慶應義塾大学学術情報リポジトリ
Keio Associated Repository of Academic resouces

Title	「テキストマイニングを使った人材の360度評価」に関する考察
Sub Title	
Author	兵藤, 誠(Hyodo, Makoto) 林, 高樹(Hayashi, Takaki)
Publisher	慶應義塾大学大学院経営管理研究科
Publication year	2014
Jtitle	
JaLC DOI	
Abstract	
Notes	修士学位論文. 2014年度経営学 第2990号
Genre	Thesis or Dissertation
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40003001-00002014-2990

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the KeiO Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

慶應義塾大学大学院経営管理研究科修士課程

学位論文（　　2014　年度）

論文題名

「テキストマイニングを使った人材の 360 度評価」に関する考察

主　査	林 高樹
副　査	高橋 大志
副　査	渡辺 直登
副　査	

学籍番号	81331072	氏　名	兵藤 誠

論 文 要 旨

所属ゼミ	林　研究会	学籍番号	81331072	氏名	兵藤 誠

（論文題名）

「テキストマイニングを使った人材の 360 度評価」に関する考察

（内容の要旨）

本論文は、人材の評価方法として、従来から採用されているアンケートデータの分析による 360 度評価ではなく、テキストデータの分析による 360 度評価を新たに提案する論文である。実際の企業から入手したデータの分析を行い、テキストデータから導かれた評価と従来のアンケートから導かれた評価を比較し、それぞれの評価方法の特徴を明らかにした上で、実務において「どのようなケースで」「どの評価方法が適しているか」まで考察する。

目次

1. 問題意識 4
1-1. アンケートベースの 360 度評価 4
1-2. アンケートベースの 360 度評価における問題点 4
1-3. テキストマイニングを利用した 360 度評価の提案 5
2. 先行事例研究 6
3. 研究方法 6
3-1. 使用データ 7
3-2. 使用ツール・辞書 9
3-3. 想定アウトプット 10
4. 研究手順 11
4-1. 全体構成 11
4-2. 各パート詳細 11
5. 結果 20
5-1. 概要 20
5-2. 従来の評価との比較 22
5-3. テキスト評価と各コンピテンシーの関連性 25
6. 実務への応用 28
7. まとめと今後の課題 32
8. 終わりに 34
謝辞 34
参考文献 35
付録 1：多特性多評価者行列による妥当性検証 37
付録 2：本研究で作成した R のコード 39

1. 問題意識

1-1. アンケートベースの 360 度評価

筆者は、前職において人材ビジネス関連の企業で勤務しており、人材の採用や評価という行為に強く関心を持っている。そして日々の業務の中で「人の評価は、アンケートによる調査だけで本当に十分であるのか？」という問いを抱いていた。この疑問が本研究の出発点である。

現在、実務の中で多く採用されている人材評価方法は「360 度診断」である。360 度診断とは、上司・同僚・部下という自分の周りにいる全てのポジションの人物からアンケートベースの 5 段階評価（もしくは 100 点評価）の形で自分の評価を行ってもらう診断方法である。例えば、評価対象が課長だと仮定して、一緒に働いている上司・同僚・部下に「課長として求められる行動がどの程度達成できているか」を 5 段階評価で評価してもらう。アンケートベースで算出された得点は、組織内において昇進する人材を決める際の検討材料として、もしくは組織内の人間関係の把握・改善のための材料として使用されている。

360 度診断を日本企業が導入している背景は、「日経ビジネスオンライン 2007 年 6 月 18 日号」[1]に以下のように記載されている。

> 多くの企業が「360 度評価」を導入した背景には、組織のフラット化が進み、上司が以前より多くの部下を抱えるようになったことがある。上司は、通常の業務に加えて、報告や承認、他部門との調整といった仕事が増えている。そのため、部下一人ひとりの行動や能力を観察する時間が少なくなってしまった。また、成果主義の導入にともない、上司は部下の評価に差をつけなければならない。優秀な部下や問題児は、目立つので評価しやすい。一方、それ以外の大多数の部下は、評価に差をつけるのが難しい。そのため、部下や同僚など、日頃の仕事ぶりをよく知っている人たちの「目」を借りようというのがこの制度の狙いだ。この評価の利点は、上司だけの視点ではなく、より多くの人から判断材料を集めることで、より公正な評価ができると言われている。

1-2. アンケートベースの 360 度評価の問題点

従来の 360 度診断は前述した長所を持つが、全てのケースで正しい評価を計測でき

[1]日経ビジネスオンライン（2007）「イヤな上司を排除するのにも使える「不幸」な制度職場の雰囲気を乱す「360 度評価」の危険性」（http://business.nikkeibp.co.jp/article/skillup/20070612/127148/?P=2）（2014 年 6 月現在）

る万能な評価方法では無いと考える。

その理由は、あくまで「アンケートの点数」という限られた種類のデータしか得られない、という事実に由来する。正しく評価を行うためには、定量的に評価が可能な形でデータを取得することが必要で、アンケートという手段はその意味では効率的なデータの取得方法である。しかし、アンケートだと定量的な項目しか聞けない分、様々な情報が削ぎ落とされることとなり「評価者の感情が見えない」定量的データとして扱われてしまう。実務の場では、評価者は被評価者に対し怒っていることもあれば、本気で尊敬していることもあるだろう。失望していることもあるだろう。その被評価者に抱いている様々な感情の機微までは評価できず、ただ 1〜5 の 5 段階尺度で表現されることになってしまう。

1-3. テキストマイニングを利用した 360 度評価の提案

本研究では、先述した問題を解決するための評価方法として、「テキストマイニングによる 360 度評価」という方法を提案したい。

具体的には、まず、テキストの単語毎にポジティブかネガティブか、それはどのくらいの程度ポジティブ・ネガティブなのか、を+1〜−1 の範囲で定義する（−1 ならネガティブ、+1 ならポジティブ）。次に「被評価者の強み・弱み」をテキストデータで取得し、「被評価者の強み」の文章からポジティブな単語を抽出し、1 評価に含まれる全てのポジティブ単語の合計点を算出する。同じように「被評価者の弱み」の文章からはネガティブな単語を抽出し、1 評価に含まれる全てのネガティブ単語の合計点を算出する。最後に、ポジティブ単語から算出した得点とネガティブ単語から算出した得点を合計し、その点数を「評価者が被評価者を評価した点数」とする方法である。

本手法を採用することで、先述した問題が解決できる可能性がある。テキストデータは人間の喜怒哀楽の感情が書き込まれている、生々しいデータである。アンケートの項目のみを分析に利用した評価ではなく、評価者の強み・弱みをテキストで自由回答してもらい、そのデータを分析に使用することで、「評価者に抱く感情」という観点も考慮に含めた評価が可能ではないか、という考えを持った。（この点は本研究にご協力頂いたディアグロー株式会社の内山社長も強く問題意識を持っており、その結果、データの提供や全面バックアップを頂くことができた。）

要約すると、本研究は「360 度のアンケート評価」という従来から使用されている方法ではなく「360 度のテキスト評価」という新しい方法を提案し、両方の結果を比較しながら、テキストでの評価方法が従来の方法と「何がどのように」異なるのか考察する。さらに、実務において、それぞれの評価方法に適したシチュエーチョンも考察し、人材評価の方法に新しい示唆を与えることを目指す。

2.先行事例研究

筆者が先行事例を調査した限りでは、「人事評価」を「テキストマイニング」で行う研究は見られない。ただし、「他の要素」を「テキストマイニング」で分析する事例、「人事評価」を「テキストマイニング以外の方法で」分析している事例は多く見られたので、以下に紹介する。

①「他の要素」を「テキストマイニング」で分析している事例

国内における類似研究としては、企業の研修についての自由記述の感想をテキストマイニングで分析する方法（2013.石橋）がある。但し、実際のデータを分析してそこから示唆を導く部分は記載が無く、テキストの頻度分析等の方法論の記載が主となる論文である。

また、海外においては優秀なプロジェクトマネジャーの資質をテキストマイニングで測る研究が存在する（2014.Nishtha Langer)。インドの 209 人のプロジェクトマネジャーに、仕事上で起こりうる重大な出来事に対してどう反応するかをテキストで聞き、文書内の出現頻度や共起表現を踏まえて、人事評価の高いプロジェクトマネジャーに頻出される単語を抽出する研究である。結果として、成果の高いプロジェクトマネジャーのテキストに良く見られる単語は approach/discuss/talk/understand などが見られ、高いコミュニケーション力を持つマネジャーは高い成果を出し易い、という示唆を出すことに成功している。

②「人事評価」を「テキストマイニング以外の方法で」分析している事例

一般的に、人事評価は 360 度診断のアンケートで実施する例や、実際の仕事の場面で使用する様々なスキルを判定するアセスメントセンター型の評価がメインである（2012.廣瀬)。具体的には、経営幹部を選抜するために、ペーパーテスト、面接演習、グループ討議、プレゼンテーション演習など、1 人の人材を仕事上における様々な切り口から評価する「インバスケット試験」が有名である。

3.研究方法

ある企業の実際の 360 度診断のデータを用いた実証研究を行う。本研究に協力して頂いているディアグロー株式会社では、360 度診断を商品の 1 つにしており、その診断では評価者が被評価者の「リーダーシップ」「マネジメントスキル」等の要素を 5 段階評価アンケートで評価した定量的な数値データと、被評価者の強み・弱みを 200

文字で記入するコメントデータの両方を取得している。ディアグロー社ではアンケートデータの評価を点数化して顧客に提出しており、テキストデータの定量化は実施していない。テキストデータは被評価者が自分の強み・弱みを確認して今後の仕事に役立ててもらうために参考情報として特に手を加えず、テキストデータのまま顧客に提示している。本研究では、アンケートを用いた被評価者の評価と、テキストデータを用いた評価の両方を定量化し、比較を行う。

3-1. 使用データ

データの検証にあたっては、2 社の企業に協力して頂いた。まずはディアグロー株式会社である。先述したように、360 度診断を自社商品として取り扱っている企業である。次に、某自動車部品メーカー（名前は非公開）である。この企業はディアグロー社の 360 度診断を人材評価方法として採用しており、本研究はこの企業から実際に得られたデータを元に検証を行っている。使用データの概要は以下に記載する。

データ概要（詳しくは資料 1 を参照のこと）

- 調査日時：2013 年 11 月
- 被評価者：10 名（中堅クラス社員。10 人はそれぞれ異部署の社員。）
- 評価者：61 名（被評価者 1 人につき、上司・同僚・部下がそれぞれ 1～3 人評価している。）

- データに含まれる項目
  - 被評価者氏名
  - 被評価者部署
  - 評価者氏名
  - 評価者部署
  - 被評価者との関係（上司 or 同僚 or 後輩）
  - 評価者のアンケート評価点※
  - 被評価者の強みに関するコメント
  - 弱みに関するコメント

※被評価者の評価は、評価者が 40 問のアンケート（5 段階評価）を回答し、その結果をディアグロー社が詳細な 8 項目のコンピテンシーに分解し、各項目の点数を算出している。（個人特性【表層面】・個人特性【深層面】・対人スキル【アウトプット】・対人スキル【インプット】・リーダーシップ【思考】・リーダーシップ【行動】・マネジメント【仕事】・マネジメント【人】）。また、その 8 項目の点数の算術平均を評価

