

ソフトウェアの定量的開発管理技術

Quantitative Software Development Management Technologies

森 俊樹

■ MORI Toshiki

ソフトウェアの大規模・複雑化に伴い、ソフトウェアの定量的開発管理がますます重要になっている。定量的開発管理の導入により、開発状況の可視化や、問題の早期発見、データに基づく客観的な判断、意思決定の根拠の明確化、などの効果が期待される一方、データ計測の負荷や収集データの精度など、解決すべき課題も多い。

これらの課題を解決するためには、既存の計測の枠組みを再構築するアプローチやデータ及び活用方法の標準化が重要であり、東芝は、ソフトウェアの定量的開発管理技術を研究・開発し、ソフトウェア開発現場への展開を推進している。

Quantitative software development management is becoming essential as the scale and complexity of software products increases. The introduction of quantitative development management brings in visualization of development, early detection of problems, and rational decision making based on quantitative data. Various issues still remain, however, including the high cost of measurement activities and low reliability of measurement data.

To overcome these problems, Toshiba is promoting a practical approach that involves restructuring of the existing measurement framework and standardization of data and utilization methods, toward the optimization of quantitative software development management.

1 まえがき

製品開発におけるソフトウェア開発の比重は高まっており、製品に搭載されるソフトウェアの大規模・複雑化が進んでいる。デジタル家電製品などの組込みソフトウェアの開発現場では、ソースコード規模が数百万行に達する大規模プロジェクトを、常時、数か月のサイクルで回していることも珍しくない。このような状況から、最近、ソフトウェアに起因する品質トラブルも増加する傾向にある。一時期ニュースをにぎわせた銀行ATM（自動預払機）のシステム障害などは、ソフトウェアの不具合が大きな社会的影響を及ぼした一例として記憶に新しい。

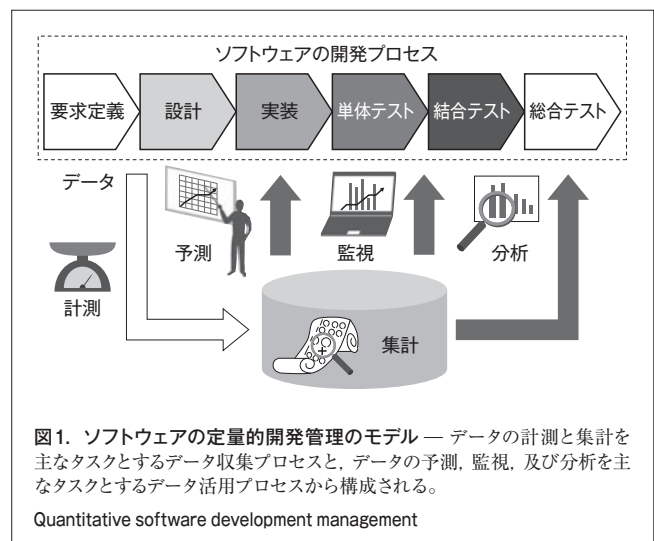
ソフトウェアの大規模・複雑化に伴い、ソフトウェアの開発管理はますます重要になってきている。しかし、ハードウェアの製造工程と異なり、ソフトウェア開発は作業状況がよく見えないことから、従来、定量的管理の導入が困難であった。多くのソフトウェア開発プロジェクトでは、いまだに、管理者の技量に頼った、勘と経験に基づく開発管理が実施されているのが現状であり、その結果として、意思決定の根拠があいまい、大規模プロジェクト管理のノウハウが十分に継承されない、などの問題が生じている。

ここでは、これらの問題を解決するため、東芝が開発しているソフトウェアの定量的開発管理技術の概要と実現へのアプローチについて述べる。

2 ソフトウェアの定量的開発管理

ソフトウェアの定量的開発管理は、開発プロセスとその成果物を測定することで得られる様々な定量的データを用いたソフトウェアの開発管理手法であり、データに基づく客観的な判断や開発状況の可視化などを目的としている。

ソフトウェア定量的開発管理のモデルの例を図1に示す。データの計測と集計を主なタスク（作業）とするデータ収集プロセスと、データの予測、監視、及び分析を主なタスクとするデータ活用プロセスから構成される。



