビュー強化策の検討例>

(1) 目的（活用シーン）

主に開発組織のマネジャー層が、次のような観点から、プロジェクト・マネジメントの改善を推進するシーン（PMO及び品質マネジメント推進部門が、マネジメントを支援するシーンを含む）

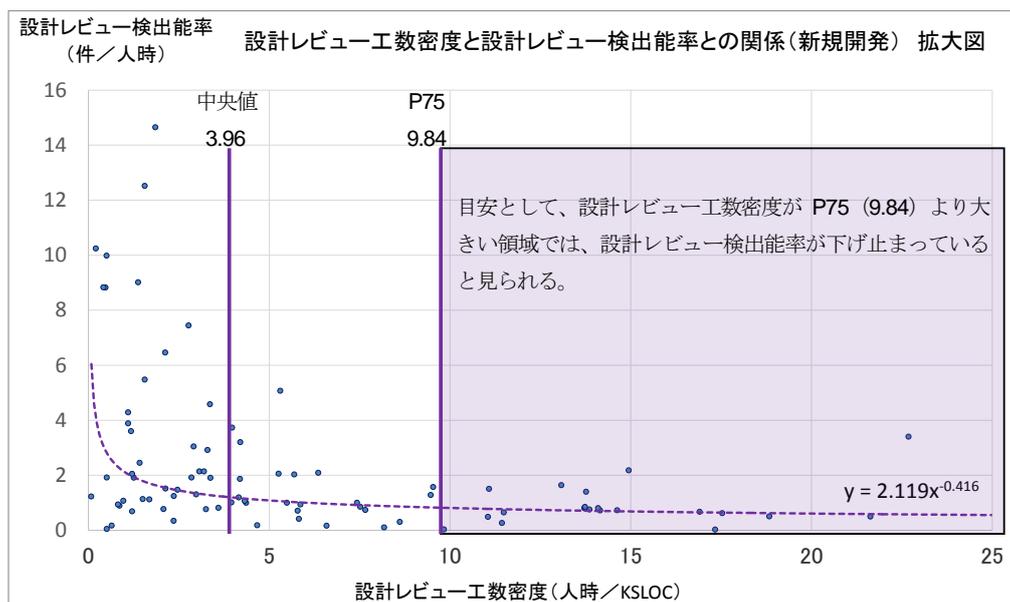
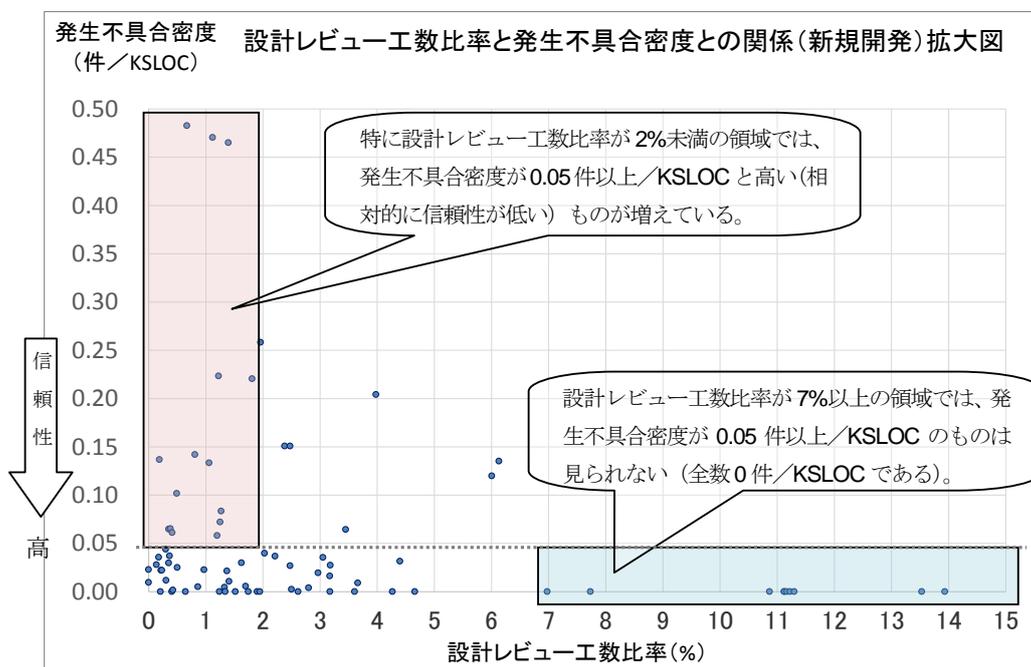
◇設計レビュー強化の重点マネジメント方針に沿って、自組織の設計レビューの現状を眺みながら、近未来の現実的な設計レビュー強化目標を検討し設定する。検討には、品質確保の観点だけでなく、経済性の観点を加える。

(2) ベンチマーク

公開ベンチマーク（「ソフトウェア開発データが語るメッセージ」）を参照すると、設計レビュー工数に関し

て以下のことが示されているのでヒントとして参考にする。

- ◇「不具合を十分になくすための設計レビュー工数はどの程度か」という品質確保の観点で見ると、設計レビュー工数比率が7%以上になると、発生不具合密度が0に近づいている。一方、2%未満では発生不具合密度が高いプロジェクトが多数存在している。
- ◇レビュー工数の投資対効果という経済性の観点で見ると、設計レビュー工数を増やすほど設計レビューで指摘される件数の割合が低下していく。(設計レビュー工数密度を高くするほど、設計レビュー検出能率が低下していく。)特に設計レビュー工数密度のP75あたり(9.8人時/KSLOC)で設計レビュー検出能率が下げ止まる傾向が見られることから、そこが経済性の高い設計レビュー工数のかけ方の目安になる。なお、設計レビュー工数密度 P75 (9.8 人時/KSLOC) は、設計レビュー工数比率のおよそ5.5%に相当する。



<参照箇所>

詳しくは、「ソフトウェア開発データが語るメッセージ」の「レビューのコントロール」に関する記事を参照されたい。

(3) ベンチマーキング方法

① 設計レビュー強化に向けて、次のことから、近未来の設計レビュー工数密度の目標を10人時/KSLOC以上に設定する。

- ・設計レビュー工数比率7%以上は、現状とのギャップが大きくハードルが高すぎるので、近未来の目標としては現実的ではないと判断した。
- ・公開ベンチマーク（ソフトウェア開発データが語るメッセージ）に経済性の高い設計レビュー工数のかけ方として、設計レビュー工数密度のP75（9.8人時/KSLOC）が目安として示されていることから、設計レビュー工数密度がこれより低い現状にあれば、この目安を参考にして設計レビュー強化を図ることとする。

(注1) 内部ベンチマーク中に同様の知見が示されている場合、内部ベンチマークにおける設計レビュー工数密度のP75の値を採用する方向で検討することが望ましいと考えられる。

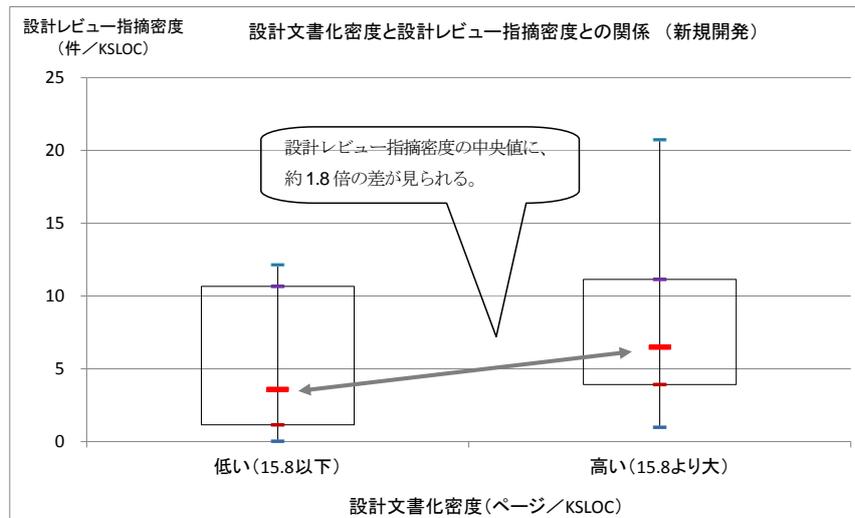
(注2) 計画したレビュー項目のレビューが済んでいない場合は、この限りではない。

(備考) 設計レビュー能力のある要員が不足していると、設計レビュー工数を増やすこともままならない。そういう状況であれば、設計レビュー能力のある要員や有識者を増やすべく体制を強化することも、併せて検討する必要がある。

② 設計レビュー強化のためには、設計レビュー工数を増やすだけでなく、設計レビューのパフォーマンス向上も図る。

設計レビューのパフォーマンスに関連して、公開ベンチマーク（「ソフトウェア開発データが語るメッセージ」）に次のことが示されている。

◇設計文書量を増やすことによって、設計レビュー効果が向上する傾向が見られる。具体的には、設計文書化密度が相対的に低いもの（中央値約15.8ページ/KSLOC以下の集合）と高いもの（中央値より大の集合）とでは、設計レビュー指摘密度の中央値に明らかな差が見られる。（新規開発では約1.8倍の差、改良開発では約5.1倍の差が見られる。）設計文書化密度が低い現状にある場合、設計文書化密度をこの中央値約15.8ページ/KSLOCを目安にして増強することをお勧めする。



<参照箇所>

詳しくは、「ソフトウェア開発データが語るメッセージ」の「レビューのコントロール」に関する記事を参照されたい。

このことを参考にして、自組織の設計文書化密度とレビュー効果の相関を確認し、次に設計文書化密度の高低の要因を現物確認した上で、設計文書化密度の基準値を設定する。

(備考) 設計文書量が増えると、より具体的で漏れの少ない設計文書になるので、設計レビュー効果が向上する（設計レビューで見つけるべき不具合の多くを検出できる）と考えられる。一方、設計文書化が不足していると、設計レビューで摘出すべき不具合を見つけられないケースが増えることが懸念される。従って、上記の「経済性を勘案した設計レビュー強化策」を成功させるためには、前提として適切な水準まで設計文書化しておく必要があると考えられ、設計文書化密度 20 ページ/KSLOC 以上を目安として設計文書量を確保することをお勧めしている。

なお、設計レビュー効果を向上させるには、設計文書量を増強する以外にも、次のような設計関連標準類の改良/整備が有効と考えられる。

- ・設計書作成標準の改良/整備
- ・レビュー・チェックリストの改良/整備
- ・レビュー手法等の改良/整備
- ・解析ツールの改良/整備

<経済性を勘案したテストの強化策の検討例>

(1) 目的 (活用シーン)

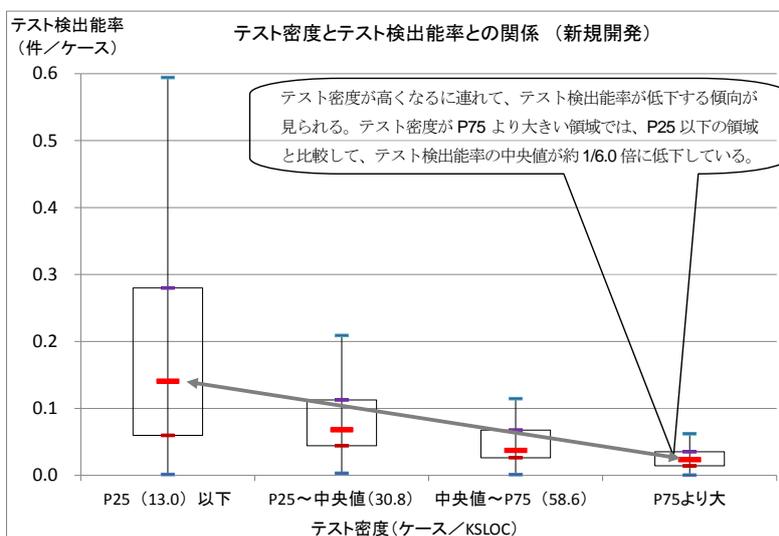
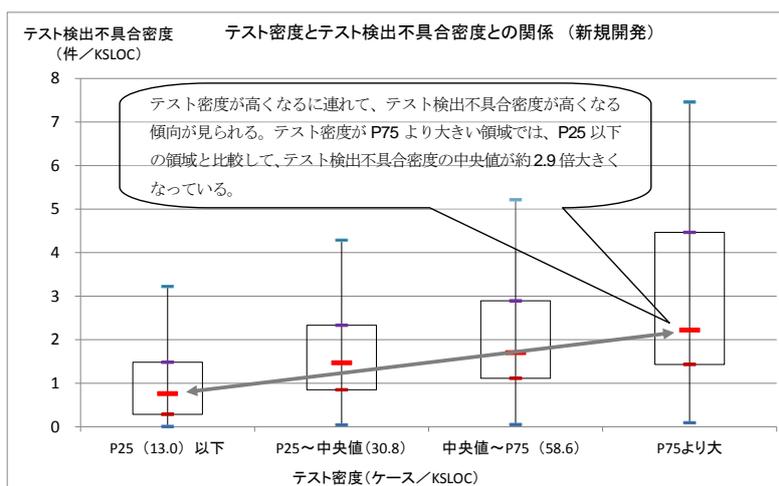
主に開発組織のマネジャー層が、次のような観点から、プロジェクト・マネジメントの改善を推進するシーン (PMO 及び品質マネジメント推進部門が、マネジメントを支援するシーンを含む)

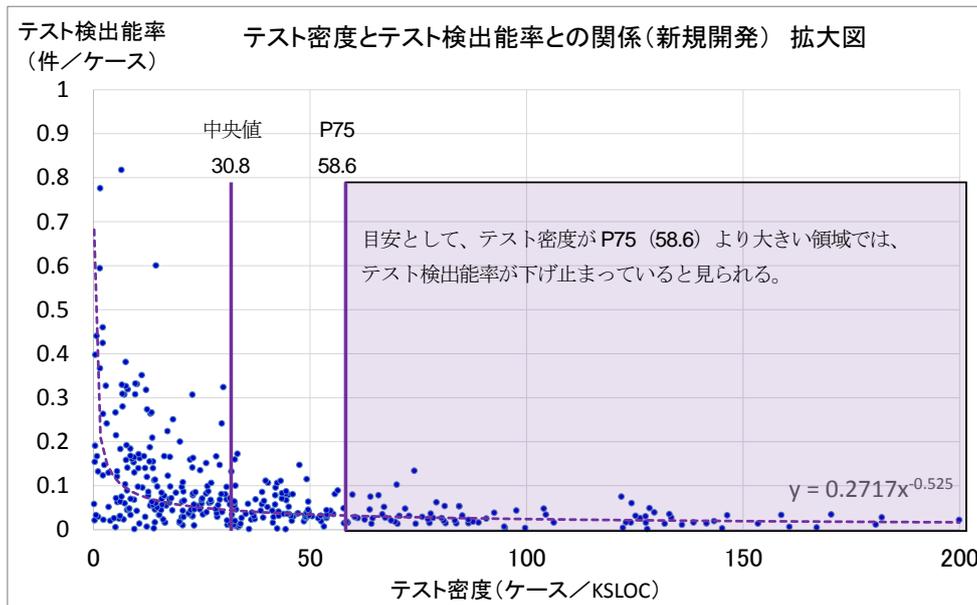
◇テスト強化をマネジメント方針にする場合、自組織のテストの現状を眺みながら、近未来の現実的なテスト強化目標を検討し設定する。検討には、品質確保の観点だけでなく、経済性の観点を加える。

(2) ベンチマーク

公開ベンチマーク (「ソフトウェア開発データが語るメッセージ」) を参照すると、テスト密度 (テストケース数÷開発規模) に関して以下のことが示されているのでヒントとして参考にする。

◇テストケース数の投資対効果という経済性の観点で見ると、テストケース数を増やすほどテストで検出される不具合件数の割合が低下して行く。(テスト密度を高くするほど、テスト検出能率が低下して行く。) 特にテスト密度の P75 あたり (58.6 テストケース/KSLOC) でテスト検出能率が下げ止まる傾向が見られることから、そこが経済性の高いテストケースの積み方の目安になる。





<参照箇所>

詳しくは、「ソフトウェア開発データが語るメッセージ」の「テストのコントロール」に関する記事を参照されたい。

(3) ベンチマーキング方法

テスト強化に向けて、次のことから、近未来のテスト密度の目標を 60 テストケース/KSLOC 以上に設定する。

- ・公開ベンチマーク（「ソフトウェア開発データが語るメッセージ」）に経済性の高いテストケース数の増やし方として、テスト密度の P75（58.6 テストケース/KSLOC）が目安として示されていることから、テスト密度がこれより低い現状であれば、この目安を参考にしてテスト強化を図ることとする。

