

ソフトウェアプロセスアセスメント手法の評価

堀田勝美 稲田 満 浜畑 諄

日本電信電話株式会社 ソフトウェア研究所

あらまし ソフトウェア開発プロジェクトの自己改善や外注先選定を目的に米国カーネギーメロン大学ソフトウェア工学研究所 (SEI) で開発されたソフトウェアプロセスアセスメント手法をNTTの実際の6プロジェクトに試行適用した。

NTTの組織や用語に合わせて、オリジナルの85項目からなる設問をブレイクダウンしてその曖昧性を除去するとともに、プロジェクトサイトでのインタビュー、証拠資料の確認などを行い、SEIのアセスメント手法に忠実な形態で実施した。本報告では、試行したアセスメントの実施方法、結果を述べ、アセスメント手法を考察する。

和文キーワード ソフトウェアプロセスアセスメント、ソフトウェア工学、ソフトウェア品質保証、ソフトウェア生産性、プロジェクト診断

Evaluation on Software Process Assessment

Katsumi HOTTA Mitsuru INADA and Yoshimi HAMABATA

NTT Software Laboratories
1-9-1 Kohnan Minato-ku Tokyo 108, Japan

Abstract

An experimental software process assessment was carried out on actual 6 software development projects within NTT. The assessment program, called SPA (Software Process Assessment), developed by Carnegie Mellon University's SEI (Software Engineering Institute) was selected for this experiment as it has widely used in US. Each of the 85 original criteria questions in the program was broken down into several questions to improve clarity of the original questions, and on-site hearing from the current development project leaders to put on the correctness of assessment result was held.

In this paper, the implementation of SEI's SPA, lessons learned from the experimental assessment and some topics of assessment program itself are discussed.

英文 key words software process assessment, software engineering, software quality assurance, software productivity, project evaluation

1. はじめに

NTTでは、社内業務用のシステムなど毎年多くのシステムの開発を行っており、ソフトウェアの開発規模も増大している。言うまでもなく、ソフトウェアの生産性と品質向上のための社内の委員会等を通じて、自己改善目標の設定や改善努力をしてきたが、方法論、開発ツール、設備、研修など分野が多岐にわたり、第三者の目で見ると真にバランスのよい目標設定ができていたのか、改善の達成度がどの程度（プロジェクト横並び、及び内外他社との比較）なのかが不明であり、客観的評価手段が必要であると考えていた。また、NTTは開発プロセスの多くの部分を外注業者に依頼しており、これらに対しても評価・指導できる客観的枠組みの必要性を感じていた。

このような背景から、内外のアセスメント手法を検討した結果、改善達成度の段階付けや改善の道しるべを与えていること、手法が公開されていることなどから、SEIのプロセスアセスメント手法（以下、SEI手法）^{[1][2][3]}を選定してNTTの開発プロジェクトに対するアセスメントを行い、NTTのプロジェクトの客観的評価とアセスメント手法自体の評価を試みることにした。

本報告では、SEI手法の概要、アセスメント実施方法、アセスメント結果を述べ、アセスメント手法を考察する。

2. SEI手法の概要

試行した診断手法は、ソフトウェア開発組織の自己改善や外注先選定を目的にSEIで開発されているもので、米国で適用実績のある手法である。診断用技術資料は1987年に初版が公開され、現在改定が行われている最中である。本手法の背景となっている概念モデルは、成功した他の産業分野の組織の生産プロセスの進化過程と同様の過程でソフトウェア開発組織のプロセスも成熟するという仮説にもとづいたものである。このモデルでは、初歩的な段階から究極の段階までの5段階の組織の成熟段階を定義し、それぞれの段階での組織の状態と要件を規定している。アセスメントは、被診断組織の開発プロセス（技術管理、組織構造など）の成熟度を調査するための85の質問をプロジェクト構成員に行い、開発能力の診断結果を5段階

のいずれの段階にあるかを示すとともに、次の段階に成長するための改善勧告を行う（図2-1）。改善勧告にもとづく改善を実施することによって組織が成長し、生産物の品質と生産性の向上が期待できる。

SEIと類似の標準としてISO標準（ISO 9000-3）^[8]がある。ISOの項目は、SEIの3レベルまでに相当すると考えられる。ISO、SEIとも診断項目が公開されており、外注先選定の公平な手段となり得る。SEI手法は、改善の過程とその目標も示されている点で、自己改善用としての適用性が高い。

