

## 80

# GSN を活用した技術者能力計測手法の提案<sup>1</sup>

## 1. 概要

本編では、技術者の思考過程を GSN<sup>[3]</sup> (Goal Structuring Notation) 形式で可視化及び数値化することで、技術者の思考過程に依存する能力 (リスクを抽出する能力等) を短時間で計測できるか試みた結果を報告する。

### 1.1. 背景

JAXA では、ロケット、人工衛星、国際宇宙ステーション等の宇宙機の研究や開発を行っている。これらのシステムの特徴は、システムの動作環境が過酷であり、基本的に部品交換等の修理ができない点である。そのため、設計段階で不具合の混入や検証段階で不具合の見逃しを防ぐことが重要であり、品質向上に向けた様々な活動をしている。客観的に品質を担保する取り組みの一つとして、「技術、組織、資金」の独立性を確保した第三者の技術者によるソフトウェア検証及び妥当性確認の活動[1,2] (IV&V 活動: Independent-Verification and Validation) を行っている。

IV&V 活動は、従来の開発担当によるレビューやテスト等の V&V 活動に加えて、ソフトウェア開発成果物 (設計書やソースコード等) に対し、開発担当と異なる技術者が開発と並行して評価を行いフィードバックする活動である。不具合の要因となる問題点の修正提案や重大な問題点の有無をステークホルダーへ報告することも含まれている。そのため、IV&V 活動の重要な要素には、ソフトウェア開発や利用に関連する「リスク」を抽出していくことである。(図 80-1)

IV&V 活動に従事する技術者 (IV&V 技術者) は、ドメイン知識、ソフトウェア設計や検証技術の知識に加えて、ソフトウェアが要因で不具合となるようなリスクを抽出する能力が求められている。しかし、リスクを抽出する能力がある技術者の適性は、過去の作業結果や同じ業務の経験者の意見から把握するに留まっていた。そのため、IV&V 活動に適性がある技術者かどうか判断をするために、実際に業務を経験する必要があるが、他に事例が少ない IV&V 活動ではその適正を事前に把握できない課題があった。

---

<sup>1</sup> 事例提供：国立研究開発法人宇宙航空研究開発機構 (JAXA) 梅田 浩貴 氏

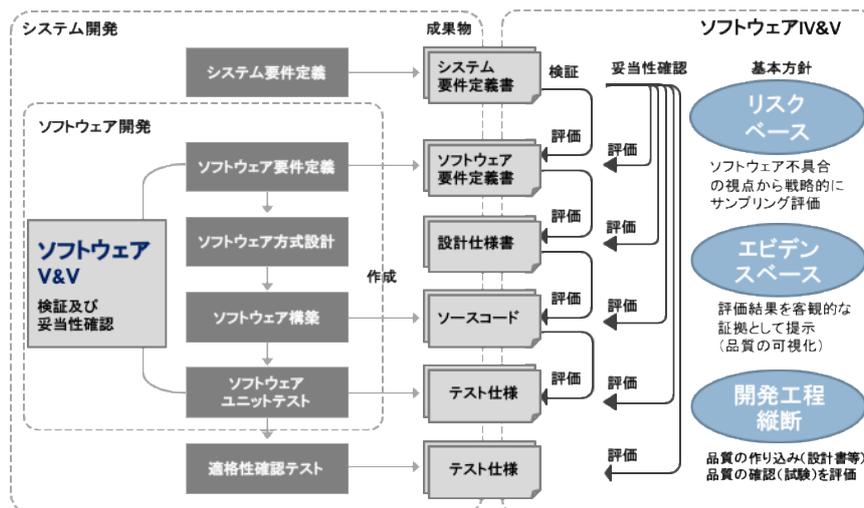


図 80-1 ソフトウェア開発や利用に関連する「リスク」を抽出の流れ

## 1.2. 能力計測手法としての特徴比較

短時間で技術者の能力を計測する方法としては、通常、選択式試験、記述式試験、小論文試験等がある。これらの方法では知識や経験の有無は計測できるが、どのような視点や思考過程によって結論を導いたのか、を計測することは困難である。一方、提案する GSN 形式の試験では、そのモデルの特性から思考過程を可視化できるため、思考力に依存する能力を計測できる。IV&V 活動では、システムやソフトウェアの特性に合わせて分析する能力が重要であるため、本方法を用いて IV&V 活動に適性のある技術者かどうか判断できる。提案手法と従来手法の違いを表 80-1 と表 80-2 に示す。

表 80-1 技術者能力計測の提案手法と従来手法の違い (その 1)

種別	計測手法	回答方式	計測可能な対象	試験運用の容易性
提案手法	GSN 形式試験	GSN 形式の成果物	<ul style="list-style-type: none"> <li>論理的な思考力</li> <li>設問に対する解釈</li> </ul>	<ul style="list-style-type: none"> <li>GSN モデリングルールや特徴を活かした採点が可能</li> </ul>
従来手法	小論文式試験	制限されていない文字数の文章	<ul style="list-style-type: none"> <li>文章を書く能力</li> <li>設問に対する解釈</li> </ul>	<ul style="list-style-type: none"> <li>査読になるため、有識者の判定が必須。</li> </ul>
	選択/記述式試験	複数の選択肢から選択、制限された文字数	<ul style="list-style-type: none"> <li>設問にある知識の有無</li> </ul>	<ul style="list-style-type: none"> <li>機械的に採点が可能</li> </ul>

表 80-2 技術者能力計測の提案手法と従来手法の違い（その2）

No	計測手法	特徴	強み	弱み
1	選択式試験	<ul style="list-style-type: none"> <li>出題範囲を限定して、類似概念/用語の理解度を判定できる。</li> <li>「正解、誤り」による機械的な判定ができる。</li> <li>大規模な実施が可能で、能力判定の前提として利用可能で、応用範囲が広い。</li> </ul>	<ul style="list-style-type: none"> <li>概念や用語の違いを理解しているか計測できる。</li> <li>機械的に計測ができる。</li> <li>問題範囲や重み付けができる。</li> <li>出題プロセス（Exam ガイドなど）を策定することが可能で運用が容易。</li> <li>広く（大規模な人数）実施することが可能。</li> </ul>	<ul style="list-style-type: none"> <li>問題数に一定の量が必要。同一問題をストックする場合は、出題数の x 4~5 の問題数が必要（定期的メンテナンスを前提）。</li> <li>偶然正解する可能性が他より高い。</li> <li>正確な判定を行うためには、高い問題作成能力が問われる。</li> <li>出題プロセスの作成には高い能力が必要。</li> </ul>
2	記述式試験	<ul style="list-style-type: none"> <li>問題文に対し制限のある字数で回答。</li> <li>大設問、小設問の組み合わせが可能。</li> </ul>	<ul style="list-style-type: none"> <li>選択式より精度の高い能力判定が可能。</li> <li>設問の設定により測定範囲を特定（誘導）できる。</li> </ul>	<ul style="list-style-type: none"> <li>文章を書く能力が必要。</li> <li>複数の回答が考えられ、採点が複雑。</li> </ul>
3	小論文式試験	<ul style="list-style-type: none"> <li>問題文に対し、文字数制限が少ない。</li> </ul>	<ul style="list-style-type: none"> <li>設問に対して読解能力を判定できる。</li> <li>1、2 より更に精度の高い能力判定が可能。</li> <li>課題形成、分析、対策、予測を、論理的に思考すること、他者理解力を文章能力で優劣を測定できる。</li> </ul>	<ul style="list-style-type: none"> <li>文章を書く能力、構成する能力が必要。</li> <li>採点が査読となり、有識者による判定が必要。</li> <li>自由度が高く、設問の本質と異なる主旨の回答となる可能性がある。</li> </ul>
4	GSN 形式試験	<ul style="list-style-type: none"> <li>思考過程が論理的に表現できる。</li> <li>目的に対し、前提、根拠、視点の関係を構造化して表現できる。</li> </ul>	<ul style="list-style-type: none"> <li>問題文（対象）をどのように解釈したか計測できる。</li> <li>GSN の作成能力を判定できる。</li> </ul>	<ul style="list-style-type: none"> <li>知識の有無は計測できない。</li> <li>GSN の習得が必須。</li> </ul>

## 2. 前提知識の説明

本章では、提案手法の適用先である IV&V 活動と手法構成の核となる GSN に関する説明をする。

### 2.1. IV&V 活動とは

#### 2.1.1. IV&V 活動の概要

開発担当が設計レビューやテストとして行う、V&V (Verification and Validation) がある。V&V とは、要件定義、設計、製作などのプロセスが正しく行われていること、また各プロセスの成果物も正しく作られていることを、検証と妥当性確認という 2 つの視点から評価することである。<sup>[1]</sup>

JAXA では、検証とは、「各プロセスの開発中間成果物が、前のプロセスからの入力情報に照らし、正しく作られているかを確認すること」としている。そのため、Verification を「評価対象ソフトウェアのライフサイクルを通し、上流プロセスの成果物から下流プロセスの成果物への要求トレースが可能であり、それらが内容を含め整合していることを検証すること」と定義している。

また、妥当性確認とは、「各プロセスの開発成果物がユーザの期待するとおりに作られていることを確認すること」としている。そのため、Validation を「評価対象ソフトウェアが最上位要求であるミッション要求、安全要求、運用要求から求められる機能、性質及び品質を満足していることを確認すること」と定義している。

なお、妥当性確認は、開発が完了したソフトウェア最終製品に行うのではなく、ソフトウェア要件定義やソフトウェア方式設計といった、より上流プロセスでも行う必要がある。上流工程で埋め込まれた欠陥は、ソフトウェア最終製品に、より重大な品質問題を与えてしまうことや、より多大な修正工数がかかるためである。このため妥当性確認は、ソフトウェアライフサイクルの全般を通して、常にソフトウェア製品（中間成果物を含む）が最上位要求に合致しているかを評価することが重要である。

