

# 未来技術動向予測のための技術文書マイニング

## Technical Document Mining Method for Future Technology Projections

西山 莉紗\*<sup>1</sup>      竹内 広宜\*<sup>1</sup>      渡辺 日出雄\*<sup>1</sup>      那須川 哲哉\*<sup>1</sup>      前田 潤治\*<sup>2</sup>  
 Risa Nishiyama      Hironori Takeuchi      Hideo Watanabe      Tetsuya Nasukawa      Junji Maeda

倉持 俊之\*<sup>2</sup>      林口 英治\*<sup>2</sup>  
 Toshiyuki Kuramochi      Eiji Hayashiguchi

\*<sup>1</sup>日本アイ・ピー・エム (株) 東京基礎研究所  
 IBM Research, Tokyo Research Laboratory

\*<sup>2</sup>アイ・ピー・エム ビジネスコンサルティング サービス (株)  
 IBM Business Consulting Services

A future technology projection tool for strategy consultants, which consists of several components utilizing text mining technologies, is introduced in this paper. The purpose of this tool is to help consultants and their clients in discovering technological capabilities in patent documents that may contribute to the business of the clients. Our tool automatically extracts phrases suggesting possibilities created by new technologies and estimates how long until products will appear and how large the business impact may be. A preliminary trial showed some effectiveness in utilizing the information in some sections of the documents and provided statistical information about keywords in the set of documents for estimates.

### 1. はじめに

近年新技術の社会的影響力がますます強まっており、新技術によってどのようなイノベーションが可能になるか、ということについて多くの経営者が興味を持っている。そのため、現在から数年後に渡る将来技術動向を分析・把握し、新規ビジネスの創出や企業形態の変革などを提言する戦略コンサルティングビジネスの需要が高まっている。

このようなコンサルティングを行う際に重要な情報源となるのは、経済産業省の技術戦略マップ [経済 06] や野村総合研究所のITロードマップ [野村 06] のような、専門家のグループがいくつかの技術分野について動向をまとめたテクノロジー・マップである。また、例えば特許文書を対象として自動的にテクノロジー・マップ (特許マップ) を作成する技術 [新井 03, 市村 01, 渡部 03] も既に多く研究されている。これらは特許文書の出願日や本文中に出現するキーワード、及び本文中に書かれた課題や解決手段などの目的に応じた軸で全特許文書の傾向をまとめ、ユーザーに提示するものである。

以上に挙げた既存のマップは戦略コンサルティングに貢献するものであるが、これらは全体を概観するための要約情報を提示するものか、または特定の有力技術を詳説するものである。そのため、まだ取り組みが少ないがビジネス価値のある技術の発見や、各々の技術の実現難易度をそこから直接知ることが難しい。これらの既存のマップから得ることが難しい情報は、将来実現可能になる技術を用いた新規ビジネスを創出・提案する戦略コンサルティングにおいて非常に重要な役割を果たす。そのため著者らは、特許文書を始めとする技術文書から当該技術が実現する事柄を示す表現を抽出し (以後この表現を可能性表現と呼ぶ)、提示することで既存のツールを補完する、新しい技術文書マイニング手法の研究に取り組んでいる。

本稿ではまず著者らの手法が用いられるコンサルティング手法である BVT (Business Value from Technology) 手法につい

て説明し、そこで用いられるテクノロジー・マップの要件を挙げる。次に、それらの要件を満たすテクノロジー・マップを出力するべく提案された本ツールについて詳しく説明し、実際に出力されるテクノロジー・マップのサンプルを紹介し、結果について議論する。

### 2. 本研究の背景 BVT 手法

BVT 手法とは、顧客企業のイノベーション実現に対して提供されるコンサルティング手法の一つである。

基本的な考え方は、技術が持つ可能性 (テクノロジー・ケーパビリティ) に対してビジネスが要求する可能性 (ビジネス・ケーパビリティ) を照らし合わせることにある。例えば金融業界の顧客企業を仮定し、技術としてハイスピードコンピュータを考えてみる。この場合のテクノロジー・ケーパビリティは明らかに「超高速計算能力」である。一方金融業界においては個人顧客の取り込みは大きなビジネスバリューであることから、ビジネス・ケーパビリティとしては「個人単位での投資のモンテカルロシミュレーション」が考えられる。モンテカルロシミュレーションは高い計算機負荷をかけるため、このビジネス・ケーパビリティの実現には高速計算機環境が必要不可欠である。この時「超高速計算能力」というテクノロジー・ケーパビリティと「個人単位での投資シミュレーション」というビジネス・ケーパビリティは交錯し、「個人投資家単位の投資分析環境を提供することで個人顧客を取り込む」というイノベーションが考えられる。