(注 1) 内部ベンチマーク中に同様の知見が示されている場合、内部ベンチマークにおけるテスト密度の P75 の値を採用する方向で検討することが望ましい。

(注 2) 計画したテストケースのテスト実施が済んでいない場合や、不具合検出の推移に収束傾向が見られない場合はこの限りではない。

(注 3) テスト強化のためには、単にテスト密度向上を図るだけでなく、テストケース設計の品質向上を図ることも大切である。

4.1.3. IT 企業での具体例 1

<効率的な品質改善に向けた CMMI 成熟度レベル別の要因分析>

(1) 目的（活用シーン）

内部ベンチマーク中の「成熟度レベル別の品質変動要因及び品質改善のポイント」をヒントとして参考にし、自組織の開発プロセスの現状と対比しながら品質向上に向けた自組織の重点強化領域を特定し、品質マネジメント方針を検討する。

(2) 考え方

ソフトウェアが組み込まれた情報システムは拡大の一途を辿り、今日ではソフトウェアは日常生活になくはならない重要な社会基盤要素となっている。それに伴ってソフトウェアの不具合に起因する障害やサービス中断が社会に与える影響が拡大してきている。このようにソフトウェアが重要な役割を担うにつれて、ソフトウェアの品質をより一層向上させることが重要となってきている。品質向上に向けた重点強化領域は組織の成熟度レベルによって異なると考えられるので、自組織の成熟度レベルに適した重点強化領域を特定し、効率的に品質向上を図るための品質マネジメント方針を検討したい。

(3) ベンチマーク

<内部ベンチマークの導出方法>

各プロジェクトから、次のデータ項目について、実績データを収集した。

通番	データ項目	単位
1	開発規模（計画値）	ライン数
2	開発規模	ライン数
3	全工数	人時
4	上工程工数	人時
5	レビュー工数	人時
6	テスト工程工数	人時
7	全バグ数	件
8	上工程バグ数	件
9	テスト工程バグ数	件
10	テスト項目数	件
11	上工程バグ率	%

出荷後バグ基準を達成しているプロジェクト（達成）と達成していないプロジェクト（未達）の差異の要因をロジスティック回帰分析によって分析した結果、組織の成熟度レベルによって寄与している要因が次のように異なることが分かった。

成熟度レベル (ML)	達成と未達の差異要因	達成の可能性との関係
ML1	テスト工程バグ数/KL	低いと達成の可能性が高くなる
	開発規模	低いと達成の可能性が高くなる
ML2	上工程バグ率	高いと達成の可能性が高くなる
	レビュー工数/KL	低いと達成の可能性が高くなる
	開発規模	低いと達成の可能性が高くなる
ML3	テスト工程工数/KL	低いと達成の可能性が高くなる
	レビュー工数/KL	高いと達成の可能性が高くなる
	テスト工程バグ数/KL	低いと達成の可能性が高くなる

さらに未達プロジェクト群の傾向を分析した結果を踏まえて、成熟度レベル別に、効率的な品質改善のためのポイントを次のように纏めた。
これを内部ベンチマークとする。

<内部ベンチマーク>

ML	未達の要因	直近の対策	品質改善のポイント
ML1	・開発規模が大きくなると制御不可 ・プロジェクト遂行の差が大きく、特に未達はテスト不十分	・大規模開発 PJ の重点管理 ・基本事項の実行徹底 ・テスト工数の確保	基本的な開発・管理技術の整備と徹底
ML2	・開発規模が大きくなると制御不可 ・レビュー品質の良し悪しが出荷後品質に影響	・大規模開発 PJ の重点管理 ・基本事項の実行徹底 ・レビュー内容の確認	レビュー手法の標準化とビジネス特性に応じた最適化
ML3	・実績を基にした適切なマネジメントが実施されていない	・週次の実績データ収集とマネジメント会議開催	実績に基づく週次のマネジメント

<参照箇所>

詳しくは、ソフトウェア品質シンポジウム 2015 (SKiP2015) の B1-2 「効率的な品質改善に向けた CMMI 成熟度レベル別の要因分析」(柳田 礼子(日本電気(株)))を参照されたい。

<http://www.juse.jp/sqip/symposium/timetable/day1/>

本稿 (IT 企業での具体例 1) は、ベンチマーキングの観点から、その発表資料を抜粋したものである。

(4) ベンチマーキング方法

- ① 自組織の成熟度レベルは 2 なので、自組織の開発プロセスの現状を上記(3)の成熟度レベル 2 (ML2) の内容に照らしてみる。その結果、次のことが分かった。

「レビュー工数をかけている割にはレビューの品質 (パフォーマンス) が悪く、テスト工程でも多くのバグが検出されテスト工数もかかっているプロジェクトが多い。レビュー品質の良し悪しが出荷後品質に影響していると判断できる。」

- ② 上記のことから、「レビュー手法の改良と標準化」を、自組織に適した (効率的な) 品質マネジメントの重点方針とする。

参考ポイント

信頼性変動要因と品質向上に向けた重点強化領域は組織の成熟度レベルによって異なることから、自組織の成熟度レベルに適した重点強化領域を特定し、効率的／効果的に品質向上を図るための品質マネジメント方針を検討することは、理に適ったより良いアプローチと言える。

この例は、信頼性変動要因と品質改善のポイントとを組織の成熟度レベルごとに分析・検討したものをベンチマークとして採用し、それをヒントにして自組織の成熟度レベルに見合った最適な品質マネジメント方針を検討するためのベンチマーキング方法の事例として参考になると考えられる。

4.1.4. IT 企業での具体例 2

<ベンチマーキングによるプロセス改善事例>

(1) 目的（活用シーン）

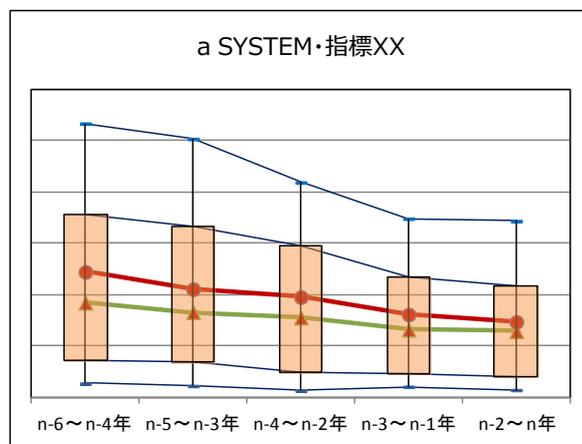
成熟度が高いシステム開発組織のプロセス指標群をベンチマークにして、自組織の課題を発掘し、成熟度向上に向けた改善アクションプランを検討・策定する。

(2) 考え方

当事例は改良開発中心の組織の事例である。改良開発が中心の場合、一般に同じシステムにおいては同じメンバーで対応することが多い。このため、知識やノウハウの継承がしやすいという長所があるが、いったん作業スタイルが確立すると、特に大きな問題が発生しない限り、自分たちのスタイルを意図的に変えることは少ない傾向がある。しかし、現場のメンバーは、課題を認識しており、それらの課題は往々にして当たっていることが多い。必要なのは変化へのトリガーであろう。定量データは客観的かつ多面的に仕事のやり方を見直すことができる有効なトリガーになる。

(3) ベンチマーク

この企業では、プロセス改善の推進、品質指標の基準値策定等を目的に、品質管理部門が各部門の案件実績データを取りまとめ、この企業独自の「白書」による開発プロセスの可視化に取り組んできた。「白書」には、指標単位のシステム別最新測定結果や経年変化の状況を掲載している。（下図はイメージ）



かねてより、自部門の仕事の進め方に課題を感じていた K 部部長は、定量データによる他部門との比較を通して、組織課題の発掘と改善を行いたいと考え、品質管理部門がそれを支援することになった。

そこで、成熟度が高い a システムをベンチマークにして、K 部所管の b システムと比較することにした。

(4) ベンチマーキング方法

① カルテの作成

実際の指標値や分布は「白書」で確認できるので、開発プロセス全体を俯瞰してベンチマークと比較する「カルテ」を作成した。なお、a システムと K 部所管の b システムは同じ開発標準を適用しており、開発環境も同じである。

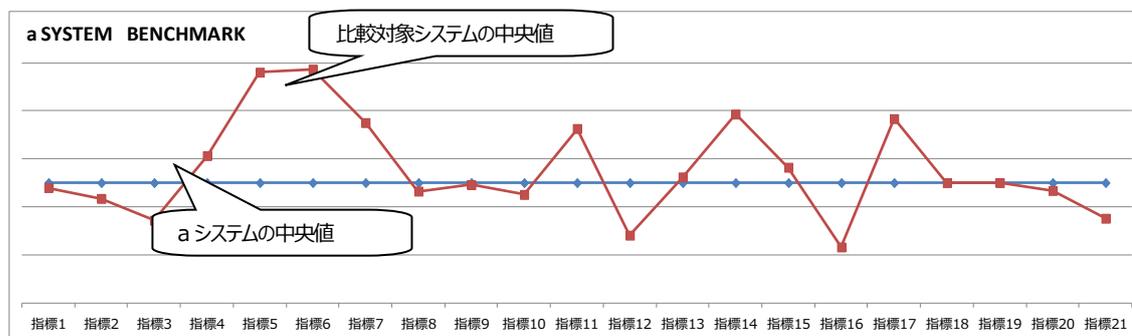
比較は中央値で行うこととし、俯瞰して比較できるように、各指標の中央値を変換した。a システムの各指標の中央値をすべて一定の値に変換し、比較対象システムの各指標の中央値を一定値との相対比で表した。

(例) a システムの中央値を 100 とした場合

	実数値		変換後	
	a システム	比較対象システム	a システム	比較対象システム
指標 1 の中央値	0.10	0.12	100	120
指標 2 の中央値	25.00%	35.00%	100	140
:	:	:	:	:

このように変換した指標の中央値を開発工程順に並べ可視化するとともに、比較から想定できる開発プロセスの特徴等をコメントし、「カルテ」として K 部へレポートした。このとき、特に強調したのは「指標が高い・低い、仕事の進め方が良い・悪いに直結するものではない。客観的かつ多面的に分析することで、仕事のやり方を振りかえってほしい。」ということである。

【カルテのイメージ】



② カルテに基づく自組織の課題の分析

この「カルテ」を受けて b システムを担当する K 部のラインマネージャーは、指標の差と、自チームの作業実態から次のように分析した。

◆ b システムのマネージャーの分析

【着目した指標】

指標	a システム比
設計レビュー率 (設計工程におけるレビュー工数の割合)	やや低い
レビュー指摘密度 (設計資料ページあたりの指摘数)	かなり高い
検証比率 (製作工程工数に対する検証工程工数の比率)	かなり高い
検証レビュー率 (検証工程におけるレビュー工数の割合)	高い
テスト密度 (KLOC あたりのテストケース数)	高い

【分析結果】

- ・設計レビューにかける時間が少なく、レビューでの指摘件数が多い。

サブマネジャーのプレレビュー⁴が不十分であり、成果物の完成度が低いケースがある。本来はレビューまでに解消すべき不備を、レビューの場で指摘している可能性がある。担当 SE やサブマネジャーにレビューで指摘してもらえばよい、という意識があり、成果物の完成度を高める努力が足りないのではないかと。

・検証工程に時間がかかりすぎている。

テスト方針やテスト方法のすり合わせが不十分なままテストを実施しており、テストのやり直しや過剰なテストを行っている。検証レビュー率が高いのは、やり直しによる再レビューの結果の可能性が高い。担当 SE、サブマネジャーとも余分な作業の結果、検証工程の工数が多いのではないかと。

③ アクションプランの策定

上記の分析結果に基づいて、自組織に適した改善アクションプランを検討・策定した。

【改善に向けたアクションプラン】

・設計工程におけるプレレビューの徹底

サブマネジャーがプレレビューを承認した設計資料のみをレビュー対象とする。

・レビューアによる設計資料の事前チェック

レビューアが事前に設計資料を確認し、一定のレベルに達したものをレビュー対象とする。

・指摘数の上限設定

レビューで指摘数が一定数を超えた場合、いったんレビューを中断し、再レビューとする。

・テスト方針・方法の認識合わせ

担当 SE とサブマネジャーの間で、よく話し合っ、テスト方針や目的、テスト方法、結果チェックの視点を合わせる。例えば、テスト方法であれば、DB の加工方法、エビデンス取得範囲、デグレード検証の方法等を確認しあう。

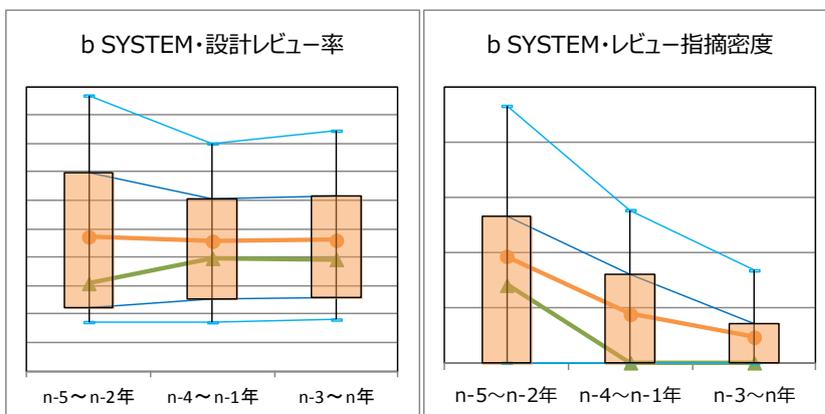
・テストレビュー⁵の強化

担当 SE とサブマネジャーの認識やテスト方針、方法の齟齬がないことを確認するとともに、テスト範囲や視点の適切性を精査し、過剰検証を抑制する。

④ 効果の確認

【アクションプラン実施後の評価】

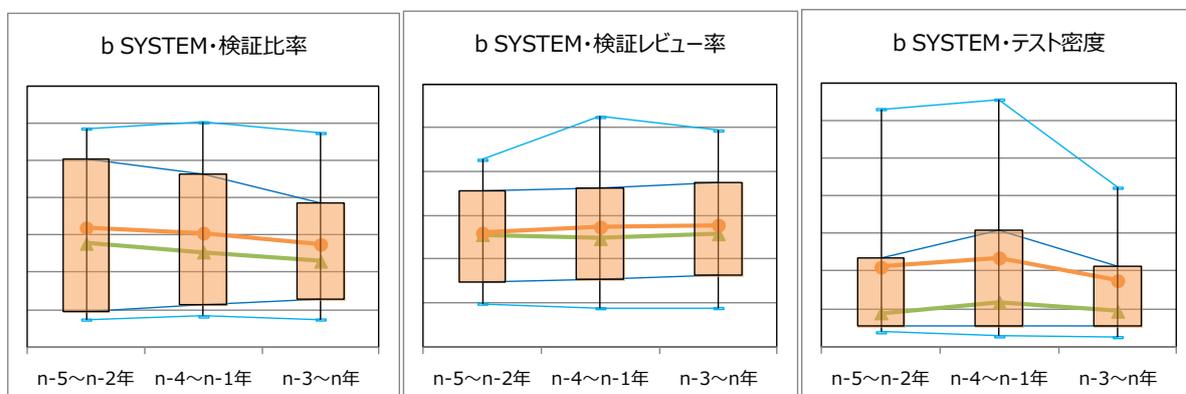
※箱ひげ図について 下端：P10、上端：P90 ●：平均値 ▲：中央値



⁴ サブマネジャーが担当 SE の成果物（設計資料やプログラム仕様書）を本レビューの前にレビューする。

⁵ 基本的に設計レビュー時にテスト計画やテストケースレビューを行う。また、テスト終了後にテスト結果レビューを行う。

設計レビュー率は、当初、低い部分に集中していたが、中央値が約 5 ポイント増加し、バランスがとれてきた。また、プレレビューの徹底により、成果物の品質が向上し、レビュー指摘密度は大きく減少、中央値は 0 でレビューの半数で指摘は発生していない。



検証比率は中央値、平均値とも、1 ポイント以上減少し、全体的に工数の圧縮が図られている。一方、検証レビュー率は、テストのやり直しに伴うレビューは減ったものの、テストレビュー自体の中身の充足を図ったことで、検証工程に占める割合に大きな変化は見られなかった。また、テスト密度については、中央値、平均値とも大きな変化はなかったが、全体のばらつきは小さくなっており、「念のためのテスト」の是正には効果があったと考えている。

参考ポイント

この例は、良い成果を挙げている成熟度の高い部門の仕事のやり方（開発プロセス）をベンチマークとし、それと自組織の現状とを比較することによって自組織の課題を分析し、その分析結果に基づいて自組織に適した改善アクションプランを検討・策定したものである。

この例は、このように典型的なベンチマーキング方法の好例として参考になると考えられる。

また、定量的管理、特にベンチマーキングを次のように仕事のやり方を見直すトリガーにするという点でも参考になると考えられる。

「いったん作業スタイルが確立すると、特に大きな問題が発生しない限り、自分たちのスタイルを意図的に変えることは少ない傾向がある。しかし、現場のメンバーは、課題を認識しており、それらの課題は往々にして当たっていることが多い。必要なのは変化へのトリガーであろう。定量データは客観的かつ多面的に仕事のやり方を見直すことができる有効なトリガーになる。」

