

# SPI活動と連携した定量的プロジェクトマネジメントの実践

TOiNX(トインクス) 東北インフォメーション・システムズ株式会社 開発運用本部 業務管理部 業務統括課 主査 佐藤 浩明

ソフトウェア開発プロジェクトを成功に導くためには、自社のメトリクスを活用した「定量的プロジェクトマネジメント」の実践が有用である。一方で、メトリクスの解析と活用には、多くの検討事項がある。本稿では、定量的プロジェクトマネジメントの実践に向けたSPI活動の教訓や、メトリクスの活用事例について紹介する。

## 1 はじめに

ソフトウェア開発プロジェクトを成功に導くためには、自社の実績データによるメトリクスを、計画段階から活用する「定量的プロジェクトマネジメント」の実践が有用である。勘と経験だけでは、高精度な計画を策定し、プロジェクトのQCDを正確に監視・制御することは困難である。

一方で、組織の標準プロセスが定着しない状態では、データのばらつきが大きく、メトリクスの精度向上は期待できない。SPI(Software Process Improvement)活動と、データ解析活動は表裏一体の関係である。開発プロセスの成熟度向上には、この2つの活動が連携している必要がある。

本稿では、東北インフォメーション・システムズ株式会社(以下、当社)におけるSPI・データ解析活動の経緯やメトリクスの活用事例を、データ解析に携わる実務者の観点から紹介する。

## 2 SPI・データ解析活動の経緯

当社は、2001年度からCMMI®(Capability Maturity Model Integration)に基づくSPI活動を開始した。2003年度には成熟度レベル2を、2008年度にはレベル3を達成している。また、2010年度には、一定の実績データが蓄積できたため、データ解析を本格化するための「メトリクス分析支援ツール」を開発した。以降、毎年100パターンを超えるメトリクス解析結果を、プロジェクトに提供している。

約15年間にわたるSPI活動であるが、標準プロセスの定着を実感できたのは、活動開始後10年以上が経ってからである。そして、この契機は、メトリクスの公開と活用にある。定量的プロジェクトマネジメントの実践には、様々なアプローチがある。しかし、当社の経験から言えることは、いかに厳格な計測・収集・解析のプロセスを定義しても、SPI・データ解析活動は浸透しない。重要なことは、プロジェクトが要求するメトリクスを、常にフィードバックし続けることである。

## 3 メトリクス解析結果とは

当社では、プロジェクトに対して、以下のメトリクス解析結果を提供している。以降、各メトリクスの概要を紹介する。

### 3.1 「プロセス実績ベースライン」と「プロセス実績モデル」

プロセス実績ベースライン(以下、基準値)と、プロセス実績モデル(以下、予測モデル)は、最も初歩的なメトリクスである。これらは、図1で示すように、社内のポータルサイトで参照できる。

基準値とは、各メトリクスの幾何平均値と上下限値を、開発タイプ別に解析したものである。主に、開発タイプごとの傾向把握に利用するが、後述する予測モデルを確立できない場合、暫定的に予測にも用いる。

予測モデルとは、見積もりや妥当性評価に用いる予測ツールである。開発タイプ別の層別による単回帰分析で、標準的な予測値と任意の信頼率による上下限値を算出する。べき乗モデル( $y = b_0 \times x^{b_1}$ )を採用しており、プロジェクト規模を反映した予測が可能である。本モデルは、寄与率、係数 $b_1$ のp値、残差の傾向などを確認し、一定の条件を満たしたもののだけが公開の対象となる。

### 3.2 メトリクス応用ツール

メトリクス解析結果の応用として、表1に示す各種のツールを開発し、プロジェクトに提供している。本ツールは、以下の①から③のタイプに分類することができる。

#### 3.2.1 プロジェクト特性に応じた重回帰分析(①)

タイプ①は、OSや顧客などのプロジェクト特性を組み合わせた層別に対して、複数の説明変数(FP、SLOC、テストケース数など)を用いた重回帰分析の仕組みである。各種の特性を反映しており、予測モデルより格段に精度が高い。代表的なツールとして、工数見積もりの妥当性を評価するための「工数評価ツール」がある。