者から見た被評価者の総合点数、と定義している。

資料１：入手したデータ

＜5＞評価結果一覧　（高評価観察者順位）

								評価者の評価							
								個人特性		対人スキル		リーダーシップ		マネジメント	
順位	被評価者氏名	部署1	評価者氏名	部署1	観察者評価	被評価者との関係	観察者平均	表層	深層	インプット	アウトプット	思考	行動	仕事	人
1	被評価者氏名	MJ G110 乗用車グルー	評価者氏名	生産センター	95.0	同僚	74.1	95.0	95.0	95.0	95.0	95.0	95.0	95.0	95.0
2		N020 生産計画物流管理		生産センター	95.0	部下/後輩	81.6	95.0	95.0	95.0	95.0	95.0	95.0	95.0	95.0
3		H110 第1吸気システムG		テクニカルセンター	93.0	部下/後輩	74.6	89.0	92.0	95.0	92.0	92.0	95.0	95.0	95.0
4		MJ G110 乗用車グルー		生産センター	91.0	部下/後輩	74.9	89.0	85.0	88.0	92.0	92.0	92.0	95.0	95.0
5		H530 CAE解析技術開発		テクニカルセンター	88.5	上司	76.7	86.0	86.0	92.0	92.0	89.0	85.0	86.0	92.0
6		NT50 栃木第2製造G T50		生産センター	88.1	部下/後輩	67.6	89.0	92.0	89.0	85.0	89.0	86.0	86.0	89.0
7		NT50 栃木第2製造G T50		生産センター	87.9	部下/後輩	67.6	83.0	83.0	86.0	92.0	86.0	92.0	86.0	95.0
8		D031 埼玉品質保証G		生産技術センター	87.1	同僚	61.0	89.0	92.0	86.0	86.0	86.0	86.0	83.0	89.0
9		N020 生産計画物流管理		生産センター	86.3	上司	81.6	79.0	86.0	95.0	89.0	80.0	83.0	89.0	89.0
10		N020 生産計画物流管理		生産センター	85.8	同僚	81.6	85.0	86.0	89.0	85.0	85.0	78.0	86.0	92.0
11		H720 材料研究Gr.		テクニカルセンター	85.5	部下/後輩	65.3	89.0	72.0	95.0	95.0	86.0	74.0	78.0	95.0
12		H530 CAE解析技術開発		テクニカルセンター	84.2	同僚	76.7	95.0	65.0	95.0	80.0	84.0	95.0	84.0	76.3
13		NT50 栃木第2製造G T50		生産センター	83.8	同僚	67.6	89.0	86.0	83.0	86.0	80.0	83.0	80.0	83.0
14		H530 CAE解析技術開発		テクニカルセンター	82.6	部下/後輩	76.7	86.0	83.0	80.0	80.0	80.0	80.0	86.0	86.0
15		H110 第1吸気システムG		テクニカルセンター	82.3	部下/後輩	74.6	82.0	79.0	82.0	89.0	86.0	79.0	83.0	78.0
16		D050 生産前品質保証G		品質保証センター	81.8	上司	64.3	89.0	79.0	71.0	78.0	89.0	76.0	86.0	86.0
17		MJ G110 乗用車グルー		営業部	81.5	部下/後輩	74.1	86.0	85.0	80.0	83.0	76.0	79.0	80.0	83.0
18		N020 生産計画物流管理		生産センター	81.1	部下/後輩	81.6	83.0	68.0	86.0	83.0	83.0	83.0	80.0	83.0

＜7＞コメント一覧　被評価者別（評価者コメント）

被評価者	評価者	被評価者との関係	強みに関するコメント	啓発（改善）に関するコメント
被評価者氏名	評価者氏名	上司	物事を的確に判断し、指示等も迅速に行う事が出来る。昨年中途で入社した部下に対しても、客先監査の経験を積ませる事を積極的に行い飛躍的に成長させた。	部下に業務指示を与える際に、具体的な道筋まで示した方が良いと思うケースが見受けられる。部下に自ら考えさせてやらせる事も必要だが業務内容によっては最短で対応させるために心がけてほしい。また、パート員に対しての言動もきつい時があるため注意が必要。
		同僚	物事を論理的に捉え、それを基に他者に上手く説明する事が出来ます。これは顧客でのゲストエンジニア体験を通じて多くの事を学ばれた成果と思われます。	どうしたら皆が同じ方向を向いて仕事が出来るか？　個々の調整も時には必要で有効ですが、時にはチームでオープンな議論の基で結論を導くこともリーダーとして必要であり、この点について更なる飛躍を期待します。
		同僚	部下の仕事に対し、納得いくまで、妥協せず根気よく時間に関係なく付き合い、部下の成長を促している。	強みの裏返しであるが、稟意先提出資料等にて時間の制約はあるが、あまり遅くならないよう、また体調管理に気を付けてください。
		上司	設計部門での経験を活かして、物事を客観的に分析して判断できる事が最大の強みです。 今後も幅広く知識と技術（経験）を習得し、更なるレベルUPに努めて下さい。	グループ内のコミュニケーションは問題有りませんが、他部署との連携が必要な場合、あと一歩押しが弱い部分があります。 今後は今迄以上に折衝業務が増える事が予測される事から、率先して折衝業務の経験を積んで下さい。
		部下/後輩	問題に直面した際に過去の経験や対策を忘れずにフィードバックできる姿勢が感じられます。 業務を進める上で役割分担を的確に把握し、業務を分配し効率的に遂行することができる。	正論が多いため、業務に広がりを持たせ相手の身になった行動、言動を期待します。（部下の育成についても） 他責にする傾向がうかがえる。
		部下/後輩	設計あがりであり、製品の知識は豊富で的確なアドバイスをしていただき、とても頼りになります。	相談しにくい雰囲気がありますので、その点は改善していただきたい。
			何事にもフットワークが軽く、大車輪の活躍を見ていて非常に頼もしく思っています。 下位者への指導も加わり大変だと思いますが、小野寺さんであれば大きな	開発部署との人脈作り、製品開発プロセス関連の知識向上、英語力向上に注力され、スキルアップを目指していただければと思います。

3-2.使用ツール・辞書

本研究では、複雑な統計処理を行うためのツールとして「R」を利用した。また、テキストマイニングの処理は「R」の中のパッケージ「RMeCab」を利用した。

「R」とは統計処理を行うためのソフトウェアであり、フリーで公開されている。統計処理を行う際にコードを記載しながら進める必要はあるが、Excel にプリインストールされている機能では実施できない高度な統計処理が可能である。また、R のもう一つの長所は、「パッケージ」と呼ばれるアドインツールが簡単に WEB 上でダウンロード・活用できることである。パッケージを上手に活用することで、コードを 1 から作成すると時間と手間がかかる統計処理でも、簡単に実施することができる。

次に、テキストマイニングで使用する様々な処理を実施するために「RMeCab」という R のパッケージを使用した。これは徳島大学の石田基広教授が作成したパッケージであり、日本語の形態素分析や共起表現の発見などを簡単に行うことが可能になる（形態素分析などの専門用語については後述する）。

また、ネットワーク図の作成には「Cytoscape」というソフトウェアを使用した。「Cytoscape」は、主に遺伝子分野でネットワーク図を作成するために用いられるソフトウェアであるが、本研究においても応用可能である。

テキストマイニングの処理を行う際には、2 種類の辞書を利用している。まず、MeCab に標準搭載されている IPA 辞書と呼ばれる辞書である。情報処理振興事業協会（IPA）が作成した、品詞体系をベースに日本語約 24 万語が登録されている辞書である。この辞書は一般に公開されており、自由な利用が可能である。

次に、単語が持つポジティブ・ネガティブの程度を判別するために東京工業大学高村大也氏が公開している辞書「単語感情極性対応表」（資料 2 参照）も利用している。この評価辞書の説明[2]を高村氏のホームページから引用して以下に記載する。

> 日本語および英語の単語とその感情極性の対応表を、 研究目的の利用に限り公開します。 感情極性とは、その語が一般的に良い印象を持つか（positive） 悪い印象を持つか（negative）を表した二値属性です。 例えば、「良い」、「美しい」などは positive な極性、 「悪い」、「汚い」などは negative な極性を持ちます。
>
> 感情極性値は、語彙ネットワークを利用して自動的に計算されたものです。 もともと二値属性ですが、-1 から+1 の実数値を割り当てました。 -1 に近いほど negative、+1 に近いほど positive と考えられます。 リソースとして、日本語は「岩波国語辞書（岩波書店）」を、 英語は WordNet-1.7.1 を使わせていただきました。

[2] 高村大也「単語感情極性対応表」（http://www.lr.pi.titech.ac.jp/~takamura/pndic_ja.html）（2014 年 7 月 12 日）

資料 2：単語感情極性対応表（1 部分のみ抜粋）

```
優れる:すぐれる:動詞:1
良い:よい:形容詞:0.999995
喜ぶ:よろこぶ:動詞:0.999979
褒める:ほめる:動詞:0.999979
めでたい:めでたい:形容詞:0.999645
賢い:かしこい:形容詞:0.999486
善い:いい:形容詞:0.999314
適す:てきす:動詞:0.999295
天晴:あっぱれ:名詞:0.999267
祝う:いわう:動詞:0.999122
功績:こうせき:名詞:0.999104
賞:しょう:名詞:0.998943
嬉しい:うれしい:形容詞:0.998871
喜び:よろこび:名詞:0.998861
才知:さいち:名詞:0.998771
徳:とく:名詞:0.998745
才能:さいのう:名詞:0.998699
素晴らしい:すばらしい:形容詞:0.998617
芳しい:かんばしい:形容詞:0.998578
称える:たたえる:動詞:0.998576
適切:てきせつ:名詞:0.998406
崇める:あがめる:動詞:0.998357
```

当辞書も公開されており、研究目的に限り使用が可能である。

3-3. 想定されるアウトプット

先述したデータ・ツールを利用することで、人材の評価方法として 2 種類の値を導くことが可能になる。

1 つ目は、従来から 360 度評価として採用されていた、アンケートの 5 段階評価を元に算出される得点である。ディアグロー社が行っているアンケートの質問（40 問）は被評価者の「個人特性・対人スキル・リーダーシップ・マネジメント」などの 8 種類のコンピテンシーを計る質問に分類され、それぞれのコンピテンシーが 100 点満点で算出される。また、算出された 8 種類の得点の算術平均を「被評価者に対する評価得点」として定義する。

2 つ目は、テキストデータから算出される得点である。具体的には、強みコメントからポジティブワードの総得点を算出し、弱みコメントからはネガティブワードの総得点を算出し、その合計得点を「被評価者に対する評価得点」と定義する方法である。

この 2 種類の得点から、評価関係を示すネットワーク図をそれぞれ作成し、「テキストマニングで評価した時の点数」と「360 度診断のアンケート点数で評価した時の点数」の比較・考察を行う。

最後に「どのようなコンピテンシーを持っていれば、テキストマイニングで算出された他者からの評価が高くなるのか？」という検討を加える。具体的には、被評価者のテキスト評価点数と 360 度診断のアンケートから取得した被評価者の 8 種類のコンピテンシー点数[3]の、それぞれの相関係数を観察する。「テキスト評価点数と関係性が強いコンピテンシーは何か？」を明確にすることで、職場間の評価関係において重要となる要素の特定を行う。

4. 研究手順

4-1. 全体構成

2-3 に記載したゴールイメージに至るためのプロセスは「①データ・辞書の準備」「②評価辞書に登録されていない単語の点数化」「③強み・弱みのコメントの得点化」「④総合得点の算出」「⑤ネットワーク図で評価関係を図示化」の大きく 5 パートに分けることができる。また、それぞれのパートの中でもさらに作業が分解されるので、以下では分けて進め方を記載する。