4.2. テスト工程完了評価関連標準類の見直し（品質マネジメント関連）

4.2.1. ベンチマーキング方法の手順例

ベンチマーキング方法の基本的な手順の例を以下に示す。

(1) ベンチマーキングの目的を確認する。

ベンチマーキングの目的が、「ベンチマーク中に「テスト完了評価方法に関する知見」があれば、自組織の現状と対比しながら必要に応じて、品質マネジメントのテスト工程完了評価関連の標準類に反映することを検討すること」であることを確認する。

(備考) ベンチマーク中の「テスト完了評価方法に関する知見」は、ポストプロセス計測データ（完了プロジェクト群の実績データ）を用いた評価方法であって、インプロセス計測データを用いたリアルタイムな評価方法ではない。（インプロセス計測データを用いた評価方法については、必要に応じて「定量的品質予測のススメ」及び「続 定量的品質予測のススメ」を参照されたい。）

一般に、テスト完了評価方法では、インプロセス計測データに基づく評価基準を含めた複数の評価基準から成る総合評価方式を採っている。（各評価基準は総合評価における必要条件の一つ一つに相当する。）例えば、次のような評価基準から成る。

前提として、テスト関連文書がテスト関連標準類に沿って作成され、それらのレビューが実施されていること。かつ、レビューコメントの処置が完了していること。

前提として、テスト密度（テストケース数÷開発規模）が基準値を満足していること。

前提として、未テスト項目、未修正項目が残っていないこと。

テスト密度とテスト検出不具合密度（テスト検出不具合数÷開発規模）との関係、あるいはテストケース数とテスト検出不具合数との関係等が、管理基準を満足していること。

障害／誤り件数の推移に、収束傾向が認められること。

（インプロセス計測データ使用）

障害内容（障害の重大度、障害の発生条件）の推移に、収束傾向が認められること。

（インプロセス計測データ使用）

障害／誤り件数の推移から予測した残存誤り密度が信頼性目標を満足するものであること。

（インプロセス計測データ使用）

ベンチマーク中の「テスト完了評価方法に関する知見」は、主に上記の④に該当する。また、開発組織のマネジャー層やPMO及び品質マネジメント推進部門が、ポストプロセス計測データから簡便にテスト完了評価する方法に関するものと言える。

(2) ベンチマーク中に「テスト完了評価方法に関する知見」があれば、自組織の現状のテスト完了評価方法と対比しながら、テスト工程完了評価関連の標準類（テスト工程完了評価基準等）に反映（追加）すると良いかどうかを検討する。

(3) 上記の検討結果を、品質マネジメントのテスト工程完了評価関連の標準類（テスト工程完了評価基準等）に反映する（必要に応じて追加する）。

4.2.2. 白書等を用いた具体例

<ゾーン分析によるテスト完了評価を追加検討する例>

(1) 目的 (活用シーン)

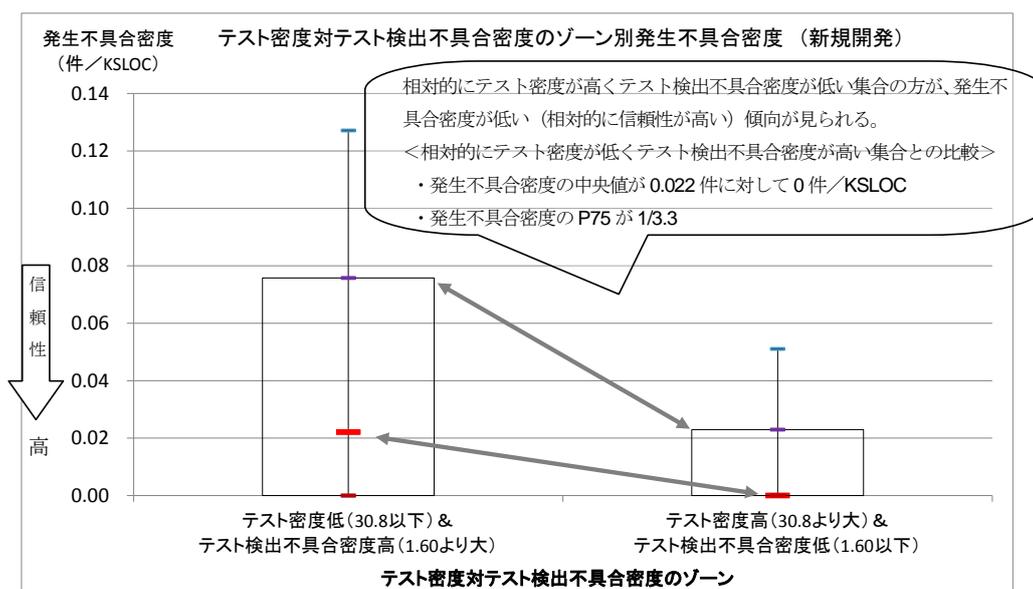
主に開発組織のマネジャー層が、次のような観点から、プロジェクト・マネジメントの改善を推進するシーン (PMO 及び品質マネジメント推進部門が、マネジメントを支援するシーンを含む)

◇ベンチマーク中に「ゾーン分析に関するテスト完了評価方法の知見」があれば、自組織の現状と対比しながら必要に応じて、品質マネジメントのテスト工程完了評価関連の標準類に反映することを検討する。

(2) ベンチマーク

公開ベンチマーク (「ソフトウェア開発データが語るメッセージ」) を参照すると、テスト評価方法に関して以下のことが示されているのでヒントとして参考にする。

テスト密度が高くてテスト検出不具合密度が低いのは相対的に信頼性が良い兆候の一つである。一方、テスト密度が低くてテスト検出不具合密度が高いのは相対的に信頼性が良くない兆候の一つである。この見方をテストの評価項目の一つとして採用することをお勧めする。



(備考) 考察

「テスト密度が高くてテスト検出不具合密度が低いのは相対的に信頼性が良い (稼働後の不具合発生数が少ない) 兆候の一つである」ことの主な要因として、次のことが考えられる。

◇テスト密度が高いのにテスト検出不具合密度が低いものは、テスト開始時点の出来が良い (潜在不具合の密度が低い)、つまりいわゆる作込み品質が良いものと考えられる。

◇テスト検出不具合密度が高いものは発生不具合密度 (稼働後の不具合密度) も高い傾向が見られる。また、(テスト密度が低くなくて) テスト検出不具合密度が低いものは発生不具合密度 (稼働後の不具合密度) も低い傾向が見られる。つまり、テストに至るまでの良し悪しがテストによって逆転するケースは少ないと言える。信頼性を高めるには、やはり作込み品質向上を目指すことが王道であ

り、テストによって挽回しようという作戦の成算は薄いと考えられる。

<参照箇所>

詳しくは、「ソフトウェア開発データが語るメッセージ」の「テストのコントロール」に関する記事を参照されたい。

(3) ベンチマーキング方法

自組織の現状のテスト完了評価方法を眺みながら、当方法をテスト工程完了評価関連の標準類（テスト工程完了評価基準等）に反映（追加）すると良いかどうかを検討する。

（備考1）ゾーンについて

2×2の4ゾーンに分ける方法、3×3の9ゾーンに分ける方法等がある。

（備考2）各ゾーンの見方と善後策について

- ・テスト密度が低くてテスト検出不具合密度が低いものについては、テスト不足の恐れがあり、現状で良否を判断するのは難しい。
- ・テスト密度が低くてテスト検出不具合密度が高いものについては、作込み品質が低い恐れがある。テスト密度を高めればさらに不具合を検出できる可能性が高いと考えられる。
- ・テスト密度が高くてテスト検出不具合密度が高いものについても、作込み品質が低い恐れがある。テストによって不具合を検出し尽くせたかどうかを現状で判断するのは難しい。

上記の各ゾーンについては、追加テストを実施し、その結果に基づいて再評価するなどの善後策が必要と考えられる。

上記の検討結果を、品質マネジメントのテスト工程完了評価関連の標準類（テスト工程完了評価基準等）に反映する（必要に応じて追加する）。

<テスト検出能率によるテスト完了評価を追加検討する例>

(1) 目的（活用シーン）

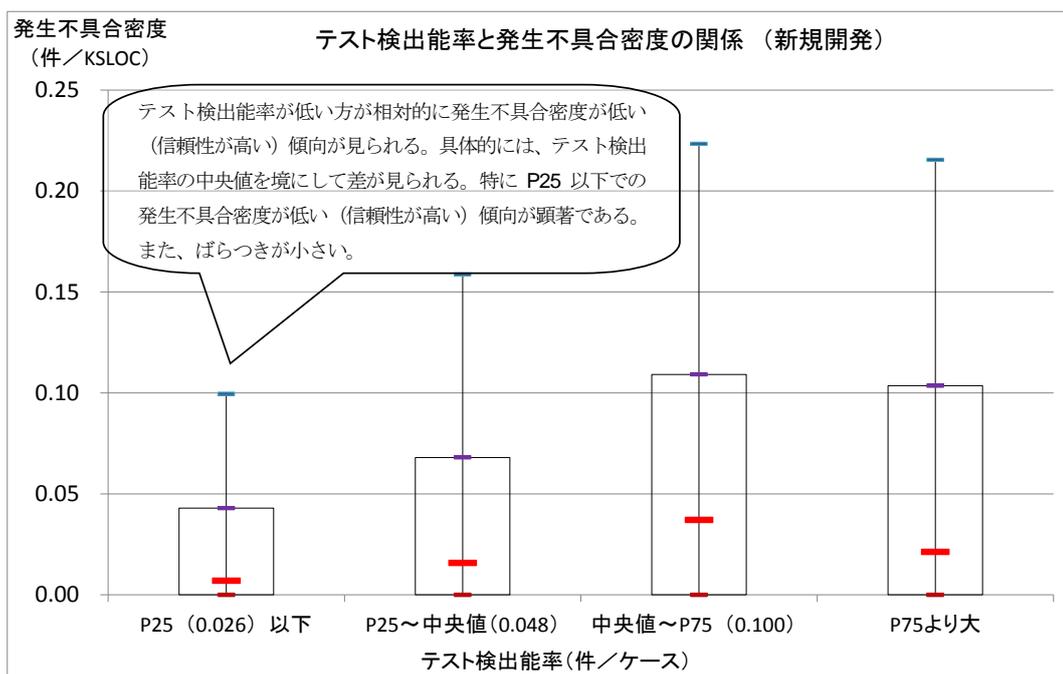
主に開発組織のマネジャー層が、次のような観点から、プロジェクト・マネジメントの改善を推進するシーン（PMO及び品質マネジメント推進部門が、マネジメントを支援するシーンを含む）

◇ベンチマーク中に「テスト検出能率に関するテスト完了評価方法の知見」があれば、自組織の現状と対比しながら必要に応じて、品質マネジメントのテスト工程完了評価関連の標準類に反映することを検討する。

(2) ベンチマーク

公開ベンチマーク（「ソフトウェア開発データが語るメッセージ」）を参照すると、テスト評価方法に関して以下のことが示されているのでヒントとして参考にする。

「テスト検出能率がある一定のレベルまで低下しているか否か」を、テストの評価項目の一つとして利用することをお勧めする。また、追加テストの収束性評価に利用することをお勧めする。



(備考) 考察

この結果からは、テスト検出能率が 25 パーセンタイル値を目安として下回ることが望ましいと考えられる。新規開発の場合は、テスト検出能率の 25 パーセンタイル値は約 0.026 件/テストケース。これは、およそ 40 ケースのテストに対して 1 件の不具合検出に相当する。上記の結果は、テスト工程全体での累積値に基づくものであることに留意されたい。テスト終盤におけるテスト検出能率を評価できれば、そのテスト検出能率は上記の結果よりも低くなるはずである。

テスト終盤におけるテストの収束性を評価するシーンにおいても、テスト検出能率による評価が有用と考えられる。(テスト検出能率がテストの進捗に連れて低下していくということは、潜在している(残存している)不具合が減少していくということを意味する。テスト検出能率によってテストの収束性を評価することは、理に適っていると考えられる。)

<参照箇所>

詳しくは、「ソフトウェア開発データが語るメッセージ」の「テストのコントロール」に関する記事を参照されたい。

(3) ベンチマーキング方法

- ① 自組織の現状のテスト完了評価方法を眺みながら、当方法をテスト工程完了評価関連の標準類 (テスト工程完了評価基準等) に反映 (追加) すると良いかどうかを検討する。

(備考) テスト終盤での収束性評価や、追加テストの評価にも有用と考えられる。

- ② 上記の検討結果を、品質マネジメントのテスト工程完了評価関連の標準類 (テスト工程完了評価基準等) に反映する (必要に応じて追加する)。

4.3. ユーザの関与（ステークホルダー・マネジメント関連）

4.3.1. ベンチマーキング方法の手順例

ベンチマーキング方法の基本的な手順の例を以下に示す。

(1) ベンチマーキングの目的を確認する。

ベンチマーキングの目的が、「ベンチマーク中に「ユーザの関与に関する知見」があれば、自組織の現状と対比しながら必要に応じて、信頼性／生産性の向上に向けてプロジェクト・マネジメント（特にステークホルダー・マネジメント）の改善を検討し、検討結果を踏まえて、ステークホルダー・マネジメント関連の標準類を見直すこと」であることを確認する。

(備考) ステークホルダー・マネジメントにおいては、ステークホルダー（特にユーザ）を効果的に関与させることが、プロジェクトを成功に導く重要な要素となっている。また、スコープ・マネジメントにおいては、成果物スコープ（成果物とその要件）のベースラインを確立しステークホルダー間で合意形成するプロセス（要件定義に相当）が重要な要素となっている。例えば、ユーザ担当者の要求仕様への関与度合いが高いと、要求仕様が早期に明確になる、要求仕様の変更／紆余曲折が少なくなり手戻りが少なくなる等の効果が期待できるので、生産性、信頼性両面で有利と考えられる。
文書化、レビュー、テスト等の品質保証プロセスの工夫・努力によって信頼性向上を図ることは主に開発者自身が工夫・努力すれば良いことであるが、当件の場合はユーザの関与度合いを高めるべく、開発組織のマネジャー層がユーザの理解と協力を取り付けることが重要かつ効果的と考えられる。

(2) ベンチマーク中に「ユーザの関与に関する知見」があれば、自組織の現状のステークホルダー・マネジメントと対比しながら、ステークホルダー・マネジメント関連の標準類に反映（追加）すると良いかどうかを検討する。

(3) 上記の検討結果を、ステークホルダー・マネジメント関連の標準類に反映する（必要に応じて追加する）。

4.3.2. 白書等を用いた具体例

<ユーザ担当者の要求仕様関与度合いの向上の例>

(1) 目的（活用シーン）

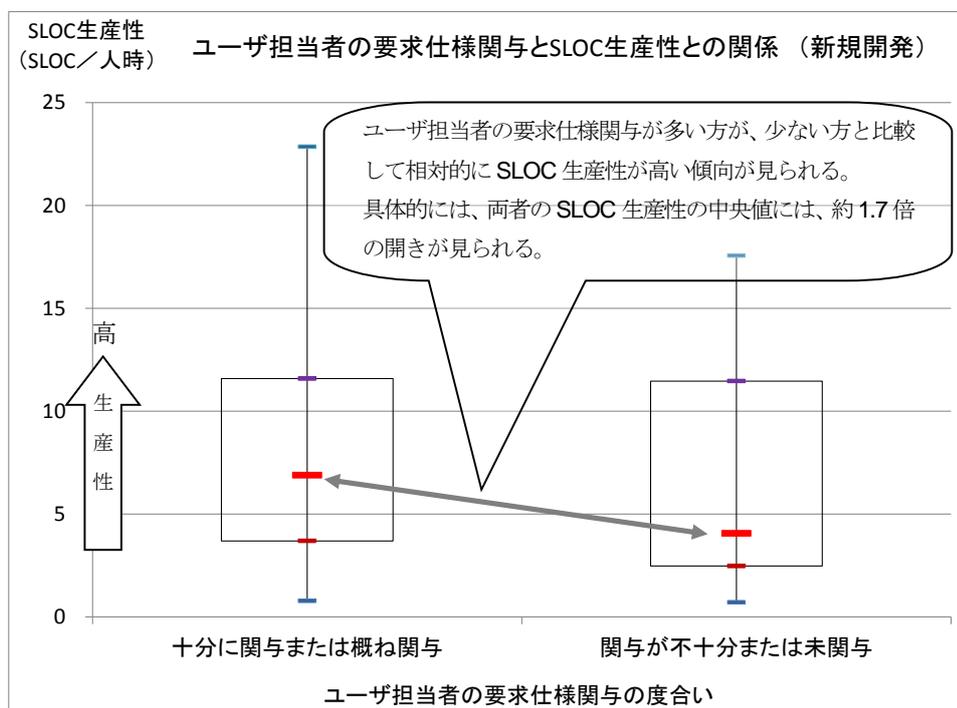
主に開発組織のマネジャー層が、次のような観点から、プロジェクト・マネジメントの改善を推進するシーン（PMO 及び品質マネジメント推進部門が、マネジメントを支援するシーンを含む）