### 3.2.2 複数メトリクスの統合(②)

レビューやテストなどは、単一の指標で品質を判断することが難しい。このため、複数の基準値を組み合わせた「ゾーン分析ツール」が有効になる。タイプ②である本ツールは、開発タイプ、対象工程、欠陥・指摘の重要度に基づく9つのゾーンを設定し、今、どのゾーンに該当するのかが確認できる。

### 3.2.3 プロジェクト特性に応じた基準の再設定(③)

プロジェクトは、多様な特性があり、必ずしも組織の基準値に合致するとは限らない。タイプ③は、工程進行中の実績情報から、プロジェクト特性を反映した基準値を再設定し、管理するものである。上流工程でのレビューを管理する「u管理図作成ツール」や、統合テスト工程以降で利用する「信頼度成長モデル」などがある。

**プロセス実績モデル (予測モデル)**

- 見積もり・妥当性評価に利用するツール。  
(例) FPを入力すると工数を予測するなど
- べき乗モデルによる単回帰分析。 $(y = b0 \times X^{b1})$
- 新規開発、エンハンスといった開発タイプごとに用意。
- 予測値のほかに $\pm 1\sigma$ の上下限値も算出。

**プロセス実績ベースライン (基準値)**

- 開発タイプごとに幾何平均値と上下限値を提示。
- 開発タイプごとの傾向把握に利用。
- 予測モデルを確立できないメトリクスは、暫定的に幾何平均を用いた予測を推奨。

メトリクス解析結果は、ポータルサイトで社内に公開。全社員が参照可能。

図1 基準値・予測モデル

表1 メトリクス応用ツール

No.	カテゴリ	ツール名称	タイプ	機能概要
1	工数・規模	工数評価ツール	①	PJ特性に応じた工数と、その上下限値を予測。PJ初期段階の工数見積りの妥当性評価に利用。
2		工数生産性評価モデル	①	PJ特性に応じた生産性を評価。
3		FP値 (IFPUG) 推定	—	画面や帳票数などの限定された情報からFP値を推定。
4	WBS/EVM	ライン数カウント	—	空白行・コメント行を除いたライン数をカウント。修正行のみカウントすることも可能。
5		WBSツール	—	WBSとEVMの作成を支援。
6		統合EVMグラフ	—	複数のWBS/EVMツールのEVMグラフを統合。
7		WBS要員展開ツール	—	要員ごとの計画・実績工数を月別(日別)に展開。
8	品質	検出欠陥予測ツール	①	統合テストの総検出欠陥数をPJの特性に応じて予測。
9		ゾーン分析(レビュー)ツール	②	要件定義〜PGM設計のレビューのデータから工程完了時の品質を予測。
10		テスト進捗管理表(ゾーン分析ツール)	②	テスト工程の進捗管理とゾーン分析を組み合わせたツール。
11		u管理図作成ツール	③	各種レビューや単体テストのデータからu管理図を作成。
12		信頼度成長モデル	③	指数型モデルを用いた統合テスト以降の定量的品質管理を支援。

## 4 SPI・データ解析活動の教訓

当社のSPI・データ解析活動では、多くの失敗を経験した。そして、このほとんどが、活動初期における初歩的な考慮不足が原因である。ここでは、本活動の主な教訓を記載する。

### 4.1 開発プロジェクトとの合意形成

データ計測の主体は、あくまでプロジェクトである。正確な実績データの蓄積には、開発現場の協力が不可欠である。しかし、計測が習慣となっていない組織では、計測を不要な作業と捉え、敬遠する傾向がある。

データ解析の最大の難関は、計測開始からメトリクスを提供するまでに、ある程度の期間を要する点である。このため、活動初期にメトリクスの活用で「何ができるのか」「どう役立つのか」を、開発現場と共有することが重要である。将来のビジョンを具体的に伝え、計測に対するモチベーションを維持することがねらいである。

当社の場合、この共有が不十分であった。このため、初期に収集したデータは、欠損が多い状態となった。現在は、後追いではあるが、プロジェクトが求める（または使用頻度が低い）メトリクスを毎年調査し、要求に応じたメトリクスを拡充している。