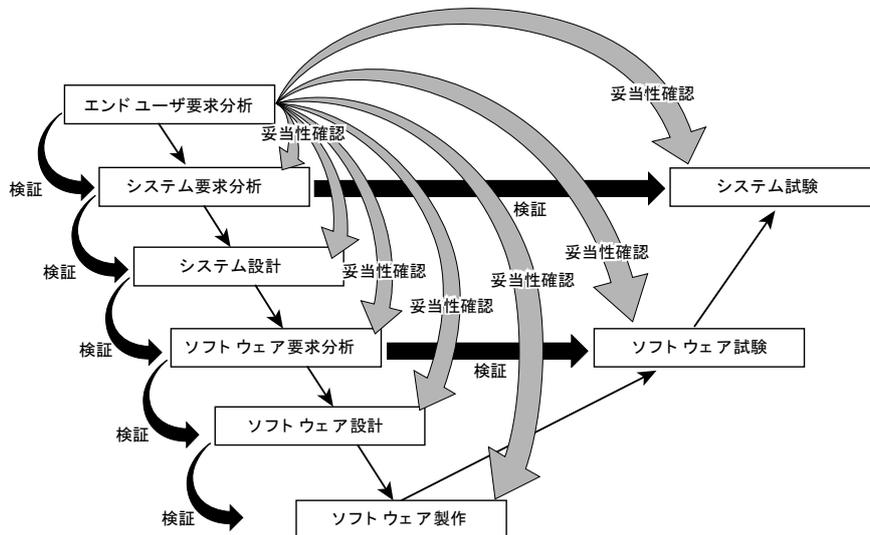


図 80-2 検証と妥当性確認(V&V)のイメージ

IV&V (Independent Verification and Validation) とは、「ソフトウェアの独立検証と妥当性確認」である。V&V を、独立 (Independent) した技術、組織、資金によって行うことである。ソフトウェア開発組織自らが行う V&V では、ソフトウェア発注者（開発プライム組織）の利益とソフトウェア開発組織の利益が一致しな

いことが多い。なぜならば、ソフトウェア開発組織自らが行う V&V では、開発活動と V&V 活動との間で、人的リソースや資金的リソースの奪い合いが起こり易いことや、設計者が根拠なく問題ないと思い込んだ場合、十分な V&V 活動の実施が妨げられる可能性があるからである。「技術、組織、資金」の3つの独立性のうち、組織的独立と資金的独立は、この利益相反による問題を防ぐために、特に重要な要素である。

なお一般的に、独立性を持った組織が実施する評価の意義としては、「品質に対する説得力の向上」、「問題が発生した場合の責任所在の明確化」といったことがある。

### 2.1.2. IV&V 活動の目的

JAXA IV&V 活動の目的は、IV&V 活動の発注者に対して、ソフトウェア開発組織が開発するソフトウェアの製品品質に関するセカンドオピニオンを提示することにより、システムが致命的な状況に陥る可能性等のリスクを低減させることである。このことから、JAXA IV&V における分析視点について重要な点は、「製品に含まれ得るリスクに対して、論理上あり得ることをなるべく網羅すること」である。

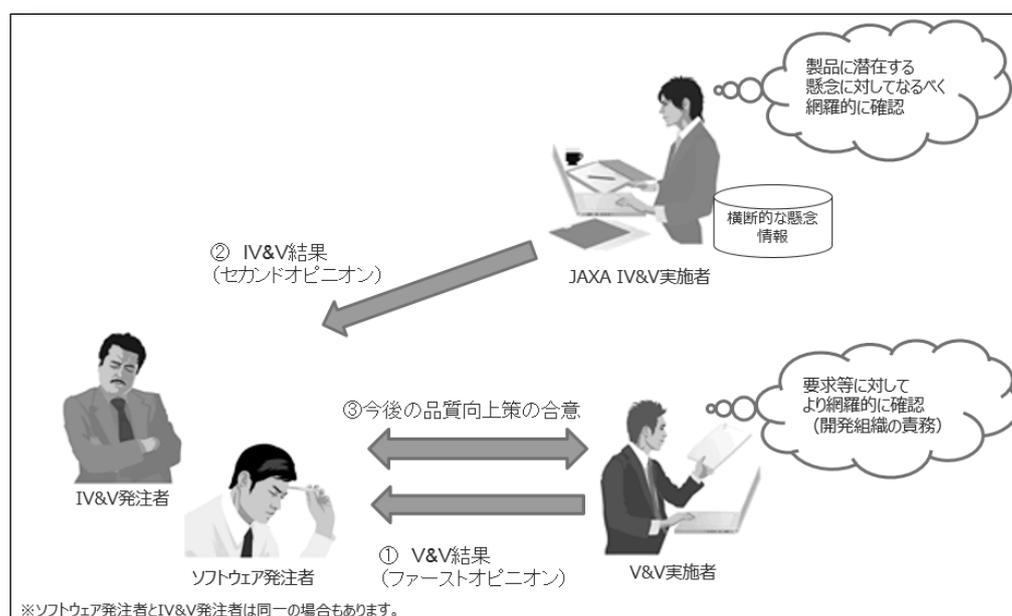


図 80-3 IV&V 仕事の基本構成

### 2.1.3. IV&V プロセス

IV&V プロセスは、IV&V ならではの価値を創出すること、IV&V 活動が拡大することに対応できること、技術者に取って実践可能であること、の3つの構築方針で作成された<sup>[2]</sup>。

1つ目の IV&V 活動の価値創出は、ソフトウェア発注者に安心感を与える説明責任、開発担当へ負荷を与えることなく低減可能なリスクとその対処の提案をするといった点である。2つ目の IV&V 活動の拡大は、多種多様な開発組織のソフトウェアを評価する必要がある、評価対象であるシステムやソフトウェアの特性に合わせて IV&V 活動の目的や評価方法をカスタマイズ又は新規構築していく点である。3つ目の「技術者が実践可能」とは、評価手法だけでなく、仕事としての企画、ステークホルダーへの説明や作業マネジメントを含め、技術者がどの位置づけで何をすべきか明確にする点である。

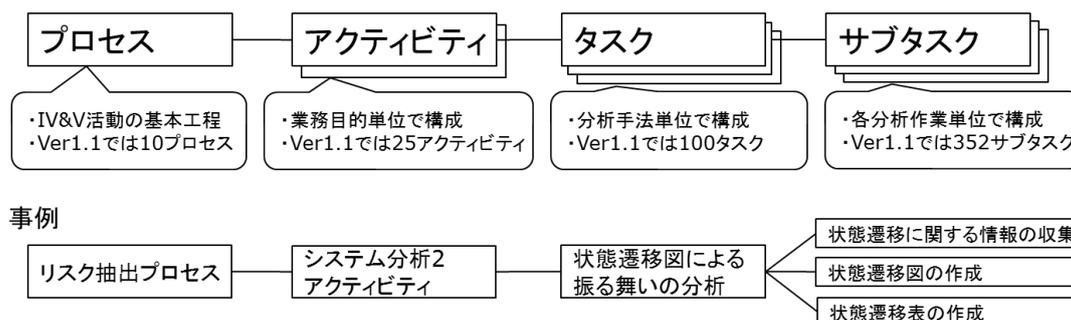


図 80-4 IV&V 活動の分析の経緯

前述を踏まえ、IV&V プロセスは、評価対象単位に管理する「プロセス」、業務や仕事として管理する単位の「アクティビティ」、分析作業や手法として管理する「タスク」の 3 層構造を設けている。

表 80-3 IV&V の 3 層構成の例

項目	プロセス	アクティビティ	タスク
説明	IV&V 活動の全体を構成する基本工程。 全部で 11 のプロセスがある。	プロセスの構成要素。 目的ごとに、一つのまとまった成果物を作成する作業群のこと。	アクティビティの構成要素。 作業者が成果物を作成するための具体的な作業のこと。
単体	—	同じ目的で実施する一連の作業の単位。	・ 中間成果物を作成する単位 ・ 入力情報を活用する単位 ・ 一担当で実施する作業の単位
テーラリング対象	× IV&V 活動では、全てのプロセスが必ず存在する。	○ IV&V 活動ごとに、取捨選択する。	○ 同左
用途	開発プロジェクトや IV&V マネージャーが、IV&V 活動の状況を把握すること。	IV&V リーダーが、各担当の作業進捗や成果物を、把握・確認すること。	各担当が、作成する成果物や実施する作業内容を把握すること。
上位構成	開発プロセスの工程 (要求、設計、実装、試験)	単一のプロセス	単一のアクティビティ

1 層目となるプロセスを構築するにあたり、IV&V 活動がもつ「独立」の特性を活かし、1 つの開発組織が持ち得ないソフトウェアに関連するリスク情報を横断的に活用できるようにするため、採用したプロセスモデルは PDCA サイクルとなっている。そのプロセスの中でも、IV&V 活動はリスクのある箇所を分析しサンプリングで評価を行うため、どんなリスクをどのように評価するのかを決める「計画立案プロセス」が最も重要になる。「計画立案プロセス」を強化するためには、不具合や検出した問題を蓄積し、「計画立案プロセス」で活用できる情報へ変換する「蓄積改善プロセス」が如何に機能するかがポイントとなる。

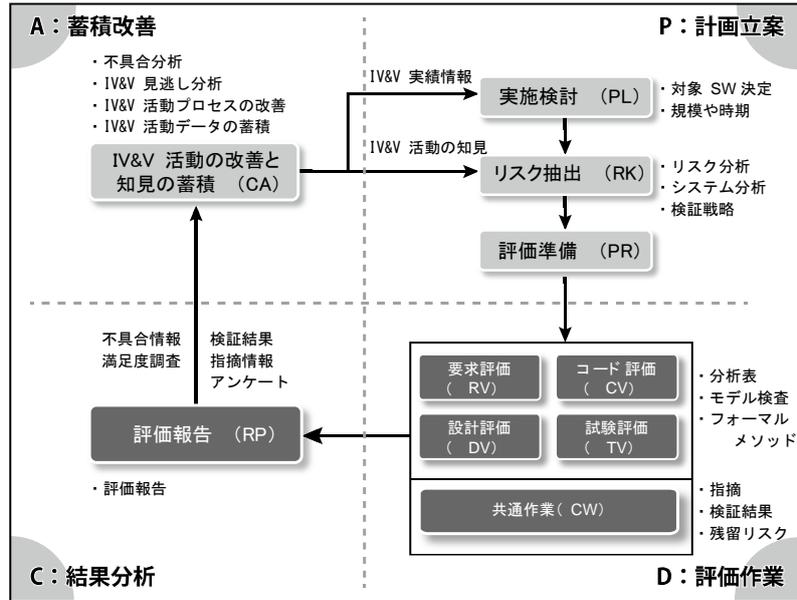


図 80-5 蓄積改善プロセスのサイクル

2層目となるアクティビティは、マネジメントと手法を統合的に扱えるよう IV&V 技術者が行う作業の位置づけを定義している。

IV&V 活動を開発の初期段階から行い、より上流工程で問題点を検出することで効果を上げることができる一方、開発後期で仕様変更が発生すると今までの IV&V 活動が無効となる場合がある。そのため、IV&V 活動を有効に機能させるためには、開発の進行状況に合わせたマネジメントが必須となる。