SEIでは、公式なアセスメントができる機関として現在、米国内の9組織を指定している。

3. SEI手法の試行

試行評価手順を図3-1に示す。

3.1 評価基準のブレイクダウン

（1）SEIの評価基準（85項目）は、そのままでは曖昧性が高いため、ブレイクダウンが必要である。ブレイクダウンにあたっては、SEIの基本概念を崩さないことを原則として、評価基準を「親質問」としてそのまま無加工で採用した。ただし、被診断プロジェクトは「親質問」を直接回答の対象としない。

（2）各「親質問」を質問内容により2～10項目の詳細評価項目にブレイクダウンした。その総数は約400項目になり、これを「子質問」として、被診断プロジェクトに直接「YES」「NO」で回答してもらう。

ブレイクダウンにあたっては、プロジェクト側の理解を助けるために、NTTの組織構成、ソフトウェア開発用語などを可能な限り使用した。

（3）「親質問」と「子質問」を対応付けて配置し、質問票を作成した（図3-2）。

3.2 プロジェクトの選定

（1）アセスメントの試行対象プロジェクトを選定するにあたっては、組織評価も行なう目的から、SEI手法に則り、6つのプロジェクトを選定した。

（2）合わせてSEI手法を評価するため、性質の異なるプロジェクトを選定することとし、①分野 ②プロジェクトの規模、③工程時期、④外注の有無、などの違いに分けて選定した。

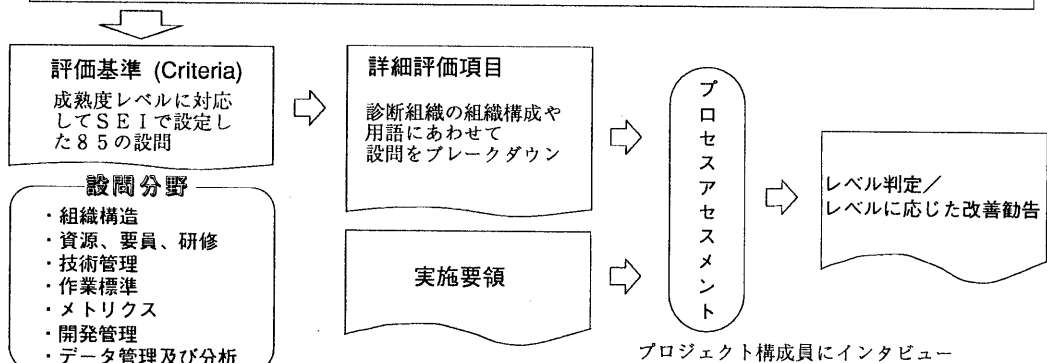
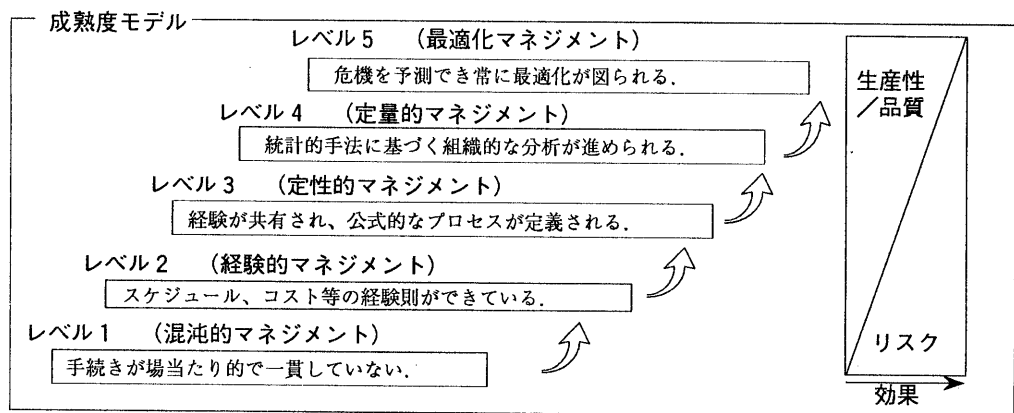