データの計測では、プログラミング、レビュー、テストなどのソフトウェア開発工程と、ソースコードやドキュメントなどの中間・最終成果物に対して、必要な属性の定量化と記録を行う。小さなタスクやイベントレベルでのリアルタイムのデータ収集が要求され、構成管理ツールなど、管理系ツールとの連携が重要である。また、計測後なるべく早いタイミングで、データが正しいかどうかの検証作業を実施することが望ましい。

“データ集計”では、生の収集データを定められたルールに従って集計し、必要なデータがすぐに取り出せるように蓄積し保管する。その際、データの分類として、測定可能な属性をある特定の測定方法によって定量化した一次データである“計測データ”と、複数の計測データの関数として定義された“導出データ”，データの解釈のための判断基準や分析モデルを含んだ“メトリクス（指標）”を区別する必要がある。

データ活用プロセスでは、集計及び蓄積されたデータを用いて、“データ予測”，“データ監視”，及び“データ分析”を実施し、開発プロセスの計画や、制御、改善などを行う。

データ予測では、規模、コスト、スケジュール、品質などの見積り作業や、プロジェクト計画の策定、プロジェクト制御のための判断基準及びベースラインの設定などを行う。更に、定量的データがある程度蓄積されれば、ヒストリカルデータに基づく予測モデルの構築もできる。

データ監視では、予定と実績の差異を見て、プロジェクトの進捗（しんちよく）を把握し、状況に応じたプロジェクト制御を行う。基準値やベースラインとの乖離（かいり）を監視して、もし、異常値や異常パターンが検知されたら、速やかに原因を調査し、適切な対策をとる必要がある。そのためには、リアルタイムのデータ分析と開発者へのフィードバックが重要となる。

データ分析では、相関分析や、トレンド分析、仮説検定など、統計的手法を含む各種分析手法を駆使して、集計及び蓄積された定量的データを詳細に分析し、節目のレビュー時や出荷時の合否判断、及びプロジェクト完了後のプロセス改善などを行う。データ監視での分析はリアルタイム性が重視されるのに対して、ここでは、意思決定の客観的根拠や将来の改善に向けた視点などが要求される。

ソフトウェアの定量的開発管理の導入により、開発状況の可視化や、問題の早期発見、データに基づく客観的な判断、意思決定の根拠の明確化、などの効果が期待される。また、蓄積されたデータから今まで気づかなかった課題が見えてくる可能性もある。

その一方、定量的開発管理を実現するうえで様々なハードルが存在することも事実である。データ収集プロセスでは、データ計測の負荷が高すぎるため活動が長続きしない、データ定義のあいまいさや計測方法の属人性により信頼性の高いデータが集まらない、などのリスクがある。また、データ活用プロセスでは、収集の目的が不明確なためデータが有効活用され

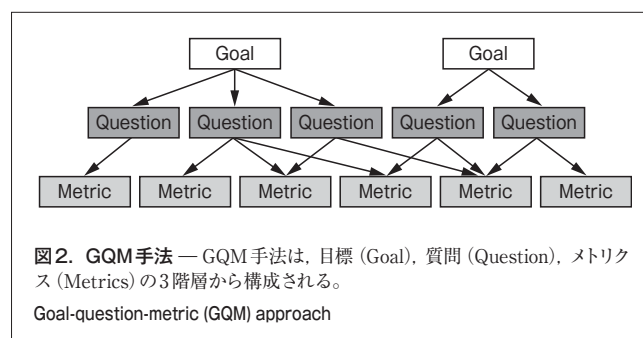
ない、適切なフィードバックが実施されず担当者のモチベーションが低下する、などのリスクが存在する。定量的開発管理の導入では、これらのリスクに対して注意を払う必要がある。

3 定量的開発管理実現へのアプローチ

必要なメトリクスのセットと、その計測の仕組み、活用方法などを定めた枠組みを“計測フレームワーク”と呼ぶ。ソフトウェアの定量的開発管理を実現するために、組織の課題や目標、ソフトウェアの開発環境、開発プロセス、及びリソースなどを十分に考慮して、計測フレームワークを構築することが重要である。

3.1 GQM (Goal-Question-Metrics)⁽¹⁾

計測フレームワークを構築するための代表的な手法として、1984年にBasiliとWeissによって提案されたGQM手法がある。GQM手法は、ソフトウェアのメトリクスをトップダウン方式で決定するアプローチであり、図2に示すように、目標（Goal）、質問（Question）、メトリクス（Metrics）の3階層を持つ。