他にも、従来の V&V 活動を行っていた IV&V 技術者は、目の前に提示されたソフトウェア開発成果物に対し問題点がないかに注力しがちである。しかし、ステークホルダーの関心事項は、ソフトウェアが最終製品として利用や運用に悪影響を与えないかという点である。そのため、各技術者が捉えているリスク情報をより明文化し議論することが重要であるため、リスク抽出プロセスのアクティビティ数が強化されている。

一方、アクティビティが導入されたデメリットとして、熟練者からすると特定の手法をリスク抽出から評価作業まで一貫して適用していた場合、各アクティビティに合わせて中間成果物を作成しなければならないため、作業効率が低下する懸念がある。但し、過去の作業ファイルを再利用しやすくなることや成果物テンプレートの整備が促進することで作業効率の低下を補完することができる。

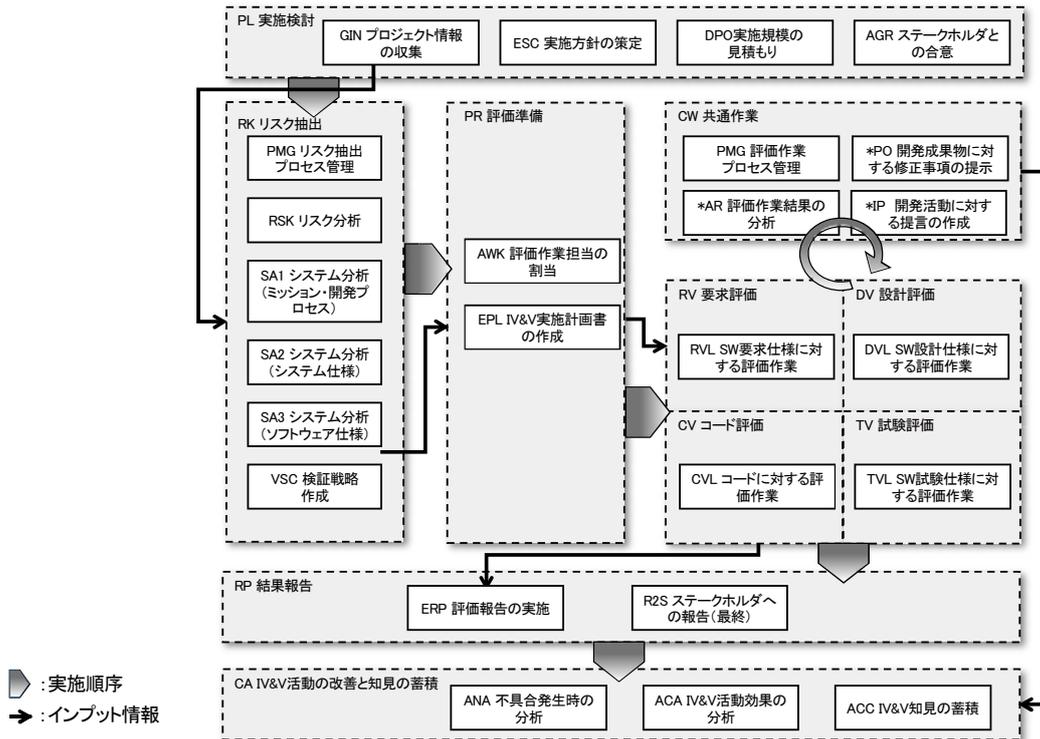


図 80-6 IV&V アクティビティ構成例

3層目となるタスクは、技術者の作業自由度を確保すること、過去の適用実績や他の技術者の具体的な分析結果を参照し手法の発展を促すこと、担当していない他のプロジェクトやソフトウェアの評価内容や業務及び作業を把握できる単位として定義している。従来、プロジェクト横断的に活用できるよう抽象的な評価観点として整備していたが、具体的な作業のノウハウを記録できる枠組みが必要となったため、分析手法の単位で定義した。よって、各タスクの名称や粒度は、その統一性よりも、タスクの修正や拡張のしやすさを重視した単位となっている。ここでいうタスクの修正や拡張の方法は、経験者が記憶に辿って作成することや、有識者が集まって議論によってその成功体験から作成することよりも、具体的な事実（作業時の失敗事例や該当成果物の受け手の発言等）と作業者がその場で記録した改善すべき考察を基に、「タスク」の修正や拡張を行うことを想定している。そのため、作業成果物に対し具体的な改善点（ステークホルダーからのコメント等）を記録できること必要があり、作業成果物が明確になるようタスクの下位層にサブタスクを定義した。

他にも IV&V 活動結果からより改善されるようタスクの運用方法として、成果物 ID の割り当て方針を下記のように設定した。

- ・プロジェクト横断的に分析事例を引き出せるようにした。タスク定義書は文章を中心として解説するため、そのノウハウを記録することに限界があることから事例による検索を推奨している。よって、IV&V 技術者が過去の成果物を容易に検索できるよう、作業成果物の ID を「プロセス、アクティビティ、タスク ID」を割り当てた。
- ・全体の IV&V 活動における作業成果物の一意性を確保するため、先頭に IV&V 活動の最小単位である「共通 ID」を導入した（共通 ID+プロセス ID+アクティビティ ID+タスク ID）。

- ・タスク ID がなく新規である分析方法の成果物は、成果物の識別子を特殊な ID となる方針とした（4 桁であるタスク ID の先頭番号が Z とする等）。

■アクティビティのタスク構成例

タスクID	タスク名
1000	コンポーネント図の分析
2000	機能関連図の分析
3000	FTAによる重要機能の分析
4000	状態遷移図による振る舞いの分析
5000	STPAによる非安全なシナリオの分析
6000	システム制約の抽出
7000	運用シナリオによる分析
続く	

+ タスク記述 (様式 Ver-140116)

項目	内容	
タスクID	RK-SA2-5000	
タスク名称	STPAによる非安全なシナリオの分析	
改訂日・Ver#	2014/8/18 改訂	Version 0.2
目的	IV&Vの対象SWと他コンポーネントの相互作用や協調動作に注目し、どのような非安全状態(リスク)に陥る可能性があるのか分析すること。	
参照/利用する技術リソース (IDと名称)	PEB-13094	IV&VにおけるSTPA適用手順書
入力情報 (文書IDと名称)	—	運用シナリオ、システム設計書、ユースケースなど、システムの利用方法が記載されている上位文書
	RK-SA2-1000	コンポーネント図
出力情報 (成果物IDと名称)	RK-SA2-5020	アクティビティ図
	RK-SA2-5030	コントロールストラクチャ
	RK-SA2-5040	制御操作一覧の作成
	RK-SA2-5050	非安全な制御操作一覧
完了基準	以下の成果物が作成されること。 ・アクティビティ図 ・制御関連図(コントロールストラクチャ) ・制御操作一覧 ・非安全な制御操作一覧	
特記事項	STPAの考え方や詳細な点は、Nancy G. Leveson 著 "Engineering a Safer World"を参照すること。	
サブタスク		
10	ID	RK-SA2-5010
	名称	分析範囲及びハザードの定義
	内容	■STPAの分析範囲を確認する。分析範囲を決める際には、RK-SA2-1000のコンポーネント図を参照し、対象SWの周りにどのようなコンポーネントが

図 80-7 アクティビティのタスク構成と記述例

2.1.4. IV&V 技術者の能力

IV&V 活動の業務特徴から人材に求められる能力の基本要件は「システム特性の分析により、IV&V の目的を明確にステークホルダーに示すことができ、その目的を達成するための手段を、十分な根拠に基づき戦略を立て実施することができる技術者」としている。仕事の基本構成(図 80-3)をもとに IV&V 活動の業務特徴を反映し、IV&V 技術者の能力要件を定義した。その分析した経緯を GSN (図 80-4) に示す。このように GSN を用いて可視化すると、その思考過程が明確になり抜けている内容の確認することや、前提である環境や方針が変更された場合にどの内容に影響あるのか明確になる。

表 80-4 IV&V 活動の業務の特徴

No	IV&V 業務の特徴
①	IV&V 活動は、開発企業が行うソフトウェア製品の製造や保証ではなくステークホルダーに安心感を与える説明責任を果たすこと。
②	IV&V 活動は、V&V に対する技術的な独立性確保や差別化をするため、ソフトウェアに関連するリスクに着目した検証活動を行うこと。
③	SW 開発成果物(開発文書やコード等)の内容、開発プロジェクトの状況やシステムの特性に応じた分析方法を採用すること。

④ IV&V 活動の対応による開発企業の負荷が増加し V&V の質低下とならないよう、第三者として客観的なエビデンスを取得ができる、プロダクトに対する評価を中心とすること。