図2-1 SEIのプロセスアセスメント手法

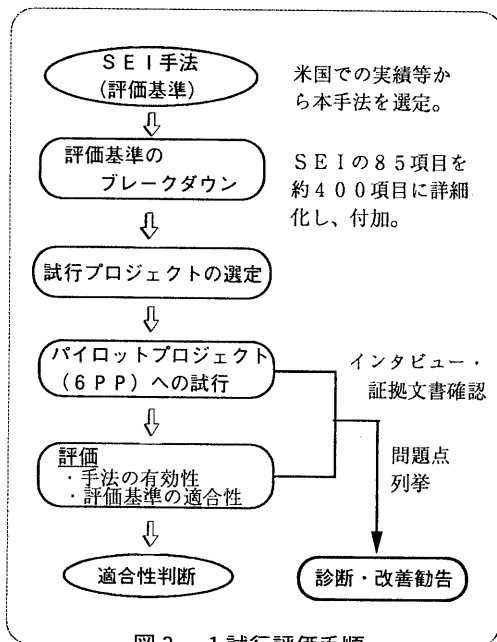


図3-1 試行評価手順

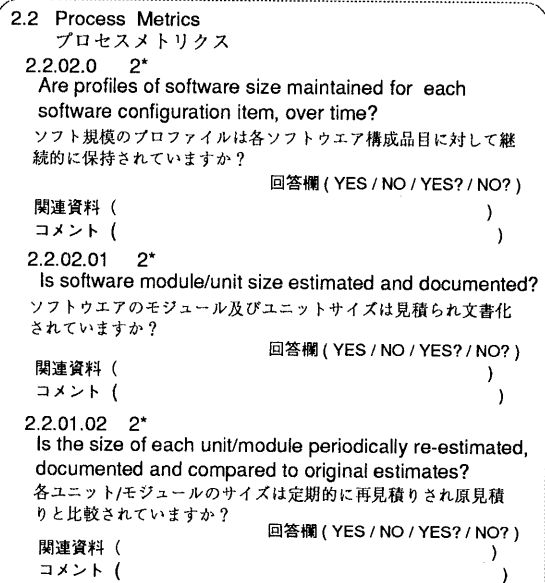


図3-2 評価基準のブレイクダウン例

なお、SEIの思想では、“問題のあるプロジェクトか、あるいは優秀なプロジェクトを選ぶうとしてはいけない”ということがあり、上記選定にあたってはこの点も考慮にいれて決定した。

3.3 アセスメントの試行

(1) 事前説明

アセスメントは、プロジェクトマネージャ、プロジェクトリーダー、およびスタッフ（上級技術者）へのプレゼンテーションからスタートし、アセスメントの原則やスケジュール、基本ルールについて説明した。事前の説明には1プロジェクト当たり1日を充てた。ここでは、あらゆる疑問点や関心事が示され、アセスメントに対する相互理解を図ることができた。

(2) インタビューの実施

アセスメント実施の上で、当該プロジェクトを正しく理解し、適切な結論を引きだし、意味のある作業勧告を行うためには、インタビューは必要不可欠なことである。ここでは、あらかじめ配布した質問票に対して、プロジェクトで回答を記入しておいてもらい、当日項目毎に回答の確認を行う。確認は、関連すると思われる証拠文書（オブジェクトエビデンス）を要求し、その内容を確認することにより行う。インタビューはアセスメントの客観性を確保するため、原則に則り、常時3名以上のアセッサで臨んだ。インタビューには多くの時間がかかり、質問の重点化を図ったものの、1プロジェクトあたり正味2日間を要した。

(3) 結果の集計（判定会議）

インタビューの結果から、アセッサは各質問に対する「YES」「NO」を判定するための会議を開き、納得のいく結論と改善勧告事項を導き出さねばならない。今回はアセッサ3名以上での並行評価を行ない、評価結果の合意を形成した。各「子質問」に対する判定を行った後、「親質問」に対する判定を行う。「親質問」の判定基準としては、厳格さを期すため、全ての「子質問」が「YES」の時のみ「親質問」が「YES」とした。以上の結果を、図3-3に示すレーティングアルゴリズムに照らし合わせてレベル判定を行なった。