GQM手法の実施手順を以下に示す。

- (1) 対象となる組織やプロジェクトにおけるビジネス目標と、それに結び付く計測目標を設定する。
- (2) 計測目標に関連する一連の質問リストを作成する。ここで、質問とは、目標の達成を確認するために答えるべき質問であり、目標が達成されたか否かを判断するための具体的な基準となるものである。
- (3) 質問に対して定量的な答えを得るのに必要なメトリクスを定義する。一つの質問に対して複数のメトリクスが必要になる場合もある。
- (4) データ収集方法を決定して、実際にデータの収集と分析を行い、目標が達成されているかどうかを判定する。

GQM手法は、これからデータ計測を始める新しいソフトウェア開発組織にとって非常に有効である。しかし、ある程度成熟した組織では、既になんらかの形でデータ計測を行っている場合が多く、つど、一から計測フレームワークを作り直すというのもあまり現実的ではない。既存の資産をなるべく生かして、中長期的な視点でデータを蓄積できるように、計測フ

フレームワークを再構築するアプローチが重要と思われる。

3.2 計測フレームワークの再構築

組織の課題は、時間の経過とともに徐々に変化する。それに伴って、計測フレームワーク自体も、課題との乖離が生じるため、ある一定期間ごとの段階的な見直しが必要となる。ここでは、GQM手法などのトップダウンアプローチと、既存の計測フレームワークの分析に基づくボトムアップアプローチを融合した、計測フレームワークの再構築手法を示す(図3)。

- (1) まず、関係者のヒアリング調査や組織のコスト構造分析などを実施して、組織の重要課題を洗い出し、そのうえで、GQM手法などのトップダウンアプローチにより、課題に対応した主要メトリクスを決定する。
- (2) ボトムアップアプローチとして、既存の計測フレームワークの棚卸し調査を実施する。すなわち、すべての収集データに対して、その単位、定義、使用するツールと帳票類、計測と集計の頻度、粒度(集計の細かさ)、データ精度、及びデータの活用方法などを詳細に調査し結果をまとめる。
- (3) 棚卸し調査に続いて、データ収集プロセスの評価及び検証として、実際にデータ集計表を埋めてみる。データ精度はどうか、欠損値はどの程度あるか、どれだけの頻度や粒度でデータ収集できるか、データ収集の負荷はどのくらいかなど、組織の成熟度を知ることができる。
- (4) トップダウンアプローチとボトムアップアプローチの両方の結果を合わせて、計測フレームワークの再構築を実施する。データの集計粒度や活用方法などを見直すとともに、課題との対応関係や、将来の利用可能性、データ精度、計測負荷など、コストと効果の両面から評価を行い、収集データの追加や削除を決定する。その際、不要なデータの計測は中止する一方、将来的に必要なデータは残すなど、中長期的な視点に立った判断も必要となる。
- (5) 新しい計測フレームワークを定着させるには、開発者へのデータ分析結果のフィードバックが重要である。既

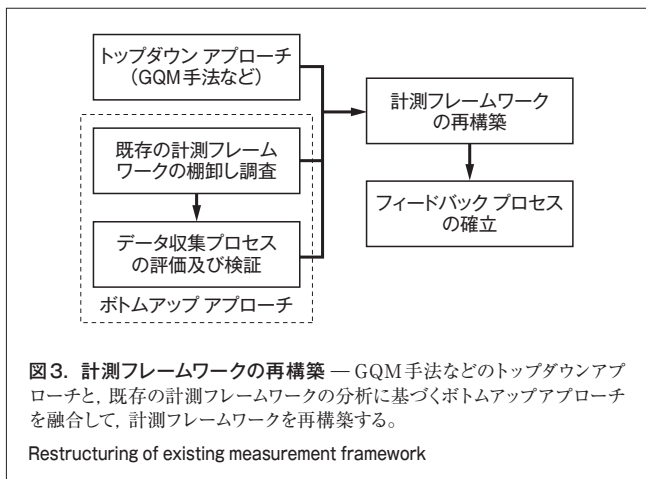


図3. 計測フレームワークの再構築 — GQM手法などのトップダウンアプローチと、既存の計測フレームワークの分析に基づくボトムアップアプローチを融合して、計測フレームワークを再構築する。