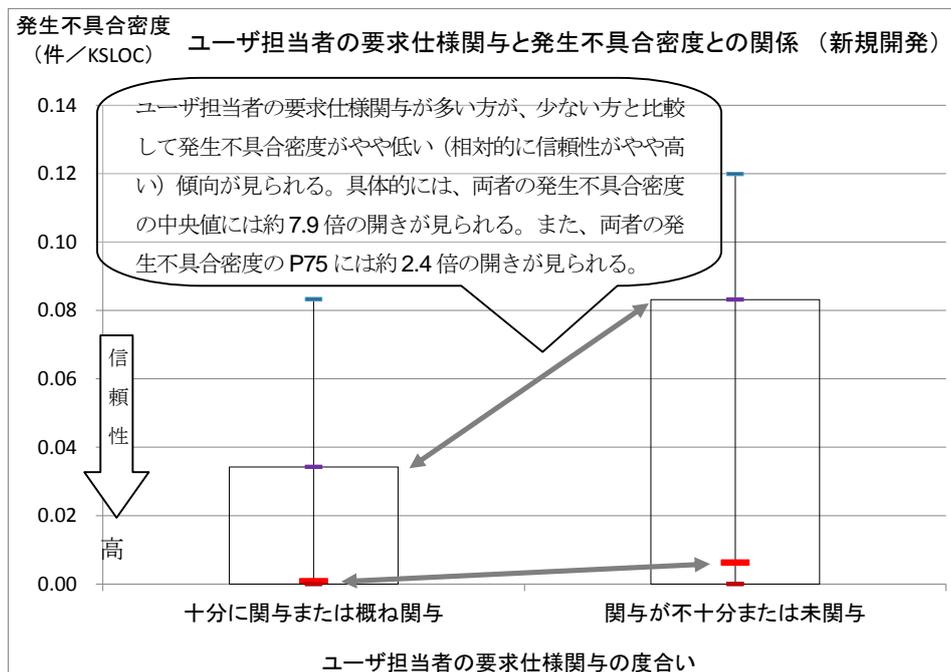
◇ベンチマーク中に「ユーザの関与に関する知見」があれば、自組織の現状と対比しながら必要に応じて、信頼性／生産性の向上に向けてプロジェクト・マネジメント（特にステークホルダー・マネジメント）の改善を検討する。検討結果を踏まえて、ステークホルダー・マネジメント関連の標準類を見直す。

(2) ベンチマーク

公開ベンチマーク（「ソフトウェア開発データが語るメッセージ」）を参照すると、「ユーザの関与に関する知見」の一つとして、「ユーザ担当者の要求仕様関与度合い」について以下のことが示されているのでヒントとして参考にする。

ユーザ担当者の要求仕様関与が多い方が、少ない方と比較して相対的に生産性及び信頼性が良い傾向が見られる。生産性向上及び信頼性向上に向けて、ユーザ担当者の要求仕様関与度合いを高めるようユーザと調整することが望ましい。





<参照箇所>

詳しくは、「ソフトウェア開発データが語るメッセージ」の「ユーザの関与」に関する記事を参照されたい。

(3) ベンチマーキング方法

- ① 自組織の現状の標準類を眺みながら、ステークホルダー・マネジメントに関して、例えば次のことを検討する。
 - ・ユーザ担当者の要求仕様関与度合いを高めるべく、ユーザとの関係の築き方、ユーザの参画度の高め方、ユーザとの役割分担の決め方等を検討する。
 - ・各プロジェクトにおいて、ユーザ担当者の要求仕様関与度合いを高めるべく、開発組織のマネジャー層がユーザの理解と協力を取り付ける。
- ② 上記の検討結果を、ステークホルダー・マネジメント関連の標準類に反映する (必要に応じて追加する)。

4.4. 成果物スコープの明確化（スコープ・マネジメント関連）

4.4.1. ベンチマーキング方法の手順例

ベンチマーキング方法の基本的な手順の例を以下に示す。

(1) ベンチマーキングの目的を確認する。

ベンチマーキングの目的が、「ベンチマーク中に「要求仕様の明確化に関する知見」があれば、自組織の現状と対比しながら必要に応じて、信頼性／生産性の向上に向けてプロジェクト・マネジメント（特にスコープ・マネジメント）の改善を検討し、検討結果を踏まえて、スコープ・マネジメント関連の標準類を見直すこと」であることを確認する。

(備考) 要求仕様を早期に明確化すると、要求仕様の変更／紆余曲折が少なくなり手戻りが少なくなる等の効果が期待できるので、生産性、信頼性両面で有利と考えられる。スコープ・マネジメントにおいては、成果物スコープ（成果物とその要件）のベースラインを確立しステークホルダー間で合意形成するプロセス（要件定義に相当）が重要な要素となっている。また、ステークホルダー・マネジメントにおいては、ステークホルダー（特にユーザ）を効果的に関与させることが、プロジェクトを成功に導く重要な要素となっている。要求仕様の早期明確化に向けて、要件定義に注力するとともにユーザの要求仕様への関与度合いを高めることが重要と考えられる。

文書化、レビュー、テスト等の品質保証プロセスの工夫・努力によって信頼性向上を図ることは主に開発者自身が工夫・努力すれば良いことであるが、当件の場合には要求仕様の早期明確化に向けて、自組織のマネジャー層が要件定義に関するユーザの理解と協力を取り付けることが重要かつ効果的と考えられる。

(2) ベンチマーク中に「要求仕様の明確化に関する知見」があれば、自組織の現状のスコープ・マネジメントと対比しながら、スコープ・マネジメント関連の標準類に反映（追加）すると良いかどうかを検討する。

(3) 上記の検討結果を、スコープ・マネジメント関連の標準類に反映する（必要に応じて追加する）。

4.4.2. 白書等を用いた具体例

<要求仕様の明確化の例>

(1) 目的（活用シーン）

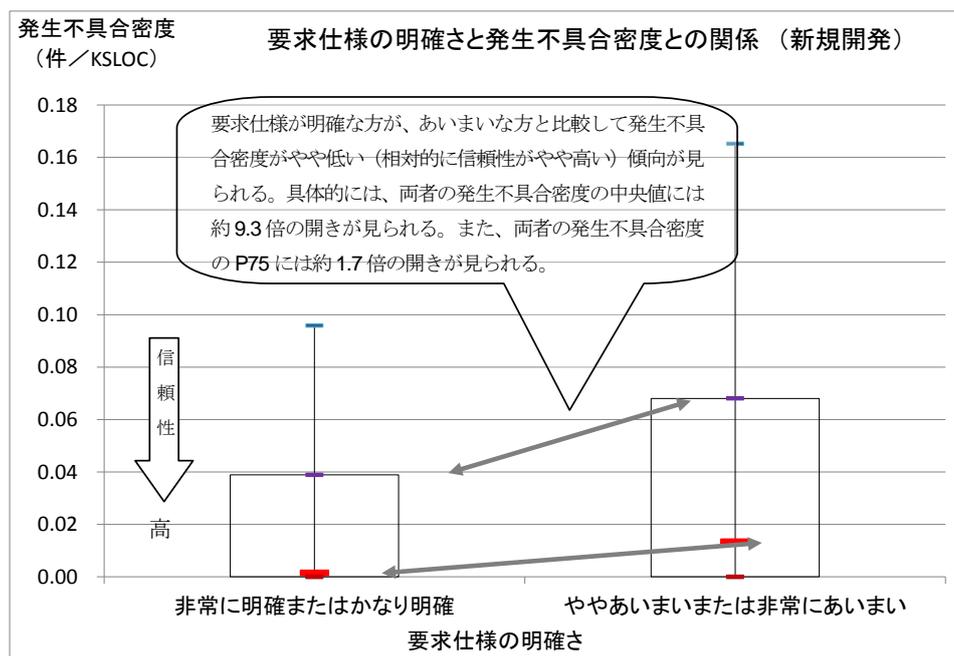
主に開発組織のマネジャー層が、次のような観点から、プロジェクト・マネジメントの改善を推進するシーン（PMO 及び品質マネジメント推進部門が、マネジメントを支援するシーンを含む）

◇ベンチマーク中に「要求仕様の明確化に関する知見」があれば、自組織の現状と対比しながら必要に応じて、信頼性／生産性の向上に向けてプロジェクト・マネジメント（特にスコープ・マネジメント）の改善を検討する。検討結果を踏まえて、スコープ・マネジメント関連の標準類を見直す。

(2) ベンチマーク

公開ベンチマーク（「ソフトウェア開発データが語るメッセージ」）を参照すると、「要求仕様の明確化に関する知見」に関して以下のことが示されているのでヒントとして参考にする。

要求仕様が明確な方が、あいまいな方と比較して発生不具合密度がやや低い（相対的に信頼性がやや高い）傾向が見られる。実現性の高い開発計画策定に向けて、要求仕様を早期に明確化するようユーザと調整することが望ましい。



<参照箇所>

詳しくは、「ソフトウェア開発データが語るメッセージ」の「要求仕様の明確化」に関する記事を参照されたい。

(3) ベンチマーキング方法

自組織の現状の標準類を眺みながら、スコープ・マネジメントに関して、例えば次のことを検討する。

- ・要求仕様を早期に明確化すべく、スコープ・マネジメントの実施方法（スコープ計画プロセス、スコープ定義プロセス等）や要件定義工程のプロセスを検討する。
- ・各プロジェクトにおいて、要求仕様を早期に明確化すべく、開発組織のマネジャー層がユーザの理解と協力を取り付ける。

(備考) 要求仕様の早期明確化に関しては、一般に開発ベンダ側、ユーザ側双方にいくつかの問題が存在することがある。これらの問題を解決できないプロジェクトの場合、これらをリスクとしてマネジメントせざるを得ないこともある。

- ・ベンダ側要員の業務経験・業務知識が乏しくて、ユーザ要件を的確に把握できない。
- ・ユーザ側要員の業務経験・業務知識が乏しいなど、ユーザ側の体制が弱くてユーザ要件を詰められない。
- ・ユーザが要件を小出しにする。設計工程やテスト工程においても仕様の追加／変更が絶えない。
- ・ユーザ側の情報システム部門が現場部門のニーズを的確に把握できていないなど。

上記の検討結果を、スコープ・マネジメント関連の標準類に反映する（必要に応じて追加する）。

5. 開発組織のマネジャー層用のベンチマーキング（組織の改善）

5.1. 組織体制の整備

5.1.1. ベンチマーキング方法の手順例

ベンチマーキング方法の基本的な手順の例を以下に示す。

- (1) ベンチマーキングの目的を確認する。
ベンチマーキングの目的が、「ベンチマーク中に「組織体制の整備に関する知見」があれば、自組織の現状と対比しながら必要に応じて組織体制の改善を検討し、検討結果を踏まえて、組織体制関連の標準類や年度計画／中長期計画を見直すこと」であることを確認する。
- (2) ベンチマーク中に「組織体制の整備に関する知見」があれば、自組織の組織体制整備状況と対比しながら、組織体制関連の標準類や年度計画／中長期計画に反映すると良いかどうか検討する。
- (3) 上記の検討結果を、組織体制関連の標準類に反映する。また、年度計画／中長期計画を見直す。

5.1.2. 白書等を用いた具体例

<品質保証体制の例>

(1) 目的（活用シーン）

主に開発組織のマネジャー層が、次のような観点から、組織の改善のためのマネジメントを推進するシーン（PMO 及び品質マネジメント推進部門が、マネジメントを支援するシーンを含む）

◇ベンチマーク中に「組織体制の整備に関する知見」があれば、自組織の現状と対比しながら必要に応じて組織体制の改善を検討する。検討結果を踏まえて、組織体制関連の標準類や年度計画／中長期計画を見直す。

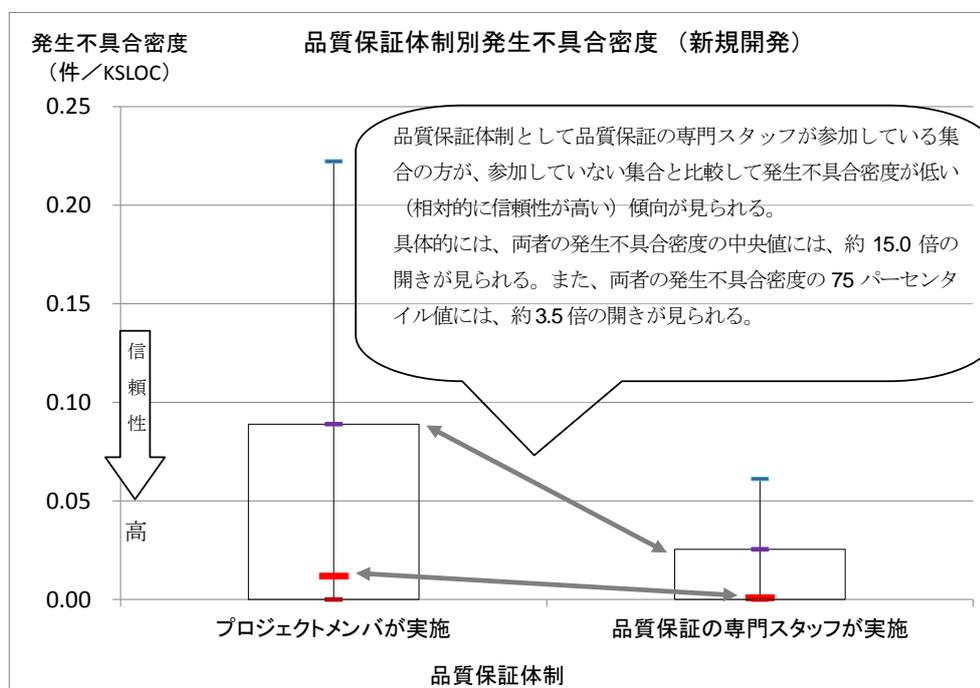
(2) ベンチマーク

公開ベンチマーク（「ソフトウェア開発データが語るメッセージ」）を参照すると、「組織体制の整備に関する知見」に関して以下のことが示されているのでヒントとして参考にする。

品質保証体制として品質保証の専門スタッフが参加しているプロジェクト群の方が、参加していないプロジェクト群と比較して発生不具合密度が低い（相対的に信頼性が高い）傾向が見られる。

信頼性向上に向けた組織体制整備の一環として、次のことをお勧めする。

- ・より多くの開発プロジェクトに品質保証の専門スタッフが参加できるように、組織体制を強化する。
- ・開発要員に対して、ソフトウェアの品質保証及びプロジェクト・マネジメントのスキル向上を図る。



<参照箇所>

詳しくは、「ソフトウェア開発データが語るメッセージ」の「組織体制の整備」に関する記事を参照されたい。

(3) ベンチマーキング方法

自組織の組織体制整備状況を眺みながら、品質保証体制の強化に向けて、組織体制関連の標準類や年度計画／中長期計画の見直しを検討する。

(備考) 考察

品質保証の専門スタッフがプロジェクトに参加することによって、専門スタッフのスキル／知識／経験が、主に次のような場面で、信頼性向上に向けて有効に働くと考えられる。

◇リスクの少ない（実現可能性の高い）開発計画の策定支援

種々の成功／失敗の事例を知っていることと、プロジェクト・マネジメントのスキルが高いことから、種々のプロジェクト・リスクを予見するとともにリスク対策について助言できる。

◇レビュー及びテストに向けた効果的な助言

種々の不具合事例に基づいた弱点（陥りやすい誤り、漏れやすい箇所）に関する知識が豊富なので、レビュー及びテストに対して効果的に助言できる。

◇精度の高い信頼性予測・評価

信頼性予測・評価スキルが高いので、開発時データ（文書化、レビュー、レビュー指摘、テスト、テスト時障害・誤り等のデータ）から、中間成果物及び成果物の信頼性／収束性を精度高く予測・評価できる。

◇定量的管理の推進

定量的管理のスキルと豊富な経験から、プロジェクトの定量的管理を指導／支援できる。特に値を読み解く力があるので、定量データの分析結果をより適切なマネジメント・アクションに繋げることができる。

上記の検討結果を、必要に応じて組織体制関連の標準類に反映する。また、年度計画／中長期計画を見直す。

5.2. 信頼性変動要因分析と信頼性向上に向けた重点強化領域の特定

5.2.1. ベンチマーキング方法の手順例

ベンチマーキング方法の基本的な手順の例を以下に示す。

(1) ベンチマーキングの目的を確認する。

ベンチマーキングの目的が、「ベンチマーク中に「信頼性変動要因に関する知見」があれば、自組織に当てはまるかどうかを検討し、必要に応じて信頼性向上に向けた重点強化領域を見直し、その結果を踏まえて信頼性向上方策を策定し、開発作業標準類、管理標準類、年度計画／中長期計画等を見直すこと」であることを確認する。