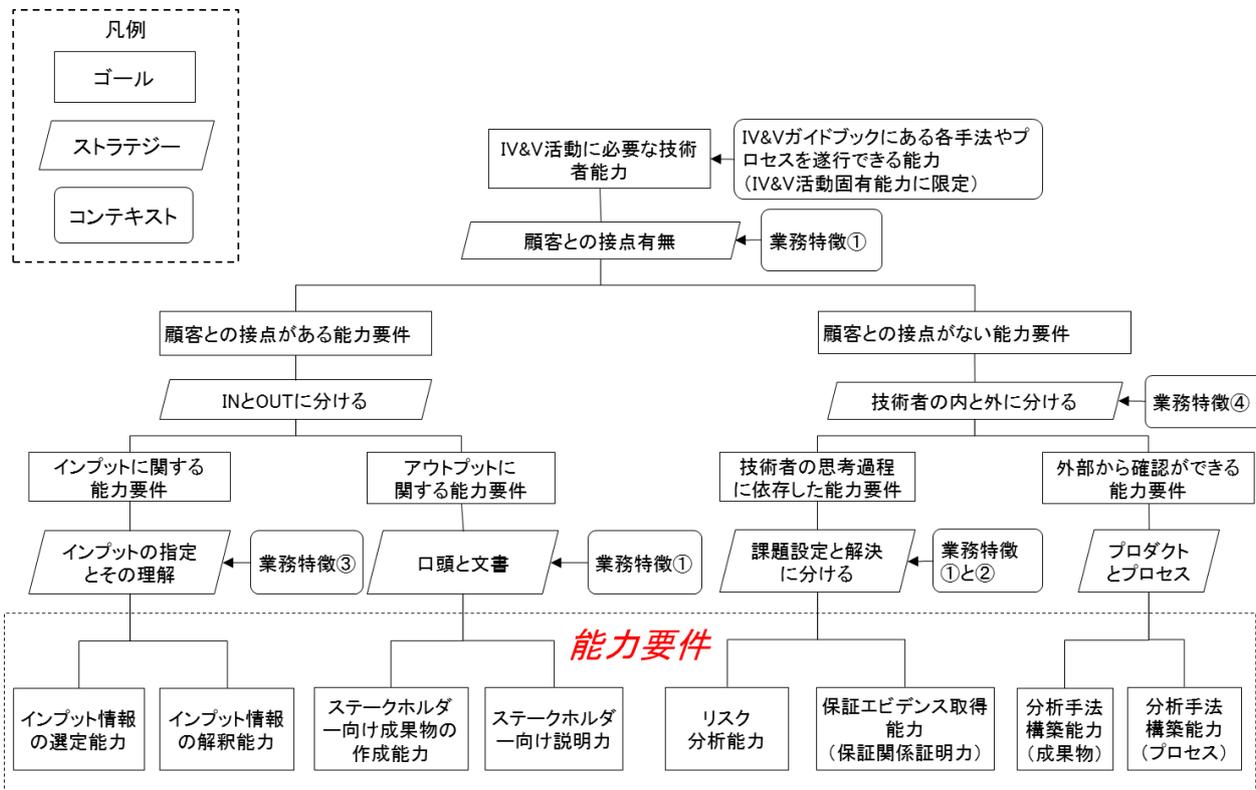


図 80- 8 IV&V 活動に必要な技術者能力 (GSN)

ドメイン知識やソフトウェア関連の知識は、他の手段でも習得できるため、IV&V 活動固有となる能力に特化して定義した能力要件を表 80-5 に示す。

表 80-5 IV&V 技術者の能力要件

能力要件	解説	参考
基礎知識	IV&V 活動をする上で必要な概念や用語等の知識を理解していること。	N/A
インプット選定能力	IV&V 活動に必要なインプットを指定する能力のこと。該当ソフトウェア製品の特徴に合わせて、他のシステムと連携する情報や過去の不具合情報等を選定する必要がある。	開発成果物によるリスク分析タスク
インプット解釈能力	提示されたインプットを解釈する能力のこと。該当ソフトウェアのドメイン知識に加えて、ソフトウェア開発や検証技術やプロセスの知識や不具合情報の知識も必要となる。	不具合及び IVV 指摘情報によるリスク分析
ステークホルダー向け	IV&V 活動の依頼元に対し、該当ソフトウェア製品の特徴やリスク、又は IV&V 活動の内容を説明する成果物を作成する能力のこと。IV&V 活動自身	評価作業プロセスにあるタスク

成果物作成能力	の「客観性」を確保するために、記録された作業結果に加え、ステークホルダーへ説明した成果物が重要である。	
ステークホルダー向け説明能力	IV&V 活動の依頼元に対し、該当ソフトウェア製品の特徴やリスク、又は IV&V 活動の内容を説明する成果物を説明する能力のこと。IV&V 活動の依頼元は、必ずしもソフトウェア関連知識に詳しいとは限らないため、相手の関心事項に合わせた説明をする必要がある。	開発活動に対する提言の作成タスク
リスク分析能力	該当ソフトウェア製品の「リスク」を抽出する能力のこと。なお、「リスク」とは、不具合要因となる「発生可能性」と、製品が稼働した際に不具合が発生し外部に与える「影響」のことである。	リスク分析アクティビティあるタスク
保証エビデンス取得能力	「リスク」が十分に低減されており問題がないことを保証するために、どのようなエビデンスを収集すればよいのか検討できる能力のこと。開発プロセスの V 字モデルとなっている場合、取得方法として設計レビューと試験ケースの有無に大別され、「リスク」の内容に合わせて「保証」のしやすさを検討することが必要である。	リスクに対する検証戦略の作成タスク
分析プロダクト構築能力	該当ソフトウェア製品の特徴やリスクに合わせて分析方法の品質を確保するため、その中間成果物や最終成果物を決める、又は作成する能力のこと。IV&V 活動の客観性を確保し、ステークホルダーに対する説明力を向上させること、IV&V 活動における分析者の知見共有として必要である。	テスト観点の有効性分析タスク
分析プロセス構築能力	開発成果物に記載されているソフトウェア製品の内容の情報収集や解釈した後、該当ソフトウェア製品の特徴やリスクに合わせて分析方法を決める能力のこと。重要箇所を選択して効果を最大化する IV&V 活動では、該当ソフトウェア製品の特徴やリスクに合わせて最も効果的且つ効率的な分析アプローチを構築できる必要がある。	ソフトウェア情報にもとづく変更点分析タスク

## 2.2. GSN とは

### 2.2.1. GSN の概要

GSN とは、自分の主張や考えを論理的なツリー構造で図式化する記法である。

GSN の特徴として、ゴール（目的）を分割しているロジックツリーに対し、ゴールの分割視点を表現した「ストラテジー」、ゴールの前提情報を明記した「コンテキスト」、ゴールを達成している事実情報を追加した「エビデンス」のノードを追加した記法である。アシュアランスケースの 1 つである D-CASE<sup>[4]</sup>としても記法が拡張されて活用されおり、顧客との合意形成<sup>[5]</sup>として使われている事例がある。近年、IoT (Internet of thing) 時代を迎え、多数の機器が接続された場合に、どのように品質を保証しているのか説明責任を果たす方法論としても注目を浴びている<sup>[6]</sup>。

## 2.2.2. IV&V 活動における GSN の活用

IV&V 活動は、開発企業とは異なる依頼元（経営層やプロジェクトマネージャー等のステークホルダー）が存在しており、限られた人的リソースで最大の効果を出すためにはステークホルダーと IV&V 活動の内容を合意することが重要である。そのため、IV&V 活動では「どんなリスクを評価するのか」と「該当リスクをどのように低減していくか又はしたのか」の 2 点について説明責任を果たす必要があり、「リスク導出経緯」と「検証戦略」という GSN を計画プロセス時に作成する成果物を定義している<sup>2)</sup>。

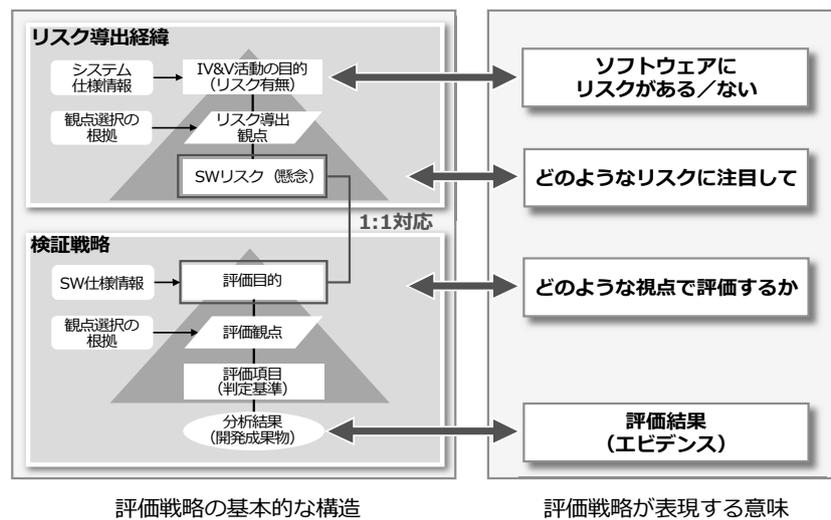


図 80-9 IV&V 評価戦略の構成

IV&V ガイドブック【導入編】 Ver.1.1 (JAXA 発行)より引用

## 3. GSN を活用した技術者能力計測手法

### 3.1. 手法の概説

提案手法は、技術者が GSN を作成する工程（試験部）、GSN を計測する工程（計測部）、計測結果を能力値に変換する工程（変換部）の 3 つの工程で構成される（図 80-10）。

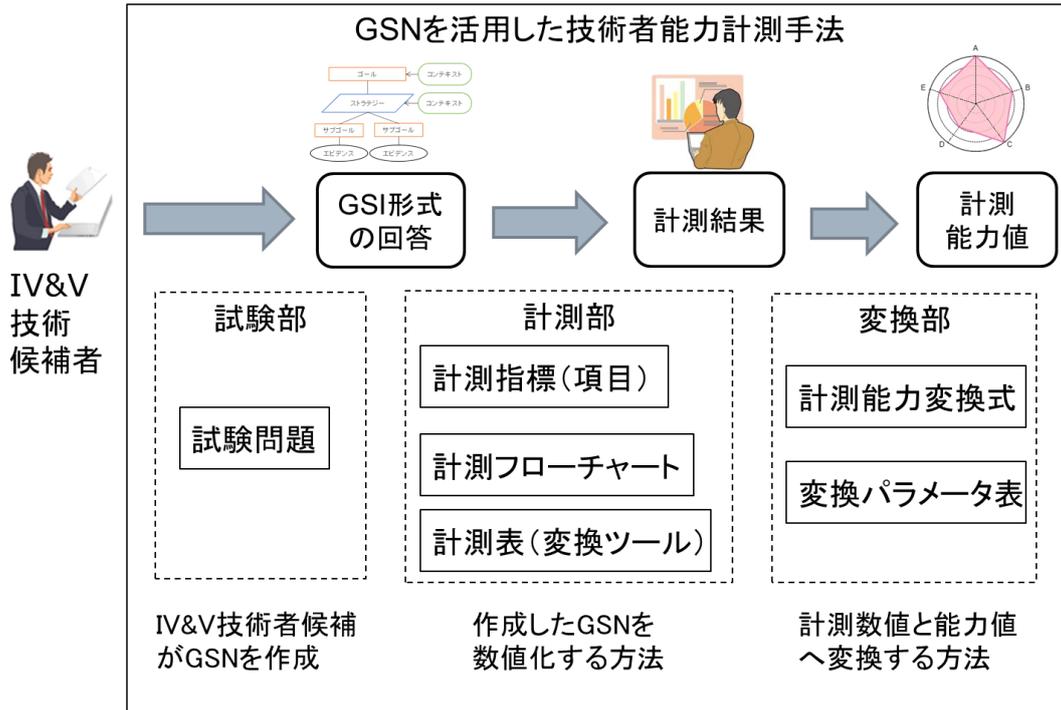


図 80-10 GSN を活用した技術者能力計測手法

試験部は、技術者が GSN を作成する工程である。本工程では、技術者にどのような試験問題を提示すれば、各技術者の能力差が判定可能、且つ低コストな方法を実現できるかが検討のポイントである。主な検討要素は、試験問題、試験時間、GSN 作成ツール、計測指標や項目の説明である。本手法では、ドメイン知識やソフトウェア設計及び検証知識の有無を計測することが目的ではないため、技術者のバックグラウンドに依存しないような試験問題が望ましい。また、試験時間やツールの習得も本業に影響を与えないよう配慮する必要がある。