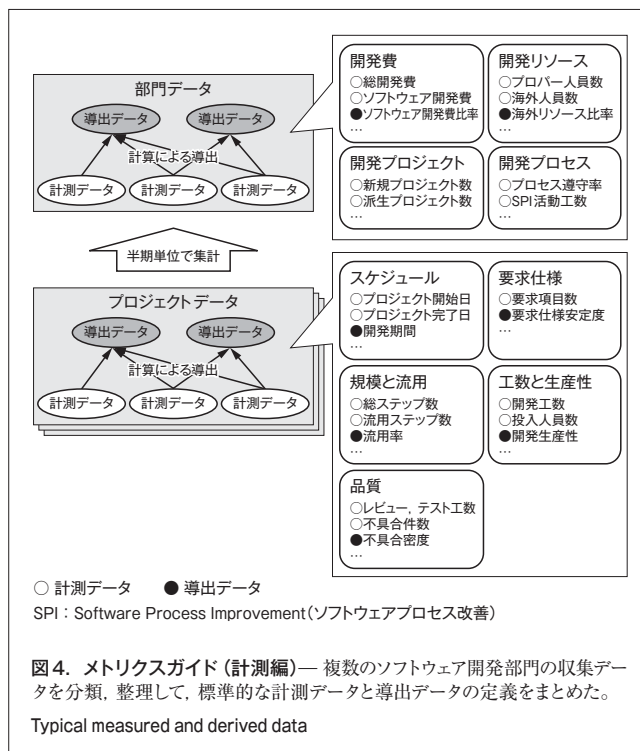
存のフィードバック プロセスの有効性を評価して、適切な改善を実施する。

3.3 収集データと活用方法の体系化

ビジネス環境や開発プロセスなどの違いから、組織の課題がそれぞれ異なっているように、計測フレームワークも組織ごとのカスタマイズが必要となる。しかし、基本的な収集データやデータ活用方法などについて、共通部分も多く見られる。中長期的な視点で考えると、このような共通部分に関して、知識やノウハウを体系化し、標準化を推進することが重要である。それにより、データの収集と活用の標準環境の整備、知識やノウハウの共有による活動のレベルアップ、及び定量的開発管理技術の研究・開発の加速などにつながる事が期待される。

まず、収集データに関して、複数のソフトウェア開発部門を調査したところ、ソースコード規模や、開発工数、テスト項目数、不具合件数など、だいたい似たようなデータを計測していることがわかった。ただし、同じデータ名であっても、部門によって微妙に定義が異なっているものもある。例えば、プログラミングの生産性は、一般に、ソースコード規模を工数で割ったものとして定義されるが、部門によって、分子が総ステップ数だったり、新規や変更ステップ数だったりする。このような問題を解消するために、各部門の収集データを分類して整理し、標準的な計測データと導出データの定義をまとめた“メトリクスガイド(計測編)”を作成した(図4)。

更に、データの活用方法に関しても、現在、過去事例や文献などを調査して、知識やノウハウの分類と整理を行っている。その際、対象データや、利用シーン、分析方法などをパッ



ケージ化した“メトリクス活用ストーリー”というものを分類の基本単位として、予測、監視、及び分析などのデータ活用プロセスに沿った体系化を進めている。将来的には、“メトリクスガイド(活用編)”としてまとめる予定である。

4 ソフトウェアタグ技術によるデータ収集の自動化

前述のように、ソフトウェアの定量的開発管理を実現するうえで様々なハードルが存在するが、その中でもっとも高いハードルは、おそらく、データ精度の確保と収集の負荷と思われる。現在、多くのソフトウェア開発組織において、工数データなどの一部のデータはいまだに自己申告で、その他のデータについても計測担当者が様々なツールやドキュメントから苦労してかき集めている状況であり、データ収集プロセスがボトルネックとなっている。中央集権的な管理ツールを中心に置いて、その他のツール群や帳票などから必要なデータを集めてくるというアプローチもあるが、システムの柔軟性や拡張性などの面で課題も多い。

