GQM モデルに基づく
設計工程完成度計測手法の提案

松村知子，森崎修司，玉田春昭，大杉直樹，門田暁人，松本健一
奈良先端科学技術大学院大学　情報科学研究科

1. はじめに

EASE (Empirical Approach to Software Engineering) プロジェクト[1]では，ソフトウェア開発プロジェクトをリアルタイムにモニタし，「プロジェクト遅延」や「障害対応遅延」の状況，または遅延に繋がる原因を早期に検出する手法の研究・開発を進めている．問題を早期に発見し，適切な対処を行うことで，進行中のプロジェクトの納期遅延，コスト超過，品質低下を防止することが可能になる．

著者らは，既に GQM (Goal/Question/Metric) パラダイム（分析目的を明確にした上でトップダウンに計測を行うフレームワーク）[2]を用いて，「プロジェクト遅延の主な要因」を検出するモデルを作成した [3]．このモデルでは，EASE プロジェクトで開発したツール EPM(Empirical Project Monitor) [1]の導入を前提とし，EPM 収集データから計測されるメトリクスを用いる．しかし，EPM が収集するデータは，あくまでコーディング工程以降の情報，つまりプログラムコードの変更履歴とテスト時の欠陥データをベースとするため，下流工程（コーディング工程以降）で問題を発見することしかできない．

一般的に「欠陥は，要求や設計工程で多く混入し，後工程での修正には多くのコストを要する」[4]と言われるとおり，より上流の工程でのプロジェクトの遅延要因となる品質上の問題（欠陥）発見が，コストの面からも望ましい．

本報告では，上記の GQM モデル[3]を上流工程に適用し，設計工程における成果物，すなわち仕様書や設計書の完成度を計測する方法を提案する．上流工程における仕様書・設計書の変更履歴やレビューで見つかる指摘事項（欠陥）を同様のモデルで分析し，完成度を推定することで，『次工程への移行を留保して品質の再確認を行う』，『下流工程での仕様変更や設計変更のリスクに予め備える』，などの対策が可能になる．上流工程での早期の対策により，プロジェクトの納期遅延，コスト超過を防止できると考えている．

本稿では，具体的な仕様書・設計書の変更履歴や欠陥から計測可能なメトリクスによる完成度の計測方法を提案するとともに，各メトリクスの下流工程に

---

[1] http://www.empirical.jp/top.html

おける品質との関係を統計的に分析した結果について述べる.

2. 提案する計測手法

図 1 に，上流工程における GQM モデルを示す.

Goal として「仕様書の変更履歴とレビュー欠陥データから仕様書変更パターンを解析し，企業における特定のプロジェクトにおいて，プロジェクトマネージャの視点から，仕様書の完成度について評価する」としている. [3]の下流工程を対象とした GQM モデルでは， Goal の 1 つは「構成管理データ（CVS）と欠陥管理データ（GNATS）からプログラム変更，仕様変更パターンを解析し，企業における特定のプロジェクトにおいて，プロジェクトマネージャの視点から，要求の不安定さについて評価する」となっており，Goal についてはほぼ流用した.

計測対象は，上流工程で計測可能なメトリクス，仕様書・設計書の変更量，変更回数，レビューでの指摘数(欠陥数)，レビュー時間などから選択した. 図 2 および図 3 は，図 1 のモデル作成時の計測案の模式図である. 図 2 では，レビュー工数や発見欠陥数，残存欠陥数などのレビュー毎の推移を表している. 図 2 の（a）は完成度が高い場合の推移で，レビュー時間や発見される欠陥数が収束している. (b) では，レビュー時間が増加し続け対象文書のページ数も増加しているため，まだ完成度は低いと考えられる. 図 3 は，仕様書・設計書の変更量や変更工数のバージョン毎の推移例だが，(b)ではページ数や変更量，変更工数などが増加し続けていることから，設計作業が終了に向かっているとは考えにくい.

上流工程の仕様書・設計書の変更履歴とレビュー情報に基づく計測は，メトリクスの計測方法やその手順・ツールなどが整備されていないため，実際には収集できたデータを確認しながら，試行錯誤で行った. まず，このようなレビューや変更に関する計測値は，レビュープロセスや文書のバージョン管理方法に大きく依存するため，仕様書や設計書，レビュー議事録等のバージョン管理方法について統一した手順を規定し，それに従った管理をプロジェクト管理者および開発者に徹底した（3.1 参照）. 次に，完成度を示すと思われるメトリクスを選択し，計測を行った（3.2 参照）. 最後に，下流工程での GQM モデルの適用では，週次で分析・フィードバックを行った[5]が，上流工程については詳細設計終了時の品質評価会議で，分析結果を各社にフィードバックしインタビューを行って評価した.

なお，最終成果物の品質に与える影響の分析として，上流工程での計測メトリクスと，上流工程で混入し下流工程で発見された欠陥の有無や欠陥数との関係を統計的に分析した.