計測部は、前工程で作成された GSN を一定の指標に基づき数値化する工程である。本工程では、作成された GSN が業務の目的や求めている価値にどれだけ合っているかを示す計測指標の定義と、その計測指標を数値化する水準を決める計測項目の定義がポイントである。また、GSN で表現された内容を読み取る際、計測結果に対する属人性をいかに排除していくかもポイントである。技術者は必ずしも GSN に習熟しているとは限らないため、GSN の各ノードにある情報を読み解かなければならない場合がある。また、ツリーモデルである GSN の構造によっては、他のノード情報との重複を考慮しながら計測しなければならない必要がある。そのため、GSN にある情報量を重複なく効率的に計測するため、計測結果を記録しやすい表形式にツールで変換して計測した。計測結果の属人性を排除するため、計測フローチャートを用意し、GSN の 1 ノードの情報をどのように読み取るのか規定した。その際、GSN 形式によって表現が一定の範囲に限定されているが、GSN ノード内の情報では自由度が高いため、計測対象が増えると情報の粒度や内容が多種多様に渡る。その結果、計測フローチャートの分岐が増えていくが、試験問題が限られた情報量であるため、一定の分岐数で計測フローチャートは固定されるため問題とならない。試験問題毎に、計測フローチャートを作成する方が具体的事例を条件分岐に反映できるため、計測者によって計測結果の幅がある場合は、試験問題毎に計測フローチャートを作成する。

変換部は、前工程の計測結果を能力値に変換する工程である。業務の目的や価値の視点から数値化された GSN の計測結果を能力要件の数値に変換する。通常、特定視点から計測した水準を満たす成果物を作成した

場合、複数の能力を駆使している。そのため、計測項目と能力要件は1対1とならない前提で、計測項目毎に能力要件の関与度合いを設定した。1つの能力計測手法で、該当業務で必要とする全ての能力要件を計測することはできないためことを前提に、各計測項目はどの能力要件を保有していると達成できるのか、その関与度合いから変換テーブルを作成した。能力要件、又は計測指標（項目）が変更されれば、変換テーブルは変更される。

### 3.2. 能力計測手法（試験部）

試験問題は、計測したい能力要件を使用する試験問題を作成する必要がある。そのため、試験問題の内容は、該当業務を模擬したテーマに設定することを推奨する。但し、計測対象者が該当の業務を未経験の場合、計測対象者が理解できるテーマに再設定する必要がある。計測したい主な能力要件は、技術者の思考過程に依存する「リスク分析能力」であるため、リスクを抽出した思考過程を GSN 形式で作成する試験問題とした。なお、実際の業務である IV&V プロセス<sup>2</sup>では、リスク導出経緯（RSK-VSC-2000）と該当し、数々の分析結果を基に作成する必要があるが、試験として提示する情報量や時間の制約から、各種分析は指定していない。

試験部を構築する際に、検討する項目は下記がある。

- ・試験問題は、技術者が理解できるか
- ・試験問題は、技術者によってその結果が異なるか
- ・試験問題は、GSN を作成できる情報量があるか
- ・試験時間は、業務への影響を許容できる負担であるか
- ・試験回答は、どのような形式（紙、電子データ）で行うのか
- ・試験回答のため、使用するツール（ソフトウェア、ホワイトボード、付箋等）は何とするか

各項目は、実際に試行しながら修正していく項目と、後工程で負担を減らすための項目、試験のしやすさを向上する項目の3つに分かれる。

<JAXA IV&V 試験部の実装例>

- ・試験問題：宇宙機ドメインの技術者であるため、普及しはじめだったウェアラブル端末とした  
リスクを検討する源泉として、複数の過去の不具合情報を設定した
- ・試験分量：10 ページ以下（試験概要、計測指標、GSN の事例含む）
- ・試験時間：2 時間（説明は別途 30 分）
- ・試験場所：試験監督がいる部屋
- ・試験回答：Visio 形式（ツールを提供）

開示情報

1. 製品開発への要求事項

新製品となるウェアラブル端末を開発した。前モデルは、初期開発もあって、市場からのクレーム(【解決すべき市場クレーム/不具合の一覧】)が寄せられている。中にはソフトウェアのアップデート(ランニングチェンジ)で対応したものもある。今回は、経営層からの強い姿勢で、市場クレームの激減と市場での製品評価を向上させることを求められている。将来的には、クラウドサービスに接続しライフログからさまざまな生活支援のサービスを提供することが予定されている。

クラウドサービス (将来)

ライフログ

スマートフォン

ウェアラブル端末

スマートフォン連携

エクササイズ

「脈拍計測」

- ◎運動強度と消費カロリー
- ・LEDと整合でリアルタイムにお知らせ(本体)
- 運動強度、脈拍、時間
- ・リアルタイムで脈拍数を表示(スマートフォン)
- ・一日の活動時間や歩数を表示(スマートフォン)
- ◎自動で睡眠検知
- ◁前提条件) 起床や体調により正しく検知できない場合がある

製品仕様

時計表示 有り(24時間表示)

外形 35mm×35mm×7mm

質量 35g

対象年齢 14歳～19歳

脈拍計測連続時間 36時間

待機時間 96時間

データ転送 スマートフォン

Android 4.2.1以上 iOS 7.0.1以上 Bluetooth 2.0以上

図 80-11 開示情報

製品仕様情報

2. 解決すべき市場クレーム/不具合の一覧

市場クレーム/不具合	事象	事象の原因	原因	発生要因
異常な発熱	使用中の発熱で停止状態になる。	熱センサーが異常発熱を感知、通知をシャットダウンした。	バッテリーのソフトウェア管理の不トラブル	バッテリー低下状態で充電すると発熱
仕様が厳しくなる利用時間	カタログの仕様では、96時間であるが、実稼働状態から消費利用で半時間で停止する。	バッテリー低下状態を感知して、シャットダウンした。	特定の処理状態になると、CPU負荷が異常に上昇するトラブル	バッテリーとの接続で発熱
正確な脈拍が取得できない	脈拍が異常に低い表示をする。	センサー部の圧着度が低いと脈拍検知ができない	・使用上の説明不足 ・センサー入力不正時の処理が不発	リストバンドが緩いと発熱
バッテリーが切れる	スマートフォンとの接続が原因で停止する。	バッテリーのプロトコルで、以前のバージョンがサポートできていない	・バッテリーの消費を抑えるための省電力モードのみを有効とした	利用条件の記載誤り

図 80-12 製品仕様情報

3.3. 能力計測手法 (計測部)

計測部は、計測したい能力要件を考慮し業務の価値から作成する「計測指標」、数値化するための「計測項目」、作業効率の向上とその過程を記録する「計測表」、計測に対する属人性を下げる「計測フローチャート」から構成される。

IN: ・GSN形式の回答  
・計測指標(項目) + 計測フローチャート

➔

OUT:  
・各計測項目の数値

**内容**

回答(GSN)

**計測指標**

懸念点、論理性、検証実現性分析(独自性、多角性)

計測指標	計測項目
(分析) 多角性	・評価観点や視点(戦略)の種類がどれだけあるか? ・仕様情報を仮説しているか?(分析視点が暗示)

**計測フローチャート**

GSNは自由度が高いためノード内情報が多岐に渡る → 計測値の客観性確保

**判定部**

**計測部**

自動抽出で漏れなく効率化

計測する前提を満たしているか

指標に対する水準を計測

技術者の回答結果		判定部			計測部		
GSNノードID	ストラテジー(リスク導出観点)	分析の視点が明確か	仕様情報を転記しているか	他に同じストラテジー	2項対立の場合	3要素以上で論理的MECEがある場合	仕様を抽象化している場合(反例の記入)
S1.1	ウェアラブル端末の内部・外部に分ける。	明確	独自	新規	内部と外部		

多角性の計測表(イメージ)

図 80-13 多角性の計測表 (イメージ)

(1) 計測指標

計測指標は、作成された GSN が業務の目的や求めている価値にどれだけ合っているか、という目的の

視点と、GSN の特徴を活かした視点の 2 点から作成できる。

IV&V 活動の目的は、リスクの提示やリスクの低減根拠を提示することであるため、リスクが抽出されているか、そのものを計測する「懸念点」とリスクの低減根拠となるエビデンスを収集できるかを計測する「検証実現性」を計測指標として定義した。また、IV&V 活動の特徴として、開発企業が見逃す可能性があるリスクを抽出しなければならないため、開発成果物に記載されていないことを独自に分析することや、より多くの視点から評価することが求められる。そのため、開発企業の分析視点や方法とは異なるかを計測する「分析の独自性」と評価する視点の種類が多数あるかを計測する「分析の多角性」を定義した。

GSN モデルの特徴として、前提（コンテキスト）と視点（ストラテジー）を記載することで上位ゴールに対する下位ゴールの網羅性を表現し、ゴール間の矛盾や抜けを確認することができる。その特徴を活かし、下位ゴールの抜けや矛盾を計測する「論理性」を定義した。その指標定義を下記表に示す。

