



Skills Management Association

スキルの形式的評価手法開発部会 活動報告

2012年6月12日

一般社団法人 組込みスキルマネジメント協会

■ スキルの形式的評価手法開発部会の目的

ETSSのスキル評価を客観的に実施可能にする手法を開発する

■ 活動成果の目標

1. スキル特性推定手法の研究・開発

- 収集したスキル分布からスキルの特徴を統計的に把握する手法の研究
- スキル分布の推定の信頼性に関する手法の研究・開発

2. スキル評価用設問形式の研究・開発

- スキルを形式的に推定するための設問の形式、問題数、設問時間の関係、パラメータに関する研究・開発

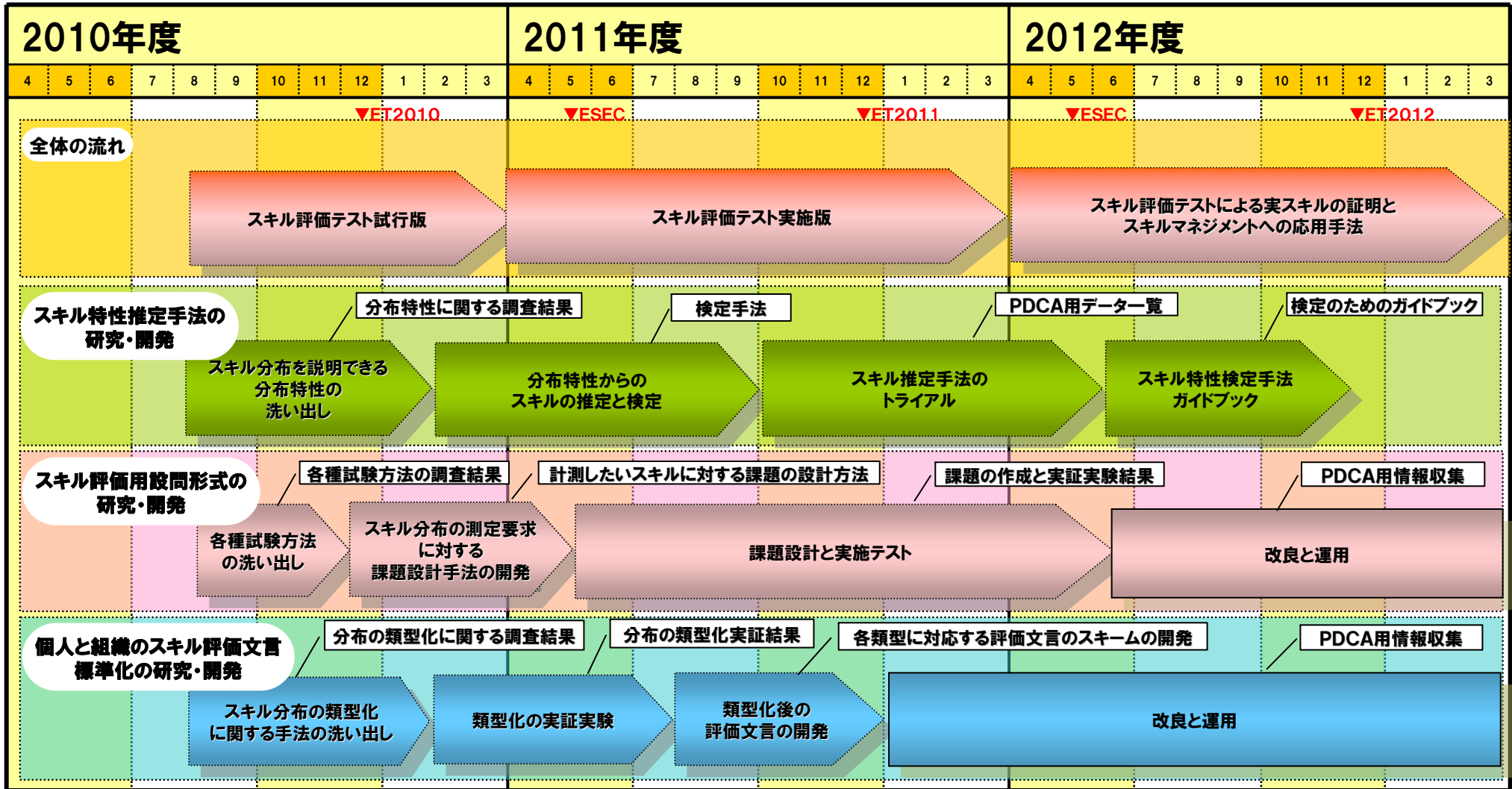
3. 個人と組織のスキル評価文言標準化の研究・開発

- スキル分布の類型化(例:分布のモーメントから類型化)に関する研究
- 類型化したスキル分布ごとに標準的な評価文言を開発する研究・開発

活動スケジュール(3カ年計画)

■ スキルの形式的評価手法開発部会の3カ年スケジュールを以下に示す。

□ は各工程の成果物



■ アクションアイテム

- スキルの形式的評価手法の要件の確認
- 評価ツールの仕様設計
- 評価対象者および評価内容の決定
- 評価試験項目の開発
- 評価試験項目作成要綱の開発
- サンプル試験評価
- 実試験評価
- 多変量解析によるスキル評価パターンの分析