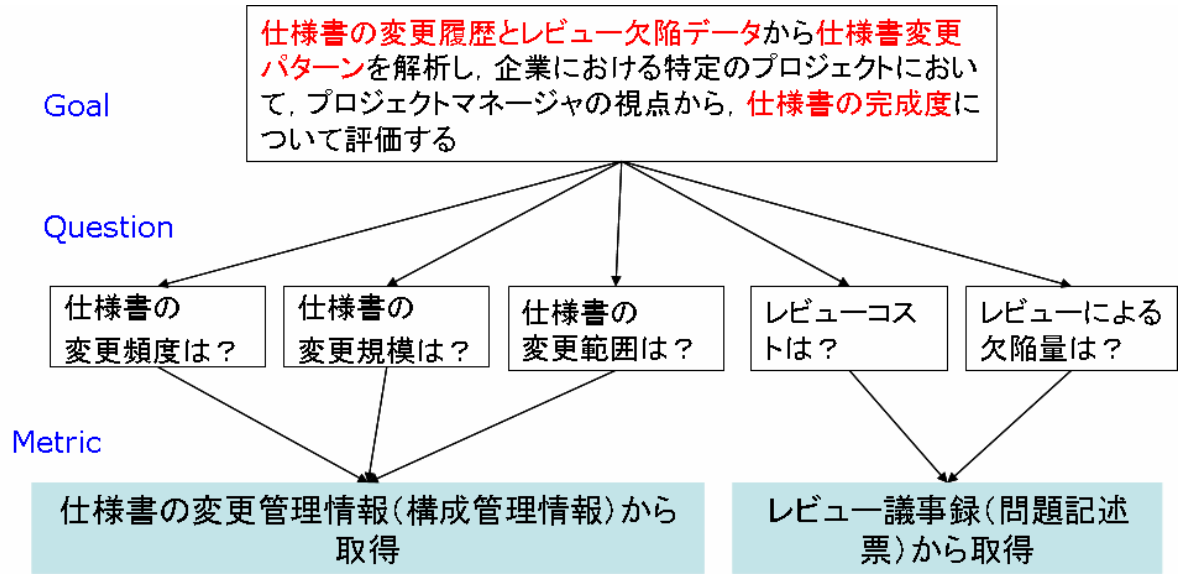
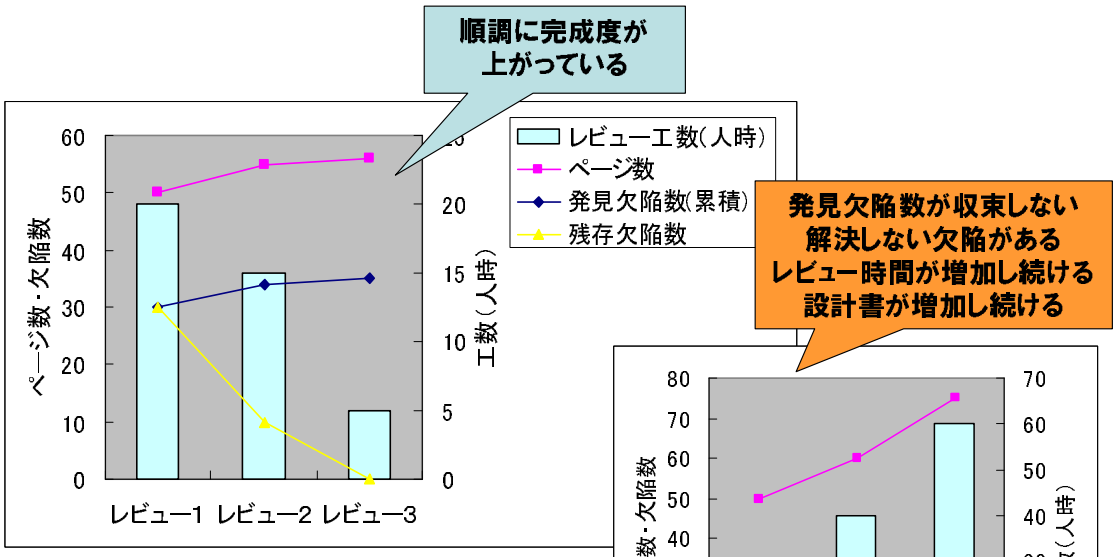


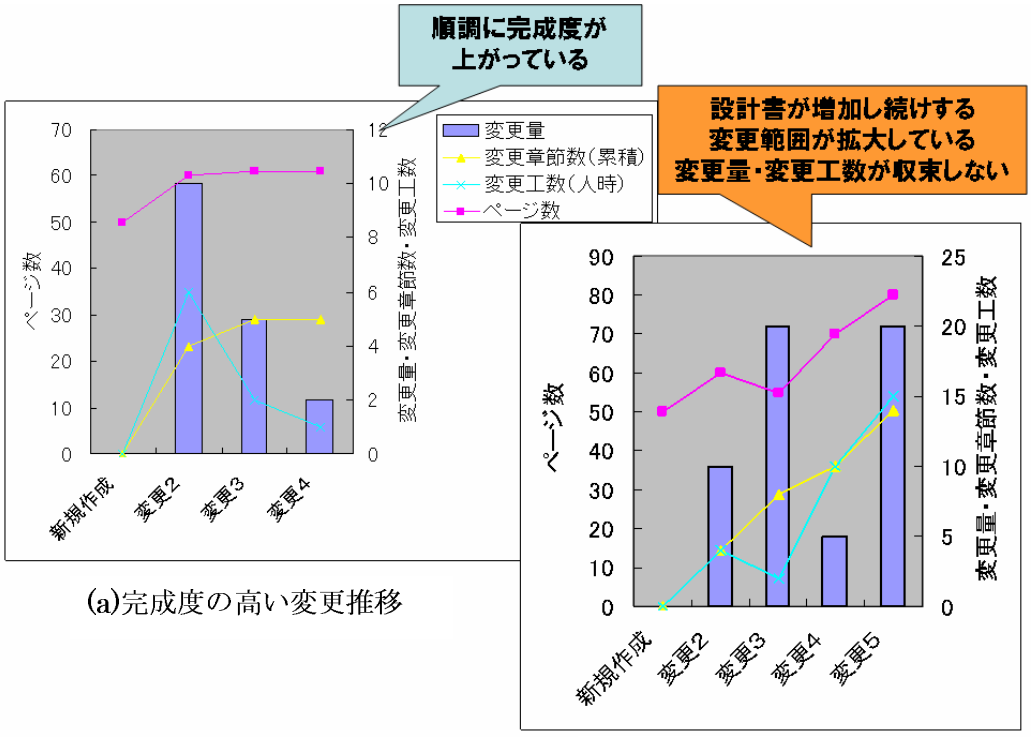
図 1 「仕様書の完成度評価」GQM モデル



(a)完成度の高いレビュー推移

(b)完成度の低いレビュー推移

図 2 レビュー特性値の推移模式図



(a)完成度の高い変更推移

(b)完成度の低い変更推移

図 3　成果物（仕様書・設計書）特性値の推移模式図

3.　計測手順

3.1　データ収集方法

仕様書・設計書のバージョン管理には，構成管理システム CVS を用いた．図 4 に仕様書・設計書の変更とレビューの運用ルールを示す．仕様書・設計書は，レビュー前に Ver0.1（CVS 上は Ver1.1）を CVS に登録する．Ver0.1 に対するレビュー指摘事項が修正された場合，その修正版を Ver0.2（CVS 上は Ver1.2）として登録する．最後にすべての修正が完了した仕様書・設計書をバージョン Ver1.0 として登録し，更新完了となる．ほとんどの仕様書・設計書は，MS-Word もしくは MS-Excel によって作成される．

レビュー議事録（以降，問題記述票と呼ぶ）は，MS-Excel の帳票形式を用いた．ファイル名は，「レビュー日付＋文書名(省略可)」という命名規則を設け，文書別にレビュー毎に起票するようにする．指摘事項がない場合も，レビューを実施した場合，レビュー日時・レビュー時間等のデータ収集のため，問題記述票を起票する．