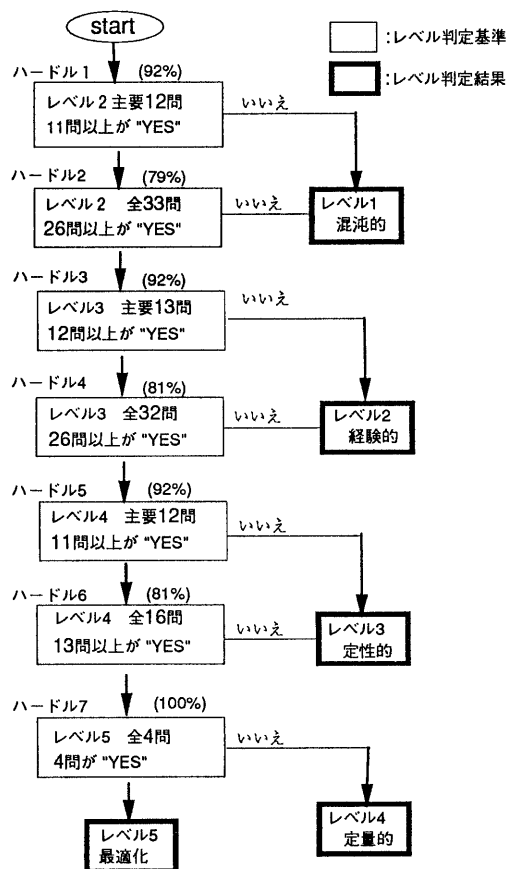


図3-3 SEIのレーティングアルゴリズム

(4) プリーフィングと結果の修正

判定の結果と明らかになったプロジェクトの問題点/課題等をまとめて勧告案を作成し、プロジェクトに対してプリーフィングを行う。この議論を通じて勧告内容が妥当であるかどうか判断する。もしプロジェクト側からみて納得がいかなければ、その点について十分議論をつくり、互いの合意を形成する。改善項目はそのプロジェクトに閉じてできるものもあれば、組織的に取り組まなければならない事項もある。そのため、勧告は幹部に対しても行われなければならない。インタビュー以降のこのような慎重な会議を経て得られた結論は最終的に公式な報告会（エグゼクティブサマリー）をもって終了した。

4. 試行結果

4. 1 試行アセスメント結果

4. 1. 1 判定レベル

今回の試行の結果、選定したプロジェクトはすべてレベル1であった（図4-1）。すなわち、SEIの定義によれば、場あたりのな対処を行っているという状況である。しかし、プロジェクトによってかなりバラツキがあり、レベル2に近いプロジェクトもあれば、半分程度しか満たしていないプロジェクトもある。概して大規模なプロジェクトほどポイントが高く、また、リーダーの資質に依存している傾向がある。また、レベル3の得点がレベル2と同等以上のプロジェクトもあり、成熟度の観点からすると必ずしもバランスのよい改善実施がされていないことがわかる。

試行診断を実施したプロジェクトのすべてが、

能力レベル1

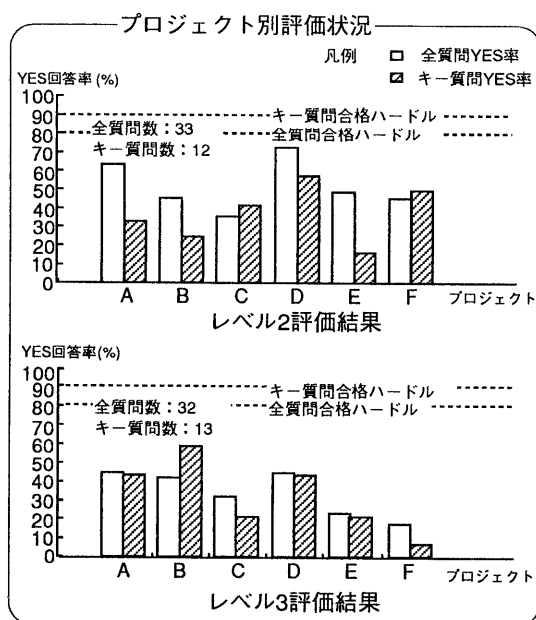


図4-1 試行アセスメント結果

4. 1. 2 改善勧告

改善勧告事項はプロジェクトのレベルによって若干の違いはあるが、概ね類似した内容のものになった。すなわちいずれも同様な弱点をか

かえている。この中には社の特質を越えて、日本的な習慣やものの考え方に根ざしたものもあると思われ、おそらく他社においても同様の問題を抱えているところも多いのではないだろうか。以下に主な改善勧告事項を述べる（表4-1）。