■ スキルの形式的評価手法の要件の確認

- 実際のスキルを評価した個人のスキル分布を真の値と仮定
- スキル評価を行った同一の対象者の試験による分布を測定
- 試験による分布から実スキル分布を推定する手法を開発
- 推定したスキル分布に固有の特性を抽出し、スキル評価につなげる手法を開発
- この手法を用いた試行実験を行ない、実用化の準備に入る。



■ 評価ツールの仕様設計

□ 基本仕様(要求仕様)

- ✓ C/S形式、Stdアローン(現行はVBScriptで、実行形式で作成)
- ✓ 不正ができないように、時間の操作・キー無効などを実装
- ✓ 時間記録(記録する時間)
 - 試験開始(試験開始時間)
 - 各設問回答時間(最初の入力時間)
 - 各設問確定時間(最終の入力時間)
 - *問題回答時間(=最終の入力時間-試験開始時間)
 - 問題の出題順(一般試験にするために、擬似ランダム?とか、複数問題から抽出) → 今回は、問題は変えない。
- ✓ 問題時間設定(0-60Sec)を外部から設定変更ができること
- ✓ 回答方式設定、四者択一/四者選択を外部から設定変更ができること
- ✓ 出題されてから、問題時間設定で自動的に次の出題に移行すること。前の問題に戻れないこと。先の問題にも故意に移動できないこと。

□ セキュリティ

- ✓ 回答の保存、改ざん対応
- ✓ セキュリティを考えた試験環境の条件を検討(CBTとした場合に、基準)

□ その他

- ✓ 試験条件、試験環境(受験環境)の設定
被験者が同じ環境で受験することが必要。途中退出禁止(データ無効扱い)、作業の割り込みが無い環境で実施(試験に**集中できる環境**で実施)

■ 評価対象者および評価内容の決定

□ 実際のスキルを評価した個人のスキル分布を真の値と仮定

- ✓ 被験者の個人評価軸(ETSS、社内評価)を真の値として仮定して、エンジニアのスキルレベルを決定
- ✓ 入社年度、経験、評価別に受験対象者を選定
 - 入社2-3年目 : 評価 上位/中位
 - 入社4-7年目 : 評価 上位/中位/(下位)
 - 入社8年目以上 : 評価 上位/中位/(下位)

■ 評価試験項目の開発

□ スキル評価を行った同一の対象者の試験による分布を測定

- ✓ 評価手法に準じた問題作成 → 180問 (90分)
- ✓ 評価環境の調整
- ✓ 評価の実施
- ✓ 評価(サンプリング)
- ✓ 評価の回数 (例えば、同じグループで2回試験を実施する。異なるグループで、同じ試験を実施する。など) 評価試験問題作成ガイドライン

■ 評価試験項目作成要綱の開発

スキルの形式的評価の試験問題の作成にあたり、スキルの形式的評価試験システムの特性に合わせた試験問題の作成ガイドラインが必要と仮定し、下記の基準を検討。

□ 適正な問題の記述文字数

1題30秒の前提で、出題文字制限をかけて実施

□ デシジョンテーブルなど、理解するのに時間が必要と思える出題について

デシジョンテーブルとコードカバレッジの問題で確認

□ 使用する辞書について

社内で普及している用語のみでなく、被験者は顧客環境で作業を行っているために、用語辞書の準備が困難。JSTQBやISO、IEEEなどで定義されているが、日常的に利用しない用語の使用を確認する。

■ サンプル試験評価

□ 適正な問題の記述文字数

40～80文字、80～120文字、160文字以上のサンプル問題を実施。

(結果)特に変化は見られないが、160文字以内が妥当。

□ デシジョンテーブルなど、理解するのに時間が必要と思える出題について

デシジョンテーブルとコードカバレッジの問題で確認

(結果)計算と論理的思考問題は、不適。

□ 使用する辞書について

社内で普及している用語のみでなく、被験者は顧客環境で作業を行っているために、用語辞書の準備が困難。

JSTQBやISO、IEEEなどで定義されているが、日常的に利用しない用語の使用を確認する。

(結果)業務で利用している点もあり、特に問題が検出されず。

■ 実試験評価

□ 90分180問で実施した結果 > 被験者(受験者)からの不満が噴出。

✓ 試験時間が長い(飽きる)。時間が長いのではなく、淡々と問題が過ぎてゆくのが辛い。

✓ 緊張感が続かない。

次年度に向けた課題となる。

■ 評価用のテストツールの開発

汎用的に利用できるテストツールの開発

- 90分間で適切なタイミングで中断／再開できる機能
- 1／3サイズの範囲で評価を実施できるように開発
- 未搭載のセキュリティ機能の実装

■ サンプルテストの応用

他スキル領域(キャリアモデル)への適用検討

■ 試験の仕組みの検討と実施

実施結果の分析と仮説の証明

実用に向け、評価制度などを整備