仕様書・設計書と問題記述票の対応をとるため，仕様書・設計書を CVS へ登録する際のコメントとして，更新理由となる問題記述票のファイル名を記載す

るようにした．ただし，プロジェクト管理上（最終的な仕様書・設計書の形式上）仕様書・設計書にも同様の変更履歴の記入が行われるため，CVS のコメントへの記入は任意とした．CVS 上のコメントや仕様書・設計書への変更履歴の書式は，例えばレビューでの指摘に対する修正の場合，「“問題記述票のファイル名”の指摘番号１～６番の修正」と取り決めた．

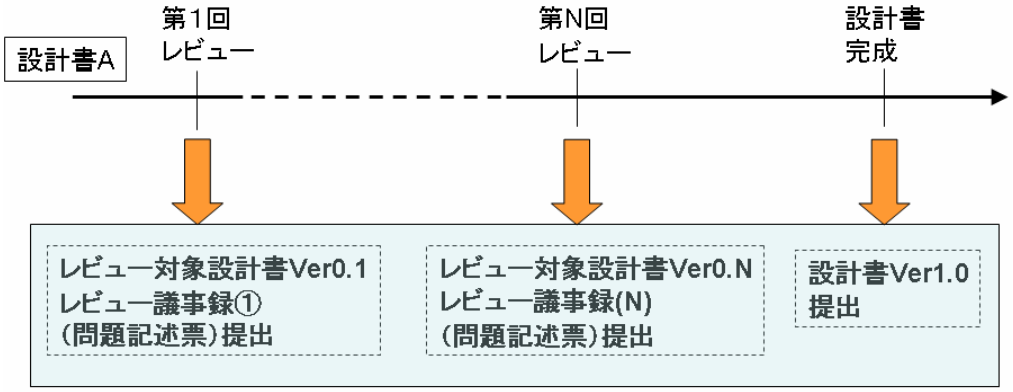


図 4　ドキュメント（仕様書・設計書）変更ルール

図 5　文書間差分例

3.2　計測メトリクス

問題記述票は，レビュー単位で作成され，レビュー情報として対象文書名，対象文書のページ数，レビュー時間，レビュア人数，レビュー日時などが記入される．各指摘事項（以降，欠陥と呼ぶ）については，指摘内容，解決方法，欠陥種類，原因，重要度，解決日時，重要度が大・中の場合修正工数などが記入されている．

仕様書・設計書は，CVS リポジトリからチェックアウトすることでバージョンごとの文書を取り出して，計測を行う．MS-Word の場合，ページ数はファイルを開いたときのページ数を用いる．この時，表紙，履歴，目次の 3 ページはページ数に含む．MS-Excel の場合，シート数をページ数とする．変更量は，市販ツール DiffDoc[2]を用いて計測した．DiffDoc による出力結果を図 5 に示す．結果には変更・追加・削除箇所数と変更箇所それぞれの内容が示される．この出力結果から，Excel-VBA のツールなどを用いて，表 1 に示すメトリクスを抽出する．

表 1　変更量メトリクス一覧

メトリクス	詳細	備考
変更箇所数	ヘッダに表示される Changes Found	
非重要変更箇所数	変更箇所で，表紙・変更履歴・目次等，仕様・設計内容に関連のない箇所数	変更箇所を目で確認
追加箇所数	ヘッダに表示される Additions Found	
非重要追加箇所数	追加箇所で，表紙・変更履歴・目次等，仕様・設計内容に関連のない箇所数	追加箇所を目で確認
削除箇所数	ヘッダに表示される Deletions Found	
非重要削除箇所数	削除箇所で，表紙・変更履歴・目次等，仕様・設計内容に関連のない箇所数	削除箇所を目で確認
変更文字数	赤字の文字数（取り消し線なし［変更後］と取り消し線付き［変更前］の合計)	Excel VBA ツールで計測
追加文字数	緑字の文字数＋赤字の取り消し線なし文字数（変更箇所の変更後文字数)	Excel VBA ツールで計測
削除文字数	青字の文字数＋赤字の取り消し線付き文字数（変更箇所の変更前文字数)	Excel VBA ツールで計測

[2] http://www.marroncraft.com/diffdoc/

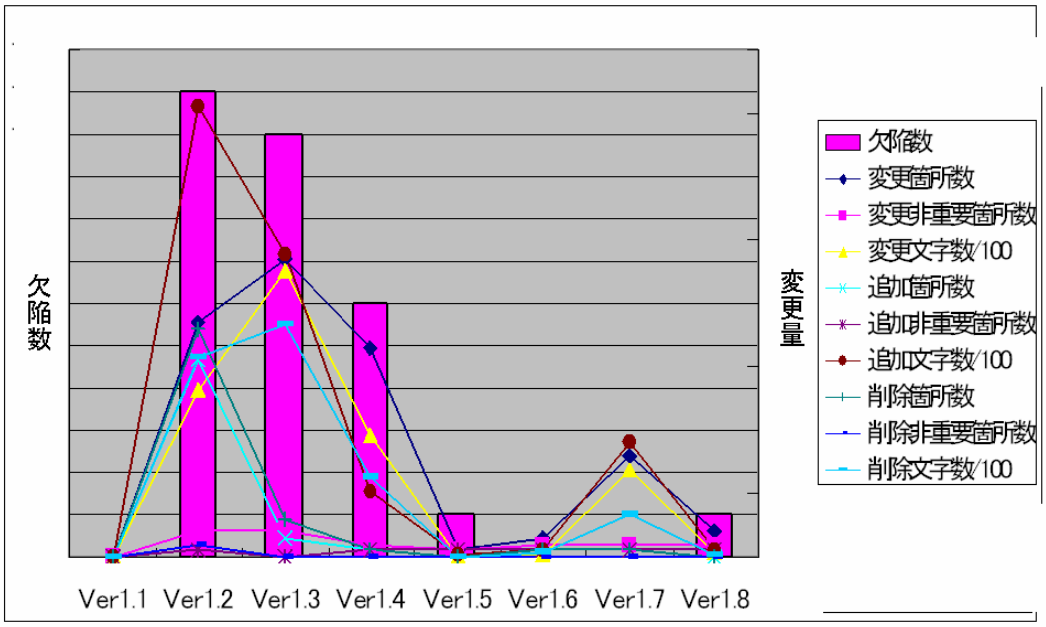


図 6　詳細設計書分析結果例 1

3.3　レポート例とその見方

表 1 に示したメトリクスの計測結果の可視化例を図 6 に示す. Ver1.1 は, CVS 上に登録された最初のバージョンを示し，以降更新ごとに Ver1.2，Ver1.3 と示される．Ver1.2 の変更量（折れ線グラフ）は Ver1.1 と Ver1.2 の差分から計測された値である．Ver1.2 の欠陥数は，Ver1.2 への更新時に修正が行われた欠陥数である．この値は，Ver1.2 のコメントから取得できる．例えば，「20060501_文書 A_問題記述票.xls の欠陥 1,3,5-9 の修正」と書かれている場合は，欠陥数が 7 となる．これは，通常は Ver1.1 の文書に対して指摘された欠陥数と考えられる．ただし，Ver1.1 に対するレビューで見つかった欠陥が Ver1.3 まで持ち越されて修正されるケースもあり, その場合は Ver1.3 の欠陥数として計上される.