表 80-6 計測指標の定義

計測指標	計測内容
(目的) 懸念点	対象の製品やソフトウェアの不具合要因（ソフトウェア内部、外部、利用時の品質等の低下）となる懸念点を提示できるか。
(分析) 独自性	開発活動内で行われる妥当性確認や検証活動と異なり、独自のアイデアで分析できるか。
(分析) 多角性	リスク導出や検証観点（戦略）がどれだけ多角的であるか。
(主張) 論理性	どれだけ抜け漏れがないか。
(証明可能性) 検証実現性	検証（証明）可能となる内容まで分解されているか。

計測指標を設定する際、その数や内容について下記の 3 点を考慮した。

- ① 計測指標の数は多ければ多い程、多数の数値を取得でき、能力要件の度合いを推定する精度が高くなる。一方、GSN にある各ノードの組合せ以上の計測指標になると、ノード内にある情報に対し判定することになるため、計測者の採点負荷が高くなることや、ノードパターンに着目して変換ツールによる作業効率化ができなくなる恐れがある。
- ② 計測指標の独立性を担保し、計測の網羅性を確保する方法もあるが、GSN のみで計測できる能力要件は限定的であると予想したため、業務価値に関与している、計測したい能力要件、且つ、思考過程に対し計測可能と考える計測指標を設定した。
- ③ 計測指標の内容が曖昧である場合、計測項目の作成やその水準を設定することが困難になり、計測者への依存度が高くなり、GSN による計測結果の精度が低くなる恐れがある。よって、1 つ以上のトップダウンのアプローチで確実に計測項目とその水準が設定できるか検討した上で計測指標を設定した。

## (2) 計測項目

計測項目とは、計測指標を数値化するため、水準判断がしやすい単位に分割した項目である。

計測項目を作成する際に、計測指標から導出するトップダウンアプローチと、複数の計測対象（GSN）に記載された情報を比較し、指標に対しその水準を決めるべきかどうか検討しながら決めるボトムアップ

アプローチがある。

トップダウンアプローチでは、設定した計測指標が GSN のモデリング上、どのノードに記載されるのか、又はどの複数のノードの組合せで計測できるか検討する。数値化できない計測項目となった場合、業務に対する GSN のモデリングルールを曖昧になっている場合がある。例えば、GSN 上のノードとして「コンテキスト」は「前提や根拠」となっているが、IV&V 活動における GSN ノードの「コンテキスト」は、ゴールに付随する場合は開発から提示された「仕様情報」、ストラテジーに付随する場合は IV&V 技術者が考えた「懸念情報」というように定義している。

ボトムアップアプローチでは、トップダウンアプローチで計測項目を 1 つ以上設定した上で、複数人の技術者が作成した GSN を試行しながら計測項目を作成する。その際、指標に対して水準が異なると判断した事例を収集し、計測の前提条件や計測項目を検討する。また、計測項目を多数増やすと計測作業の負荷が高くなってしまいうため、計測したい能力要件に関与するか検討した上で設定する必要がある。

表 80-7 計測指標の条件と項目

計測指標	計測条件 (前提条件)	計測項目
(目的) 懸念点	・懸念情報は、起きて欲しくない事象や状態であること	懸念情報は、一般的にありうる懸念情報であるか
	・懸念情報は、「対象」が明確であること ・懸念情報は、製品の特徴に合わせ具体性があること	懸念情報は、対象製品 (SW 直接関与なし) にあり得る懸念点であるか
		懸念情報は、対象製品 (SW 直接関与あり) にあり得る懸念点であるか
		提示された懸念情報は、製品の利用者へ影響があるか
(分析) 独自性	仕様情報をそのまま転記していないこと	図や表を作り懸念を探しているか
	システム分析を示している場合 (独自に図や表を作成しているか)	分析視点を明示しているか。(外部概念を活用している等)
		分析視点の網羅性を担保しているか
	仕様を仮説設定している場合	該当製品や SW にありえる仕様を仮説しているか
該当製品や SW にありえると言えない仕様を仮説しているか		
(分析) 多角性	仕様情報をそのまま転記していないこと	評価観点や視点 (戦略) の種類がどれだけあるか?
	仕様情報を抽象化している場合	仮説した仕様を提示しているか? (提示した仕様にはない要素があるか)
(主張) 論理性	ストラテジーとして成立していること。 (ストラテジーは、MECE となる視点を 用いていない場合等)	上位のゴールに対して、下位のゴールの抜けや重複が発生しているか?
(証明可能性) 検証実現性	ストラテジーとして成立していること。 (1 つのストラテジーに 1 つのサブゴールは成立していると言えない=ゴールを 具体化しただけ)	保証エビデンスを取得する (検証戦略) の作成範囲 (又はその内容) が明確か。

### (3) 計測表

計測表とは、GSN として表現されている試験結果に対し計測項目で採点した結果を記録しやすい表形式に変換したものである。計測表を作成する考え方は、下記の 2 通りある。

- ・案 1：計測表を計測指標単位で作成する場合  
※計測者からすると、計測項目を固定して、該当する試験回答の情報を採点していく。
- ・案 2：計測表を GSN ノードの抽出パターン毎に作成し、1 つの表で複数の計測指標を計測する場合  
※計測者からすると、試験回答の情報を固定して、該当する計測項目を採点していく。

計測指標を覚えて複数の情報を採点する作業となるため、案 1 の方が計測者にとって負担が低い。一方、案 1 の指標毎に計測表を作成する場合、GSN の各ノード情報を読みとりながら、どの計測指標に該当する情報か判断する必要があるため、GSN に対して網羅的に情報を抽出したかどうかの確認が困難になる。案 2 の場合は、モデリングルールからなる抽出パターンに沿って機械的に表へ変換するため、採点の抜け漏れの確認はしなくてよい。GSN のノード数が多ければ多い程、案 2 の方が優位となることや、計測経緯を記録することで後述の計測フローチャートの修正をすることを考慮し、案 2 で計測表を作成した。他にも、重複したノード情報（例：ストラテジー）を採点しなかった記録を残す方法としても、案 2 である GSN ノードの抽出パターンだと 1 箇所記録できるが、案 1 の計測指標毎だと複数の箇所に記録されることになってしまうデメリットがある（実際の試験結果では、GSN の 1 要素であるロジックツリーの要件を満たしていない場合、重複したストラテジーやゴールが多数発生した）。

表 80-8 計測単位

計測指標	計測単位	参考情報
(目的) 懸念点	全コンテキスト	
(分析) 独自性	全コンテキスト	
(分析) 多角性	ストラテジー	ストラテジー付随コンテキスト
(主張) 論理性	ストラテジー	ストラテジー付随コンテキスト
(証明可能性) 検証実現性	末端ゴール	ストラテジー

なお、計測方法はあくまでも GSN を参照しながら、その採点結果を計測表へ記入する方式である。周辺のノード情報を読み取って判定する必要があるため、計測表にある情報だけで判定することは困難である。

ツリー型モデルである GSN は、ロジックツリーとして MECE 且つ上位から下位にかけて抽象度の包含関係が整理されていれば問題ないが、試験解答は必ずしもそうになっていない。そのため、ノード数は多いが試験回答としての情報量は少ない、ということがあり得る。よって、重複した情報を排除することや、計測項目の前提条件を満たしていない場合、該当するノードを計測しないようにする「判定部」を計測表に導入した。計測表の構成単位と構成概念は図計測単位の検討経緯概要となる。

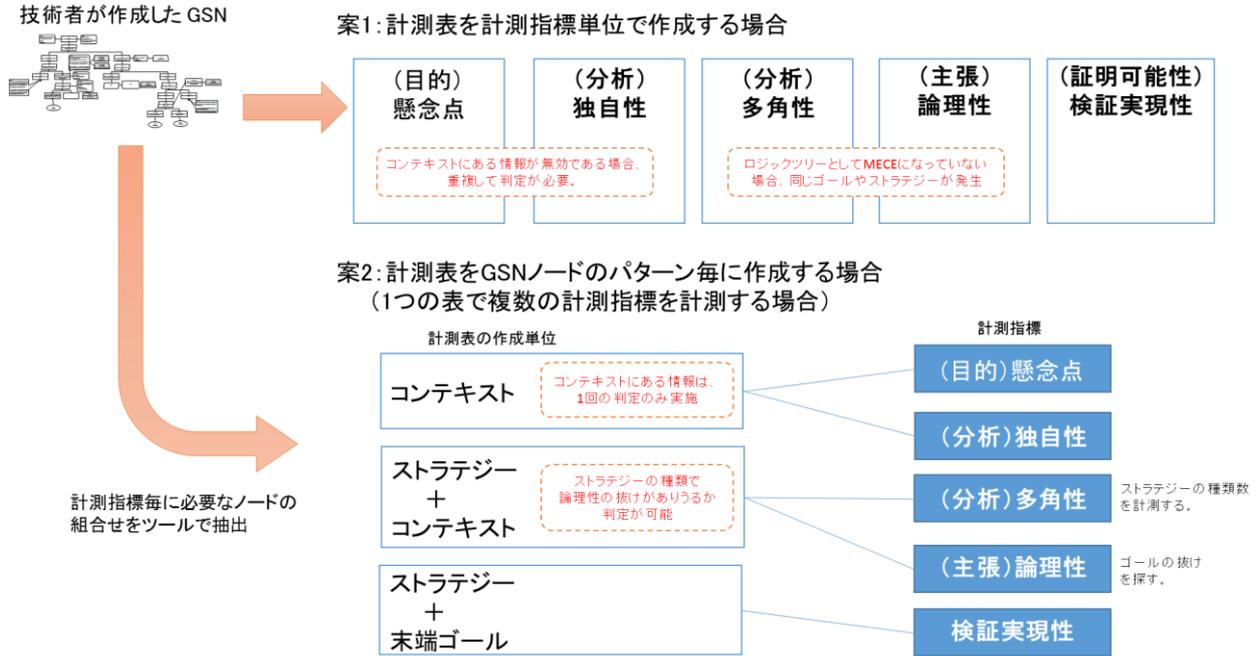


図 80-14 計測表の構成単位と構成概念は図計測単位の検討経緯概要

(4) 計測フローチャート

計測フローチャートとは、計測表にある計測単位 (GSN ノードパターン) の情報をどのように判定するのか、その判断基準をフローチャート化した図である。計測フローチャートは、判定における属人性を排除することを目的としているが、判定の実績が少ない場合、新たな計測項目やその前提条件が必要ないか検討する際に活用する。例えば、最終的にリスクとなる懸念情報を書いているが、試験問題で IV&V の対象はソフトウェアと指定しているにも関わらず、ソフトウェアと全く関係のないリスクが試験解答であった場合、題意に沿っているとは考えられず、計測項目を分割する等の修正が必要である。また、計測表の講師が GSN ノードパターンを優先したため、計測表にある判定部の結果によって以降の計測部の記入範囲が変わることもあり、計測者の負荷が高まる。計測者が採点に慣れていない場合は、計測フローチャートを参照しながら途中の分岐点の結果を判定部、最終的に辿りついた結果を計測部へ記入する採点方法とした。