以上に述べたようなビジネス・ケーパビリティとテクノロジー・ケーパビリティの発見とマッピングを行うためのツールとして、テクノロジー・マップが利用される。このとき、テクノロジー・マップは以下の要件を満たす必要がある。

- 収録技術に網羅性があること
- 技術全体を見渡せる一貫性があること
- ビジネスへの活用可能性を検討するための記述を含んでいること

連絡先: 西山 莉紗, 日本アイ・ピー・エム (株) 東京基礎研究所, <http://www.research.ibm.com/trl/people/lisa>, [lisa@jp.ibm.com](mailto:lisa@jp.ibm.com)

本ツールでは以上3つの要件を満たすテクノロジー・マップの作成を支援することを目的としている。次章ではその詳細について説明する。

### 3. 未来技術動向予測ツール

本ツールは特許文書を入力とし、以下に示す3段階のプロセスで技術文書进行处理して、例えば「web2.0」や「音声認識」など、注目している技術分野に関連した可能性表現(前章におけるテクノロジー・ケーパビリティ)を抽出し、マップの形式で出力する。

1. 注目技術分野の可能性表現抽出
2. 可能性表現の特性推定
3. テクノロジー・マップに整形

なお、現在はステップ2までの自動化に取り組んでいる。以下に入出力の形式及び各々のステップでの処理について詳しく説明する。

#### 3.1 入力

入力は図1に例を示すような公開特許公報をXMLで表記したものを利用する。公開特許公報は定型文書であり、いくつかのセクションから成るが、本ツールでは「要約」中の「課題」「解決手段」セクション、そして「発明の開示」中の「発明の効果」セクションをこの後の処理に利用する。なお簡単のために、以降では処理に利用する公開特許公報を特許文書と呼び、「課題」セクションを課題記述部、「発明の効果」セクションを効果記述部と呼ぶ。

<p>【発明の名称】 移動ロボットの制御装置</p> <p>【課題】 自身の発する作動音に障害されずに音声認識を行うことのできる移動ロボットの制御装置を提供する。</p> <p>【解決手段】 可動部の運動を制御する行動制御手段(7)、少なくとも...</p> <p>...</p> <p>【発明の効果】 本発明によれば、運動状態検出手段の信号から作動音を発するタイミングを予測でき、これによってロボット自身が発する作動音を把握することができる。そして...</p>
--

図 1: 入力特許文書の例

#### 3.2 ステップ 1: 注目技術分野の可能性表現抽出

ステップ1では前章で述べたテクノロジー・ケーパビリティを示す表現(可能性表現)を抽出する。コンサルタントを含めた事前調査の結果、そのような表現は一般に表1に示すような手がかり語を持つことが分かった。これを利用し、全ての入力特許文書中の課題及び効果記述部から、いずれかの手がかり語で終わる、ある深さ以内の係り受け範囲の表現をバッチ処理で抽出しておき、これを可能性表現とする。そして、要求に応じて対象分野に関連することが期待される特許文書を絞り込み、それらに含まれる可能性表現を抽出する。例えば音声認識分野の未来技術動向予測を行う場合には、「音声認識」という文字列を含む文書を絞り込む、ということが考えられる。この絞り込みに利用する文字列はユーザーによって任意に指定可能なものとする。

表 1: 可能性表現抽出に利用する手がかり語と抽出される表現の例

手がかり語	抽出される表現例
~を向上する	ユーザの使い勝手を向上する
~を高める	光の利用効率を高める
~を改善する	初期の劣化を改善する
~に優れる	冷熱サイクル性に優れる
~の向上	処理速度の向上
~の...化	全体の薄型化
~の改善	コントラストの改善
~を防止する	画像の劣化を防止する
~を抑制する	変動による影響を抑制する
~を低減する	消費電力を低減する
~を抑える	半導体素子の電流を抑える
~を可能にする	強度を確保することを可能にする
~することができる	カウントすることで把握することができる
~を容易にする	チャンネルを選択することを容易にする