4.　適用事例

4.1　対象プロジェクト

本報告の対象プロジェクトは，今日の日本における情報システム開発の典型とも言える「マルチベンダ開発」である．具体的には，経済産業省の支援を受けて進められた中規模の情報システム開発プロジェクト「プローブ情報システム開発プロジェクト」を対象とした．本プロジェクトは，2 年にわたって新規開発(一年目)，処理の共通化と機能追加（二年目）が行われたが，2006 年度はその二年目にあたる.

このプロジェクトでは，EASE プロジェクトとソフトウェアエンジニアリング

技術研究組合(以降，COSE と呼ぶ)[3]，情報処理推進機構（IPA）下に設置されたソフトウェアエンジニアリングセンター（以降，SEC と呼ぶ）[4]の 3 者が連携して，データ収集体制の構築とデータ分析を行った．COSE はユーザ的立場の会社とプロジェクト管理担当会社，開発担当５社から成り，ユーザ的立場の会社が要求を出し，開発担当５社が分担してサブシステムの開発を行い，プロジェクト管理担当会社が 5 社を統括した全体のプロジェクト管理を行う．開発はウォーターフォールプロセスにより進められ，具体的には，要件定義を受けて基本設計を全社の協働で行い，以降各社で詳細設計，プログラム設計，製造，単体試験，社内結合試験が行われた．その後，各社のプログラムを集め，組合全体として社間結合試験，総合試験が行われた．各社で分担して作成した基本設計書については，全社の担当者によるレビューが行われたが，以降社内結合試験までは週次の進捗報告(進捗率)と各工程完了時の品質評価報告のみがプロジェクト管理担当会社に提示される．このような限定されたプロジェクト管理方法となった要因の１つには,開発対象が協調（公開）領域と競争（非公開）領域が混在したものになっており，社間のみならずプロジェクト管理担当会社に対してもプロダクトや社内情報に関する機密保持が求められたからである．

本研究で用いた計測手法は，対象プロジェクトのような研究組合型マルチベンダ開発の特徴である限定された情報によるプロジェクト管理に対して，有効な支援を行う方法として導入された．社間で統一されるべきデータ収集方法などの取り決めは，開発が始まる前に各社の開発担当者と協議して決定した．

4.2　結果例

文書の計測例を図 6～図 8 に示す．

図 6 は，典型的なパターンの１つである．仕様書・設計書に対して修正した欠陥数が，バージョンごとに順調に減少していき，変更量も Ver1.2 では追加量が多く，Ver1.3 では変更量が多いが，以降は修正量が収束している．ただし，Ver1.7 での修正（非重要箇所でない修正）がなぜ入っているか，Ver1.8 での修正欠陥がどの時点で見つかった欠陥なのか，またこれまで修正されなかった理由は何か，などがチェックのポイントとなる．

図 7 は，欠陥自体は少ないが，欠陥対応以外の修正が多い例である．このような場合，仕様や設計方針の頻繁な変更が発生している，あるいは設計内容がなかなか決められずに後回しにされている，などの問題があると考えられる．