計測フローチャートの弱点として、情報の有無や過去に計測された情報は採点の属人性が排除されると言えるが、「論理性」のように計測者が抜けている反例を定時する計測項目はカバーすることができない。よって、計測の難易度が高いと考える計測項目は、2人以上で採点する方法とした。例えば、論理性の抜けを採点する場合、GSN の記法としての特徴から、コンテキストが明確であることや特定のストラテジーであればゴールの抜けが発生しないが、独自のストラテジーを設定している箇所のみ 2人以上で採点する。なお、抜けが発生しない特定のストラテジーとは、対立概念 (例: 外と内、ON と OFF) や「設計や検証」における論理的な展開パターン (例: 異常対処は検知、分離、復帰) が該当する。

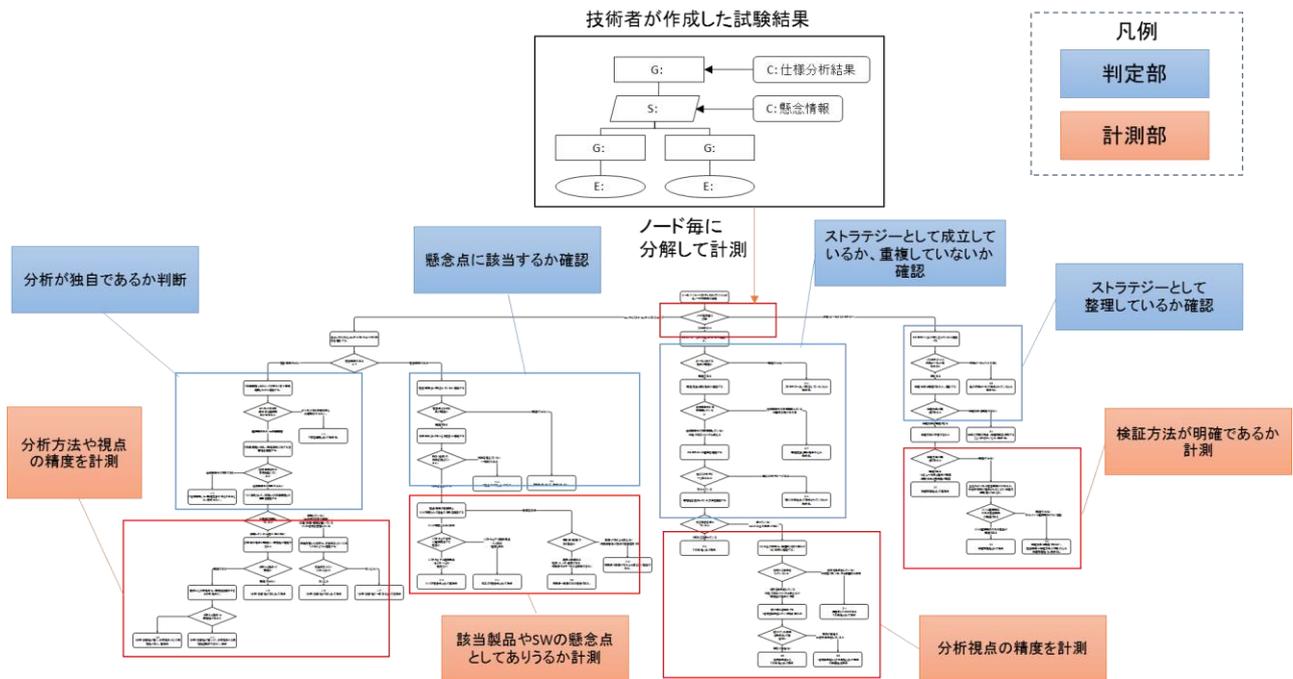


図 80-15 計測フローチャート

### 3.4. 能力計測手法（変換部）

変換部とは、計測部で数値化した結果を各能力要件の数値へ変換することである。変換した結果の数値を「計測能力値」とし、次の計算式で定義した。

$$\text{計測能力値} = \text{計測項目の数値} \times \text{計測指標間の傾斜} \times \text{能力要件への関与度合いの合計}$$

「能力要件への関与度合い」とは、該当する計測項目がどの能力要件があると推定されるのか、その関与度合いを3段階の水準（直接、間接的に関与する、関与しない）として設定し、計測能力値変換表を作成した。（表 80-9）本手法の構築当初にあった計測項目は14であったため、該当すると推定される能力要件は6つであったため、合計84パラメータに対し3段階の水準を設定した。

表 80-9 計測能力値変換表（抜粋）

計測指標	計測項目	能力要件			
		インプット 解釈力	ST 向け成果物 作成	リスク分析 能力	分析手法構築力 (プロセス)
独自性	モデル（図や表等）を独自に作成しているか。	直接関与 (2点)	直接関与 (2点)	間接関与 (1点)	間接関与 (1点)
多角性	リスク導出（検証）の視点は何種類あるか。	関与なし	関与なし	直接関与 (2点)	間接関与 (1点)

論理性	上位に対し下位のゴールの抜け(反例の提示)があるか?	関与なし	間接関与 (1点)	間接関与 (1点)	関与なし
-----	----------------------------	------	--------------	--------------	------

「計測指標間の傾斜」とは、1つの計測指標にある計測分類間の重み付けを設定できるよう導入したパラメータである。例えば、本パラメータは指標「分析独自性」における計測対象として GSN のコンテキストノードが該当する。その際、コンテキストに記載された内容が「抜けている仕様を文章で提示した(仕様仮説)」よりも「モデル(図や表の作成)を作成して網羅的に提示した」方がその説得力や効果が高く、高い数値と判定できるようにするためである。(例: 状態遷移のパターンを箇条書きにするより、状態遷移表や図で網羅性を確保する方が労力と効果が高い。)

①: 「計測能力値」とは  
回答(GSN)から推定される能力値

②: 同じ指標内にある計測項目の重み付け  
例: 分析独自性の指標内で、分析図を作成している場合は「2」、  
分析結果のみ文字で表現している場合は「1」と重み付け。

**計測能力値**

= (計測項目の数値 × 指標間の傾斜 × 能力への関与度) 各計測指標の合計

③: 計測項目が、どの能力要件を計測していることに該当するのか、  
その関与度合いを直接と間接の2段階を、84パラメータ(=6能力要件×14計測項目)に設定。

計測指標	計測項目	ST向け成果物作成	リスク分析能力
多角性	リスク導出(検証)の視点は何種類あるか。	関与なし	直接関与 (2点)
論理性	上位に対し下位のゴールの抜け(反例の提示)があるか?	間接関与 (1点)	間接関与 (1点)

図 計測能力値変換表(イメージ)

図 80-16 計測能力値変換表 (イメージ)

## 4. 提案手法の妥当性確認結果

### 4.1. 妥当性確認の方法

JAXA IV&V 活動では、過去の経験や実績によって技術者の能力を判断していた。そのため、各技術者が IV&V プロセスとして整備されている各分析手法をどれだけ習得しているかと等価である(2.2 項を参照)。IV&V 活動は、対象となるシステムやソフトウェアの特性に合わせて各種の分析手法を選択および修正して適用する必要があることから、技術者の実績や経験値を数値化は、IV&V 活動の従事年数ではなく、IV&V プロセスにある「サブタスク(分析手法の作業工程)」の習熟度合いとした。

「経験能力値」 = サブタスクの経験水準 × サブタスクの難易度 × 必要とする能力要件のその度合い

(1) IV&V 技術者の経験調査（サブタスクの経験水準）

各技術者が IV&V 活動で使う手法の習得度合いを調査するため、各サブタスクの経験水準（表 80-10）のいずれかに該当するの技術者にアンケートによる調査を行った。過去のサブタスクの成果物を分析して客観的に調査することも可能であるが、調査コストの負担が大きいことと、実際の IV&V 活動でどの技術者がどの程度の分析や結果を出しているか把握できているため、算出された経験能力値が技術者間の相対的な関係が妥当であるか大凡の判断ができると判断した。

表 80-10 サブタスクに対する経験水準

経験水準	水準の内容
レベル 0	該当サブタスクの実施方法を知らない
レベル 1	該当サブタスクの実施方法を知っている
レベル 2	該当サブタスクを他者の支援を受けて実施したことがある
レベル 3	該当サブタスクを独力で実施したことがある
レベル 4	該当サブタスクで成果（SW リスク、問題点、提言等）を出したことがある

(2) 経験能力値の作成

(1)項にある経験の調査結果を各能力要件の数値に変換する。経験能力値は、サブタスク経験水準（表 80-10）× 難易度（表 80-11）× 能力要件に対する必要度（表 80-11）の合計とした。「サブタスクの難易度」とは、各タスクにある分析手法や作業の難易度を反映するパラメータである。

表 80-11 能力要件に対する必要度

難易度	水準の内容	事例
水準 1	該当サブタスクは、1 回の経験で習得でき、知識や用語を理解していれば誰でも同じ水準の成果物作成や効果を発揮できる。	・テンプレートがある情報収集タスク
水準 2	該当サブタスクは、数回の経験で習得でき事例を参照し、経験さえあれば誰でも同じ水準の成果物や作業効果を発揮できる。	・テンプレートがある分析図の作成タスク
水準 3	該当サブタスクは、属人性が高く、経験回数に依存せず技術者によって成果物品質や作業効果に明らかな差が発生する。	・モデリング手法 ・ガイドワード分析