4-2. 各パート詳細

① データ・辞書の準備

ディアグロー株式会社から評価者・被評価者のデータを入手し、CSV の形で整理する。この時点で、名前データは匿名化し、セキュリティに配慮する。

資料 3：CSV の形で整理したデータ

被評価者氏	ユニークナ	被評価者と	被評価者部	被評価者の	被評価者の	評価者のコ	観察者評価	本人評価平	被評価者平	平均評価キ	個人_表層_	個人_深層_	対人_インプ
評価者1	被評価者8評価	上司	D04	設計部門での	グループ内の	他者の立場や	70.4	67	64.3	-2.7	52	68	72
評価者10	被評価者1評価	部下/後輩	生産センター	相談を持ち掛	私が感じる限り	基調なアドバイ	95	66.1	81.6	15.5	64	72	85
評価者11	被評価者6評価	同僚	生産センター	自分の立場を	問題解決に際	現在デイリーで	59	56	67.6	11.6	64	40	64
評価者12	被評価者4評価	部下/後輩	テクニカルセン	ひとつの出来事	気さくに話すこ	声が通るので、	78.6	67.3	74.6	7.3	56	60	72
評価者13	被評価者7評価	上司	テクニカルセン	だれとでも気さ	後輩に対する	緩慢になるのを	49.9	54.9	65.3	10.4	48	56	72
評価者14	被評価者8評価	同僚	H40	物事を論理的に	どうしたら皆が	組織、リーダー	52.5	67	64.3	-2.7	52	68	72
評価者15	被評価者10評	部下/後輩	営業部	得意先との交	芯が強い半面、	より周囲に注意	45.5	69	53.7	-15.3	68	76	80
評価者16	被評価者10評	同僚	営業部	英語が話せる	電話の応対に	今後、落ち着い	47	69	53.7	-15.3	68	76	80
評価者17	被評価者6評価	部下/後輩	生産センター	いつもアドバイ	いつもアドバイ	今回の課題を	87.9	56	67.6	11.6	64	40	64
評価者18	被評価者4評価	上司	テクニカルセン	計画的に物事	少し背伸びする	上司への依存	63.8	67.3	74.6	7.3	56	60	72
評価者19	被評価者2評価	上司	テクニカルセン	周囲との調和を	周囲との調和を	的確なご指摘	63.8	64	76.7	12.7	75	63	71
評価者2	被評価者7評価	同僚	テクニカルセン	業務における	感情の起伏が	長い間海外で	43.3	54.9	65.3	10.4	48	56	72
評価者20	被評価者9評価	同僚	生産センター	製品知識が他	グループの総括	個別内容の難	69.4	78.9	61	-17.9	76	76	80
評価者21	被評価者9評価	同僚	生産技術センタ	テクニカルセン	御忙しいと思い	おっしゃる通り	87.1	78.9	61	-17.9	76	76	80
評価者22	被評価者9評価	部下/後輩	品質保証センタ	品質保証という	ご本人の抱え	現在の業務に	40.8	78.9	61	-17.9	76	76	80
評価者23	被評価者8評価	部下/後輩	C11	問題に直面し	正論が多いた	相手の立場に	35.9	67	64.3	-2.7	52	68	72
評価者24	被評価者6評価	同僚	生産センター	業務の中での	現場の改善事	自分で判って	83.8	56	67.6	11.6	64	40	64
評価者25	被評価者1評価	部下/後輩	生産センター	小野寺さんは、	小野寺さんは、	基調なアドバイ	81.1	66.1	81.6	15.5	64	72	85
評価者26	被評価者1評価	部下/後輩	生産センター	小野寺さんは	自覚は無いと	基調なアドバイ	63.9	66.1	81.6	15.5	64	72	85
評価者27	被評価者8評価	同僚	D05	部下の仕事に	強みの跳ね返	業務の入り口と	80.9	67	64.3	-2.7	52	68	72
評価者28	被評価者3評価	同僚	MJ G110	業務内容として	直属の上司や	ご助言ありがと	68.9	43	74.9	31.9	44	48	60

次に、先述した R および RMeCab にてコメントデータの形態素分析を行う。
形態素分析とは、文章を意味のある単語に区切り、辞書を利用して品詞や内容を判別

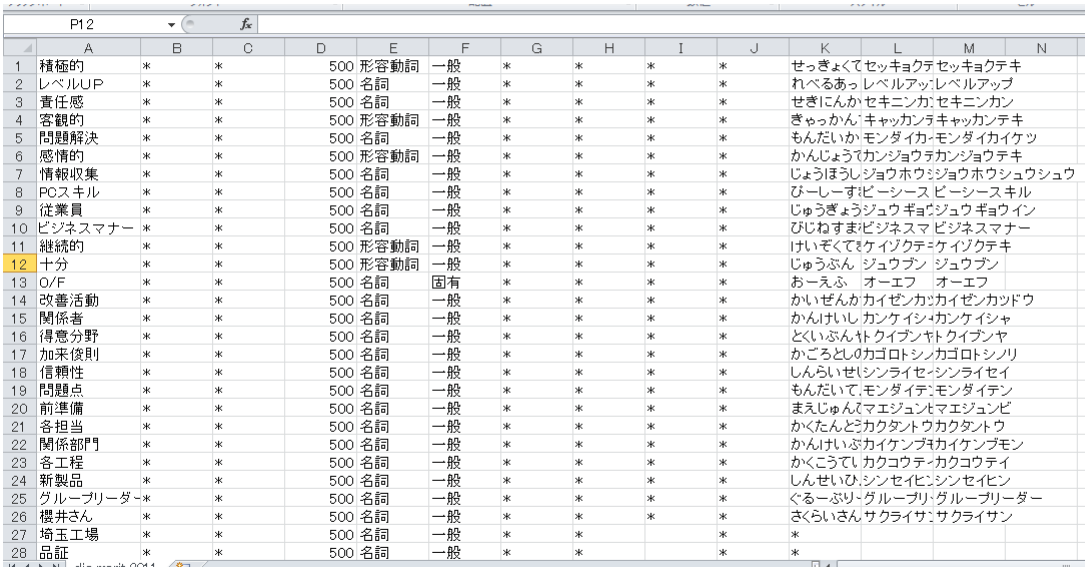
[3]個人特性（表層）・個人特性（深層）・対人スキル（アウトプット）・対人スキル（インプット）・リーダーシップ（思考）・リーダーシップ（行動）・マネジメント（仕事）・マネジメント（人）の 8 種類の点数

することである。「形態素」とは、「文章の要素のうち、意味を持つ最小の単位」のことである。例えば、「積極的」という単語であれば、「積極」という名詞と「的」という名詞に分解できる。また辞書は、RMeCab に標準で搭載されている辞書の他、ユーザーの好みで別の辞書データを使用することも可能である。当処理によって、単語を品詞レベルで分解することが可能になる。

次に、テキストの解読精度を向上させるため、RMeCab 内の IPA 辞書の編集を行う。例えば、分析者が「積極的」という単語を「形容動詞」として解釈したい場合でも、辞書が無編集の状態だと「積極」と「的」という「名詞」として解釈される。このような場合に分析者の意向を反映させるため、MeCab では抽出する単語に重み付けを与えられる機能が存在する。例えば先述の状況に対応するためには、「積極的」という「形容動詞」を、重み付けを強めに設定すれば、「積極」「的」と分割した名詞で認識することなく、「積極的」という形容動詞の形で読み込んでくれる。重み付けの設定は Excel 上でデータを作成し、そのデータを MeCab で読み込むことで反映される（資料 4 参照）。

本研究ではレコード数が少なく、目で全てのテキストデータを確認できる程度のボリュームなので、形態素分析を実施した後に全ての内容を目で確認し、名詞・形容動詞を中心に不自然に分割されている単語は辞書を使って修正している。

資料 4：辞書編集用の Excel データ



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	積極的	*	*	500	形容動詞	一般	*	*	*	*	せっきょくてセッキョクテ	セッキョクテキ		
2	レベルUP	*	*	500	名詞	一般	*	*	*	*	れべるあっ	レベルアップ	レベルアップ	
3	責任感	*	*	500	名詞	一般	*	*	*	*	せきにんかん	セキニンカン	セキニンカン	
4	客観的	*	*	500	形容動詞	一般	*	*	*	*	きゃっかんて	キャッカンテキ	キャッカンテキ	
5	問題解決	*	*	500	名詞	一般	*	*	*	*	もんだいかい	モンダイカイ	モンダイカイケツ	
6	感情的	*	*	500	形容動詞	一般	*	*	*	*	かんじょうてき	カンジョウテキ	カンジョウテキ	
7	情報収集	*	*	500	名詞	一般	*	*	*	*	じょうほうしゅう	ジョウホウシュウ	ジョウホウシュウシュウ	
8	PCスキル	*	*	500	名詞	一般	*	*	*	*	ぴーしーすきる	ピーシースキル	ピーシースキル	
9	従業員	*	*	500	名詞	一般	*	*	*	*	じゅうぎょういん	ジュウギョウイン	ジュウギョウイン	
10	ビジネスマナー	*	*	500	名詞	一般	*	*	*	*	びじねすまなー	ビジネスマナー	ビジネスマナー	
11	継続的	*	*	500	形容動詞	一般	*	*	*	*	けいぞくてき	ケイゾクテキ	ケイゾクテキ	
12	十分	*	*	500	形容動詞	一般	*	*	*	*	じゅうぶん	ジュウブン	ジュウブン	
13	O/F	*	*	500	名詞	固有	*	*	*	*	おーえふ	オーエフ	オーエフ	
14	改善活動	*	*	500	名詞	一般	*	*	*	*	かいぜんかつどう	カイゼンカツドウ	カイゼンカツドウ	
15	関係者	*	*	500	名詞	一般	*	*	*	*	かんけいしゃ	カンケイシャ	カンケイシャ	
16	得意分野	*	*	500	名詞	一般	*	*	*	*	とくいぶんや	トクイブンヤ	トクイブンヤ	
17	加来俊則	*	*	500	名詞	一般	*	*	*	*	かごろとしのり	カゴロトシノリ	カゴロトシノリ	
18	信頼性	*	*	500	名詞	一般	*	*	*	*	しんらいせい	シンライセイ	シンライセイ	
19	問題点	*	*	500	名詞	一般	*	*	*	*	もんだいてん	モンダイテン	モンダイテン	
20	前準備	*	*	500	名詞	一般	*	*	*	*	まえじゅんび	マエジュンビ	マエジュンビ	
21	各担当	*	*	500	名詞	一般	*	*	*	*	かくたんとう	カクタントウ	カクタントウ	
22	関係部門	*	*	500	名詞	一般	*	*	*	*	かんけいぶもん	カイケンブモン	カイケンブモン	
23	各工程	*	*	500	名詞	一般	*	*	*	*	かくこうてい	カクコウテイ	カクコウテイ	
24	新製品	*	*	500	名詞	一般	*	*	*	*	しんせいひん	シンセイヒン	シンセイヒン	
25	グループリーダー	*	*	500	名詞	一般	*	*	*	*	ぐるーぷりーだー	グループリーダー	グループリーダー	
26	櫻井さん	*	*	500	名詞	一般	*	*	*	*	さくらいさん	サクライサン	サクライサン	
27	埼玉工場	*	*	500	名詞	一般	*	*		*	*			
28	品証	*	*	500	名詞	一般	*	*		*	*			

② 評価辞書に登録されていない単語の点数化

単語毎のスコア付与には、2-2 で言及した極性辞書（評価辞書）に登録されている単語の点数をそのまま使用する。しかし、「評価辞書には登録されていないが、評価

コメント内で頻出する単語」も存在する。例えば「面倒見」や「リーダーシップ」などの単語は評価辞書内に登録されていないが、多くの評価者がこれらの単語を使用しており、既に存在している評価辞書に合わせた形で点数化を行わないと精度が低くなることが懸念された。

これらの単語の得点化はKBS学生に対してアンケートを通じて行った。具体的には、資料5のようなWEBアンケートフォームを作成し、KBS学生M36/M37に単語の印象を入力してもらった。対象となった単語は全体で20単語であり、大きく2つに大別される。1つ目は評価辞書に登録されておらず、かつ全体のテキストの中で出現頻度が2回以上の単語であり、これが15単語、次に、既に評価辞書に入っている単語をランダムで抜き出した5単語である。既存の評価辞書の単語をアンケート内で同じように評価してもらうことで、未知の単語のアンケート評価と既存の単語のアンケート評価を接続させることが可能となる。また、アンケートは2014年6月12日（木）～6月16日（月）で実施し、57人の回答が得られた。