上で述べた手がかり語を用いた手法により、特許文書から網羅的に可能性表現を抽出することが可能である。しかしこれらの可能性表現は膨大な量となるため、テクノロジー・マップの要件の1つである、マップの一覧性を低下させる恐れがある。そこで当該技術分野に固有の用語を集めたキーワードリストを用意し(以降、これを対象分野キーワードリストと呼ぶ。)、その中のいずれかの単語を含む可能性表現のみを抽出する。このキーワードリストは、当該技術分野と高い相関で現れる体言を集めることで作成できる。このようにして得られるキーワードリストを使うことにより、例えば「実現を可能にする」等の分野に依存しない一般的な表現をフィルタアウトして、対象分野により深く関連する技術表現のみを抽出することが可能になる。また同時に、例えば「音声認識を実現する」と「音声認識の実現が可能となる」というような、共有しているキーワードの数が多技術表現が同一特許文書から抽出された場合には「同じ内容を示している」として一つにまとめる。このような可能性表現のフィルタリングおよびマージは、マップの可読性の向上に貢献すると考えられる。

次のステップでは抽出された各々の可能性表現について、その表現が持つ特性を推定する。

#### 3.3 ステップ 2: 可能性表現の特性推定

ステップ2では、前ステップで抽出された可能性表現が持つ特性として、実現にかかる時間の長短、実現された場合のビジネスインパクトの大小、そして技術の流れを考え、これを抽出された表現が持つキーワードや、抽出された元のセクション情報を利用して推定する。

##### 1. 実現時間推定

一般に、課題記述部にはこの文書に書かれた技術が直接的に可能にする事柄が端的に述べられており、対して効果記述部には、この文書に書かれた技術が実現されることで間接的に可能となる、さらなる効果が記述されている場合が多い。このようなセクションごとの内容の違いを利用し、本手法では、課題記述部から抽出された可能性表現について、その実現までの時間は短いと推定し(近未来)、効果記述部から抽出された可能性表現については長いと推定する(未来)。

## 2. ビジネスインパクト推定

実際のビジネスインパクトは様々な要因から与えられるが、ここではその一つである、特に技術の斬新な応用や活用によって与えられるインパクトに注目する。ここでは各々の可能性表現中における、当該技術分野と関連の浅い語の出現に注目する。このような語の出現は、当該技術分野とはこれまで結びつきが考えられなかったような技術応用の方向性を示唆するものであると考えられるためである。

具体的には図2に示すように、3.2節のフィルタリングで利用した対象分野キーワードリスト（対象分野と関連の深い語が集められていることが期待できる）と「発明」など全文書で頻出の用語をあつめた汎用キーワードリストを利用する。対象となる可能性表現内に出現する体言のうち、いずれのリストにも無いもの（図2の濃色部）を対象技術分野と関連の浅い語とし、そのような関連の浅い語を含む表現に高いビジネスインパクトを与える。

## 3. 流れ推定

例えば「音声認識」から「感情認識」へと遷移していくというような、技術の流れについても情報提供できることが望ましい。本ツールでは同一特許文書の課題記述部から抽出された可能性表現（実現時間は「近未来」と推定されている）と、効果記述部から抽出された可能性表現（実現時間は「未来」と推定されている）が一つの技術の流れを形成しているとみなす。なぜなら、これらはその特許文書で描かれている一つの技術が実現されることによって近未来と未来に可能になる事柄を示していると言えるためである。