表4-1 主な改善勧告事項と対応するレベル

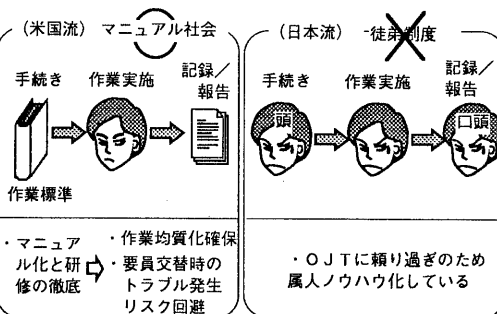
主な改善勧告事項	レベル	
	2	3
① ソフトウェアエンジニアリング活動の手続きの文書化が不十分	○	○
② 構成管理が不備	○	○
③ 責任分担体制が不明確（開発、テスト、構成管理、SQA）		○
④ ドキュメント化がモジュール単位で行われていないため、追跡可能性が低い		○

(1) 手続き等の文書化

開発のリスクを減少するためには、実施すべきことが明確であり、反復可能である必要がある。このためには、作業計画や作業方法が明確に記述され、結果についても記録されていることが必要である。これまで作業計画立案や見積りその他の作業方法はOJTを中心に育成が図られてきたため属人ノウハウ化しており、必ずしも文書化されていない、あるいは文書化されていても曖昧な場合が多いのが実情である。従って、記録についても同様なことが言える。

改善勧告事項(1)

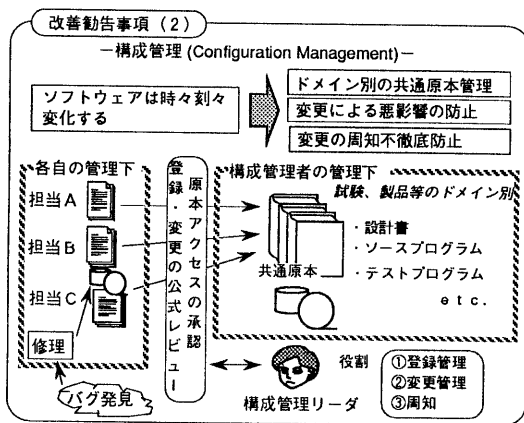
一文書と作業モデル



(2) 構成管理

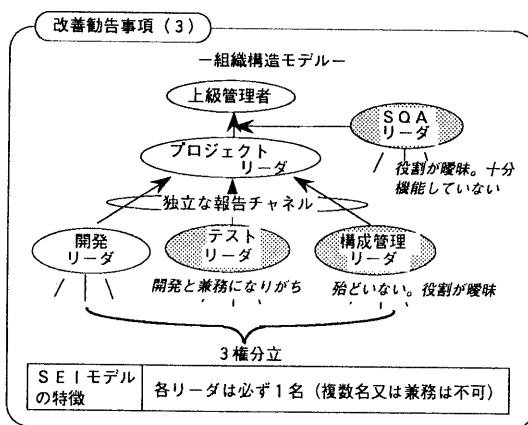
従来、変更管理としてプログラムや基本となるドキュメントについては比較的良好に管理され

てきた。しかし、この場合にもベースライン化や対象物の選定、公式レビュー等の基本となる手続きが不十分であったり、構成管理者の任命や役割が不明確な場合が多い。



(3) 責任分担体制

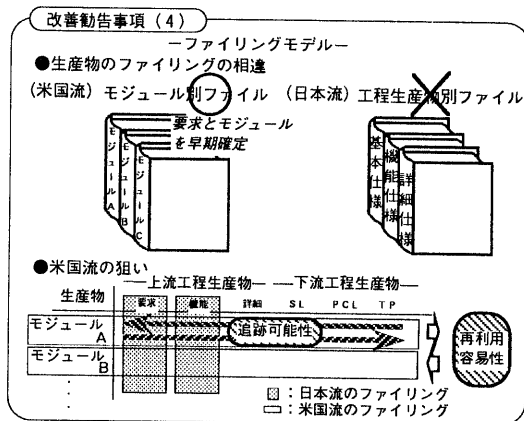
品質を保証する上では、品質保証チームを分けるだけでなく、テストチーム、構成管理チームを分離することが重要であるが、現状では、むしろ効率性を重視し、開発と兼務にしている場合が多い。