資料5：実施したアンケート

言葉が持つポジティブ具合・ネガティブ具合に関するアンケート

*必須

以下の言葉について抱く印象を、5段階で回答してください。 *

深く考えず、直感で回答してください。

	とてもネガティブ		ネガティブ		何も思わない		ポジティブ		とてもポジティブ
面倒見	○	○	○	○	○	○	○	○	○
的確	○	○	○	○	○	○	○	○	○
成長	○	○	○	○	○	○	○	○	○
上手く	○	○	○	○	○	○	○	○	○
弱い	○	○	○	○	○	○	○	○	○
明確	○	○	○	○	○	○	○	○	○
纏め	○	○	○	○	○	○	○	○	○
もう少し	○	○	○	○	○	○	○	○	○
リーダーシップ	○	○	○	○	○	○	○	○	○
周り	○	○	○	○	○	○	○	○	○
取り組む	○	○	○	○	○	○	○	○	○
良く	○	○	○	○	○	○	○	○	○
お願い	○	○	○	○	○	○	○	○	○

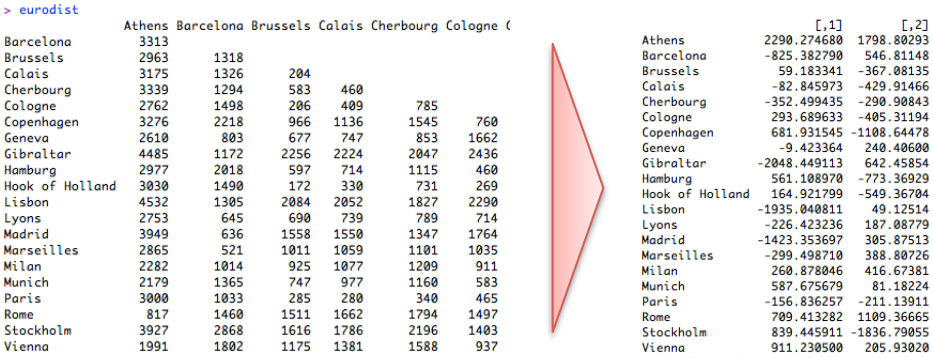
取得したアンケートデータは、まず 1（とてもネガティブ）～9（とてもポジティブ）の数字で定量化を行い（資料 6 参照）、次に多次元尺度法という統計手法を使用して単語間の距離を算出した。

資料 6：整理後のアンケートデータ

	A	B	C	D	E	F	G	H	I	J
1		学生1	学生2	学生3	学生4	学生5	学生6	学生7	学生8	学生9
2	[優れる]	9	8	9	9	9	8	9	7	9
3	[積極的]	8	7	7	7	9	9	9	8	9
4	[成長]	7	7	9	7	9	7	9	8	7
5	[リーダーシッ	8	8	5	9	9	9	9	9	8
6	[明確]	8	8	9	9	9	7	9	8	9
7	[豊富]	8	8	9	9	9	8	7	7	8
8	[温和]	7	6	5	5	7	9	9	6	6
9	[的確]	7	8	9	9	7	8	9	8	7
10	[良く]	7	7	5	7	7	6	7	6	7
11	[上手く]	7	8	7	7	7	9	9	8	7
12	[面倒見]	7	7	3	7	9	9	9	9	5
13	[取り組む]	6	5	7	5	7	9	9	6	6
14	[まとめ]	6	7	5	5	5	7	9	5	6
15	[お願い]	7	6	7	5	5	5	5	5	6
16	[周り]	5	5	5	5	5	7	5	5	5
17	[巨体]	5	6	3	5	5	5	3	3	6
18	[もう少し]	3	3	7	3	5	4	7	4	3
19	[猛獣]	3	5	3	3	3	4	3	3	4
20	[弱い]	2	2	3	1	5	3	2	2	1
21	[悪い]	1	2	7	1	1	2	1	3	1

多次元尺度法とは、個体間の距離データを低次元に集約する手法である。例えば、欧州各国間の距離データを多次元尺度法で 2 次元に集約し、それを 2 次元空間上に配置すると、地図で各国間の距離を表現したような形になる。（資料 7 参照）

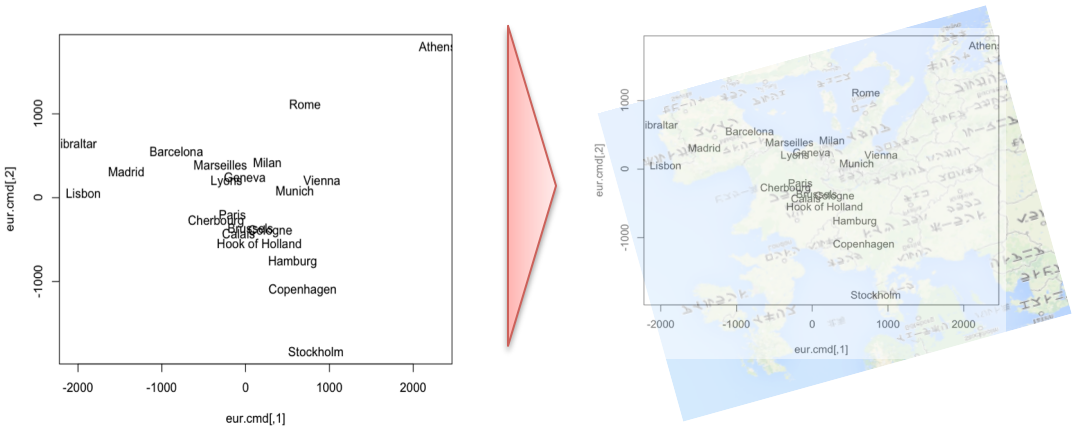
資料 7：多次元尺度法の例（欧州各国間の距離）



```
> eurodist
               Athens Barcelona Brussels Calais Cherbourg Cologne (
Barcelona        3313
Brussels         2963      1318
Calais           3175      1326      204
Cherbourg        3339      1294      583    460
Cologne          2762      1498      206    409       785
Copenhagen       3276      2218      966   1136      1545     760
Geneva           2610       803      677    747       853    1662
Gibraltar        4485      1172     2256   2224      2047    2436
Hamburg          2977      2018      597    714      1115     460
Hook of Holland  3030      1490      172    330       731     269
Lisbon           4532      1305     2084   2052      1827    2290
Lyons            2753       645      690    739       789     714
Madrid           3949       636     1558   1550      1347    1764
Marseilles       2865       521     1011   1059      1101    1035
Milan            2282      1014      925   1077      1209     911
Munich           2179      1365      747    977      1160     583
Paris            3000      1033      285    280       340     465
Rome              817      1460     1511   1662      1794    1497
Stockholm        3927      2868     1616   1786      2196    1403
Vienna           1991      1802     1175   1381      1588     937
```

```
                    [,1]        [,2]
Athens       2290.274680  1798.80293
Barcelona    -825.382790   546.81148
Brussels       59.183341  -367.08135
Calais        -82.845973  -429.91466
Cherbourg    -352.499435  -290.90843
Cologne       293.689633  -405.31194
Copenhagen    681.931545 -1108.64478
Geneva         -9.423364   240.40600
Gibraltar   -2048.449113   642.45854
Hamburg       561.108970  -773.36929
Hook of Holland 164.921799 -549.36704
Lisbon      -1935.040811    49.12514
Lyons         -226.423236   187.08779
Madrid      -1423.353697   305.87513
Marseilles    -299.498710   388.80726
Milan          260.878046   416.67381
Munich         587.675679    81.18224
Paris         -156.836257  -211.13911
Rome           709.413282  1109.36665
Stockholm      839.445911 -1836.79055
Vienna         911.230500   205.93020
```

①個体間の距離を示す行列を、任意の n 次元に整理できる（この例は 2 次元）



②2 次元空間上にプロットすると、地図で示される位置関係とほぼ同様になる。

多次元尺度法は通常、比率尺度を持つデータに対して適用されるが、順序尺度を持つデータにも対応するために「非計量多次元尺度法」という手法が存在する。本研究のアンケートデータは比率尺度ではなく順序尺度であり、かつ 2 次元にプロットすることが目的でなく、1 次元で表現される評価辞書の中に新しい単語を登録することが目的であるので、非計量多次元尺度法を利用して評価辞書に出てこない単語の数値（資料 6 参照）を 1 次元に縮尺する。上記の計算は、エクセル内には関数が用意されていないが、R では MASS というライブラリの isoMDS という関数でこの処理を簡単に行うことができる。以下に R で行った計算の結果を示す（資料 8）。

資料 8：R での計算結果

```
> library(MASS)
> cmddist <- isoMDS(score.dist,k=1) #kは次元数
initial  value 6.951617
final  value 6.267701
converged
> cmddist #以降はこの距離データをExcelで編集して、単語を得点化する。
$points
                    [,1]
 [優れる]       -6.8041620
 [積極的]       -6.0288053
 [成長]         -6.0862590
 [リーダーシップ] -6.4859069
 [明確]         -5.6914964
 [豊富]         -4.8498503
 [温和]         -4.0578173
 [的確]         -3.7444104
 [良く]         -3.9657565
 [上手く]       -3.3525094
 [固倒見]       -1.7186608
 [取り組む]     -1.0582438
 [まとめ]        0.2685122
 [お願い]        2.4400701
 [周り]          2.7934905
 [巨体]          5.8188723
 [もう少し]      7.1105341
 [猛獣]          7.8661489
 [弱い]         13.2837682
 [悪い]         14.2624818

$stress
[1] 6.267701
```

Stress とは、モデルの当てはまりを示す数値であり、一般的には 10 で「まずまず(fair)」、5 で「よい(good)」と判断することができる。

単語間の距離を算出できたので、次に評価辞書に含まれていた既存の単語の得点を参考に、未知の単語の距離を評価辞書と同じスケールに修正（資料 9）し、修正後の

距離を未知の単語の得点とした。

資料 9：距離のスケールを評価辞書の+1～-1 に合わせて修正

	A	B	C	D
1		MDS距離	ベンチマーク得点	修正後得点
2	[優れた]	-7.2428404	1	1
3	[積極的]	-6.4354503		0.871373215
4	[成長]	-6.4051934		0.866552933
5	[リーダーシップ]	-6.1067927		0.819014175
6	[明確]	-5.8407116		0.776624311
7	[豊富]	-4.6515789		0.587181417
8	[温和]	-4.1043413	0.5	0.5
9	[的確]	-3.6233073		0.475738759
10	[良く]	-3.7943553		0.484365668
11	[上手く]	-3.291689		0.459013389
12	[面倒見]	-1.7219532		0.379842814
13	[取り組む]	-0.7383078		0.330232063
14	[まとめ]	0.4595925		0.269815236
15	[お願い]	2.6760659		0.158026059
16	[周り]	3.0781294		0.137747743
17	[巨体]	5.8092901	0	0
18	[もう少し]	7.3031284		-0.490889272
19	[猛獣]	7.3308535	-0.5	-0.5
20	[弱い]	13.1213969		-0.922839562
21	[悪い]	14.1780645	-1	-1

ここまでの処理で「評価辞書で得点化されていた単語」「得点化されていないが、テキスト上で頻出する単語」双方の得点化を行うことができた。

③ 強み・弱みコメントの得点化

次に、上記で定義した単語のスコアと、強み・弱みテキストを形態素分析で分割した単語の紐付けを行った。その際、R のコードは、「R ではじめる Twitter 解析」（R ユーザ会）を参考にした。使用した R のコードは、当論文の付録として巻末に記載する。

コメントを得点化する方法は、テキストマイニングの不確実性を軽減するために工夫を加えている。コメントは「強みコメント」と「弱みコメント」の 2 種類が取得できるので、「強みコメント」からはポジティブな単語のみ抽出、「弱みコメント」からはネガティブな単語のみ抽出しそれぞれの合計点数を算出している。例えば「強みコメント」に「決断力が弱いということは全くない」というコメントがあった場合、先述した工夫を行わないと「弱い」というネガティブな単語を抽出してしまい、精度が下がってしまう。強みコメントはポジティブなことを記載しているはず、弱みコメントはネガティブなことを記載しているはずという前提を置いた上で、このような工夫を行った（資料 10 参照）。