(備考) 信頼性変動要因をコントロールできれば、信頼性が向上すると期待できる。従って、信頼性変動要因は、信頼性向上に向けて重点的に強化すると効果的な領域（重点強化領域）を特定するための有用な情報となる。信頼性向上に向けては、自組織の信頼性変動要因群を踏まえて重点強化領域を特定し、それらの領域を改善する適切な方策を立てることが一つの重要かつ効果的なアプローチと言える。

(2) ベンチマーク中に「信頼性変動要因に関する知見」があれば、自組織に当てはまるかどうかを検討し、必要に応じて信頼性向上に向けて重点的に強化すると効果的な領域（重点強化領域）を見直す。

(注) 各変動要因が自組織にとって主要な変動要因に該当するか否かを判断する必要がある。例えば、自組織において固定的になっている変動要因は、自組織にとっての変動要因には当たらない。

(3) 見直した重点強化領域を改善する適切な方策を策定する。

(注) 既に対策済みであり、更に強化する必要はない（卒業している）と判断できる重点強化領域については、その限りではない。

(4) 策定した改善方策を、自組織の標準類や年度計画／中長期計画に反映する。

5.2.2. 白書等を用いた具体例

<信頼性変動要因に基づく重点強化領域特定の例>

(1) 目的（活用シーン）

主に開発組織のマネジャー層が、次のような観点から、組織の改善のためのマネジメントを推進するシーン（PMO 及び品質マネジメント推進部門が、マネジメントを支援するシーンを含む）

◇ベンチマーク中に「信頼性変動要因に関する知見」があれば、自組織に当てはまるかどうかを検討し、必要に応じて信頼性向上に向けた重点強化領域を見直す。その結果を踏まえて信頼性向上方策を策定し、開発作業標準類、管理標準類、年度計画／中長期計画等を見直す。

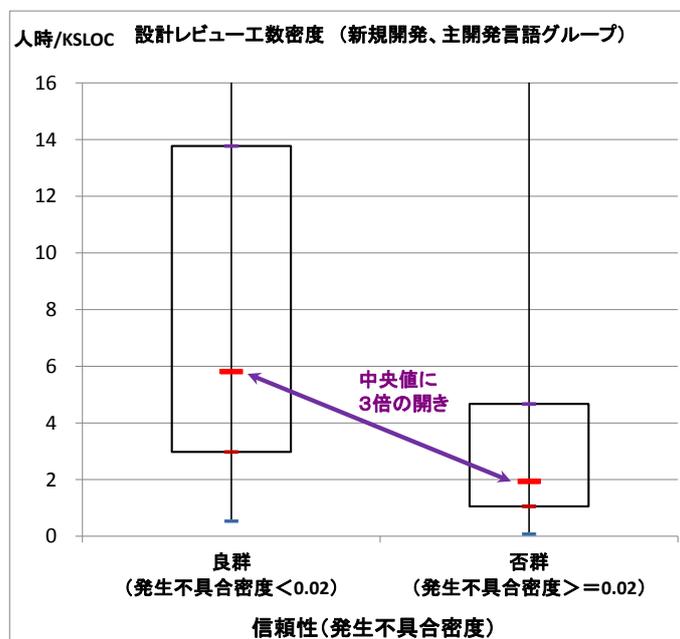
(2) ベンチマーク

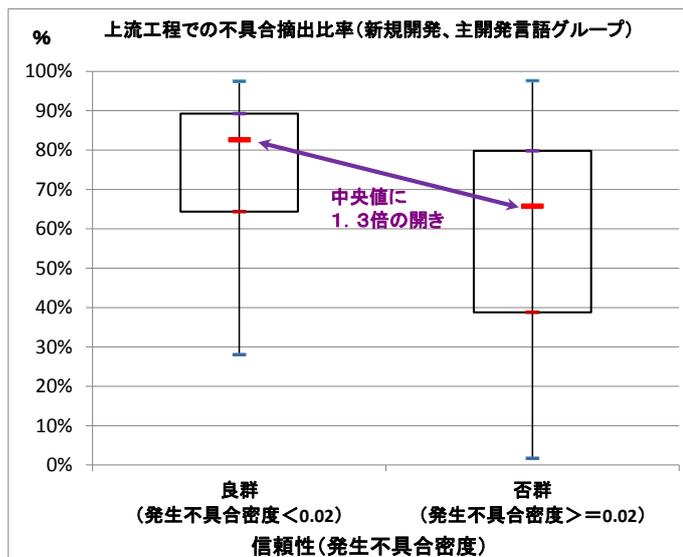
公開ベンチマーク（「ソフトウェア開発データが語るメッセージ」）を参照すると、「信頼性変動要因に関する知見」に関して以下のことが示されているので、ヒントとして参考にする。

「ソフトウェア開発データ白書」のデータに基づいた分析結果では、いくつかの要因によって信頼性（発生不具合密度）が変動する傾向が見られる。

具体的には、相対的に信頼性の良いグループ（良群）に次の傾向が見られる。

- ・レビュー工数密度が高い。
- ・上流工程での不具合摘出比率が高い。
- ・テスト検出不具合密度が低い。
- ・テスト検出能率（テスト項目当りの検出不具合数）が低い。





「レビュー工数密度を高めると信頼性が向上する」、「上流工程での不具合抽出比率を高めると信頼性が向上する」等から、設計及びレビューのプロセス改善が信頼性向上に向けた重点強化領域の一つと考えられる。

(例) レビュー工数密度を高める。

レビュー効率を高める(設計文書化密度を上げる、設計書作成標準やチェックリストを改良する等)

<参照箇所>

詳しくは、「ソフトウェア開発データが語るメッセージ」の「信頼性変動要因に関する分析」に関する記事を参照されたい。

(3) ベンチマーキング方法

- ① 上記(2)に示されたレビュー工数密度、上流工程での不具合抽出比率等が自組織の信頼性変動要因に当てはまるかどうかを検討する。(上記(2)をヒントにして、自組織の定量データを用いて自組織の信頼性変動要因を分析することが望ましい。)当てはまるのであれば、設計及びレビューのプロセス改善を、信頼性向上に向けて重点的に強化すると効果的な領域(重点強化領域)に加えることを検討するなど、重点強化領域を見直す。
- ② 見直した重点強化領域を改善する適切な方策を策定する。
設計及びレビューのプロセス改善については、レビュー工数密度(あるいはレビュー工数比率)を高めるだけでなく、レビューのパフォーマンスを高めるための次の対策も検討する。
 - ・レビュー方式の改善
 - ・レビュー・チェックリストの改良
 - ・設計文書化密度の向上
 - ・設計書作成標準の改良
- ③ 策定した改善方策を、自組織の標準類や年度計画/中長期計画に反映する。

5.2.3. IT 企業での具体例 1

<出荷後品質の良群・否群の比較>

「ソフトウェア開発データが語るメッセージ 2015」によると、発生不具合密度の低い（信頼性の高い）プロジェクトは、上流工程での不具合摘出比率が高い傾向があることが示されている。当社のプロジェクトデータを使用し、同様に出荷後の発生不具合密度（件/KL）に基づいて品質の良いプロジェクト（良群）と品質の悪いプロジェクト（否群）の傾向を分析し、その結果を踏まえて品質向上方針を検討した例である。

(1) 目的（活用シーン）

社内の品質（信頼性）変動要因に関する知見をベンチマークとし、自組織の現状をそのベンチマークと比較検討することによって、自組織の品質向上方針を検討する。

(2) ベンチマーク

① プロジェクトの層別

組織で設定した発生不具合密度の基準値を達成しているプロジェクト（良群）と、達成していないプロジェクト（否群）に層別し、中央値を比較した。

この際、良群と否群の値に明らかな違いがあるのかを判断する目安として、Wilcoxon/Kruskal-Wallis の検定を行った。その結果、有意差があるという結果が得られたデータ項目を以下に記す。

Wilcoxon/Kruskal-Wallis の検定結果

データ項目	単位	p 値
全摘出バグ	件/KL	0.0031
上工程バグ	件/KL	0.0085
テスト工程バグ	件/KL	0.0002
上工程バグ摘出率（注）	%	0.0274
テスト工程工数	人 H/KL	0.0401

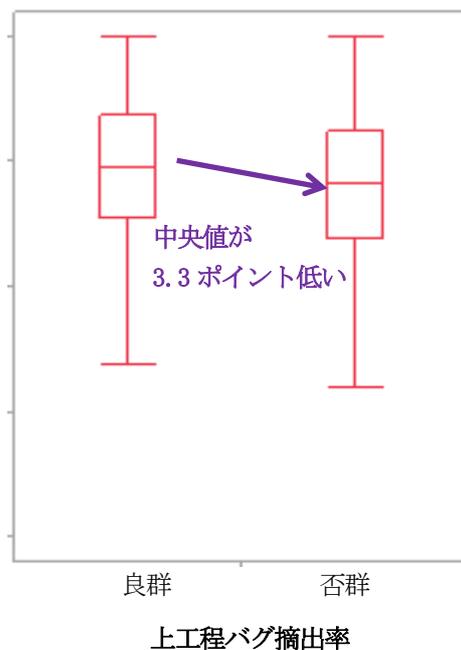
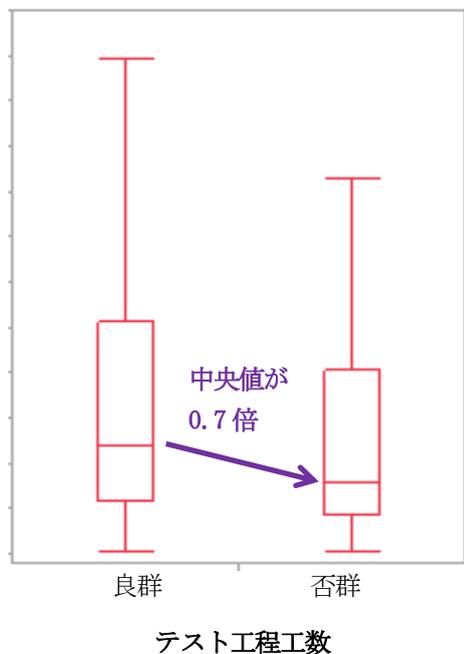
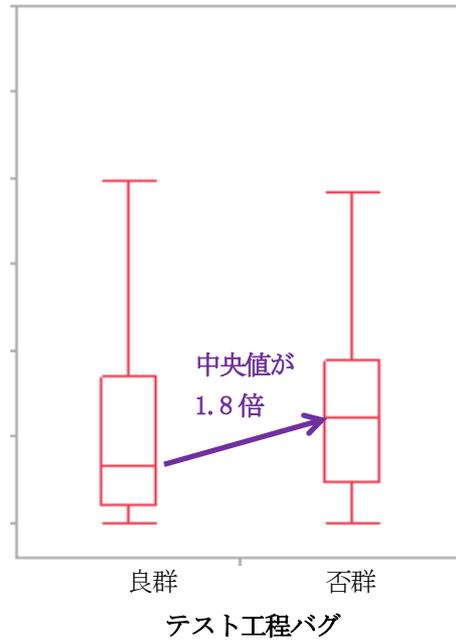
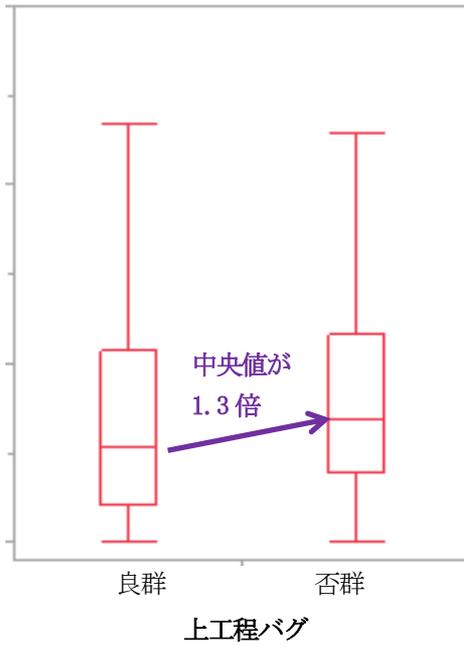
（注）当社では、V字モデルの開発プロセスにおける基本設計からコーディングまでを「上流工程」としている。

② 中央値の比較結果（ベンチマーク）

有意差ありという結果が得られたデータ項目について、良群と否群の中央値を比較した結果を分析すると、良群は否群に対して上流工程バグ摘出率が高い傾向が見られる。

一方、否群には、以下の傾向が見られる。

- ・上流工程、テスト工程のいずれでもバグが多く、もともとの設計品質が悪い。
- ・テスト工程工数が少ないにも関わらず、テスト工程バグが多い。上流工程で品質が確保できておらず、結果として上流工程バグ摘出率も低くなっている。



(3) ベンチマーキング方法（分析結果の活用）

このベンチマークから、上流工程バグ摘出率の向上が品質向上の重要な指針となることがわかる。各組織は、この指針に沿って、上流工程強化に向けて具体的に実施すべき施策を検討し、改善計画を立案する。

参考ポイント

この例は、社内の主要な品質（信頼性）変動要因に関する知見をベンチマークとし、自組織の重点的な品質向上方策を検討するという典型的なベンチマーキング方法を紹介したものであり、信頼性向上に向けたベンチマーキング方法の事例として参考になると考えられる。

なお、上流工程バグ摘出率（上流工程での不具合摘出比率）を高めることが信頼性向上のために効果的であるという指針は、5.2.2項「白書等を用いた具体例」の内容と一致している。

5.3. 生産性変動要因分析と見積り妥当性評価方法の見直し等

5.3.1. ベンチマーキング方法の手順例

ベンチマーキング方法の基本的な手順の例を以下に示す。

- (1) ベンチマーキングの目的を確認する。
ベンチマーキングの目的が、「ベンチマーク中に「生産性変動要因に関する知見」があれば、自組織に当てはまるかどうかを検討し、必要に応じて工数見積り等の妥当性評価方法を見直し、必要に応じて生産性向上のための組織の重点強化領域を特定し、生産性向上方策を検討すること」であることを確認する。
- (2) ベンチマーク中に「生産性変動要因に関する知見」があれば、自組織に当てはまるかどうかを検討し、必要に応じて工数見積り等の妥当性評価方法を見直す。具体的には、評価対象プロジェクトに該当する生産性変動要因によって生じる変動幅を勘案して、妥当性評価のための妥当と評価できる範囲を上方修正/下方修正しながら妥当性評価する方法となるように見直す。

(注) 各変動要因が自組織にとって主要な変動要因に該当するか否かを判断する必要がある。例えば、自組織において固定的になっている変動要因は、自組織にとっての変動要因には当たらない。

(備考) 工数見積り等の妥当性評価方法について

生産性の目標値や見積り工数が、妥当と評価できる範囲（例えば管理指標の25パーセンタイル値～P75パーセンタイル値の範囲）に収まっているか否かを評価することは、基本的な妥当性評価方法ではあるが、その範囲に収まっていないというだけでNG（妥当でない）と評価するのは早計である。プロジェクトの性質によっては、生産性変動要因による変動を始めとして一定の範囲に収まらなくなる合理的な理由が存在する可能性がある。評価対象プロジェクトに該当する生産性変動要因によって生じる変動幅を勘案して、妥当と評価できる範囲を上方修正/下方修正しながら妥当性評価することが望ましい。その結果においても妥当な範囲外となり、かつ生産性変動要因以外の合理的な理由がない場合には、計画や見積りを見直すことが望ましい。

- (3) ベンチマーク中に「生産性変動要因に関する知見」があれば、自組織に当てはまるかどうかを検討し、必要に応じて生産性向上のための組織の重点強化領域を検討・特定する。

(備考) 生産性向上のための組織の重点強化領域の特定について

生産性向上を進めていくためには、個々のプロジェクトのマネジメントよりも、組織の改善に向けたマネジメント・サイクルを回すことが一層重要であろう。重点的に強化すると効果的な領域を特定し、適切な方策を立てることが望まれる。そのために、組織の生産性変動要因群を把握することが効果的であり、お勧めする。

- (4) 見直した妥当性評価方法を、自組織の標準類に反映する。また、組織の生産性向上に向けて特定した重点強化領域に基づいて生産性向上方策を検討する。

5.3.2. 白書等を用いた具体例

<生産性変動要因に基づく重点強化領域特定の例>

(1) 目的（活用シーン）

主に開発組織のマネジャー層が、次のような観点から、組織の改善のためのマネジメントを推進するシーン（PMO 及び品質マネジメント推進部門が、マネジメントを支援するシーンを含む）

◇ベンチマーク中に「生産性変動要因に関する知見」があれば、自組織に当てはまるかどうかを検討し、必要に応じて工数見積り等の妥当性評価方法を見直す。また、必要に応じて生産性向上のための組織の重点強化領域を特定し、生産性向上方策を検討する。