図 8 は欠陥数・修正量ともに収束しない例である．このような場合，レビュ

---

[3] ソフトウェアエンジニアリング技術研究組合（COSE）：COnsortium for Software Engineering

[4] http://sec.ipa.go.jp/index.php

一不足やレビュー方法が不適当といった問題があると考えられる．

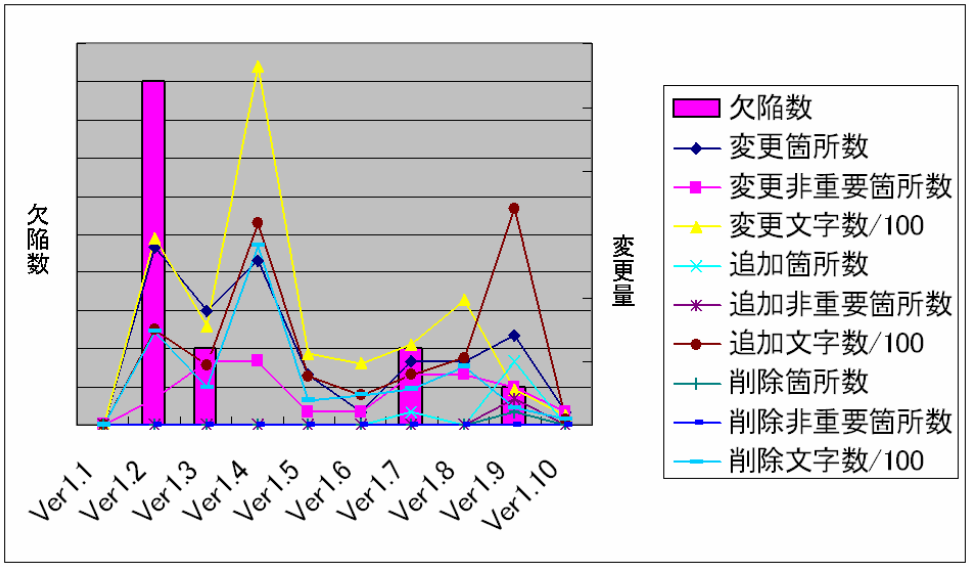


図 7　詳細設計書分析結果例 2

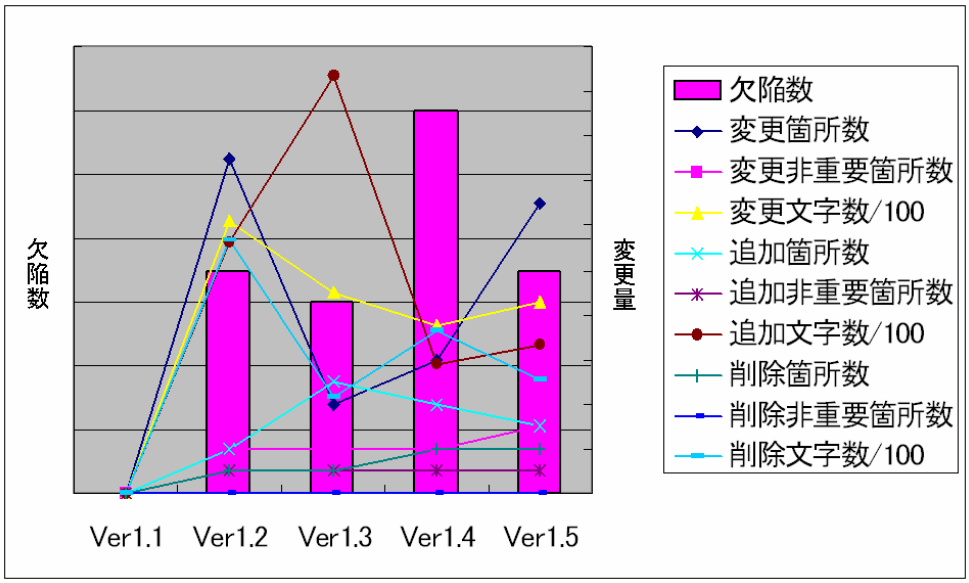


図 8 詳細設計書分析結果例 3

4.3　インタビュー結果

図 6～図 8 に示したような設計文書ごとの分析グラフを，詳細設計完了時の品質評価会議（プロジェクトマネージャ，SEC 研究員，各社開発担当者，EASE 研究員が参加）において提示し，インタビューを行った．その場で回答できない質問に関しては．質問表にて，後日メールにて回答を入手した．インタビュ

一（質問）と回答の例を示す.

例1） ［質問］xxxx 設計書で，最終 Ver1.5 での変更量が多いのはなぜか？欠陥発生が収束していないように見られるが，プログラム設計(次工程)に支障はないか？

［回答］最終レビュー参加者より、頻出する文言の変更要求があったため、変更量が増大している。プログラム設計に支障はない。

例2） ［質問］xxxx 設計書の Ver1.3 での大幅な修正理由は，何か？CVS コメント上は，「yyyy モジュールとの I/F 調整による修正、及び誤記修正」となっている．「yyyy モジュール」の欠陥，もしくは変更に伴う修正か？

［回答］I/F を yyyy モジュール側と統一するための修正で、その修正量が多く発生したためのです。特に欠陥等ではありません。

このインタビューの時点では，実際に完成度について問題がある，とプロジェクトマネージャや開発関係者が認識した事例はなかった.

5. 提案する計測メトリクスと最終成果物の品質の相関分析

5.1 評価観点

設計工程完成度を評価する観点としては，次の２つが考えられる.

- 観点１：対象工程以降で上流工程への手戻りが発生したか？すなわち，詳細設計工程中に基本設計を変更しなければいけないような欠陥が発生したか？
- 観点２：下流工程で，設計時に混入した欠陥が発見されたか？

本プロジェクトでは，下流工程で発生した欠陥により，上流工程に遡って仕様書・設計書を変更する，という管理は，仕様･設計変更管理票を用いて行われていた．しかし，基本設計書についてのみであり，詳細設計以降については，各社内での設計変更や仕様変更については各社内で管理されていた．基本設計書への詳細設計工程以降での修正は実質１，２件しか発生しておらず，社内の仕様・設計変更データについては収集対象外であったため，評価は観点２で行うこととした.

5.2 評価方法

評価は，次の２つの方法を用いて行った.

I. 各文書において，上流工程での欠陥数や修正量と，設計時に混入し下流

（試験）工程で発見された欠陥（以降，設計欠陥）の有無に関連が見られるか？

II. 各文書において，上流工程での欠陥数や修正量と，設計欠陥数に関連が見られるか？

Ⅰに対しては判別分析を，Ⅱについては重回帰分析を用いて，相関を分析する．

判別分析では，説明変数となる上流工程での欠陥数や修正量によって，設計欠陥の有無を判別したときの正答率から関連の有無を判定するとともに，判別式の係数からどの説明変数が設計欠陥の有無に強い影響を与えているか，評価することができる．重回帰分析では，説明変数となる上流工程での欠陥数や修正量による設計欠陥数を求める重回帰モデルを作成し，その決定係数（寄与率）でそのモデルでの設計欠陥数の予測精度を判定するとともに，重回帰式の係数からどの説明変数が設計欠陥数に強い影響を与えているか，評価することができる．

両分析とも，次の３つのパターンで行った．

- 基本設計・詳細設計の両方のデータを説明変数とし，基本設計・詳細設計の両方で混入した設計欠陥を目的変数とする
- 基本設計のみのデータを説明変数とし，基本設計で混入した設計欠陥を目的変数とする
- 詳細設計のみのデータを説明変数とし，詳細設計で混入した設計欠陥を目的変数とする

説明変数の一覧とその説明を表 2 に示す．

④⑤は，各工程後に行われた品質評価会議で各社から提出された品質評価シートから抽出した．⑪⑫のレビュー時間は，各文書に対して行われたレビューの総時間である．複数回行われた場合，その合計時間になる．⑬～⑱は，表 1 のメトリクスをＣＶＳの変更毎に計測し，最終バージョンまでの変更量の合計とした．⑲は，②の欠陥のうち，欠陥が混入した原因が次の 3 つのいずれかに当てはまる欠陥の数である．

✧ A．要求の確認不足
✧ B．設計条件の確認不足
✧ C．実現方式の検討不足

他の混入原因として，D．設計技術の習熟不足，E．業務知識の習熟不足，F．周知連絡の不徹底，G．表現上の配慮不足，H．修正ミスがある．これらは，問題記述票の記入者によって選択されるが，システムや設計内容の難易度