(4) ドキュメンテーション

要求条件と生産物の相互関係を明かにし、両方向の追跡可能性を高めることにより、ソフトウェアの保守性や再利用性が向上する。そのためには要求条件と生産物を構成する各要素の追跡マトリクスが必要であるが、さらに、ドキュメントをモジュール単位に整理するのが合理的

である。しかるに、これまでの習慣では工程対応にドキュメントをファイリングすることが例外なく一般化している。



4.2 SEI-SPA手法の評価

SEI手法の問題点については、いくつかの議論^{[5][6]}があるが、ここではアセスメントの試行を通じた考察を述べる。

4.2.1 評価基準内容の評価

設問分野としては効率化、高品質化のために必要なソフト開発に関連するプロセスをほぼ網羅しているが、保守に関する項目がない。また、マネジメント面が主体であり、手法やツール、すなわち技術レベルに関する評価項目は参考扱いになっている。

当初、マネジメント面では欧米文化との違いで、日本の文化になじまない項目があることが懸念されたが、いずれも合理的であり、特に不適当であるというものはない。

また、SEIの評価基準は抽象的であるため、そのままではアンケートに対して正しい回答が得られない。従ってアセッサはその内容を詳細に理解して回答者に説明するか、あるいは今回の試行のように、質問をブレイクダウンする必要がある。

4.2.2 アセスメント手順

基準の満足度の判定が難しく、また、アセッサの主観が混入しやすい。このため、経験豊富な複数のアセッサによる合議制を導入して客観

性を確保している。

また、インタビューでは回答するプロジェクト構成員が質問の意図を十分に理解し、回答結果について合意する必要がある、またオブジェクトエビデンスにより検証する等アセスメント自体に多くの稼働を要する。

4. 2. 3 レーティングアルゴリズム

現在の成熟度レベルの設定とレーティングアルゴリズムでは、レベル間の幅がありすぎ、ほとんどのプロジェクトがレベル1に落ちてしまう。これは米国での調査結果でも同様である。^[7] 自己改善用としては、レベル自体は改善の優先順位を見極める以外にあまり意味を持たないが、外注先選定用としては問題があり、レベル1をさらに細分化して評価する工夫が必要である。

また、開発分担の関係である質問内容が直接自分で行っていない場合（例えばシステム検討など）や、自動化ツールの導入等によりある作業が不用な場合、その質問自体が意味を持たないことがある（適用不可質問）。この場合、「Y

の場合）採点結果のバランスを欠くことになるので注意が必要である。外注において一部の工程を発注するような場合、本手法は網羅的であるがため、外注先評価法としては工夫の余地がある。

4. 2. 4 総合評価

（1）自己改善用としての評価

4. 1. 2で述べた通り、抽出された改善勧告事項は、いずれも的を得ており、有効である。また、レベルに対応した改善の優先順位が得られることも評価できる。しかし、これらの指摘事項は改善途上では逆に稼働が増えるのではないかという恐れを抱かせる。歴史が浅いため、これによりどの程度改善効果が得られるかを定量的に示す実例が少なく、経験を積み上げていく必要がある^[4]。また、アセスメント基準であるがために、どうあらねばならないか