### 4.2 データ計測規則の定義

通常、計測規則は活動の初期で定義するが、当社は、この定義が不完全であった。レビューやテストで検出した欠陥の重要度や、テストケースの粒度などが典型例である。

しかもこの問題は、データ蓄積時に判明せず、メトリクス解析時に初めて表面化した。例えば、「テスト密度(=テストケース数/規模尺度)」の解析で、散布図を作成した際に、予想外の集落やばらつきが散見された。この原因を調査した結果、ある部門では、テストケース数を計測すべき個所に、テストのシナリオ数を登録していたことが判明した。このように、計測の慣行は、プロジェクトや部門ごとに大きく異なることがある。

この状況を踏まえ、組織共通の計測規則を再定義し、徹底した。しかしながら、初期に蓄積した数十プロジェクトのデータは、規則に準じたものかどうかをさかのぼって確認できず、利用できないものとなった。

また、品質・工数・工期に影響を与える定性要因(質的データ)を、初期段階から計測しなかったことも大きな失敗である。量的データの解析は、層別が基本である。ところが、層別を繰り返した場合でも、いずれ壁にぶつかることがある。

この解決策は、要件の確定度や顧客の関与度といった定性要因を、量的データと併せて解析することである。当社は、定性要因をダミー変数<sup>\*1</sup>化した重回帰分析も適用しており、量的データだけよりも、高精度な予測になることを確認している。しかし、この定性要因の蓄積が遅れたため、初期のデータを解析の対象にできないという問題がある。

### 4.3 実績データの鮮度と精度の確保

当初、プロジェクトが計測したデータは、プロジェクト完了時に「完了報告」として取りまとめ、SPI担当者に提供する運用であった。しかし、この運用が、データの鮮度と精度を損なう原因となった。このため、各種の実績データを一元管理するためのシステム(以下、Co.守り)を開発すると共に、データの登録プロセスを見直した。

#### 4.3.1 実績データの鮮度の確保

当社のプロジェクトは、年々大規模化し、工期も長期化する傾向にある。このような状況で、プロジェクト完了時に、上流工程のデータをさかのぼって計測することは現実的でない。このため、Co.守りへのデータの登録を、該当工程の完了後10営業日以内とした。

また、当社では、各工程で標準プロセスに対する遵守状況を、第三者が評価する品質保証活動を実施している。この評価では、実績データの登録状況も確認し、未登録の場合は、非遵守課題として指摘し、プロジェクトに是正対応を催促している。

#### 4.3.2 実績データの精度の確保

最終工程完了時には、Co.守りを經由して、完了報告に対する同意依頼がSPI担当者に届く。SPI担当者は、登録された約300のデータ項目をすべて確認し、指摘や疑問がある場合は、この依頼を何度でも差し戻す。

この確認は、メトリクスがあるからこそ可能である。実績データは、それ単品では判断が難しい。例えば、テストケース数という基本測定量だけでは、その値が多いのか少ないのかを判別できない。しかし、メトリクス「テスト密度」で、テストケース数とFPの2変数の関係を確認することで、要注意データが特定できるのである。

## 4.4 データ解析技法の習得と効率化

当社は過去に、統計技法を理解せずに解析を行った。データの分布を確認せずに、算術平均値と標準偏差で、基準値と上下限値を設定するといった誤りである。しかし、こうして得た結果は、プロジェクトの実情と乖離し、予測や評価に利用できないものであった。

ITメトリクスは、そのままの値(真値)で正規分布に近似することが少ない。当社では、多くのメトリクスが対数正規分布に近似することを確認している。このような性質を持つデータは、事前に変数変換などの処理が必要である。統計学では周知の知識や技法を習得しなければ、解析に不要な時間を費やすだけでなく、誤った結果を導くことになる。更に、ITメトリクスの解析では、統計技法に加えて、プロジェクト特性やIT分野固有の技術的知見(FPや欠陥の性質)も考慮する必要がある。

メトリクス分析支援ツールの開発にあたっては、事前に解析手順や技法を十分に習得した。一般的なメトリクスの解析手順を、図2に示すが、本ツールは、このすべての機能を網羅している。

\*1 定量的ではない質的・属性的なデータを「0」と「1」だけの数列に変換すること。



図2 ITメトリクスの解析手順

## 5 メトリクスの活用事例

メトリクスの活用形態は以下の3段階を経た。以降、図3で示す、統合テスト工程を例に、当社の品質管理の変遷を紹介する。

- ① 第1段階(自社メトリクス非活用)
- ② 第2段階(自社メトリクスの初期活用)
- ③ 第3段階(自社メトリクスの応用活用)