に依存する欠陥として，A，B，C．のみを重大原因欠陥とした．

表 2　説明変数一覧

	変数名	説明
①	更新回数	設計書の更新回数（ＣＶＳ上のバージョン数）
②	欠陥数	各文書に対してレビューで見つかった指摘の総数
③	重要欠陥数	上記の指摘のうち，重要度が高・中のものの総数
④	新規ページ数	新規作成ページ数（品質評価会議資料からのデータ）
⑤	流用ページ数	再利用ページ数（品質評価会議資料からのデータ）
⑥	全ページ数	新規ページ数＋流用ページ数
⑦	欠陥密度(/新規ページ数)	欠陥数/新規ページ数
⑧	欠陥密度(/全ページ数)	欠陥数/（新規ページ数＋流用ページ数）
⑨	重要欠陥密度(/新規ページ数)	重要欠陥数/新規ページ数
⑩	重要欠陥密度（/全ページ数）	重要欠陥数/（新規ページ数＋流用ページ数）
⑪	レビュー時間(/新規ページ数)	レビュー時間/新規ページ数
⑫	レビュー時間(/全ページ数)	レビュー時間/（新規ページ数＋流用ページ数）
⑬	重要変更箇所数	更新毎の重要な変更箇所数(表 1 の変更箇所数－非重要変更箇所数）の合計
⑭	変更文字数/100	更新毎の設計書の変更文字数（表 1）/100 の合計
⑮	重要追加箇所数	更新毎の重要な追加箇所数(表 1 の追加箇所数－非重要追加箇所数）の合計
⑯	追加文字数/100	更新毎の追加文字数（表 1）/100 の合計
⑰	重要削除箇所数	更新毎の重要な削除箇所数(表 1 の削除箇所数－非重要削除箇所数）の合計
⑱	削除文字数/100	更新毎の削除文字数（表 1）/100 の合計
⑲	重大原因欠陥数	設計内容に関して指摘された欠陥数

5.3　データ前処理

設計欠陥については,テスト工程で発見された欠陥データの「エラーを作りこんだ工程」(以降，混入工程)を参照することで抽出できる．今回は，混入工程が「基本設計」と「詳細設計」のものを分析対象とした．

また,下流工程での品質を使って評価するため，下流工程までのトレースが可能な文書のみを対象とした．テスト工程で発見された欠陥データには,「コンポーネント名」が入力されており，上流の文書との関連付けが可能である．いくつかの文書は，基本設計→詳細設計→下流工程と進行するにつれて，細分化され，複数の文書やコンポーネントとなるが，その場合，基本設計書を基準にして，

細分化された以降の詳細設計書データ，欠陥数は合計して扱う．また，仕様書は存在するが，プログラムコードが作成されていない場合，その仕様書は分析対象外とした．一方，他システムや前年度開発分からの流用コードで，基本設計や詳細設計が行われていないコンポーネントも，分析対象外とした．

上記のデータの前処理の結果，分析対象となる文書は 25 となった．

5.4　データ統計情報

表 3 に基本設計書，表 4 に詳細設計書それぞれの説明変数として収集されたデータについての統計情報を示す．

表 5 は，下流工程（単体試験以降）で発見された欠陥のうち，混入工程が基本設計および詳細設計工程の欠陥数の文書別一覧である．基本設計工程で混入した欠陥は 6 件，詳細設計で混入した欠陥は 10 件だった．発見された工程は，社内試験で 12 件，社間結合試験で 3 件，総合試験で 1 件という内訳だった．

なお，上流工程（設計）にはプログラム設計工程も含まれるが，本報告では分析対象として扱わない．プログラム設計工程で作成される文書は各社によって形式が大きく異なり，また図表として記述されることが多く，修正量を単純に文字数で計測することができないからである．

表 3　説明変数の統計情報（基本設計書）

		更新回数	欠陥数	重要欠陥数	新規ページ数	全ページ数	欠陥密度(/新規ページ数)	欠陥密度(/全ページ数)	重要欠陥密度(/新規ページ数)	重要欠陥密度(/全ページ数)
度数	有効	25	25	25	25	25	25	25	25	25
	欠損値	0	0	0	0	0	0	0	0	0
平均値		5.560	6.080	3.080	8.960	11.800	0.914	0.575	0.420	0.274
中央値		5.000	5.000	2.000	8.000	10.000	0.500	0.500	0.250	0.167
標準偏差		1.474	4.071	3.214	5.078	6.325	0.885	0.357	0.610	0.325
分散		2.173	16.577	10.327	25.790	40.000	0.783	0.128	0.372	0.105
最小値		4.000	1.000	0.000	2.000	5.000	0.188	0.100	0.000	0.000
最大値		11.000	19.000	11.000	26.000	33.000	3.250	1.300	2.500	1.143
パーセンタイル	25	5.000	3.000	0.000	6.000	7.000	0.392	0.288	0.000	0.000
	50	5.000	5.000	2.000	8.000	10.000	0.500	0.500	0.250	0.167
	75	6.000	7.500	4.500	10.500	13.500	1.171	0.857	0.400	0.400

		レビュー時間(/新規ページ数)	レビュー時間(/全ページ数)	重要変更箇所数	変更文字数/100	重要追加箇所数	追加文字数/100	重要削除箇所数	削除文字数/100	重大原因欠陥数
度数	有効	25	25	25	25	25	25	25	25	25
	欠損値	0	0	0	0	0	0	0	0	0
平均値		0.403	0.178	16.560	42.606	3.400	38.542	2.120	23.948	3.240
中央値		0.114	0.091	14.000	39.940	2.000	36.740	1.000	25.000	2.000
標準偏差		0.774	0.216	11.303	22.611	3.979	17.786	3.655	10.822	3.126
分散		0.600	0.046	127.757	511.258	15.833	316.330	13.360	117.122	9.773
最小値		0.016	0.015	2.000	9.810	0.000	11.530	0.000	4.590	0.000
最大値		2.500	0.737	46.000	97.030	19.000	75.530	14.000	49.120	13.000
パーセンタイル	25	0.073	0.058	8.000	25.395	1.000	25.140	0.000	15.395	1.000
	50	0.114	0.091	14.000	39.940	2.000	36.740	1.000	25.000	2.000
	75	0.214	0.190	21.500	57.590	4.500	47.060	2.000	29.365	4.500