(2) ベンチマーク

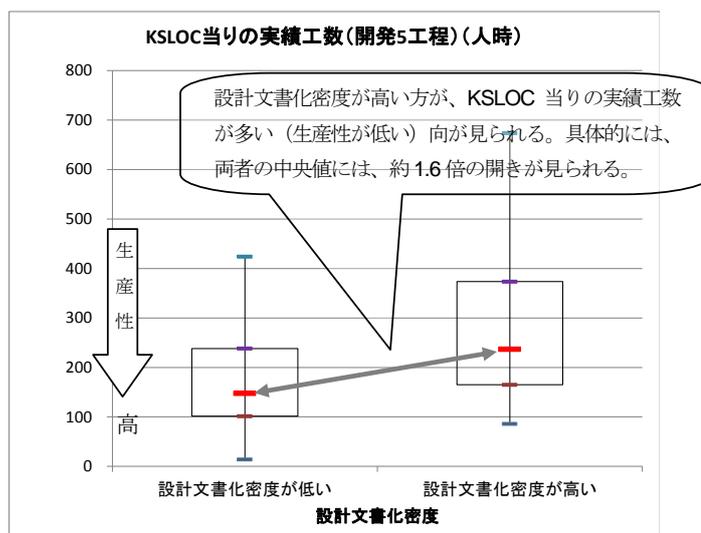
公開ベンチマーク（「ソフトウェア開発データが語るメッセージ」）を参照すると、「生産性変動要因に関する知見」に関して以下のことが示されているので、ヒントとして参考にする。

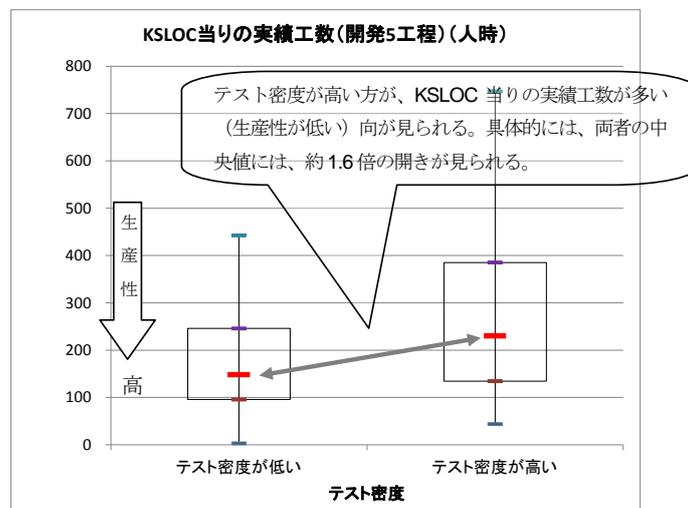
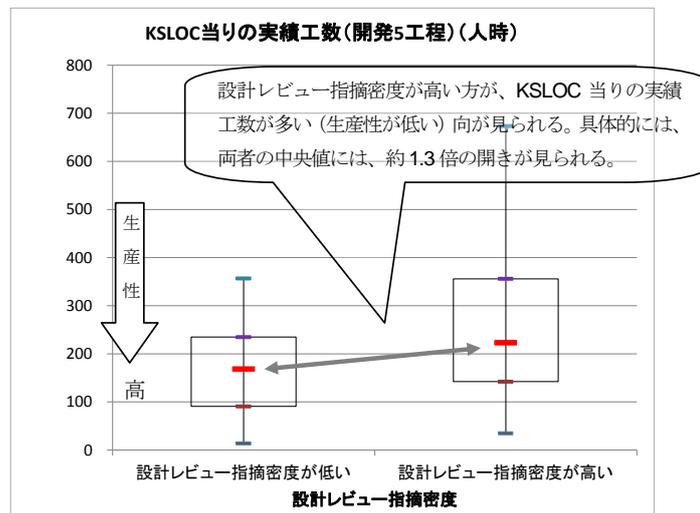
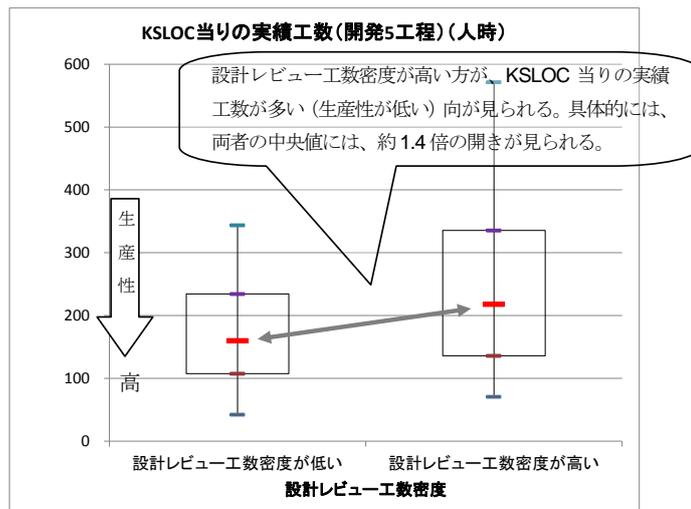
開発プロセス（主に品質保証プロセス）関連の生産性変動要因候補として、次のものがある。

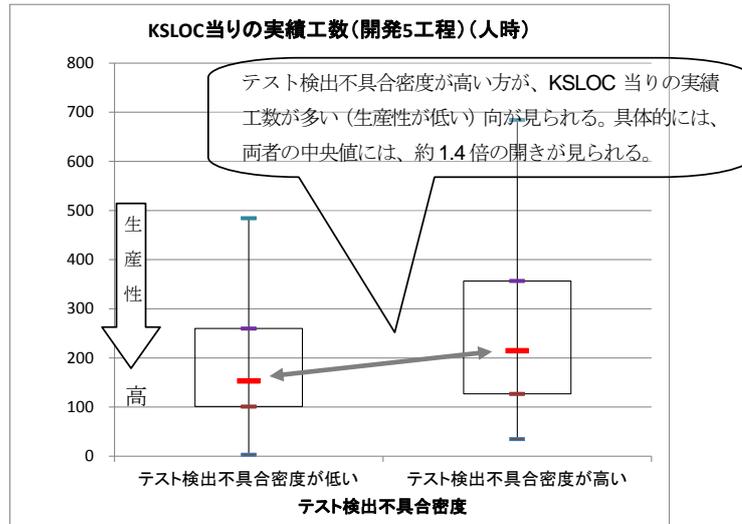
これらが高くなると、生産性が低下する傾向が示されている。

- ・ 設計文書化密度
- ・ 設計レビュー工数密度
- ・ 設計レビュー指摘密度
- ・ テスト密度
- ・ テスト検出不具合密度 等

新規開発の場合の傾向を以下に示す。







(注) 上記の分析結果に示された「両者の中央値には、約 1.4 倍の開きが見られる」などの傾向については、単一の変動要因によってそれだけの差異が生じているとは限らないことに留意する必要がある。変動要因間には一般に依存関係があり、依存関係にある他の変動要因の影響を含めた差異と見るのが妥当と考えられる。従って、「単一の変動要因による影響の度合いは、各図表に示された倍率以下」と見るのが妥当と考えられる。

(備考) 他に、システムのクリティカリティ、システムリスク、品質要求レベル、業種等も主な生産性変動要因と見られる。

<参照箇所>

詳しくは、「ソフトウェア開発データが語るメッセージ」の「生産性変動要因に関する分析」に関する記事を参照されたい。

(3) ベンチマーキング方法

- ① 上記(2)に示された生産性変動要因が自組織に当てはまるかどうかを検討する。当てはまるのであれば、工数見積り等の妥当性評価方法を見直す。具体的には、評価対象プロジェクトに該当する生産性変動要因によって生じる変動幅を勘案して、妥当性評価のための妥当と評価できる範囲を上方修正/下方修正しながら妥当性評価する方法となるように見直す。
- ② ベンチマーク中に「生産性変動要因に関する知見」があれば、自組織に当てはまるかどうかを検討し、必要に応じて生産性向上のための組織の重点強化領域を検討・特定する。
- ③ 見直した妥当性評価方法を、自組織の標準類に反映する。また、組織の生産性向上に向けて特定した重点強化領域に基づいて生産性向上方策を検討する。

5.4. 業種等のドメイン別マネジメント

5.4.1. ベンチマーキング方法の手順例

ベンチマーキング方法の基本的な手順の例を以下に示す。

- (1) ベンチマーキングの目的を確認する。

ベンチマーキングの目的が、「ベンチマーク中に「業種等のドメイン間の差異に関する知見」があれば、必要に応じて業種等のドメインに分けて自組織の信頼性／生産性をマネジメントすることを検討すること」であることを確認する。

(備考) 業種等のドメイン間の差異の例

- ・信頼性の差異
 - ・生産性の差異
 - ・信頼性／生産性の変動要因の差異
 - ・品質保証プロセスの差異
- 等

- (2) ベンチマーク中に「業種等のドメイン間の差異に関する知見」があれば、どのようなドメイン間でどのような差異が見られるかを把握する。

- (3) 自組織に該当するドメインの間で差異が見られる場合、ドメインに分けて信頼性／生産性をマネジメントすることを検討する。

(注) 単一のドメインから成る場合は、その限りではない。

(備考) すべてのソフトウェア種を十把一絡げ（じっぱひとからげ）で扱う時代は去っている。業種、業務等の分野別にマネジメントする（評価、分析、改善等を行う）ことが望ましいと考えられる。

5.4.2. 白書等を用いた具体例

<業種ドメイン別マネジメントの検討例>

(1) 目的（活用シーン）

主に開発組織のマネジャー層が、次のような観点から、組織の改善のためのマネジメントを推進するシーン（PMO 及び品質マネジメント推進部門が、マネジメントを支援するシーンを含む）

◇ベンチマーク中に「業種等のドメイン間の差異に関する知見」があれば、必要に応じて業種等のドメインに分けて自組織の信頼性／生産性をマネジメントすることを検討する。

(2) ベンチマーク

公開ベンチマーク（「ソフトウェア開発データが語るメッセージ」）を参照すると、「業種等のドメイン間の差異に関する知見」に関して以下のことが示されているので、ヒントとして参考にする。

発生不具合密度（信頼性）及び KSLOC 当り工数（生産性）は業種間で差異が見られる。また、信頼性要求レベルや開発プロセス関連の生産性／信頼性の変動要因候補においても業種間の差異が見られる。

マネジメントや分析を進める上で業種の影響を無視できないと判断できるので、業種によってドメインを分けてマネジメントすることをお勧めする。

<開発プロセス関連の変動要因候補>

- ・設計文書化密度
- ・設計レビュー工数密度
- ・設計レビュー指摘密度
- ・テスト密度
- ・テスト検出不具合密度
- ・上流工程での不具合摘出比率

(備考) 考察

特に金融・保険業では、他の業種より信頼性が高く生産性が低い傾向が見られる。また、その要因面では、他の業種より品質保証（設計文書化、レビュー、テスト）に多くの工数をかけている傾向が見られることから、システムリスクが高い（信頼性、公共性及び社会性の要求が高い）ソフトウェアの開発には、それ相応の品質保証（設計文書化、レビュー、テスト等）の工数が必要と考えられる。

<参照箇所>

詳しくは、「ソフトウェア開発データが語るメッセージ」の「業種ドメイン別マネジメント」に関する記事を参照されたい。

(3) ベンチマーキング方法

自組織に該当するドメインの間で差異が見られる場合、ドメインに分けて信頼性／生産性をマネジメントすることを検討する。

5.4.3. IT 企業での具体例 1

<業務アプリケーション別の基準値設定>

(1) 目的（活用シーン）

開発現場に納得感の高い、ドメイン別の目標基準値をベンチマークとして比較・検討することによって、自組織の品質目標値を設定する。

(2) 考え方

業種をドメインとして層別し、目標基準値を設定するケースは多い。しかしながら、同一業種であっても業務アプリケーションや顧客の特性によって、取得されたデータに大きなバラツキが発生し、開発現場として納得感のある設定値になることは少ない。また一方で、ドメインを細分化した場合、サンプルデータが少なくなり統計的な手法を用いた基準値設定が困難になる。

当社では、多少乱暴な方法ではあるが、改良開発プロジェクトを対象として、サンプルデータが少なくても、開発現場が少しでも納得感が持てることを目標に基準値提供を行った。

(3) ベンチマーク（ドメイン別の目標基準値）

<内部ベンチマークの設定方法>

① ドメインの設定

ドメインの設定は、業務アプリケーション、顧客別に設定する。

② ドメイン別の目標基準値の設定

i) サンプルが 10 件～30 件

対象となる過去プロジェクトのデータからの最小値と最大値を目標基準値として設定。

⇒過去に発生しない値は今後も発生しないとの仮定のもと、この範囲外での目標値設定は行わない。

ii) サンプルが 30 件以上

対象となる過去プロジェクトのデータから第1四分位と第3四分位の数値を目標基準値として設定。

⇒過去の実績から想定し、50%のデータが分布するエリアを目標基準の設定エリアとする。

<効果>

統計的手法を駆使したものではないものの、開発現場及びマネジement層からも各ドメインの実力から設定された目標基準値であるとの理解を得ることができた。

(4) ベンチマーキング方法

ドメイン別の目標基準値をベンチマークとして比較・検討することによって、自組織の品質目標値を設定する。

参考ポイント

ベンチマーキングにおいては、できるだけ同種のソフトウェアと比較・検討することが望ましく、典型的には同じ業種のベンチマークを利用することが推奨される。

この例では、社内において業務アプリケーション及び顧客で層別した内部ベンチマークを整備・利用することによって、開発現場等にとってより納得できるベンチマーキングを実現しており、ドメイン別マネジメントのベンチマーキング方法の事例として参考になると考えられる。

また、サンプル数が少ない場合のベンチマークの整備方法としても参考になると考えられる。

5.5. 信頼性／生産性等の経年変化に基づく組織改善計画

5.5.1. ベンチマーキング方法の手順例

ベンチマーキング方法の基本的な手順の例を以下に示す。

(1) ベンチマーキングの目的を確認する。

ベンチマーキングの目的が、「ベンチマーク中に「信頼性／生産性等の推移（経年変化）」があれば、自組織の信頼性／生産性等のマネジメントの年度計画目標／中長期計画目標を策定するにあたり、適切な目標や改善方策を検討する時の参考にすること」であることを確認する。

（備考1）自組織の信頼性の推移は、次のようなシーンで参考にすることができる。

- ・品質マネジメントの年度計画や中長期計画における目標と施策の検討等
- ・信頼性向上に向けたマネジメント活動や諸施策の総合的な効果の確認

（備考2）自組織の生産性の推移は、次のようなシーンで参考にすることができる。

- ・生産性マネジメントの年度計画や中長期計画における目標の検討
- ・生産性向上に向けたマネジメント活動や諸施策の総合的な効果の確認
- ・予算管理、単価の改定／交渉

(2) 自組織のソフトウェア・プロダクトが適用される事業のニーズ等に応じて、過去の目標対実績を勘案しながら、自組織の信頼性／生産性マネジメントの年度計画／中長期計画における目標と改善方策を立案する。

(3) 立案した目標及び改善方策の実現可能性を検討する。

- ・目標と現状とのギャップの大きさを把握する。
- ・信頼性／生産性等の向上施策面から見て、到達可能な目標か否かを検討する。
- ・内部ベンチマーク中に「信頼性／生産性等の推移（経年変化）」があれば、その推移から見て到達可能な目標か否かを吟味する。また、外部ベンチマーク中に「信頼性／生産性等の推移（経年変化）」があれば、併せて参考にする。

5.5.2. 白書等を用いた具体例

<信頼性/生産性の推移による中長期計画の実現可能性検討>

(1) 目的 (活用シーン)

主に開発組織のマネジャー層が、次のような観点から、組織の改善のためのマネジメントを推進するシーン (PMO 及び品質マネジメント推進部門が、マネジメントを支援するシーンを含む)

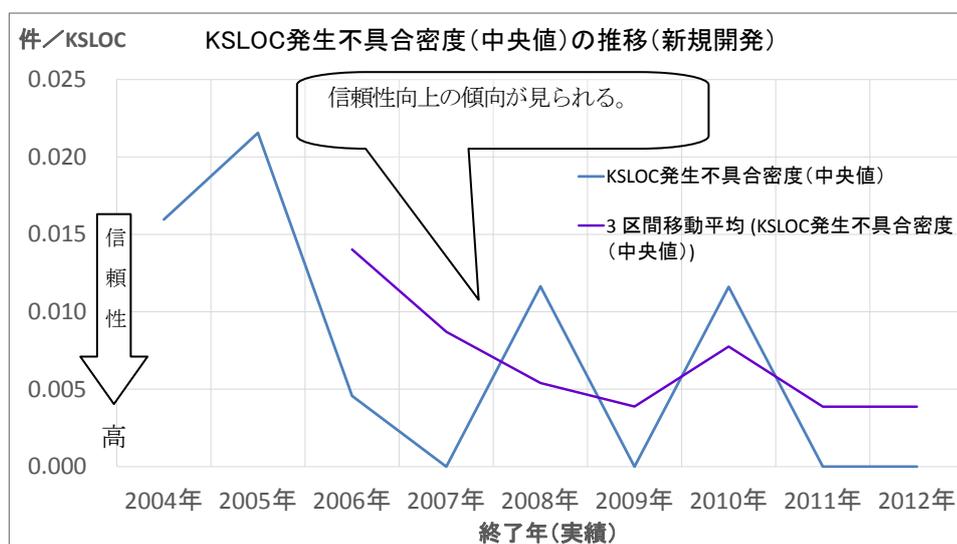
◇ベンチマーク中に「信頼性/生産性等の推移 (経年変化)」があれば、自組織の信頼性/生産性等のマネジメントの年度計画目標/中長期計画目標を策定するにあたり、適切な目標や改善方策を検討する時の参考にする。