資料 10：「強み・弱みコメントを区切って点数化」する工夫の説明

単語毎に区切って点数化
強みコメントからはポジティブワード
弱みコメントからはネガティブワードのみ抽出。

※フリー素材サイトから画像抜粋

相談を持ち掛けると、親身になってアドバイスして貰える事がとても嬉しいです。とても信頼できる上司です。

親身＝0.5点　アドバイス＝0.3点　信頼＝0.8点
強みコメント合計：1.6点

その反面、ポカミスやうっかり忘れが多いと感じる事があります。依頼事項、約束事は即メモを取る習慣をつける必要性を感じます。

ポカミス＝-0.7点　うっかり忘れ＝-0.8点
弱みコメント合計：-1.5点

合計評価：0.1点

④総合得点の算出

強み・弱みコメントから算出した得点を合計して、総合得点を計算する（資料 11 参照）。前ステップで算出した強み・弱みのスコアはそれぞれ平均・バラツキ共に異なる値のため、そのまま加算して総合評価にはせず、強み・弱みのスコアを標準化した後に加算して総得点とする。

総得点が高ければ、その評価者のテキスト評価からは強みを表現する言葉が多く見られる、もしくは弱みを表現する言葉が少ない、もしくはその両方であることを示しており、総得点が低ければ、強みを表現する言葉が見られない・弱みを表現する言葉が多い、もしくはその両方であることを示している。

資料 11：強みスコア・弱みスコア・総得点

	評価者氏名	被評価者氏名	被評価者の強み	強みスコア（標準化）	被評価者の弱	弱みスコア	総得点
2	被評価者4	評価者46	櫻井太源さんは部下	1.693183499	私が特に不満	1.321100779	3.014284278
3	被評価者1	評価者10	相談を持ち掛けると、	1.430501015	私が感じる限り	1.327279834	2.757780849
4	被評価者9	評価者22	品質保証という業務	1.776595436	ご本人の抱えて	0.799770277	2.576365714
5	被評価者5	評価者34	得意先からの無理難	2.161389203	英会話のレベ	0.338706254	2.500095456
6	被評価者1	評価者52	何事にもフットワーク	1.378636371	開発部署との	1.139436077	2.518072449
7	被評価者1	評価者25	小野寺さんは、責任	1.211277941	小野寺さんは、	0.858144962	2.069422902
8	被評価者7	評価者2	業務における社会人	2.105605581	感情の起伏が	-0.21270416	1.89290142
9	被評価者6	評価者35	困った時に相談でき	1.116553158	相手によっては	0.600930904	1.717484062
10	被評価者10	評価者6	責任感が強く、依頼	0.888623909	従来のやり方	0.768130034	1.656753943
11	被評価者8	評価者45	設計あがりであり、製	0.087786712	相談しにくい雰	1.713075855	1.800862567
12	被評価者6	評価者24	業務の中での相談先	0.834601221	現場の改善事	0.793785073	1.628386295
13	被評価者2	評価者43	グループの状況が良	0.76548595	謙虚ではあるが	0.617166813	1.382652763
14	被評価者3	評価者8	計画を立て社内を巻	1.277284613	是非、自分の考	-0.016027314	1.261257299
15	被評価者10	評価者56	新たな顧客を担当す	0.500843887	とても親切に指	0.91035406	1.411197946
16	被評価者5	評価者8	周囲や仲間にも常に	0.401476813	自分の弱みをし	0.887359258	1.288836071
17	被評価者1	評価者26	小野寺さんは周囲の	0.562549759	自覚は無いと思	0.277358284	0.839908043

次に、算出した総合得点が妥当かどうかについて、具体例を出しながら観察する。

【最高得点の強みコメント】
得意先からの無理難題に対して、異議は明確に回答し、調整出来る事は迅速に検討し、回答している。 又、得意先固有の風土や伝統を理解し、業務を的確に対応している。営業の業務以外も積極的に関わり、フォローしており関係各部署の信頼を得ている。自部署内でも後輩の業務でトラブルが発生した場合に的確なアドバイスをしており、上司、同僚からの信頼も厚い。

【最低得点の強みコメント】
英語が話せるので、海外との生産工場（ＭＥＣＮ、ＭＥＣＮ，ＭＥＣＴ等）との生産開発依頼等早急な対応が可能。

双方の文を見比べると、前者のコメントの得点が高くなり、後者のコメントの得点が低くなるのは直感的に納得できるのではないだろうか。前者は「明確」「迅速」「的確」「信頼」などのポジティブワードが多用されているのに比べ、後者はそのようなワードはあまり多く見られない。

当然、評価者によってコメントの傾向に違いはあるだろう。例えば、上記のコメントで最低評価のコメントを記載した評価者は他人を評価する際の基準が厳しく、最高評価のコメントを得られた被評価者を評価した際にも同じように低評価のコメントを記載する可能性がある。ただし、本研究では同じ評価者が複数の被評価者を評価することを行っていないため、「被評価者の変更によって評価者の評価の傾向はどの程度変動するのか」は検証できなかった。当検証は今後の課題として考えられる。

また参考までに、弱みコメントについても最高得点のコメントと最低得点のコメントを掲載しておく。最高得点・最低得点となっているのが、それぞれ直感的に納得できるのではないだろうか。

【最高得点の弱みコメント】
私が感じる限り、とても気がきく頼りになる上司ですので、指摘することなどございません。いつも優しく対応していただきとても感謝しています。

【最低得点の強みコメント】
気さくに話すことはとても良い事だと思うが、たまに、それがまわりの環境ではとても迷惑になっている場合がある。例えば、まわりで重要な話をしている時に雑談の声が大きく聞こえなかったりや、仕事でこんつまっている時に大きな声で業務外の話を

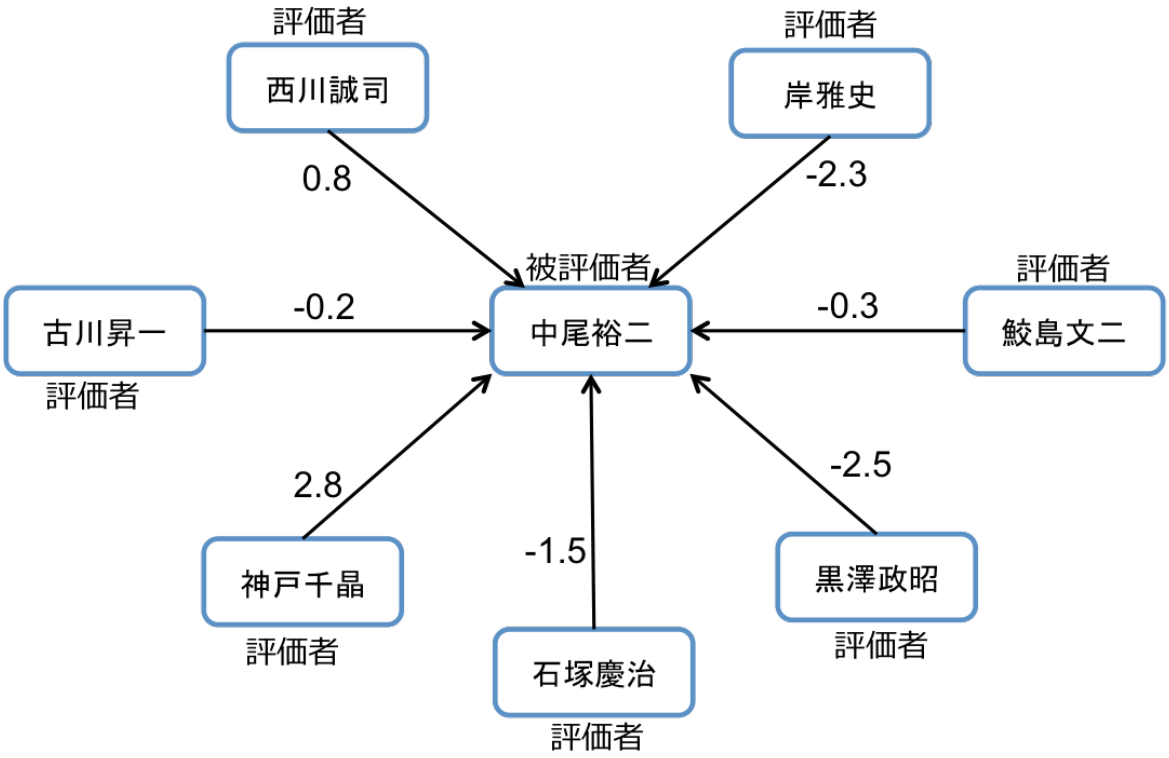
していたりなど。これから更に上の立場で仕事をされる場合、自分のグループ以外にも配慮できるようになると、より一層まわりからの信頼度も増すと思う。

⑤ネットワーク図で評価関係を図示化

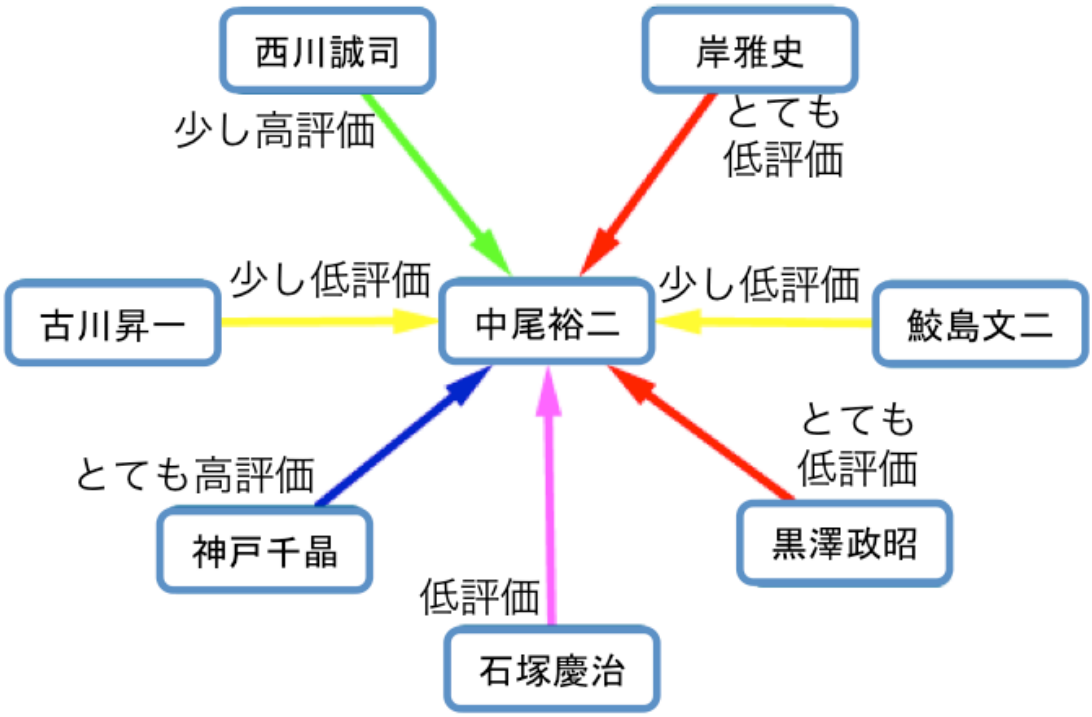
前ステップで算出した各評価者の点数を、当論文 3-2「使用ツール・辞書」のセクションで言及した「Cytoscape」を使用して、資料 13 のようなネットワーク図にした。

最初は資料 12 のような形を考えていたが、非常にわかりづらいので、資料 14 のように見やすくした。具体的には、評価の点数を点数順に 6 分割し、上位の点数群からそれぞれ「とても高評価」「高評価」「少し高評価」「少し低評価」「低評価」「とても低評価」と名前を付け、「とても高評価」は青色の矢印、「とても低評価」は赤色の矢印で表現し、少しずつ色を変えて表現する見せ方に変更した。このような工夫を加えることで「評価者から見た被評価者の評価関係」をより明確に示すことができる。