「能力要件に対する必要度」とは、サブタスクの作業をする上で必要となる能力要件の度合いを、3 段階のパラメータ（必須、推奨、無関係）で設定した（表 80-12）。その理由は、各サブタスクは一律的に同じ能力要件を必要とするのではなく、情報収集から分析モデルの作成等の多岐にわたっているためである。例えば状態遷移図や表を用いてリスク分析する場合、上級の技術者はどんな状態遷移やイベントが重要かを分析するだけでなく詳細設計やソースコードを参照し状態遷移中のイベント有無を同時に分析するが、初級の技術者はそこまで確認しない。

表 80-12 経験能力変換表（抜粋）

サブタスク ID /サブタスク名	インプット 選定能力	インプット 解釈能力	リスク 分析能力	分析手法構築 (プロダクト)	分析手法構築 (プロセス)
RK-SA2-4010 /状態遷移に関する情報収集	推奨 (1点)	必須 (2点)	無関係 (0点)	無関係 (0点)	無関係 (0点)
RK-SA2-4020 /状態遷移図の作成	無関係 (0点)	推奨 (1点)	必須 (2点)	推奨 (1点)	推奨 (1点)

## 4.2. 妥当性確認

### 4.2.1. 従来方法の結果

一般的な能力計測である「選択式試験」を実施した。本節では、その「選択式試験」の概要、試験実施方法、試験結果及び考察について記載する。

選択式試験とは、「問い」に対して「複数の選択肢」から正しいものを選択するという形式の試験である。選択式試験は、能力値を計測するために広く利用されている。例えば英語能力測定のための TOEIC などがそれに当たる。選択式試験が能力測定の方法として広く利用されていることから、その試験が IV&V 活動の能力値測定（経験量測定）に対しても有効であるかを検証する。

#### (1) 計測条件

今回、IV&V 活動の経験年数が異なる技術者 2 名（3 年以上の技術者 A と 1 年未満の技術者 B）に対して、選択式試験（95 問）を実施した。なお、選択式試験の品質を確保するため、以下を配慮して作成した。

#### <選択式試験作成の配慮事項>

- ・ 正解（正しい選択肢）は、引用元情報（IV&V ガイドブック）を基に作成した。
- ・ 不正解（誤った選択肢）は、以下を考慮し作成した。
  - 経験等に基づき誤りそうな作業内容や概念を考慮し、不正解を作成した。
  - 引用元情報の一部修正や他の引用元情報を追加し、不正解を作成した。

#### <選択式試験の事例>

表 80-13 選択式試験の事例

<p><b>【問題】</b> IV&amp;V 活動の特徴として、正しいものはどれか。次の選択肢の中から全て選びなさい。</p> <p><b>【選択肢】</b></p> <p>A. 技術面、組織面が独立していれば、IV&amp;V 活動としてみなすことができる</p> <p>B. IV&amp;V を実施すれば、V&amp;V は不要になる</p> <p>C. IV&amp;V は、ソフトウェア品質モデルにおける内部品質及びプロセス品質を評価する</p> <p>D. IV&amp;V はテスト工程だけでなく、上流工程から確認する</p>
---

## (2) 計測結果

選択式試験の実施結果を下表に示す。

表 80-14 選択式試験の正答率

受験者	正答率
中堅技術者 A	34%
若手技術者 B	36%

## (3) 考察

上表の測定結果より、選択式試験では IV&V 活動の経験量による優位な差は計測できていないことが分かる。受験者の正答率は、それぞれ中堅技術者では 34%、若手技術者では 36%であり、その差は 2 ポイント（正答数の差は 2 問程度）と小さい。測定結果が経験量に比例していないため、選択式試験は IV&V 活動に対する適性を計測する方法としては不十分である。

## 4.2.2. 提案手法の結果

### (1) 計測条件

IV&V 技術者 9 名が 2 時間の試験で GSN を作成し、各計測項目に基づき計測能力値を算出した。

なお、試験問題は、技術者のドメイン知識の豊富さによる影響を減らすため、宇宙機ではなく一般的な製品を取り上げた問題を採用した。また、IV&V 技術者 9 名のプロフィール調査結果から経験能力値を算出した。

### (2) 計測結果

計測能力値と経験能力値の相関を比較した結果を表 80-15 に示す。

表 80-15 経験能力値と計測能力値の相関

能力要件	相関値	判定
インプット解釈能力	0.632	相関あり
ST 向け成果物作成能力	0.244	低い相関あり
リスク分析能力	0.891	高い相関あり
保証エビデンス取得能力	0.823	高い相関あり
分析手法構築能力（プロダクト）	0.733	高い相関あり
分析手法構築（プロセス）	0.727	高い相関あり
各能力の合計値	0.937	高い相関あり

### (3) 考察

提案手法による計測結果は、経験値と 0.8 以上の相関を持っており、当初の目的であった思考過程が強く影響する能力「リスク分析能力」等を計測できている。但し、ステークホルダー向け成果物作成能力等の業務の相手先が関係のある能力は、GSN を用いた本手法では計測ができていると言えない。そのため、ステークホルダーへ説明する等の能力計測は、実際の業務形態に近いプレゼン等を入れた総合的な計測方法が必要と考えられる。また、ロジカルシンキングを習得していない技術者は、計測能力値でほとんど得

点が取れていないため、本手法を適用する条件としてロジカルシンキングの一定の習得が必要である。よって、計測結果から提案手法は、思考経緯に依存する能力を短時間で計測できる方法である。該当業務に従事してなく知識を保有していない新しい技術者にも能力を計測することができる。さらに、本手法の能力計測結果の精度を向上させるためには、下記の3つの施策が考えられる。

#### <適用実績による精度向上>

- ・新しく IV&V 活動に従事する技術者に対し本手法で計測し、その後の IV&V プロセス経験度合いを追跡調査し、各種パラメータを更新することで計測精度を向上することができる。

#### <参照情報の精度向上>

- ・「経験能力値への変換」の基礎となる「サブタスク」の特徴をパターン化し分析することで「経験能力値」の対応パラメータの精度を上げることができる。また、IV&V プロセスは常に更新されているため、サブタスクの追加や修正を反映して最新版の「経験能力値」を維持することが必要である。

#### <計測方法の改善>

- ・GSN 計測フローチャートを拡張し、「設問に合わせた計測パターン」や「ツールによる自動判定」を導入し、計測の効率化や計測の属人性を排除することができる。

### 4.3. 提案手法の応用

本項では、提案手法の応用先について検討した結果を述べる。

#### 4.3.1. 能力計測結果を基に体制構築や分析手法の選択

JAXA IV&V 活動では、通常、4～5名のチームを組んで、開発と並行してソフトウェアの評価を行う。多種多様なリスクを抽出するためにも、異なるバックグラウンドの技術者が議論しながら分析を進めていくことが望ましい。また、複数名で同じ範囲を分析する場合でも、思考タイプが異なる技術者同士で組合せる方がよい。他にも、特定の経験や視点に偏りがちな技術者がいた場合、自分の得意分野の分析方法を使って評価した方がよい。

- ・タイプ1：提示された情報を前提に詳細又は具体化することができる。
  - 役割分担の例：ソフトウェアへの要求に対し、設計や実装の漏れがないか評価
- ・タイプ2：提示された情報以外の情報を導出できる。
  - 役割分担の例：仕様外となりうる事象や暗黙的なソフトウェア設計の前提を評価
- ・タイプ3：特定の視点や範囲に特化している。
  - 役割分担の例：利用や運用における重要な状況からの評価
- ・タイプ4：外部概念を対象に合わせて応用している。
  - 役割分担の例：ブラックボックステスト視点となる要求に対する試験の十分性を評価
- ・タイプ5：論理的に曖昧箇所や矛盾箇所を導出できる。
  - 役割分担の例：複数条件が整理した場合、どのような振る舞いとなるのかタイミング設計を評価

### 4.3.2. 上級技術者への育成アプローチの検討

一定の経験を積んでいくと、その思考特性から技術者自身が得意となる視点の分析アプローチを採用するようになる。それに伴い、より業務の効率化や実績を重視した場合、その役割もますます固定化されていく。よって、一定の業務経験を積み特定の役割が継続している技術者は、自分の思考特性を把握し、意識的に分析アプローチを変えてみる必要がある。

- ・例1：求められる品質が異なるシステムやシステム視点から異なる役割のソフトウェアの評価を経験している技術者の場合、汎用化された懸念情報がストックされている。そのため、対象のシステムやソフトウェアの特徴を考慮し、あり得る懸念を抽出することが困難になっている場合がある。よって、複数の類似情報を参照し、その違いから将来を予測し、問題となりそうな懸念を明文化する訓練をする。
- ・例2：具体的な仕様を対象に手法や観点を固定して評価している場合、具体的な情報から抽象化ができるか、もしくは複数の抽象概念を整理して扱えるか訓練する。
- ・例3：具体的な仕様情報やソースコードではなく、抽象設計に対する評価をしている場合、抽象概念を具体化することができるか、つまり、どのようなエビデンスを取得すべきか訓練する。

## 5. まとめ

本手法は、GSN を用いて短時間で且つ該当業務が未経験な技術者を対象に思考過程に依存した能力（リスクを抽出する能力等）を計測する手法を提案した。IV&V 活動に従事している技術者に本手法を試行した結果、提案手法による計測結果が実際の業務経験値を反映しており、本手法の妥当性が確認できた。また、IV&V 活動の業務特性と同様に、業務にインプットされる内容が多岐に渡り、目的に応じて人の思考過程に依存した分析方法を構築しなければならない業務は、本手法の応用可能性があると考えられる。

参考文献

- [1] JAXA, IV&V ガイドブック～導入編～ver1.2 2017 年
- [2] JAXA, IV&V ガイドブック～実践編～ver1.2 2017 年
- [3] GSN COMMUNITY STANDARD VERSION 1 (<http://www.goalstructuringnotation.info/>)
- [4] Matsuno Yutaka, Takai Toshinori, Yamamoto Shuichiro, D-Case 入門 -Lets write a dependability case!- 2012 年
- [5] Mori Motoko D-Case 導入によるシミュレーション S/W の期待結果明確化と合意形成, Software Quality Symposium 2014 年
- [6] Kobayashi IoT 時代に求められるセーフティ設計の見える化とは ～GSN 入門～ (<http://sec.ipa.go.jp/seminar/20150730.html>) 2015 年

掲載されている会社名・製品名などは、各社の登録商標または商標です。

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 (IPA/SEC)