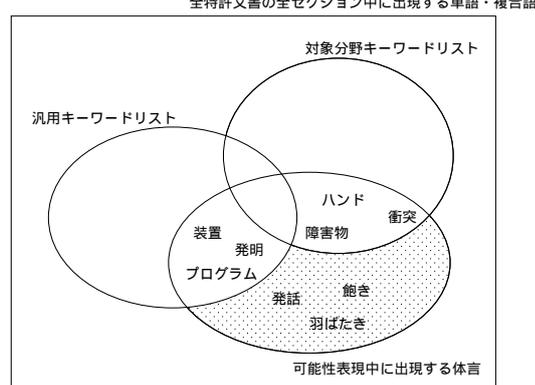


図2: ビジネスインパクトの推定に利用するキーワードリストの概念図（図中のキーワードはロボット分野における例を示す）

次のステップでは以上のように推定された実現時間やビジネスインパクト、そして技術の流れの情報を利用して、テクノロジー・マップを作成する。

## 3.4 ステップ3: テクノロジー・マップに整形

ここで出力すべきマップの形式としては様々なものを利用可能である。例えば次章に示すように、横軸を実現時間、縦軸をビジネスインパクトとした座標平面上に技術の流れを示す矢印でつないだ可能性表現を配置することで、当該技術分野において抽出された可能性表現の俯瞰図を得ることができる。ここで整形されたテクノロジー・マップが本ツールより出力される。

## 3.5 出力

本ツールによって出力されるテクノロジー・マップは、いわば雛形である。実用上は戦略コンサルタントが本ツールの出力に対して修正を加え、修正後のマップが顧客との議論に利用されることになる。具体的な可能性表現を見ることにより、まだ数件の技術文書でしか述べられていない技術の発見を可能にし、それらの顧客ビジネスへの応用についての気づきを促進する。

次章では実際にステップ1・2を実装したシステムから出力された可能性表現とそれらの推定実現時間、及びビジネスインパクトを示し、マップ例の解釈及び知見について述べる。

## 4. 抽出された可能性表現及びテクノロジー・マップの例

表2は前章で述べた方法によって実際に2005年から2006年の間に公開された65,384件の特許文書から抽出された、ロボット分野に関連すると考えられる可能性表現である（なお、ここには紙面の都合上、近未来に実現されると推定された表現のみを掲載している）。これらの可能性表現は「ロボット」という文字列を文中に含む135件の特許文書から抽出されている。そして、抽出された可能性表現からいくつかを抜粋し、テクノロジー・マップの形式で記述したイメージ図を図3に示す。

表2と図3から、本手法によって「スループットを向上する」「効率良く搬送することができる」などの、ロボット技術を利用することによってもたらされる妥当な展開については低いビジネスインパクトが与えられる反面、「家屋内を掃除することができる」「飽かせないロボットを提供する」などの、実用された場合に大きな反響を期待できるような技術に対して高いビジネスインパクトが与えられていることが分かる。

実際に「家屋内を掃除することができる」技術については2002年に（株）東芝から国内初の掃除ロボットが発表されており\*1、「飽かせないロボットを提供する」技術については2005年に日本電気（株）から対話機能によってユーザーを飽きさせないことを目指したロボットが開発されている\*2。「エレガントなモーション・パフォーマンスを簡易な編集作業により実現することができる」についても、2005年にスピーシーズ（株）より人間のように滑らかな動きを簡単に取り込むことのできるソフトが発表されているため\*3、一般に「エレガント」な動きに要求される、滑らかな動作の簡単な実現に貢献している。

このように本手法で高いビジネスインパクトを持っていると推定された技術が、実際に製品化されてメディアにも取り上げられている、つまり、実際にある程度のビジネスインパクトがあるということは、本手法によるビジネスインパクト推定の有効性を示唆していると言える。また、近未来技術と推定された技術が実際に2005年までに実現されていることから、実現時間推定方法の有効性が分かる。

\*1 [http://www.toshiba.co.jp/about/press/2002\\_09/pr\\_j0501.htm](http://www.toshiba.co.jp/about/press/2002_09/pr_j0501.htm)

\*2 <http://www.incx.nec.co.jp/robot/papero2005/>

\*3 [http://speccys.com/release/robotstudio\\_release.pdf](http://speccys.com/release/robotstudio_release.pdf)

表 2: 抽出された可能性表現と推定されたビジネスインパクト (実現時間が「近未来」と推定されたもののみ)