資料 12：最初に作ったネットワーク図



資料 13：改良したネットワーク図



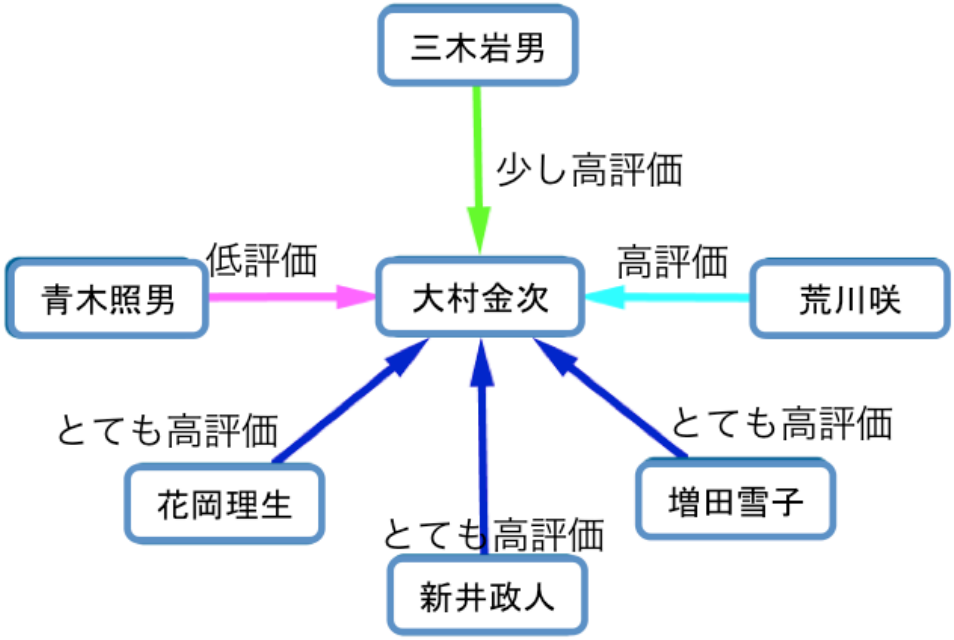
5. 結果

5-1. 結果の概要

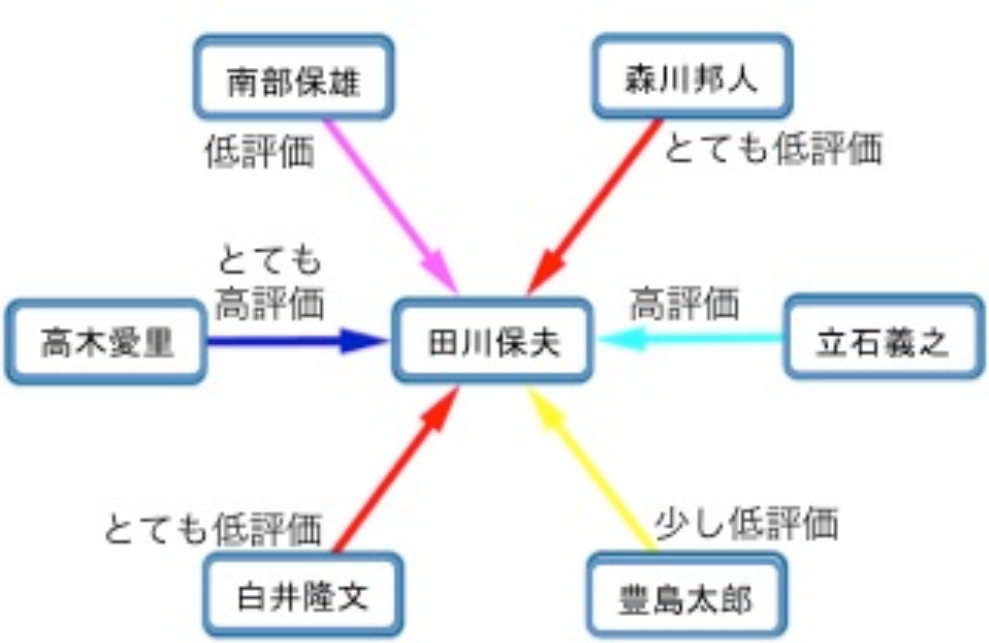
これまでの分析を通じて、テキスト情報から被評価者を取り巻く人間の評価関係を簡単なネットワーク図で図示することに成功した。また、図示したことで、各被評価者の評価の特徴も簡単に把握することが可能になった。

例として、特徴的なパターンを持つ 3 人の被評価者（資料 14～16）を比較することにする。なお、これから出てくる人名は全て偽名である。

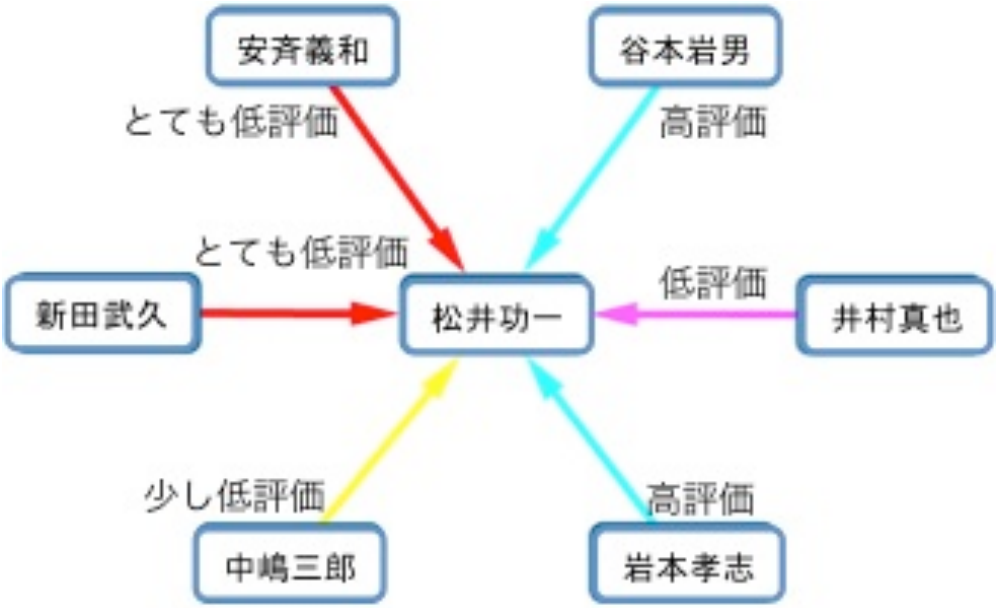
資料 14：大村さんの評価ネットワーク



資料 15：田川さんの評価ネットワーク



資料 16：松井さんの評価ネットワーク



この 3 者の例からは、大村さんは「後輩からの評価が良い」、田川さんは「同僚からの評価が良い」、松井さんは「上司・後輩からは好き嫌いが別れ、同僚からの評価は低い」という情報が読み取れる。このように、可視化を行うことで、職場間で誰がどのように評価し合っているかを簡単に表現することが可能となる。

5-2. 従来の評価との比較

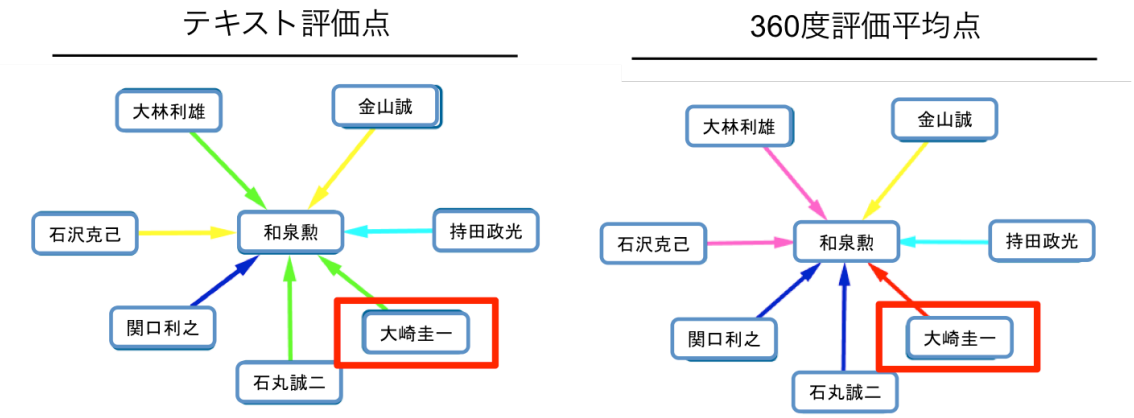
次に、アンケートで定量化した得点との比較・観察を実施する。比較しやすいようにアンケート評価の点数もテキスト評価と同じようにネットワーク図による図示化を行う。その上で、従来の評価法との「類似点」「相違点」を分けて論じる。

まず「類似点」については、テキストで評価した各被評価者のスコアとアンケートで評価した各被評価者のスコアとの相関係数を計算したところ、0.22（$P<0.01$）という値が算出された。この結果から、「わずかではあるが、テキスト評価とアンケート評価では類似性が見られる」と言う事ができる。実際に筆者が目で見ても、アンケートで高く評価されている被評価者はテキストでも高く評価される傾向が有ると判断できる。

次に「相違点」については、さらに 2 種類に分解することができる。1 つ目は「アンケートで評価した点数の方が組織内の関係性を実態に沿った形で表現しているので違いが発生しているケース」である。これは、本研究で採用しているテキストマイニング手法では文章の意味の解釈を人間と同じレベルで完璧に解釈することはでき

ないことから生じている。実際のコメント例を観察すると分かりやすい。例えば、資料 17 の和泉さん（被評価者）に対する大崎さん（評価者）の評価がそれに当てはまる。

資料 17：和泉さん（被評価者）の評価ネットワークと、大崎さん（評価者）が書いた強み・弱みコメント



【強み】
達成出来た時の喜びを常に持ち続け、意欲的にこれからも進んで頂きたい。年上の人でも叱る時はしっかり叱れる所は良いです。

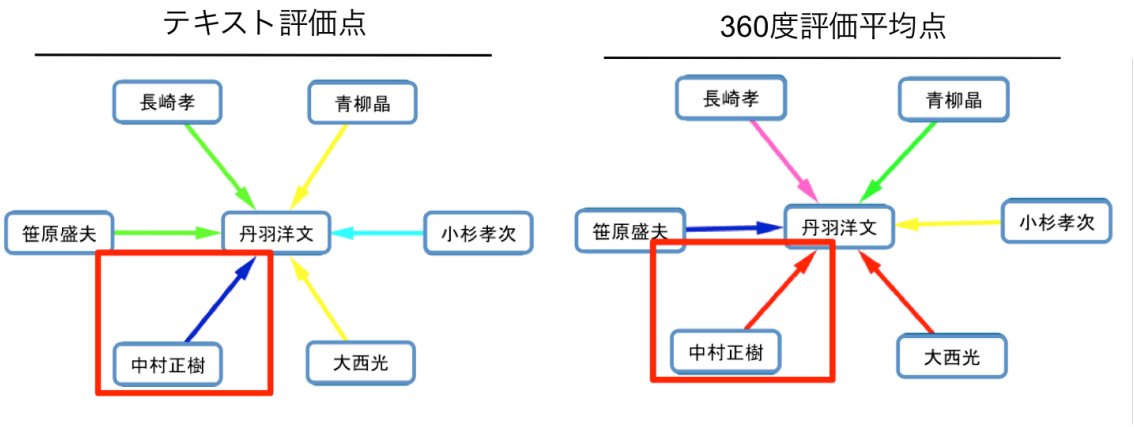
【弱み】
自分の思いを共有してみんなと一緒に取り組んでもらいたい。ワンマンプレイ的なところがあります。みんなが見ています。模範となるような言動・態度を取ってください。上司だけでなく部下ともコミュニケーション取って活動を進めてください。