第1段階におけるテスト方針書・計画書では、テストの範囲・方法・環境などを記載する。しかし、品質の評価は、工程完了時におけるテストケースの消化状況の確認や、欠陥の原因・傾向分析が中心であり、勘と経験に基づくものである。

第2段階では、計画時に予測モデルを使って組織の基準値を目標として追加する。典型的には、開発規模に応じたテストケース数を予測する「テスト密度(=テストケース数/規模情報)」や、どの程度の欠陥を検出するべきかを予測する「欠陥検出密度(=検出欠陥数/規模情報)」を用いる。

品質の評価は、工程進行中から、図4の「ゾーン分析ツール」を活用する。本ツールは、統計的な予測ではないものの、目標となるゾーン(⑤)に対して、どのように推移しているのかを確認できる。工程完了時は、第1段階の分析に加えて、組織の基準値

との乖離を分析する。しかし、評価は、あくまでも組織の基準値との比較にしかすぎない。

第3段階では、メトリクス応用ツールである「検出欠陥予測ツール」や「信頼度成長モデル」を用いる。検出欠陥予測ツールは、図5のように、規模情報と特性を入力すると、この条件を満たす最も高精度の予測式を自動選択し、欠陥数を予測する。このツールにより、自プロジェクトの特性に応じた目標の欠陥数が設定できる。

更に、この欠陥数を入力値として、図6に示す「信頼度成長モデル」が利用できる。本ツールは、計画時に欠陥数とテストケース数を入力することで、指数型モデルに応じた検出欠陥数と上下限値の基準を設定する。テスト進行中は、欠陥の検出状況を反映し、完了時の検出欠陥数、潜在欠陥数、欠陥除去率と、目標の欠陥除去率を達成するための追加テストケース数などを予測する。また、組織の基準との乖離が大きい場合、プロジェクトの欠陥検出状況から、プロジェクト特性に応じた基準を再設定できる。

このように、第3段階では、組織の基準だけではなく、自プロジェクトの特性を反映した基準で、品質が評価できる。そして、評価は工程完了時ではなく、テスト進行中から継続的に実施するのである。