（WHAT）はわかるが、具体的な改善方法が示されていないので、どうすればよいか

（HOW）を自分で考えるしかなく、対応してHOWを示した実施要領を整備するのが効果的である。

表 4 - 2 S E I 手 法 の 評 価

評価項目		特長	ウィークポイント
ク ラ イ テ リ ア	1. 適合性	①現状とのギャップはあるが、特に不適当な基準はない。	_____
	2. 有効性	①5段階成熟度モデルによりプロジェクトの弱点と改善ステップ（優先順位）が明確になる。 ②従来実証済みのソフトウェアエンジニアリング項目を、製造業等で確立された成熟度モデルに基いて体系化しており、生産性向上、リスク軽減効果が期待できる。*	①歴史が浅いため、改善効果を示す実例が少ない。 ②具体的な改善方法は示されない。
	3. 十分性	①ソフトウェア開発に関連するプロセスはほぼ網羅されている。 ②組織レベル、プロジェクトレベルを網羅している。	①保守に関する基準が不十分。
	4. 客観性	①特定組織／N T T文化色無し ②3人以上のアセスの合意による判定により主観を排除	①わかりにくい表現が多い。 ②基準の満足度の判定がむずかしいためアセッサの教育が必要。
ア セ ス メ ン ト 手 法	5. 正確性／公平性	①運用の工夫により確保 ・親質問項目の曖昧性を子質問へのブレイクダウンで回避。 ・教育された第三者ライセンスによる診断 ・インタビューと証拠資料による確認 ・アセスによる異議申し立て	①適用不可質問はYesの扱いとなり、公平性を欠く恐れがある。 ②レベル付けの方法が開発分担等、個々の事情に関係なく画一的である。 ③正確なアセスメントには相応の稼働が必要なため、体制作りが必要である。
	6. 厳格性	①質問を一般質問と主要質問に階層化 ②レベル合格に高いYes率を要求 ③子質問が全てYesの時、親質問がYes	①ハードルが高いため、ほとんどの組織がレベル1である。
総合評価		1. 自己改善用として有効であるが、具体的な改善支援や自己アセスメントを可能にしていくためには、開発実態に即した具体的な改善方法や判断基準を示す実施要領の確立が不可欠である。 2. 外注先選定評価用として、有力な候補である。ただし採用の場合は、手法のウィークポイント（特にアセスメント手法）への対策が必要である。	

*: C r o s b y の品質成熟度モデル

また、アセスメントに相応の稼働を要するが、これはやむを得ないことである。

(2) 外注先選定用としての評価

自己改善プログラムと結び付けて、相互運用することによりインセンティブ向上が図られ、業界のレベル向上に役立つ。しかし、ハードルが高すぎてほとんどレベル1になってしまうこと、適用不可質問の扱いによる不平等、等レーティング上改善の余地がある。また、外注先選定用として用いるためには第三者による客観的な評価が必要であるとともに、専門知識を備えたアセッサの育成が必要であり、日本における専門機関の設置が望まれる。さらに、ISO 9000-3^[8]との整合が望まれる。以上の評価結果を整理し、表4-2に示す。

5. あとがき

NTTでのSEIアセスメント手法の試行適用結果について述べた。今回の試行アセスメントは、SEIのアセスメントライセンスによる正式な受診ではないが、SEIの正式な受診時にとられているアセスメント体制、アセスメント手順にできるだけ忠実な形で実施した。

今回の試行アセスメントで、自己の開発プロセスの客観的な状況と弱点を認識でき、手法の有効性が確認された。開発プロジェクト側に勧告した共通的な要改善項目としては、開発とテストの独立性、構成管理、文書化されたプロセス手続き、追跡性を高めるためのモジュール単位の文書管理などであった。改善勧告を受けた各開発組織は、現在、作業標準の改定とこれに関連する研修プログラムの充実、SEIというSEPG（ソフトウェアエンジニアリング推進体制）の充実の三つの柱で自己改善に着手している。

外注業者への適用性については、改善の余地はあるものの業界の発展のためには必要と考えられ、国内に十分なアセスメントが行える第三者機関がないことが実際の運用では困ることであり、アセスメント機関の設立が重要な課題となる。

[参考文献]

- [1] Watts S. Humphrey, "Managing the Software Process", Software Engineering Institute, Addison-Wesley, 1989
「ソフトウェアプロセス成熟度の改善」、日本電気ソフトウェアプロセス研究会訳、日科技連、1991
- [2] M C. paulk, B Curtis, M B. Chrissis, et al, "Capability Maturity Model for Software", CMU/SEI-TR-24, August, 1991.
- [3] C V. Weber, M C. paulk, C J. Wise, J V. Withey, et al, "Key practices of the Capability Maturity Model", CMU/SEI-TR-25, August, 1991.
- [4] Watts Humphrey, Terry Snyder, Ronald Wills, "Software Process Improvement of Hughes Aircraft", IEEE Software, July, 1991
- [5] Bollinger, Clement McGown "A Critical Look at Software Capability Evaluations", IEEE Software, July, 1991
- [6] Watts Humphrey, Bill Curtis, "Comments on 'Critical Look'", IEEE Software, July, 1991
- [7] Humphrey, W.S., Kiston, D.H., Gale, J., "A Comparison of U.S. and Japanese Software Process Maturity", Software Engineering Institute, the proceeding of the 13th ICSE
- [8] "Guidelines for the application of ISO 9001 to the development, supply and maintenance of software", ISO 9000-3:1991(E)