表 4　説明変数の統計情報（詳細設計書）

		更新回数	欠陥数	重要欠陥数	新規ページ数	全ページ数	欠陥密度(/新規ページ数)	欠陥密度(/全ページ数)	重要欠陥密度(/新規ページ数)	重要欠陥密度(/全ページ数)
度数	有効	21	25	25	25	25	25	25	25	25
	欠損値	4	0	0	0	0	0	0	0	0
平均値		6.238	15.880	6.240	27.680	38.920	0.464	0.331	0.193	0.146
中央値		4.000	7.000	2.000	19.000	29.000	0.316	0.257	0.111	0.073
標準偏差		6.870	21.906	8.809	21.047	28.191	0.435	0.293	0.220	0.188
分散		47.190	479.860	77.607	442.977	794.743	0.189	0.086	0.049	0.035
最小値		2.000	0.000	0.000	6.000	7.000	0.000	0.000	0.000	0.000
最大値		32.000	103.000	32.000	84.000	121.000	1.800	1.108	0.914	0.744
パーセンタイル	25	2.000	2.500	0.000	9.500	0.183	0.183	0.096	0.000	0.000
	50	4.000	7.000	2.000	19.000	0.316	0.316	0.257	0.111	0.073
	75	6.500	22.000	10.500	40.500	0.633	0.633	0.494	0.310	0.254

		レビュー時間(/新規ページ数)	レビュー時間(/全ページ数)	重要変更箇所数	変更文字数/100	重要追加箇所数	追加文字数/100	重要削除箇所数	削除文字数/100	重大原因欠陥数
度数	有効	25	25	21	21	21	21	21	21	25
	欠損値	0	0	4	4	4	4	4	4	0
平均値		0.134	0.094	85.000	91.733	20.048	117.685	20.524	70.121	6.440
中央値		0.093	0.078	36.000	42.130	4.000	91.580	6.000	27.270	3.000
標準偏差		0.146	0.122	107.196	117.236	27.069	107.440	28.652	92.412	7.500
分散		0.021	0.015	11491.000	13744.251	732.748	11543.311	820.962	8539.888	56.257
最小値		0.000	0.000	0.000	0.000	0.000	1.480	0.000	0.000	0.000
最大値		0.720	0.651	403.000	470.220	96.000	455.320	105.000	364.230	28.000
パーセンタイル	25	0.067	0.042	4.000	3.635	1.500	43.125	1.000	2.195	1.000
	50	0.093	0.078	36.000	42.130	4.000	91.580	6.000	27.270	3.000
	75	0.162	0.096	166.000	165.355	33.500	172.695	35.500	130.330	12.000

表 5　文書別混入工程別設計欠陥数

文書 ID	基本設計欠陥数	詳細設計欠陥数
1	0	0
2	0	0
3	1	0
4	0	0
5	0	0
6	1	0
7	0	0
8	0	0
9	0	0
10	0	0
11	0	0
12	0	1
13	0	0
14	0	1
15	0	0
16	0	0
17	4	0
18	0	0
19	0	0
20	0	3
21	0	4
22	0	1
23	0	0
24	0	0
25	0	0

5.5　分析結果

表 6～表 9 に統計解析結果を示す．表 6 は，判別分析結果作成される判別式の係数の一覧である．ステップワイズ法[5]を用いたため，判別に役立つ説明変数のみが選択される．基本設計で混入した欠陥の有無には，追加文字数のみが選

5　ステップワイズ法では、説明変数を徐々に加えて最適な式を導き出す方法．説明変数が確定して無い場合に，予測精度が良くなるために必要な変数を，計算によって選択していく．

択された．同じく，詳細設計に関しては，3 つの説明変数が選択されているが，「標準化された正準判別関数係数」からもっとも影響の強い変数は“重要追加箇所数”であることがわかる．基本,詳細を合わせた場合，詳細設計工程の追加文字数のみが選択された． 表 7 は各判別分析の交差妥当化による正答表である．基本設計で混入された欠陥の有無についての正答率は，80.0%，詳細設計については 90.5%,両工程合わせた場合は 81.0%といずれも高い正答率を示している．

表 6　判別分析の結果（判別式係数）

目的変数	選択された説明変数	標準化された正準判別関数係数	正準判別関数係数
基本設計欠陥有無	追加文字数/100	1.000	0.062
	(定数)		-2.380
詳細設計欠陥有無	欠陥数の合計	-1.115	-0.049
	重要欠陥密度(/全ページ数)	1.099	6.217
	重要追加箇所数	1.527	0.071
	(定数)		-1.552
全設計欠陥有無	追加文字数/100(詳細設計書)	1.000	0.011
	(定数)		-1.282

表 7　判別分析正答表（交差妥当化済み）

		欠陥有無	予測ｸﾞﾙｰﾌﾟ番号		正答率(%)
			無	有	
基本設計	実績	無	18	4	80.0
		有	1	2	
詳細設計	実績	無	14	2	90.5
		有	0	5	
全設計	実績[6]	無	12	2	81.0
		有	2	5	

表 8 は，重回帰分析結果作成される重回帰式の係数の一覧である．ステップワイズ法を用いたため，欠陥数の予測に役立つ説明変数のみが選択される．基本設計で混入する欠陥数の予測には，重大原因欠陥数と重要欠陥密度(/全ページ数)が選択された．標準化係数と有意確率から重大原因欠陥数の影響がより強

6　詳細設計に関するデータが存在しないため（CVS 更新履歴がなし），分析が行われなかったケースが 4 件あるため，データ数が 21 件になっている（表 4 参照）．

いことがわかる．詳細設計に関しては，重要削除箇所数，重要変更箇所数，削除文字数，全ページあたりの欠陥数（欠陥密度），新規ページあたりの欠陥数（欠陥密度），新規ページあたりのレビュー時間，新規ページ数が選択されたが，標準化係数から，重要変更箇所数の影響が強いことがわかる．ただし，重要変更箇所数の係数は負なので，重要変更箇所数が多いほど設計欠陥数が少なくなる．一方，基本設計，詳細設計のデータを合わせた重回帰分析では，重要削除箇所数（詳細設計書），重要変更箇所数（詳細設計書），削除文字数（詳細設計書），重要削除箇所数（基本設計書），重要変更箇所数（基本設計書）が選択され，重要削除箇所数（詳細設計書），重要変更箇所数（詳細設計書），削除文字数（詳細設計書）の３つの影響が特に強いことがわかった．この場合も，重要変更箇所数（詳細設計書）の係数は負なので，重要変更箇所数が多いほど設計欠陥が少なくなる．また，表 9 から詳細設計欠陥数，全設計欠陥数の予測モデルは，予測精度が高い（R^2 ，調整済 R^2 が 0.85 以上）ことがわかる．