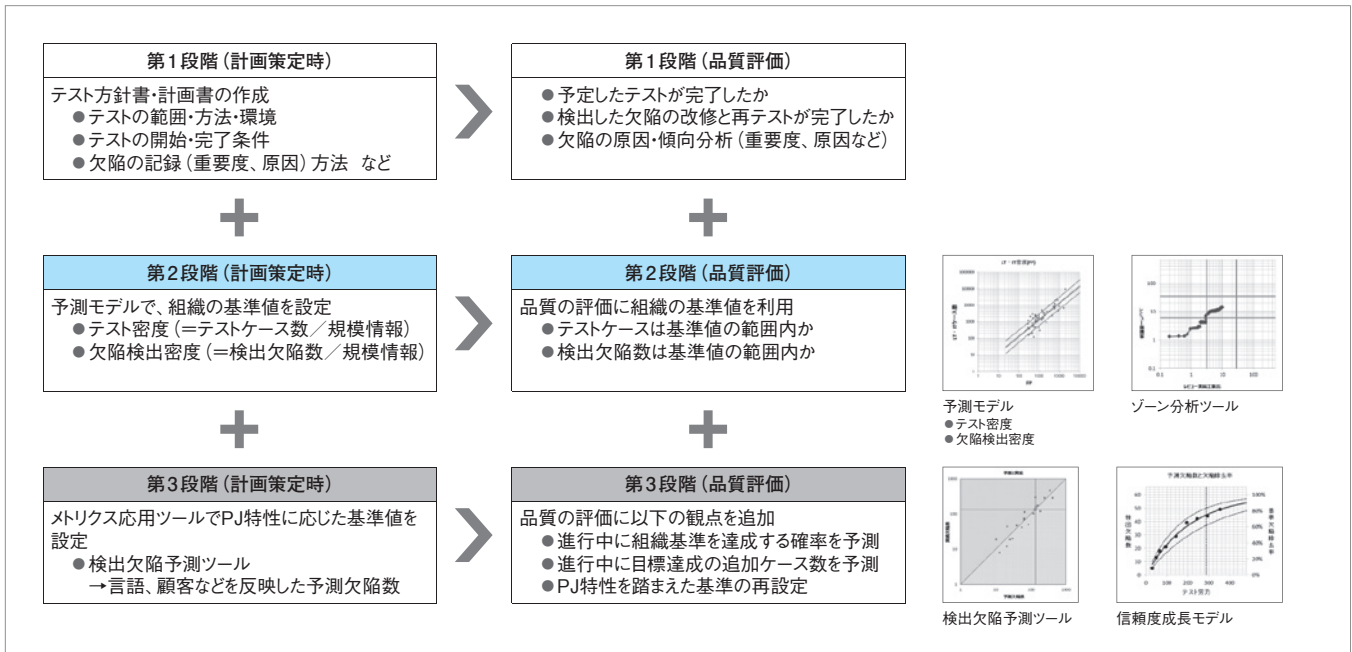


図3 メトリクスを活用した品質管理

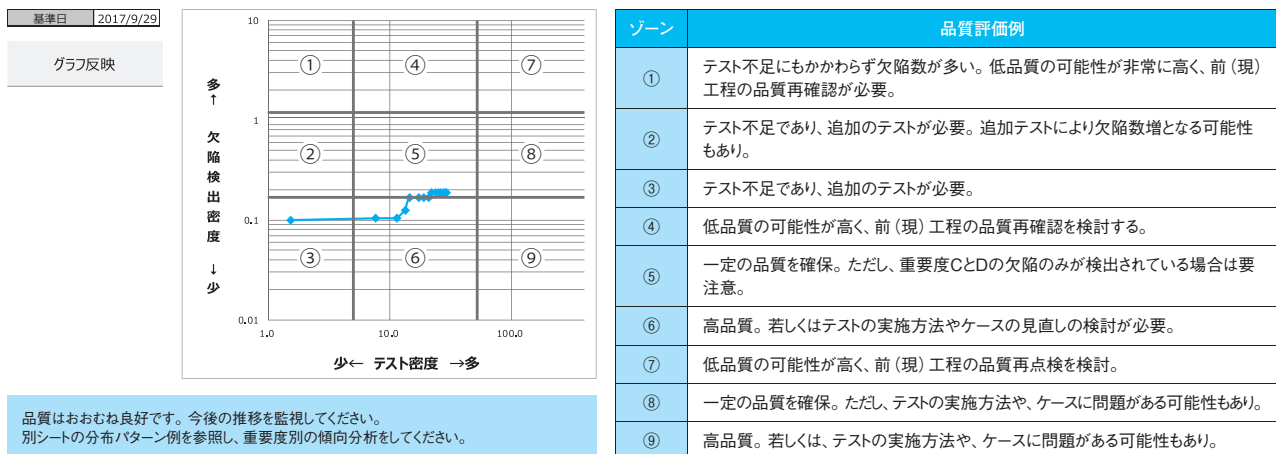


図4 ゾーン分析ツール

Ver 2.00  
TOINX 検出欠陥予測ツール(統合テスト)

【規模】  
 FP入力タイプ: 直接入力  
 直接入力FP値: 1000  
 SLOC (K): 50  
 LT数+IT数: 800

【プロジェクト特性】  
 開発タイプ: 新規開発  
 主OS: Solaris  
 主言語: Java  
 顧客: 顧客A

【予測欠陥数と上下限值】  
 予測欠陥数: 59  
 下限: 36  
 上限: 97

【選択欠陥数モデル】  
 モデル順位: 1  
 モデルID: N030F\_DM  
 開発タイプ: 新規開発  
 主OS: Solaris  
 主言語: Java  
 顧客: 顧客A  
 規模要因: F  
 データ数: 20  
 寄与率: 76.89%  
 上下限係数: 0.5

図5 検出欠陥予測ツール

SRGM for TOiNX V130

データ反映

分析対象データ **工程A1**

[計画値] **自動**

基準潜在欠陥数	80
予定テスト労力	100.0
形状パラメータ	0.023
位置パラメータ	0.000

[設定と予測]

基準設定 実績反映

SRGM予測 記録 消去

テスト労力追加 基準再設定

[実績情報]