可能性表現 (近未来)	
高 BI	インチワームロボットに関するもの
	ウェハ搬送ロボットのブレードの前垂れの度合いを計測する手段を提供する
	エンドエフェクタ上のウェハの有無を検出する
	コイル保持部材に発生する渦電流を低減することができる
	ダウンタイムを減らすことができる
	トレー部を有する
	ピン穴から取り出す
	リンクモデルで行なう
	ロボットハンドによる出し入れを行うことができる
	ロボット用サーボ機構を作る
	異物落下を防止する
	移動速度の高速化を図ることもできる
	移動対象を高速で駆動させることができる
	音声認識を行うことのできる
	可動部への適用が難しかった
	家屋内を掃除することができる
	外部環境から保護することができる
	基板搬送機構のメンテナンスを容易にする
	更新処理を行なうことができる
	高速動作が可能
	作業効率を向上させることができる
	精度よく自律的に学習する
	設備の簡略化を図ることができる
	大型空調ダクトの内面を清掃する
直線移動行程を実現できる	
物品交換用ロボットの動きを制限しない	
文を出力することを防止する	
飽かせないロボットを提供する	
無人搬送台車を提供する	
低 BI	アライメントする
	ウェハ搬送コントローラを提供する
	ステータが接触する
	スループットを向上する
	た動作を行うことができる
	ないしスループットを向上する
	ピンで支持する
	ような状態であった
	リンググラフィ装置を提供する
	リニアモータ等に用いられる
	位置ズレを補正する技術を提供する
	環境を表現する
	基板処理位置に位置決めすることができる
	形状復元する
	固定部に設置できる
	効率良く搬送することができる
	作業が同時処理できる
	作業に適した
	振動力を発することのできる
	人の顔を判定できる
	適切なタイミングで提供することができる
	等を行なう
	配線基板上に搭載することができる
	搬送することのできる
利用者に報知することのできる	
利用者の認証をすることができる	

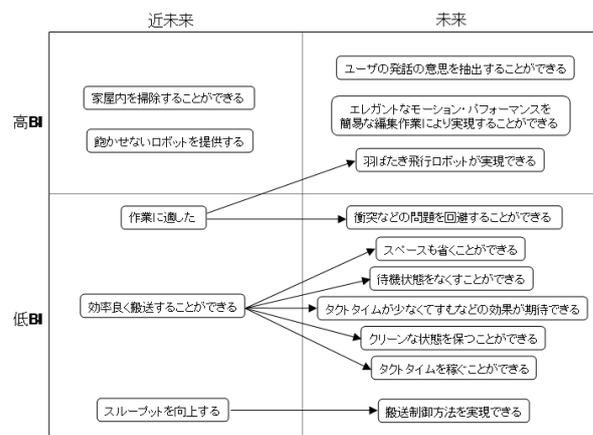


図 3: テクノロジー・マップの例：ロボット分野

## 5. おわりに

本稿ではまず BVT 手法について説明し、そこで必要となるテクノロジー・マップの要件を挙げた。次に、それらの要件を満たすテクノロジー・マップを出力するべく提案された本ツールを処理手順に沿って説明し、実際に出力される可能性表現とテクノロジー・マップのサンプルを紹介して、本手法の中の特に実現時間推定とビジネスインパクト推定の有効性について議論した。

今後の課題として、抽出された全ての可能性表現をマップに落とし込み、その一貫性についての検証を行う必要性が挙げられる。また、本手法によって推定される実現時間やビジネスインパクトの妥当性の評価を行うとともに、評価手法そのものの確立を行う必要がある。ならびに、今回は日本の公開特許広報を対象とし、その文書特有のセクション情報を利用したが、論文やニュースなど特許文書以外の技術文書にも適用可能な仕組みを考えることで、より先の技術動向予測に役立つ情報を提供できることが期待できる。

## 参考文献

- [経済 06] 経済産業省：技術戦略マップ 2006, <http://www.meti.go.jp/press/20060428011/str2006.pdf> (2007.04.05) (2006)
- [市村 01] 市村 由美, 長谷川 隆明, 渡部 勇, 佐藤 光弘：テキストマイニング：事例紹介 (<特集>「テキストマイニング」), 人工知能学会誌, Vol. 16, No. 2, pp. 192-200 (2001)
- [新井 03] 新井 喜美雄：特許情報分析とパテントマップ, 情報の科学と技術, Vol. 53, No. 1, pp. 16-21 (2003)
- [渡部 03] 渡部 勇：テキストマイニングの技術と応用 (<特集>情報の分析・解析法), 情報の科学と技術, Vol. 53, No. 1, pp. 28-33 (2003)
- [野村 06] 野村総合研究所技術調査部：IT ロードマップ < 2007 年版 >, 東洋経済新報社 (2006), ISBN: 4492580786