(2) ベンチマーク

公開ベンチマーク (「ソフトウェア開発データが語るメッセージ」) を参照すると、「信頼性/生産性の推移 (経年変化)」に関して以下のことが示されているので、ヒントとして参考にする。

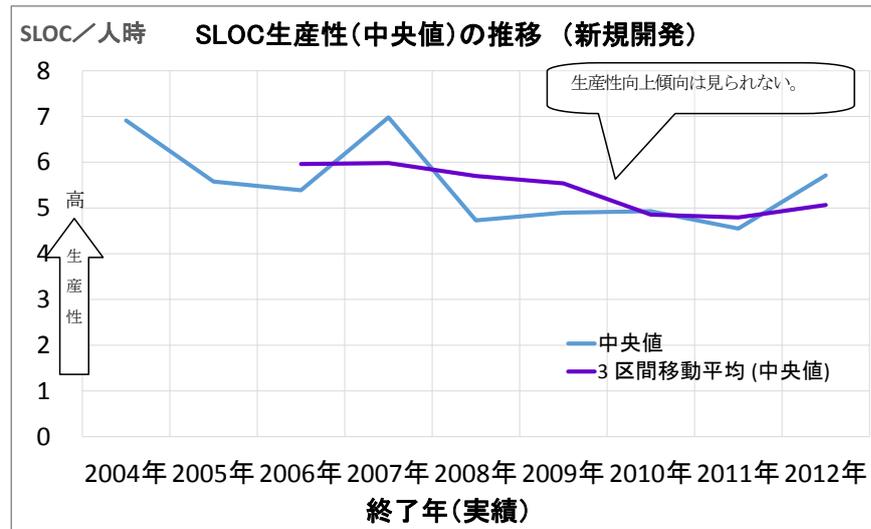
<信頼性の推移>

「ソフトウェア開発データ白書」のデータに基づく信頼性実績では、プロジェクトデータを収集開始してから最近まで (プロジェクトの実績終了年が 2004 年から 2012 年まで) の発生不具合密度の中央値の経年変化に低下傾向 (信頼性向上の傾向) が見られる。



<生産性の推移>

「ソフトウェア開発データ白書」のデータに基づく生産性実績では、プロジェクトデータを収集開始してから最近まで (プロジェクトの実績終了年が 2004 年から 2012 年まで) の SLOC 生産性の中央値の経年変化に、生産性向上 (または生産性低下) の傾向は見られない。



(備考) 考察

生産性に向上傾向が見られない要因の一つとして、次の可能性が考えられる。

開発プロジェクトの「複雑さ・難しさ」が増している一方で、開発要員のスキルレベルが追いついていない。開発プロジェクトにとっての相対的な難易度が高まっていることの影響が考えられる。少数精鋭部隊で開発していた時代は去っている。次のような状況下でも、如何にして要員のスキルレベルを確保/向上させるか、また効率的な開発方法を実現するかが課題の一つと考えられる。

- ◇開発すべきソフトウェアの総量が増大するのに伴って、ますます大勢の開発要員が必要になっている (要員の資質の平均レベルが低下する)。
- ◇大半の開発プロジェクトが改良開発 (改修・保守や拡張) になってきており、新規開発 (特にスクラッチ開発) によってスキルアップするチャンスが少なくなっている。

<参照箇所>

詳しくは、「ソフトウェア開発データが語るメッセージ」の「信頼性のトレンド」及び「生産性のトレンド」に関する記事を参照されたい。

(3) ベンチマーキング方法

立案した信頼性/生産性の目標の実現可能性を検討する。立案した目標が到達可能な目標か否かを吟味するために、ベンチマークにおける信頼性/生産性の推移 (経年変化) を一つの参考にする。その検討結果を踏まえて適切な目標や改善方策を検討する。

(備考) 考察

ソフトウェア開発プロジェクトは、低価格と短納期の強いプレッシャーに晒されることが多々ある。また、予算管理や価格交渉等の場面で、年率 5%の開発コスト削減や前年度比 10%のライン単価低減等が要求されるケースが散見される。しかし、データによる裏付け等の根拠が希薄であると、このような状況は開発プロジェクトのリスク増大や品質低下を招く恐れがある。データによる裏付けの一つとして、外部ベンチマーク中の信頼性/生産性の推移も参考にされたい。

5.5.3. IT 企業での具体例 1

<顧客満足度の経年評価による組織活動の改善>

(1) 目的 (活用シーン)

プロジェクト完了時に顧客満足度評価を行い、顧客満足度とその要因の単年度実績や経年変化をベンチマークとして、自組織の活動の見直しを行い、自組織の改善を行う。

(2) 考え方

顧客満足度の指標 (メトリクス) と基準値を設定し、複数顧客の顧客満足度評価の結果を基準値と照らし合わせ、目標が達成できていない取り組みに対して改善を図る。また、顧客満足度の指標の経年変化を監視し、下降傾向にある指標に関して分析を行い、組織としての活動を見直す。なお分析に関しては、指標毎に問題の有無を評価し、顧客のコメントから組織の活動の見直しを行う。

(3) ベンチマーク

<検証に用いる内部ベンチマークの抽出方法>

- ① 顧客満足度の指標 (メトリクス) を 5 段階評価で設定。
- ② 5 段階評価の良好から不満に変わる変化点を基準値 (3 点) として設定。

<検証に用いるベンチマーク (メトリクス) >

- ① 商談・契約 (5 段階評価)
- ② 技術力 (5 段階評価)
- ③ コミュニケーション (5 段階評価)
- ④ 納品物・サービス (5 段階評価)

(4) ベンチマーキング方法

- ① 内部ベンチマークを基に顧客満足度の指標を評価する。
問題のある指標に関しては、顧客のコメントを基に組織活動を見直す。
- ② 複数顧客の顧客満足度評価の結果を指標毎に集計し、基準値をクリアしている集計値 (指標毎の平均値) を新たな基準値として再設定する。
- ③ 次年度から①と②を繰り返し、経年での変化を分析し、改善が図られているか監視する。

参考ポイント

組織の信頼性の推移は、一般に次のようなシーンで参考にすることができる。

- ・品質マネジメントの年度計画や中長期計画における目標と施策の検討等
- ・信頼性向上に向けたマネジメント活動や諸施策の総合的な効果の確認

この例は、顧客満足度という総合的な結果指標とその要因に関する指標の推移 (経年変化) をベンチマークとして、自組織の改善に向けた組織活動を見直すとともに改善効果の監視を行うものであり、経年変化に基づく組織改善のためのベンチマーキング方法の事例として参考になると考えられる。

6. 関連ツールの紹介

IPA/SEC では、「ソフトウェア開発データ白書」を参考にしたベンチマーキングを支援する仕組みとして、以下の2つを用意している。

定量データに基づくプロジェクト診断支援ツール

白書掲載のグラフデータダウンロードサービス

本ツールを利用することによって、公開ベンチマークの一つである「ソフトウェア開発データ白書」に掲載されているグラフ上に、自社のプロジェクトデータをプロットすることができる。ベンチマーキングの一助として利用して頂けると幸いである。

6.1. 定量データに基づくプロジェクト診断支援ツール

本ツールは、2007年12月25日より無償公開し、現在稼働中である。「ソフトウェア開発データ白書」と同じデータを用いたWebシステム上で、データ白書に掲載されたデータと自社のプロジェクトデータとを視覚的に比較することができる。「ソフトウェア開発データ白書」に掲載されている様々な図表を任意に参照できるだけでなく、利用者が本ツールに入力した自社のプロジェクトデータを、画面上の散布図や箱ひげ図に重ねてプロットできるのが特徴である。本ツールを利用して、プロジェクト目標の妥当性評価や完了プロジェクトの振り返り（自プロジェクトの強み／弱みの把握等）を、定量的かつ容易に行うことができる。

本ツールの全体概要は以下のとおり。

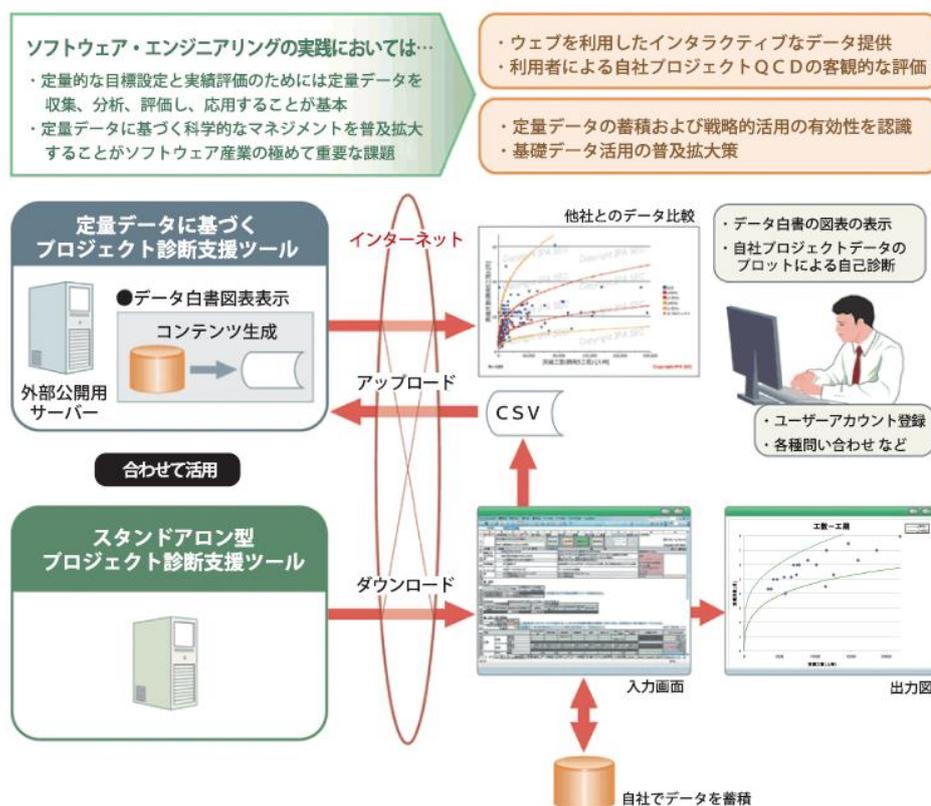


図1 プロジェクト診断支援ツールの全体概要

【自組織のプロジェクトデータの入力方法】

本ツールへのプロジェクトデータの inputs は、1件であれば画面から、複数件の場合は作成シートをダウンロードし、自社データを入力したファイルをアップロードすることにより白書記載グラフとのマッピングが可能となる（最大999件まで登録可）。これらのデータはログオフ時にすべて破棄され、サーバやログ等には一切残らないので、安心して利用できる。

プロジェクトデータ個別登録

・自社プロジェクトのデータを入力して、「登録・変更」ボタンを押してください。
・入力したデータはログアウトと同時に全て廃棄され、システムには一切保存されません。

プロフィール情報

プロジェクト名
開発プロジェクト種別
業種
アーキテクチャ
開発言語
開発対象プラットフォーム
FP実績値の計測手法

分析用基本情報

■プロジェクト規模
FP計画値
FP実績値(見積値)
SLOC実績値

■品質・信頼性

	結合テスト	総合テスト (ベンチマーク)	運用
テストケース数			
検出バグ数	検出数		
発生不具合数	検出数		

工程情報

	計画月数	実績月数	計画工数(人時)	実績工数(人時)
5 工程			-	内 外部委託工数
5 全体	-			
5 工程内訳				
基本設計	-		-	
詳細設計	-		-	
実装	-		-	
結合テスト	-		-	
総合テスト	-		-	

※5工程の実績工数(人時)には、5工程内訳(基本設計～総合テスト)に入力された数値の合計値がセットされます。

図2 1件入力画面

定量データに基づく プロジェクト診断支援ツール

IPA 独立行政法人情報処理推進機構
技術本部 ソフトウェア高信頼化センター

HOME > ソフトウェア高信頼化 > ソフトウェア・エンジニアリングの成果 > ツール&データベース一覧 > トップ > プロジェクトデータ一括登録

[お問合せ](#) [FAQ](#)

プロジェクトデータ一括登録

- ・自社プロジェクトの位置を図表上に表示するためのデータを一括登録します。
- ・一括登録用データファイルの形式は“CSV”のみ有効です。
- ・データファイルには項目名(固定)が必要ですが、列の順序が入れ替わっても問題ありません。
- ・データファイルの項目名に誤りがある場合は、データを取り込むことができません。
- ・プロジェクトデータの値に問題がある場合は未入力扱いとし、警告ウィンドウでお知らせします。
- ・警告ウィンドウは別ウィンドウで表示されますので、**ポップアップ許可の設定が必要です**。

データファイルのアップロード

データファイル(CSV)を指定して、「アップロード」ボタンを押してください。

ファイル名 C:\Users\F-tsuika\Desktop\ほとめ一括登録用データ.csv 参照...

アップロード [メニューに戻る](#)

作成シートのダウンロード

データファイルを作成するためのデータ作成シート(Excelファイル)を用意しています。
必要な方は「ダウンロード」ボタンを押してください。

[ダウンロード](#)

図3 作成シートのダウンロード画面

【機能特徴】

- (1) 自社データのプロット1件のプロジェクトデータを画面から入力、あるいは複数件のデータをファイルでアップロードし、画面に表示された散布図および箱ひげ図の上に重ねてプロットする。自社データのみの表示も可能である。
- (2) 機密保全のため、入力したプロジェクトデータはログオフ時にすべて破棄され、サーバやログ等には一切残らない。
- (3) 自社データの属性表示マウスオーバーにより、散布図上にプロットされた自社データの番号と名称を表示する。
- (4) 図表の拡大と縮小散布図を拡大、縮小する。プロットが局所に密集している場合、拡大によって、その部分の状況を詳細に確認することができる。XY軸の基点の指定も可能である。
- (5) 信頼区間の指定信頼区間が引かれている散布図の場合、信頼区間の%値を任意に指定して表示する。
- (6) 人時/人月切替え散布図および箱ひげ図の工数の単位を選択する。人月表示に切替える際は、任意の換算値を指定する。
- (7) 対数グラフの表示散布図に通常表示と対数表示がある場合、ボタンひとつで表示を切り替える。両対数変換することで、通常表示では分かりづらい傾向が見えてくる場合がある。
- (8) XY軸の反転散布図のXY軸を反転して表示する。
- (9) データ種別の選択散布図のデータに系列（種別、業種、言語など）がある場合、特定の系列に絞って表示する。
- (10) プロット可能図表一覧利用者が登録したプロジェクトデータが、どの図表上にプロットできるのかを一覧で表示する。
- (11) 図表のコピー&印刷図表はコピーしてWordやExcelなどで作成した文章に貼り込むことができる。画面をそのまま印刷することも可能である。

【ツールの利用方法】

本ツールを使用するにあたり、事前に利用者登録が必要である。

IPAのWebサイトのツール/データベース一覧からログインし、以下のトップページを表示する。

<https://sec.ipa.go.jp/login.html>

IPA Better Life with IT 情報処理推進機構

文字サイズ 標準 拡大 検索

IPAについて お知らせ一覧 サイトマップ お問い合わせ ENGLISH

HOME 情報セキュリティ ソフトウェア高信頼化 未踏/セキュリティキャンプ IT人材の育成 情報処理技術者試験 国際標準の推進

HOME > ソフトウェア高信頼化 > ログイン 本文を印刷する

ソフトウェア高信頼化

ログイン

ログイン

ユーザーID

パスワード

送信

パスワードを忘れた方は[こちらで再発行](#)します。
登録がお済でない方は[こちらで新規登録](#)をお願いします。

利用者登録

- > 新規登録
- > ログイン
- > ヘルプ
- > パスワードの再発行

> SECからのお知らせ

> 組織概要

> 重要インフラ分野のシステム障害への対策

> ソフトウェア信頼性の見える化

ログイン後は、自分が見たい項目について、選択していだけで白書に記載されているグラフにたどり着くことができる。グラフに自社データをマッピングした結果を見て、自社プロジェクトのポジションを確認することができる。

- (1) プロジェクト診断支援ツールの画面より、活用目的に沿ったものを選択する。

定量データに基づく
プロジェクト診断支援ツール

HOME > ソフトウェア高信頼化 > ソフトウェア・エンジニアリングの成果 > ツール&データベース一覧 > トップ

[お問合せ](#) [FAQ](#)