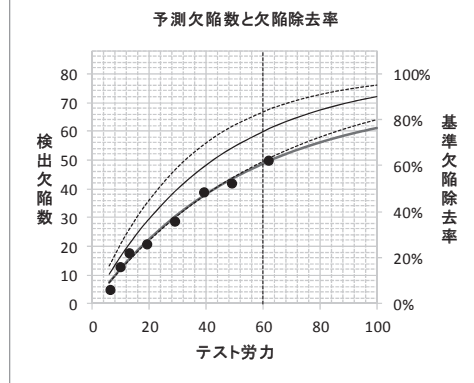
テスト労力	62
検出欠陥数	50
テスト進捗率	62.0%

[データ分割]

分割点探索 分割実行

テスト労力分割点 分割消去

指数モデル:  $m(t)=D(1-EXP(-b(t-t_0)))$



[処理結果]  
SRGM予測完了

[予測結果]

完了時基準達成確率	19%
完了時欠陥除去率	85%
潜在欠陥数	72
完了時検出欠陥数	61
信頼区間下限	55
信頼区間上限	68
現時点欠陥除去率	69%
除去率目標達成テスト労力	121.9
追加テスト労力	21.9
形状パラメータ	0.019
位置パラメータ	0.244

[テスト実績及び欠陥数管理限界]

No.	テスト 労力	検出 欠陥数	予測 欠陥数	除去率	基準欠陥		
					下限	基準	上限
1	6	5	7	9%	7	10	13
2	10	13	12	15%	12	16	21
3	13	18	15	19%	15	21	26
4	19	21	22	27%	21	28	35
5	29	29	30	38%	30	39	46
6	39	39	37	47%	37	47	55
7	49	42	43	54%	44	54	62
8	62	50	50	62%	51	61	68
9	65		51	64%	52	62	69
10	68		52	65%	53	63	70
11	71		53	66%	54	64	70
12	74		54	68%	56	65	71
13	77		55	69%	57	66	72
14	80		56	70%	58	67	73
15	83		57	71%	59	68	73
16	86		58	72%	60	69	74
17	89		59	73%	61	70	74
18	92		59	74%	62	70	75
19	95		60	75%	63	71	75
20	100		61	76%	64	72	76

初期段階では、組織の基準を利用。ただし、組織の基準との乖離が大きい場合は、指数型モデルによる基準の再設定が可能である。

テスト計画時

- 工程完了時の予測潜在欠陥数とテストケース数を入力することで、指数型モデルに応じた検出欠陥数と上下限值（組織の基準）を設定。

テスト進行中・完了時

- 検出欠陥数が組織の基準の管理限界内に入っているかどうかを視覚的に確認。
- テストの進捗が60%の時点で、テスト完了時の検出欠陥数、潜在欠陥数、欠陥除去率を予測。
- 目標となる欠陥除去率（90%）の達成に必要な追加テストケース数を予測。

図6 信頼度成長モデル

6 おわりに

メトリクスの活用によって、マネジメントのスタイルは変化する。各種の判断は、「定性・主観」から「定量・客観」となり、そのタイミングは「工程完了時」から「工程進行中」にシフトする。更に、品質やプロセスの目標は、「組織の基準」から「プロジェクト特性に応じた基準」に変わる。

開発プロジェクトでは、課題やリスクを工程完了時に特定しても手遅れである。より早期に要注意や異常な状態を特定し、是正する必要がある。事実、プロジェクトがデータ解析者に要望するのは、このようなメトリクスである。

一方で、高精度なメトリクスを提供するためには、組織レベルでのプロセス定着を推進する、SPI活動との連携が必須である。

従って、データ解析者のミッションは幅広く、蓄積した実績データを解析することは活動の一部にしかすぎない。

プロセスの改善機会は、常にプロジェクトからである。プロジェクトに有用なメトリクスを、いかにタイムリーに提供できるのかが、SPI・データ解析活動を加速するためのキーである。

参考文献・引用転載

[1] 佐藤浩明：ソフトウェア開発プロジェクトの生産性評価に関する事例—真の価値と工数削減効果を計測する手法、ユニシス技報 Vol.36 No.2 通巻129号、日本ユニシス株式会社、2016年

[2] 梶山昌之・合田英二・千野智子：ソフトウェア開発プロジェクトの計数管理フレームワークによる定量的管理、プロジェクトマネジメント学会誌 Vol.13、一般社団法人プロジェクトマネジメント学会、2011年