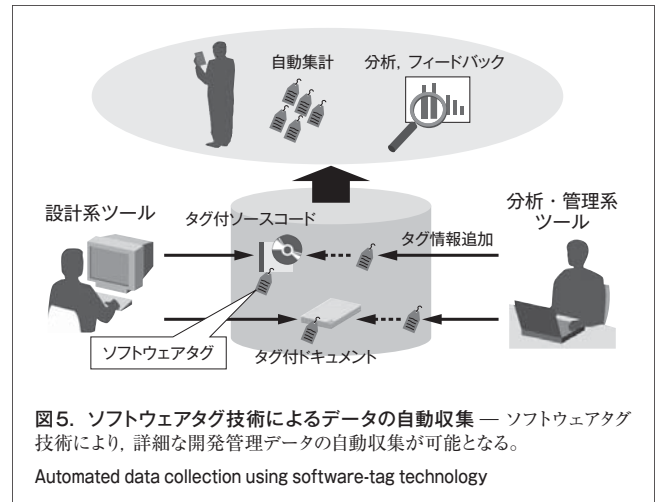
ここでは、このような課題の解決に向けた将来有望なアプローチとして、“ソフトウェアタグ技術”について述べる。

ソフトウェアタグは、文部科学省の研究プロジェクトである“StagEプロジェクト「次世代IT基盤構築のための研究開発：ソフトウェア構築状況の可視化技術の開発普及」”のなかで提案された概念である。ソフトウェア開発組織の情報や開発プロジェクトから収集した様々なデータを一定の形式でパッケージ化して、中間成果物及び最終成果物に添付したものである。ICタグによる食品などのトレーサビリティ(追跡可能性)確保の仕組みをヒントに考案され、典型的な利用イメージとしては、委託元がタグデータを見て開発状況を知る、重大問題発生時の原因究明や法的紛争時にタグデータを利用して第三者評価を行う、再利用可能なソフトウェア部品の品質保証などに使う、などを想定している。

ソフトウェアタグを実現し普及するために必要な技術として、タグの規格化と実装方法、タグデータの収集、分析、可視化のための手法とツールの開発、及びタグ利用シナリオの明確化などの一連の研究が、奈良先端科学技術大学院大学と大阪大学を中心に進められている。

StagEプロジェクトの最初の成果として、当社もメンバーに含むタグ規格技術委員会の検討を経て、ソフトウェアタグ規格書Ver.1.0²⁾が2008年10月末に公開された。そこには、ソフトウェアタグの概要及び位置づけとともに、ソフトウェアタグに含めるべきデータ項目の一覧が含まれている。タグに含めるべきデータ項目は、大きく2種類あり、システム構成や、開発手法、組織情報などのプロジェクト情報と、ソースコード規模や、作業工数、欠陥件数などの進捗情報から構成される。

ソフトウェアタグ技術に対する大きな期待の一つとして、デー



タ収集プロセスの自動化が挙げられる(図5)。将来的に、各種の設計系や分析・管理系のツールがタグの自動生成をサポートすれば、開発者がデータ計測を意識することなく、工数データなど詳細な開発管理データを自動収集できる可能性がある。また、データの絶対量と精度が大幅に向上することで、データ分析の高度化などデータ活用面の進歩も期待できる。

5 あとがき

ここでは、ソフトウェア定量的開発管理の一般的なモデルを示し、その実現に向けて、トップダウンアプローチとボトムアップアプローチを融合した計測フレームワークの再構築手法の開発、及び収集データと活用方法の体系化への取組みについて述べた。更に、将来有望なアプローチとして、ソフトウェアタグ技術の研究についても述べた。

ソフトウェアタグ技術は、開発管理データの標準化へと向かう大きな流れの一つとみなすことができる。今後、StagEプロジェクトの研究成果に注目するとともに、社内における定量的開発管理技術の研究及び開発も、ベクトルを合わせて進めていきたい。

文献

- (1) Basili, V., et al. A Methodology for Collecting Valid Software Engineering Data. IEEE Trans. Software Eng. 10, 6, 1984, p.728-738.
- (2) 文部科学省. “ソフトウェアタグ規格書 Ver.1.0, StagE”.
< <http://www.stage-project.jp/> >, (参照2009-02-28).



森 俊樹 MORI Toshiki

ソフトウェア技術センター イノベーション推進担当専事。
ソフトウェア品質技術、及び定量的開発管理技術の研究・開発に従事。
Innovation Group