IPA 独立行政法人情報処理推進機構
技術本部 ソフトウェア高信頼化センター

メインメニュー	<ul style="list-style-type: none"> ▶ データ白書の図表を探す <ul style="list-style-type: none"> ▶ データ白書の目次から探す ▶ 見積の評価に関する図表を探す ▶ プロジェクト計画の評価に関する図表を探す ▶ 品質の評価に関する図表を探す ▶ 工期の評価に関する図表を探す ▶ 生産性に関する図表を探す ▶ 自社のデータを登録する <ul style="list-style-type: none"> ▶ プロジェクトデータを1件登録する(個別登録) ▶ 2件以上のプロジェクトデータをアップロードする(一括登録) <p style="font-size: small; margin-top: 5px;">※プロジェクトデータの登録は、各図表の表示画面からも可能です</p> ▶ お役立ち情報 	<div style="border: 2px solid black; padding: 10px; width: 100px; margin: auto;"> <p>実施したい内容 を選ぶ</p> </div>
<ul style="list-style-type: none"> ▶ プロジェクトデータ個別登録 ▶ プロジェクトデータ一括登録 ▶ お問合せ ▶ FAQ ▶ アンケート ▶ 各種ダウンロード ▶ 関連サイトリンク ▶ プロット可能な図表一覧 		
サイトの説明 <p style="font-size: x-small;">本サイトの情報は、独立行政法人情報処理推進機構(IPA)技術本部 ソフトウェア高信頼化センター(SEC)提供のソフトウェア開発データ白書2014-2015Jに基いています。 なお、夜間処理のため、3:00AM ~ 4:00AMの間ご利用できません。ご注意ください。</p>		

評価項目選択画面

- (2) さらに見たい項目を絞り込む。

定量データに基づく
プロジェクト診断支援ツール

HOME > ソフトウェア高信頼化 > ソフトウェア・エンジニアリングの成果 > ツール&データベース一覧 > トップ > プロジェクト計画の妥当性評価

[お問合せ](#) [FAQ](#)

IPA 独立行政法人情報処理推進機構
技術本部 ソフトウェア高信頼化センター

メインメニュー	<ul style="list-style-type: none"> ▶ 工数と工期 <p style="font-size: x-small;">工数に対する工期の関係を確認することが可能です。見積りに対するプロジェクト計画上の全体スケジュール計画や要員計画の妥当性検証に有効です。</p> ▶ 工程別の工期・工数 <p style="font-size: x-small;">基本設計、詳細設計、製作、総合テスト、総合テスト(ベンダ確認)の各工程ごとの工期の比率を確認することが可能です。工程ごとのスケジュール計画の妥当性検証に有効です。</p> ▶ FP規模と発生不具合数 <p style="font-size: x-small;">FP(ファンクションポイント)の規模に対する発生不具合数を確認することが可能です。プロジェクト計画時の目標不具合件数の妥当性検証に有効です。</p> ▶ SLOC規模と発生不具合数 <p style="font-size: x-small;">SLOC(ソースステップ数)に対する発生不具合数を確認することが可能です。プロジェクト計画時の目標不具合件数の妥当性検証に有効です。</p> ▶ テスト工程別のテストケース数・検出バグ数 <p style="font-size: x-small;">FP/SLOC当たりのテストケース数、およびFP/SLOC当たりのバグ数を、テスト工程別(総合テスト、総合テスト)に確認することができます。プロジェクト計画時のテスト工程ごとの目標テストケース数、目標バグ数の妥当性検証に有効です。</p> 	
<ul style="list-style-type: none"> ▶ プロジェクトデータ個別登録 ▶ プロジェクトデータ一括登録 ▶ お問合せ ▶ FAQ ▶ アンケート ▶ 各種ダウンロード ▶ 関連サイトリンク ▶ プロット可能な図表一覧 		
サイトの説明 <p style="font-size: x-small;">本サイトの情報は、独立行政法人情報処理推進機構(IPA)技術本部 ソフトウェア高信頼化センター(SEC)提供のソフトウェア開発データ白書2014-2015Jに基いています。 なお、夜間処理のため、3:00AM ~ 4:00AMの間ご利用できません。ご注意ください。</p>		

比較したい項目選択画面

(3) 該当するグラフを選択する。

定量データに基づく プロジェクト診断支援ツール

IPA 独立行政法人情報処理推進機構
技術本部 ソフトウェア高信頼化センター

HOME > ソフトウェア高信頼化 > ソフトウェアエンジニアリングの成果 > ツール&データベース一覧 > トップ > 見積の妥当性評価

[お問合せ](#) [FAQ](#)

メインメニュー

- ▶ プロジェクトデータ閲覧登録
- ▶ プロジェクトデータ一括登録
- ▶ お問合せ
- ▶ FAQ
- ▶ アンケート
- ▶ 各種ダウンロード
- ▶ 関連サイトリンク
- ▶ プロット可能な図表一覧

▼サイトの説明

本サイトの情報は、独立行政法人情報処理推進機構(IPA) 技術本部 ソフトウェア高信頼化センター (SEC) 提供の「ソフトウェア開発データ白書2014-2015」に基づいています。
なお、夜間処理のため、3:00AM ~ 4:00AMの間はご利用いただけません。ご注意ください。

▶ **FP規模と工数**

FP(ファンクションポイント)の規模に対する工数の関係を確認することが可能です。FP(ファンクションポイント)の規模に対する見積りの妥当性検証に有効です。

▶ **SLOC規模と工数**

SLOC(ソースステップ数)に対する工数の関係を確認することが可能です。SLOC(ソースステップ数)に対する見積りの妥当性検証に有効です。

[SLOC規模と工数\(全開発種別、主開発言語混在\)\(信頼区間50%付き\)](#)
[SLOC規模と工数\(全開発種別、主開発言語グループ\)](#)
[主開発言語別のSLOC規模と工数\(新規開発\)](#)
[業種別SLOC規模と工数\(新規開発、主開発言語グループ\)](#)
[アーキテクチャ別SLOC規模と工数\(新規開発、主開発言語グループ\)](#)
[主開発言語別のSLOC規模と工数\(改良開発\)](#)
[業種別SLOC規模と工数\(改良開発、主開発言語グループ\)](#)
[アーキテクチャ別SLOC規模と工数\(改良開発、主開発言語グループ\)](#)

▶ **FP規模とSLOC規模**

FP(ファンクションポイント)の規模と、SLOC(ソースステップ数)の関係を確認することが可能です。FP(ファンクションポイント)とSLOC(ソースステップ数)の両方にて見積りした場合の妥当性検証に有効です。

グラフ選択画面

[グラフ表示例] 散布図



解説文

工数と工期について、回帰式で確認した結果は次のようになる。
 $工期=A \times (工数)^B$, $A=0.37$, $B=0.32$, $R=0.71$
 <注> 回帰式の利用に関しては、必ず「FAQ 回帰式利用上の注意事項」を参照すること。
 プロジェクト特性によるバラツキはあるが、工期は工数の3乗根に比例する傾向が見られる。

[グラフ表示例] 箱ひげ図 (1プロジェクトのみ表示可)

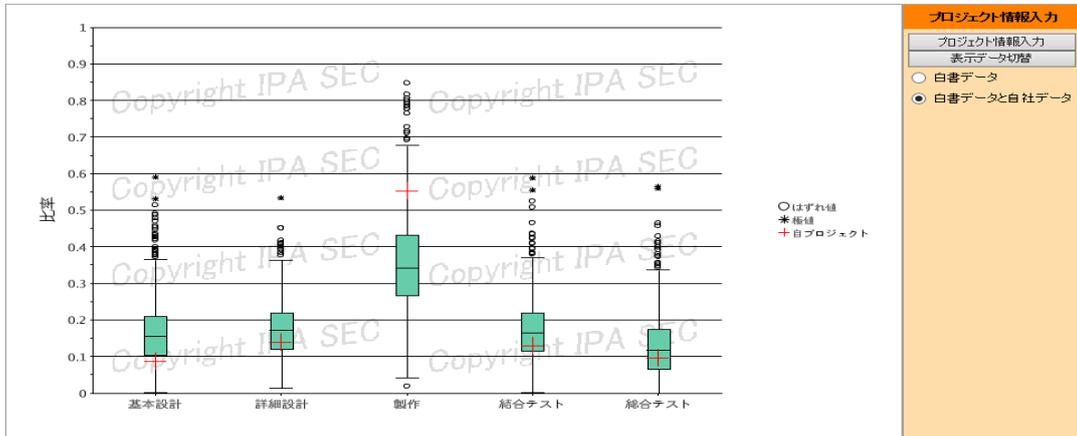
■層別定義

- ・開発工程のフェーズ有無が全て○
- ・108開発プロジェクトの種別が:新規開発
- ・工程別の実績工数に全て記入があり、各値が0より大きい

■分析・収集対象データ

- ・実績工数(総計人時)基本設計、実績工数(総計人時)詳細設計、実績工数(総計人時)製作、実績工数(総計人時)結合テスト、実績工数(総計人時)総合テスト

工程別の実績工数の比率(新規開発)箱ひげ図



基本統計量

基本統計量タイトル	N	最小	P25	中央	P75	最大	平均	標準偏差	ひげ最大	ひげ最小
基本設計	679	0.001	0.104	0.154	0.21	0.589	0.167	0.092	0.365	0.001
詳細設計	679	0.014	0.12	0.171	0.221	0.533	0.175	0.077	0.364	0.014
製作	679	0.018	0.267	0.342	0.432	0.847	0.356	0.135	0.677	0.042
結合テスト	679	0.002	0.115	0.164	0.219	0.588	0.173	0.087	0.37	0.002
総合テスト	679	0	0.067	0.117	0.175	0.564	0.13	0.085	0.336	0

単位[比率]

6.2. グラフデータ

本ツールは、2014年12月26日より無償公開し現在稼働中であり、「ソフトウェア開発データ白書」に掲載されているグラフデータをダウンロードすることにより、自社データとの比較を可能にしている。「ソフトウェア開発データ白書」に掲載されているグラフの加工ができるだけでなく、利用者が自社のプロジェクトデータをグラフに追加するだけで散布図、箱ひげ図に重ねてプロットできるのが特徴である。本ツールを利用して、プロジェクト目標の妥当性評価や完了プロジェクトの振り返り（自プロジェクトの強み/弱みの把握等）を、定量的かつ容易に行うことができる。

本ツールは、下記 URL よりアンケート回答後、ログインすることによりダウンロード可能となる。

ダウンロードは、各章別にダウンロードすることができる。

<http://sec.ipa.go.jp/publish/whitepaper>

[利用方法]

- (1) 下記ダウンロード画面から必要なグラフデータをダウンロードする。

ダウンロード

※最新版のダウンロードにはIPA/SEC WEBページの利用者登録（ログイン）が必要です。

※グラフデータは「ソフトウェア開発データ白書2014-2015」版のみのご提供です。

なお、各章のデータごとに個別のファイルになっています。

最新版

PDF版	ソフトウェア開発データ白書2014-2015  [13.9MB]
グラフデータ	第4章  [1.17MB]
	第5章  [691KB]
	第6章  [567KB]
	第7章  [461KB]
	第8章  [1.17MB]
	第9章  [639KB]
	第10章  [266KB]

- (2) ダウンロードしたグラフデータの格納形態は以下のようにになっている。
 図表番号別に、原則としてグラフデータと基本統計量が格納されている。

データ白書の記載されている4章から10章までのグラフデータおよび基本統計量などの図表が、下記イメージで入手できます。このデータをEXCEL等で加工します。

1	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
2	320	34.375	a	スタンダードアロン												
3	900	4.444444	a	スタンダードアロン												
4	950	8.421053	a	スタンダードアロン												
5	120	66.66667	a	スタンダードアロン												
6	370	21.62162	a	スタンダードアロン												
7	450	8.888889	a	スタンダードアロン												
8	313	0	a	スタンダードアロン												
9	2515	4.771372	a	スタンダードアロン												
10	325	3.076923	a	スタンダードアロン												
11	510	0	a	スタンダードアロン												
12	198	55.55556	a	スタンダードアロン												
13	250	0	a	スタンダードアロン												
14	250	4	a	スタンダードアロン												
15	629	27.02703	a	スタンダードアロン												
16	771	2.584034	a	スタンダードアロン												
17	712	2.808989	a	スタンダードアロン												
18	265	0	a	スタンダードアロン												
19	1100	0	a	スタンダードアロン												
20	473	12.88489	a	スタンダードアロン												
21	600	33.33333	a	スタンダードアロン												
22	327	18.34862	a	スタンダードアロン												
23	100	0	a	スタンダードアロン												
24	165	163.6364	a	スタンダードアロン												
25	21	0	a	スタンダードアロン												
26	860	4.651163	a	スタンダードアロン												
27	900	11.11111	a	スタンダードアロン												
28	282	0	a	スタンダードアロン												
29	112	8.928571	a	スタンダードアロン												
30	474	4.219409	a	スタンダードアロン												
31	1185	3.375527	a	スタンダードアロン												
32	650	81.53846	a	スタンダードアロン												
33	186	5.376344	a	スタンダードアロン												
34	2626	15.23229	a	スタンダードアロン												
35	240	0	a	スタンダードアロン												
36	500	50	a	スタンダードアロン												
37	300	23.33333	a	スタンダードアロン												
38	1606	0.622665	b	メインフレーム												
39	794	0	b	メインフレーム												
40	354	25.42373	b	メインフレーム												
41	1143	12.24847	b	メインフレーム												
42	2804	96.29101	b	メインフレーム												
43	2938	0.340368	b	メインフレーム												
44	188	0	b	メインフレーム												
45	322	366.4596	c	2階層クライアント/サーバ												
46	455	164.8352	c	2階層クライアント/サーバ												

基本統計量

グラフデータ

データ格納形式例

- (3) ダウンロードした Excel データを Excel のフィルター機能を用いて、必要なデータのためのグラフに編集

EXCELのフィルター機能を使用してデータを絞り込む
 (業種別、言語別、規模、アーキテクチャ、数値、開発スタイルなど)

昇順(S)
 降順(Q)
 色で並べ替え(I)
 "FP発生不具合密度[件/KFP]" からフィルターをクリア(C)
 色フィルター(I)
 数値フィルター(E)

検索

(すべて選択)
 0
 0.249563264
 0.25
 0.289939113
 0.334952269
 0.340367597
 0.346260388
 0.382848392
 0.466200466
 0.526971352
 0.550017188

OK キャンセル

指定の値に等しい(E)...
 指定の値に等しくない(N)...
 指定の値より大きい(G)...
 指定の値以上(O)...
 指定の値より小さい(L)...
 指定の値以下(Q)...
 指定の範囲内(W)...
 トップテン(T)...
 平均より上(A)
 平均より下(O)
 ユーザー設定フィルター(E)...

Excel フィルター使用例

することができる。

- (4) フィルター処理を施したデータを基に、絞り込んだ白書データのグラフを作成する。次に自社データを準備し先程作成したグラフに対し、データソースの選択を選び自社データの追加を行うことにより自社データのマッピングが可能となる。(Excelの機能)

EXCELのグラフデータ追加機能を使用して自社データのマッピングが可能です。

自社データ

ダウンロードデータ

自社データマッピング

- FP発生不具合密度(件/KFP)
- 自社FP発生不具合密度(件/KFP)

自社データマッピング例

<謝辞>

本書の「IT企業での具体例」は、IPA/SECが運営する「高信頼性定量化部会」の委員の皆様にご提供して頂いたものです。また、委員の皆様には本書のレビューにも協力して頂きました。ここに、深く御礼申し上げます。

<参考文献>

- (1) 「経済産業省 ソフトウェアメトリクス高度化プロジェクト」の成果物
経済産業省 ソフトウェアメトリクス高度化プロジェクト:ITプロジェクトの性能を客観的、共通的に分析、評価でき、他組織のデータやベンチマークと自組織のITプロジェクトに係るデータの相互運用性を高めることを目的とした産学官連携活動

<成果物のダウンロードURL>

- ① 定量的マネジメントのための公開データ利用ガイド (2009年度)
http://www.meti.go.jp/policy/it_policy/softseibi/metrics/process_metrics.pdf
- ② 定量的マネジメントのための公開データ利用ガイド付録 (2009年度)
http://www.meti.go.jp/policy/it_policy/softseibi/metrics/process_metrics_appendix.pdf
- ③ 発注者/受注者による公開データ利用方法一覧表・メトリクス関係図 (2009年度)
http://www.meti.go.jp/policy/it_policy/softseibi/metrics/metrics.xls
- ④ ITプロジェクトのベンチマーク供給者のためのガイドライン (2010年度)
http://www.meti.go.jp/policy/it_policy/softseibi/metrics/20110324process_metrics2010.pdf

- (2) 国際標準

ISO/IEC 29155

Information technology project performance benchmarking framework

ISO/IEC 15939

Measurement process

- (3) SEC BOOKS 「定量的品質予測のススメ」及び「続 定量的品質予測のススメ」