5.6　考察

設計工程で混入した欠陥の有無の判別については，特に追加文字数，重要追加箇所数で高精度の判別が可能であることがわかった．詳細設計では，欠陥数や重要欠陥密度も判別に貢献している．このことから，基本設計に関しては，追加情報量が下流での欠陥発生に大きく寄与していることがわかる．一方，詳細設計に関しては，重要追加箇所数の影響が最も大きい．欠陥数や重要欠陥密度も影響度は小さいが，影響することがわかった．ただし，欠陥数の係数が負であることから，欠陥数が多いほど下流工程での欠陥は無い方向にあるといえる．一方，重要欠陥密度の係数が正の値であることから，重要欠陥密度が高い文書では下流工程で欠陥が発生しやすいと言える．また，基本設計と詳細設計を合わせた分析では，詳細設計工程での追加文字数が下流工程での欠陥有無に影響していることがわかる．これらの結果から，対象プロジェクトにおける上流工程の品質は，主に設計の過程での追加量に依存していると考えられる．また，詳細設計では，上流工程で欠陥をより多く検出することで下流の品質が向上する，ということが言える一方，上流工程で重要な欠陥が多く検出される場合は下流の品質も悪い，という結果が出ている．

設計欠陥数の予測に関しては，基本設計の分析では，重大原因欠陥数の影響が大きい．設計内容にかかわる重大な欠陥が多く指摘されている場合，下流で基本設計に基づく欠陥の発生が多くなる，と考えられる．一方，詳細設計のみ，および基本・詳細設計を合わせた分析に関しては，重要変更箇所数，重要削除箇所数，削除文字数の順で下流の欠陥数に寄与していることがわかる．ただし，重要変更箇所数は多いほど下流での発見される設計欠陥数は少ない．つまり，

レビューなどで欠陥を洗い出し，きちんと修正を行われているケースでは，下流の品質が良くなる，と思われる．一方，削除量が設計欠陥数に影響を与えているのは，試行錯誤しながら設計が進んでいる，あるいは大幅な変更が設計中に発生している場合，十分なレビューを行うことが難しく，下流工程で発見される設計欠陥数を増大させる結果になる，と考えられる．

表 8　重回帰分析の結果（重回帰式係数と有意確率）

	選択された変数	非標準化係数	標準化係数	有意確率
		B	ベータ	
基本設計欠陥数	(定数)	-0.251		0.148
	重大原因欠陥数	0.302	1.136	0.000
	重要欠陥密度(/全ページ数)	-1.780	-0.696	0.004
詳細設計欠陥数	(定数)	-0.159		0.171
	重要削除箇所数	0.070	1.850	0.000
	重要変更箇所数	-0.038	-3.806	0.000
	削除文字数/100	0.031	2.663	0.000
	欠陥密度(/全ページ)	3.472	0.921	0.000
	欠陥密度(/新規ページ)	-1.681	-0.687	0.000
	レビュー時間(/新規ページ)	2.034	0.295	0.003
	新規ページ数	-0.015	-0.292	0.006
全設計欠陥数	(定数)	-0.034		0.841
	重要削除箇所数(詳細設計書)	0.067	1.783	0.000
	重要変更箇所数(詳細設計書)	-0.026	-2.578	0.000
	削除文字数/100(詳細設計書)	0.017	1.487	0.001
	重要削除箇所数(基本設計書)	-0.084	-0.245	0.013
	重要変更箇所数(基本設計書)	0.023	0.253	0.017

表 9　重回帰分析各モデルの寄与率・有意確率

	全設計バグ数	基本設計バグ数	詳細設計バグ数
R^2	0.903	0.563	0.973
調整済 R^2	0.871	0.523	0.958
有意確率	0.000	0.000	0.000

6. まとめと今後の課題

本稿では，上流工程での成果物およびレビュー実績などに基づき，各工程での完成度や下流工程での品質を評価する方法を検討した．具体的には，変更履歴や変更量を可視化することで，各工程のプロセス上の問題点を推定し，確認する方法を提案した．成果物の履歴管理方法や変更量の計測方法などについて，下流工程での成果物（プログラムコード）や欠陥データを分析する GQM モデル[3]を参考にした．計測結果の評価については，可視化された計測値を提示して，プロジェクト管理者や開発者に設計プロセスや完成度についてインタビューを行った．一方，下流工程での品質との相関を評価するため，基本設計・詳細設計で混入し下流工程(テスト工程)で発見された設計欠陥を抽出し，文書別に上流工程での計測データとの関連を分析した．設計欠陥発生有無についての判別分析では，追加文字数や重要追加箇所数などの影響が大きいことがわかった．一方，設計欠陥数の予測については，重回帰分析を行い，重要削除箇所数や削除文字数が多い場合下流での設計欠陥が多く，重要変更箇所数が多い場合は設計欠陥が少ないことがわかった．

適用事例では，5 社 6 拠点の異なる開発プロセスや開発体制を持つ組織がかかわっていることがあり，社別に大きな特徴が見られる計測データも存在した．設計書の登録・更新方法についてもガイドラインは設定したが，運用方法は各社で大きく異なり，結果のばらつきも大きい．従って，さらに多くの仕様書や設計書について調査し，計測方法やデータ収集のための運用ガイドラインの改善をする必要がある．

また，分析結果として可視化したグラフ上に，レビュー工数・レビュー時間，欠陥の分類（重要度，混入原因）などの情報が含まれていない．これらも，データの解釈上重要な情報なので，関連付けて参照することによって，より正確に仕様書・設計書の完成度を把握できると考えられる．

参考文献

[1] 大平 雅雄, 横森 励士, 阪井 誠, 岩村 聡, 小野 英治, 新海 平, 横川 智教, “ソフトウェア開発プロジェクトのリアルタイム管理を目的とした支援システム,” 電子情報通信学会論文誌 D-I, Vol.J88-D-I, No.2, pp228-239, 2005.

[2] 井上克郎, 松本健一, 飯田元, *ソフトウェアプロセス*, pp.112-117, 共立出版, 2000.

[3] A. Monden, T. Matsumura, M. Barker, K. Torii and V. Basili, “Customizing GQM Models for the EASE (Empirical Approach to Software Engineering) Project,” Nara Institute of Science and Technology Technical Report, 2007.

[4] A. Endres and D.Rombach, *A Handbook of Software and Systems Engineering, Empirical Observations, Laws and Theories*, Pearson Education Limited, UK, 2003.

[5] 松村 知子, 勝又 敏次, 森崎 修司, 玉田 春昭, 大杉 直樹, 門田 暁人, 楠本 真二, 松本 健一, “自動データ収集・可視化ツールを用いたリアルタイムフィードバックシステムの構築と試行,” 奈良先端科学技術大学院大学テクニカルレポート, NAIST-IS-2007001, Feb. 2007.