上記のコメントは、和泉さんを評価する他の評価者のコメントを読んで比較すると、厳しく評価しているコメントなので（あくまで筆者の目で見た判断だが）、どちらかと言えばアンケートの結果が導いている「低評価」の方が妥当で、テキストによる評価で「高評価」となっているのは納得度が低いように思える。このような差異が生まれてくるのは、本研究において採用したテキストマイニング手法の限界によるものだと考えられる。例えば、資料 17 の弱みテキストからは「ワンマンプレイ」という単語しかネガティブワードを抽出できない。しかし、実際に目で読むと「みんなが見ています。模範となるような言動・態度を取ってください。」という厳しい表現があり、この表現でネガティブさを増しているような印象を、人の目だと感じることができる。しかし、テキストマイニングを使用した評価だとその微細なニュアンスまでは汲み取

ることが難しい。

2 つ目は「テキスト評価の方が組織内の関係性を実態に沿った形で表現できているので違いが発生しているケース」である。先ほどの例とは逆で、アンケート評価での数値化の限界、つまり「感情」を読み取ることができないことから発生するケースである。例えば、資料 18 の丹波さん（被評価者）に対する中村さん（評価者）の評価がそれに当てはまる。

資料 18：丹羽さん（被評価者）の評価ネットワークと、中村さん（評価者）が書いた強み・弱みコメント



【強み】
品質保証という業務に必要とされる、製品知識（構造、機能、特性等）が非常に豊富であり、何か分からない事があって問い合わせをすると細かな事まで教えて頂き感謝しております。まだ、業務に慣れていない自分に対して、厳しい時もありますが成果が出るよう毎回、フォローして下さり助かっております。

【弱み】
ご本人の抱えている業務内容が、明確化できていないので我々が少しでもサポートできるよう明確化して頂きたい。

中村さんは、アンケートでは丹羽さんに対して低評価をつけているが、テキストを読む限りでは高く評価しているように見える。また、他の評価者とテキストの内容を比較しても、高く評価しているように見える。ここから、アンケートではなく、テキストで評価した方がより実態に沿った評価を導いていそうに思われる。

ここまでで相違点を 2 種類のケースに分類したが、その 2 種類の「アンケートで評価した点数の方が組織内の関係性をうまく表現しているので違いが発生しているケース」と「テキスト評価の方が組織内の関係性をうまく表現しているので違いが発生しているケース」の数を比較すると、だいたい 1：3 くらいの比率で後者の「テキスト評価の方がうまく関係性を表現している」ケースが多いように見受けられた。

では、結局どちらの手法が「職場間従業員の評価関係」をうまく表現できているのであろうか。そもそも「何をもってうまく表現できていると言えるのか」という問いに答えるための客観的な基準は現状存在せず、評価関係の巧拙を定量的に把握することは困難である。よって、主観的にはなるが「職場の人間関係がよく把握できている人から見た時の納得感」の高低が、評価関係を正しく表現できているかを計測する基準となるであろう。本研究ではテキストデータの提供企業に訪問し、評価者・被評価者の人物関係を把握している人事部の社員に対して、テキスト評価で作ったネットワーク図とアンケート評価で作ったネットワーク図を比較して頂いた。その結果「テキストの評価と職場内の被評価者の人物を照らして考えると、少なくともテキスト評価で表現した方が職場内の評価を正しく表現できていそうだと思う」という回答を頂いた。

この評価が本当に正しいかどうかを検証するのであれば、1 社だけで印象を確認するのでは無く、より多くの企業で同じような比較を実施し、印象を聞く必要があるだろう。ただ、今回得られた結果だけでも、テキストマイニングを使った評価を使用することで職場従業員間の関係性がより実態に沿った形で表現できる可能性がある、とは言えるであろう。

5-3. テキスト評価と各コンピテンシーの関連性

前項までで従来の評価方法と、テキストによる評価方法の違いについてある程度の示唆を出すことができた。この項では、テキスト分析によって算出された評価点を、もう一段踏み込んで研究する。具体的には、360 度診断のアンケートで取得する 8 種類のコンピテンシー点数はテキスト評価の点数とある程度関係性がある、という前提を置いた上で、テキストマイニングの各被評価者の評価点とアンケートで取った 8 種類のコンピテンシーの値のそれぞれの相関係数を算出し、テキスト評価と各項目の関連性を観察する。この観察を通じて「被評価者を評価する際に、評価者は具体的にどのような要素を確認しているのか」を明確にする。

分析する際には、評価者を上司・同僚・後輩に分類する。何故かというと、上司・

同僚・後輩と、立場が違うと評価する軸も異なってくると考えられるからである。例えば上司は、被評価者のリーダーシップ面に注目する傾向があって、部下は、自分たちと接する際のコミュニケーション能力面を注目する傾向がある等、それぞれのポジションによって着目するポイントが違うことが考えられるからである。

また、そもそも上司・同僚・後輩で分けたテキスト評価点のデータに妥当性があるのかを検証するために、心理学で妥当性を検証するために使われる手法の 1 つである MTMR 行列を利用して、妥当性が高いデータであることを検証した。同手法と結果の詳細は付録に付記する。

上記の考え方に基づいて、上司・同僚・後輩に分けて、各評価者の「テキスト評価の総合点」と「8 要素のコンピテンシー点数」の相関分析を実施した。この分析も R によって実施した（資料 19〜24 参照)。

結果として、上司のテキスト評価の点数は 360 度評価の「リーダーシップ（行動)」面の点数と関連性が強いことがわかった（ただし相関係数の P 値は 0.16 と統計的に完全に信頼が置けるわけではないことに注意)。同僚からのテキスト評価の点数は「仕事のマネジメント」「リーダーシップ（思考)」「対人能力（インプット)」の点数と関連が強い。また、後輩からのテキスト評価の点数はどの要素も関連しているとはいえない、という結果になった。

この結果から、従業員を評価する際、上司は「リーダーとしての行動そのもの」、同僚は「リーダーとしての行動に至るまでの思考」や「コミュニケーション」を評価の軸として置いていることが推測される。また後輩は、それぞれの評価者で評価をする軸が異なっている、つまり統一された評価軸がまだ定まっていないことが考えられる。

ただし、本研究の分析はサンプル数が少なく（上司からの評価 n=19、同僚からの評価 n=19、後輩からの評価 n=23)、またある 1 社の中でのデータを用いているだけであるので、正確な推定を行うにはより大きなサンプル数、様々な業種・職種・規模の企業で検証を行う必要があることを付記しておく。

資料 19：上司の評価点数と各要素の相関行列

評価者:上司	総得点	個人 表層	個人 深層	対人 インプット	対人 アウトプット	リーダー 思考	リーダー 行動	マネジ 仕事	マネジ 人
総得点	1	0.29	0.02	0.22	0.04	0.29	0.33	0.23	−0.04
個人_表層	0.29	1	0.49	0.52	0.62	0.77	0.59	0.62	0.62
個人_深層	0.02	0.49	1	0.66	0.42	0.65	0.3	0.34	0.78
対人_インプット	0.22	0.52	0.66	1	0.62	0.77	0.68	0.78	0.74
対人_アウトプット	0.04	0.62	0.42	0.62	1	0.62	0.6	0.79	0.63
リーダー_思考	0.29	0.77	0.65	0.77	0.62	1	0.6	0.81	0.7
リーダー_行動	0.33	0.59	0.3	0.68	0.6	0.6	1	0.76	0.48
マネジ_仕事	0.23	0.62	0.34	0.78	0.79	0.81	0.76	1	0.59
マネジ_人	−0.04	0.62	0.78	0.74	0.63	0.7	0.48	0.59	1

資料 20：各相関係数の P 値

評価者:上司(P値)	総得点	個人 表層	個人 深層	対人 インプット	対人 アウトプット	リーダー 思考	リーダー 行動	マネジ 仕事	マネジ 人
総得点	0	1	1	1	1	1	1	1	1
個人_表層	0.23	0	0.42	0.3	0.1	0	0.12	0.1	0.1
個人_深層	0.92	0.03	0	0.05	0.77	0.07	1	1	0
対人_インプット	0.36	0.02	0	0	0.1	0	0.03	0	0.01
対人_アウトプット	0.87	0	0.07	0	0	0.1	0.12	0	0.1
リーダー_思考	0.23	0	0	0	0	0	0.12	0	0.02
リーダー_行動	0.16	0.01	0.21	0	0.01	0.01	0	0	0.43
マネジ_仕事	0.34	0	0.15	0	0	0	0	0	0.12
マネジ_人	0.87	0	0	0	0	0	0.04	0.01	0

資料 21：同僚の評価点数と各要素の相関行列

評価者:同僚	総得点	個人 表層	個人 深層	対人 インプット	対人 アウトプット	リーダー 思考	リーダー 行動	マネジ 仕事	マネジ 人
総得点	1	0.44	0.21	0.45	0.37	0.46	0.32	0.49	0.4
個人_表層	0.44	1	0.63	0.69	0.65	0.73	0.86	0.77	0.64
個人_深層	0.21	0.63	1	0.76	0.85	0.74	0.66	0.7	0.82
対人_インプット	0.45	0.69	0.76	1	0.91	0.79	0.73	0.89	0.81
対人_アウトプット	0.37	0.65	0.85	0.91	1	0.79	0.72	0.85	0.82
リーダー_思考	0.46	0.73	0.74	0.79	0.79	1	0.76	0.89	0.6
リーダー_行動	0.32	0.86	0.66	0.73	0.72	0.76	1	0.78	0.73
マネジ_仕事	0.49	0.77	0.7	0.89	0.85	0.89	0.78	1	0.7
マネジ_人	0.4	0.64	0.82	0.81	0.82	0.6	0.73	0.7	1

資料 22：各相関係数の P 値

評価者:同僚(P値)	総得点	個人 表層	個人 深層	対人 インプット	対人 アウトプット	リーダー 思考	リーダー 行動	マネジ 仕事	マネジ 人
総得点	0	0.34	0.39	0.34	0.37	0.34	0.37	0.26	0.37
個人_表層	0.06	0	0.04	0.02	0.03	0.01	0	0	0.03
個人_深層	0.39	0	0	0	0	0.01	0.03	0.01	0
対人_インプット	0.05	0	0	0	0	0	0.01	0	0
対人_アウトプット	0.12	0	0	0	0	0	0.01	0	0
リーダー_思考	0.05	0	0	0	0	0	0	0	0.06
リーダー_行動	0.18	0	0	0	0	0	0	0	0.01
マネジ_仕事	0.03	0	0	0	0	0	0	0	0.01
マネジ_人	0.09	0	0	0	0	0.01	0	0	0

資料 23：後輩の評価点数と各要素の相関行列

評価者:後輩	総得点	個人 表層	個人 深層	対人 インプット	対人 アウトプット	リーダー 思考	リーダー 行動	マネジ 仕事	マネジ 人
総得点	1	0	0.09	0.06	0.06	0.13	0.12	0.05	0.1
個人_表層	0	1	0.89	0.87	0.87	0.91	0.91	0.91	0.86
個人_深層	0.09	0.89	1	0.89	0.91	0.9	0.94	0.89	0.89
対人_インプット	0.06	0.87	0.89	1	0.93	0.82	0.84	0.89	0.93
対人_アウトプット	0.06	0.87	0.91	0.93	1	0.87	0.9	0.9	0.94
リーダー_思考	0.13	0.91	0.9	0.82	0.87	1	0.96	0.92	0.91
リーダー_行動	0.12	0.91	0.94	0.84	0.9	0.96	1	0.93	0.92
マネジ_仕事	0.05	0.91	0.89	0.89	0.9	0.92	0.93	1	0.95
マネジ_人	0.1	0.86	0.89	0.93	0.94	0.91	0.92	0.95	1

資料 24：各相関係数の P 値

評価者:後輩(P値)	総得点	個人 表層	個人 深層	対人 インプット	対人 アウトプット	リーダー 思考	リーダー 行動	マネジ 仕事	マネジ 人
総得点	0	1	1	1	1	1	1	1	1
個人_表層	1	0	0	0	0	0	0	0	0
個人_深層	0.68	0	0	0	0	0	0	0	0
対人_インプット	0.79	0	0	0	0	0	0	0	0
対人_アウトプット	0.77	0	0	0	0	0	0	0	0
リーダー_思考	0.55	0	0	0	0	0	0	0	0
リーダー_行動	0.59	0	0	0	0	0	0	0	0
マネジ_仕事	0.82	0	0	0	0	0	0	0	0
マネジ_人	0.65	0	0	0	0	0	0	0	0

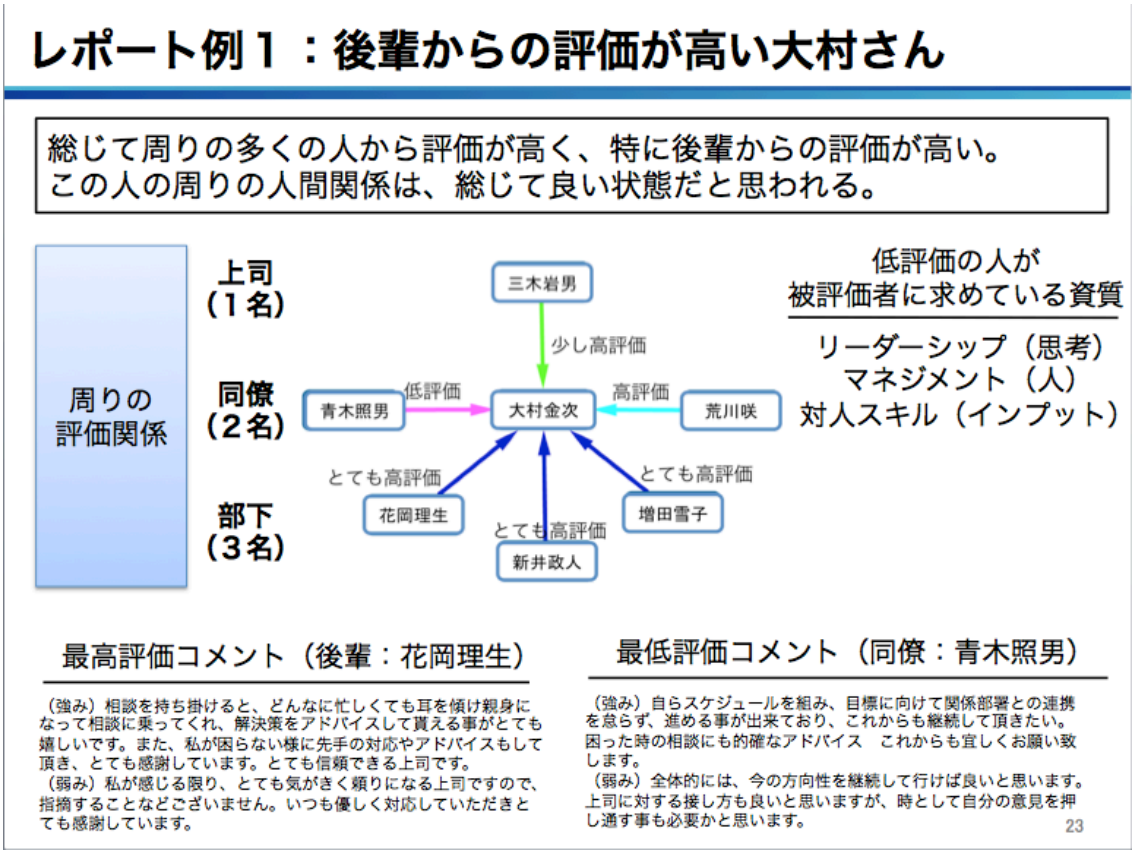
6. 実務への応用

今までの結果を踏まえて、特に職場間の従業員同士の評価に関しては、テキストでの評価は、分析精度の限界は存在するものの、ある程度実態に沿った形を表現することが可能だと思われる。では、このテキストによる人材評価の手法は、一般的な企業

の実務のシーンにおいて、どのように活用されるべきであろうか。

筆者の現時点での結論は、「被評価者の上司に評価関係をレポートすることで、上司が職場の人間関係を改善する」のが最も良い実務での活用方法ではないかと言うものである。一般的に、企業が高いパフォーマンスを発揮するためには従業員間の協力関係が必須であるが、その協力関係が円滑に築けていない部分を一目で見ることができる。この図示化を行うことで、上司に対して、マネジメントの際に「誰と誰の関係性に気をつけるべきか」という新しい示唆を与えることができると考える。また、本人に指導する際に、「後輩と接する際、もっとリーダーシップを持った方が良い」「同僚に対して必要以上に競争意識を持たず、時には協力する存在であることを忘れないように」という「具体的な」アドバイスを行うことも可能となるであろう。本研究では、筆者はデータの提供に協力頂いた企業に対して、被評価者の上司用のフィードバック資料を作成した（資料 25～27）。

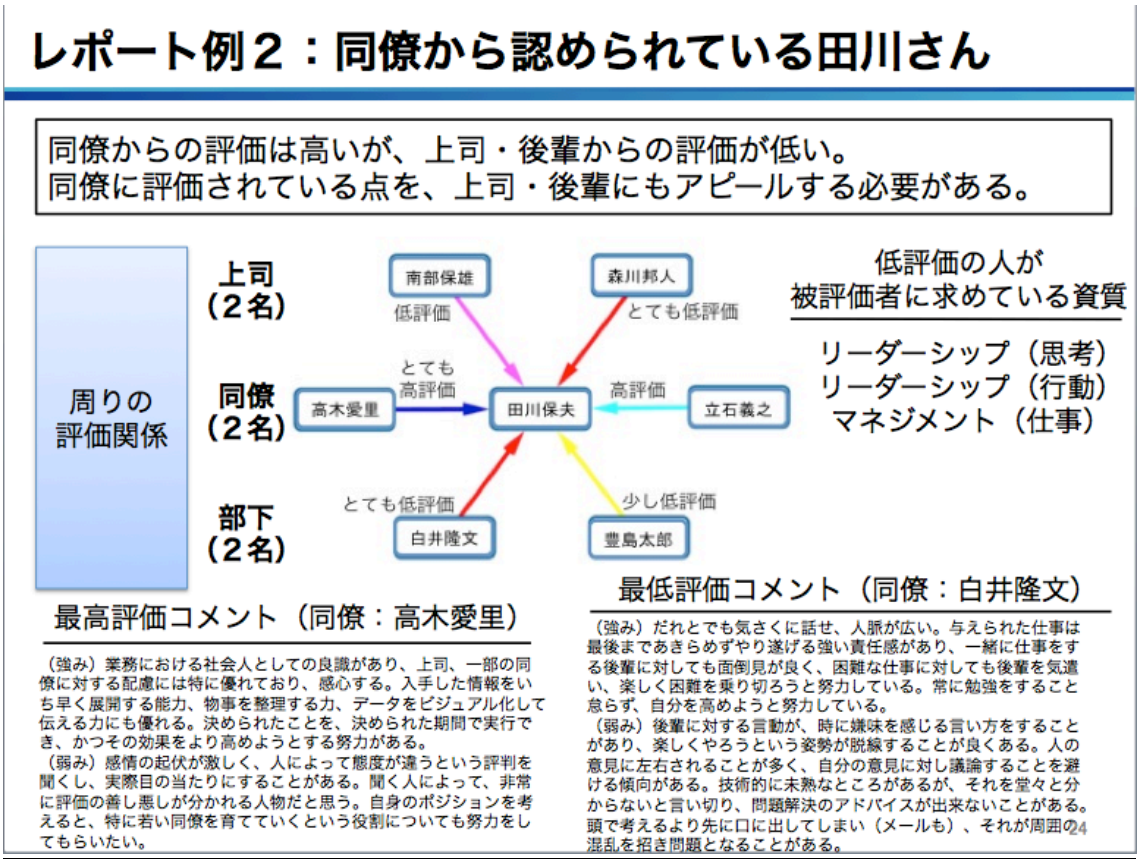
資料 25：被評価者 大村さんのレポート例



このレポートは、今まで出した評価関係のネットワーク図に加え、最高評価と最低評価のコメントを掲載することで具体的なイメージを持たせ、さらに「低評価」と評価した人が 360 度評価の 8 要素のコンピテンシーの中でどの要素を低く評価したのか

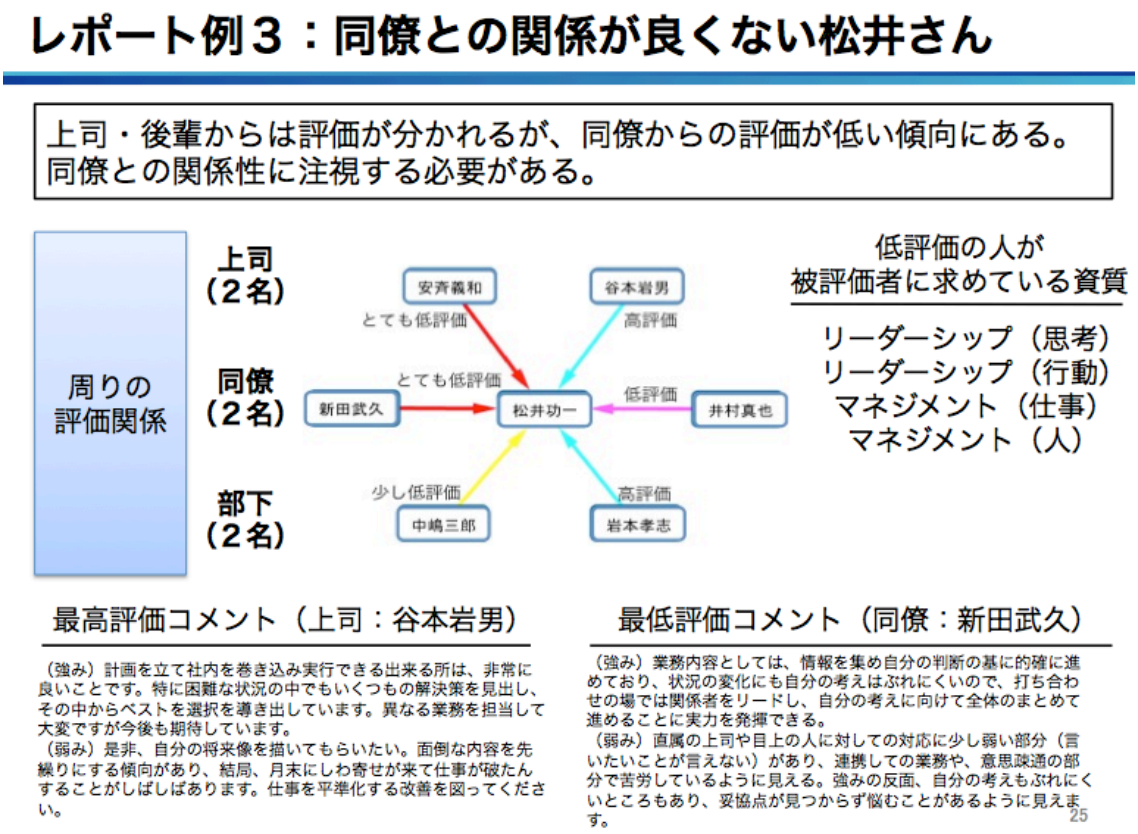
を記載している。こうすることで、具体的にどの部分の関係性に問題があり、何を修正すれば職場内の評価が上がりやすいのかを容易に確認することが可能である。例えば資料 25 だと、大村さんは青木さんからの評価が低いので、その評価を高めるためにはリーダーシップ（思考)、つまりリーダーとしての視点で物事を考える力をさらに磨く必要がある。

資料 26：被評価者　田川さんのレポート例



次に、資料 26 の田川さんの例だと、同僚からの評価は高いが上司・後輩の評価が低い。上司・後輩に対して、リーダーシップを発揮していくことが今後の課題であると言える。

資料 27：被評価者松井さんのレポート例



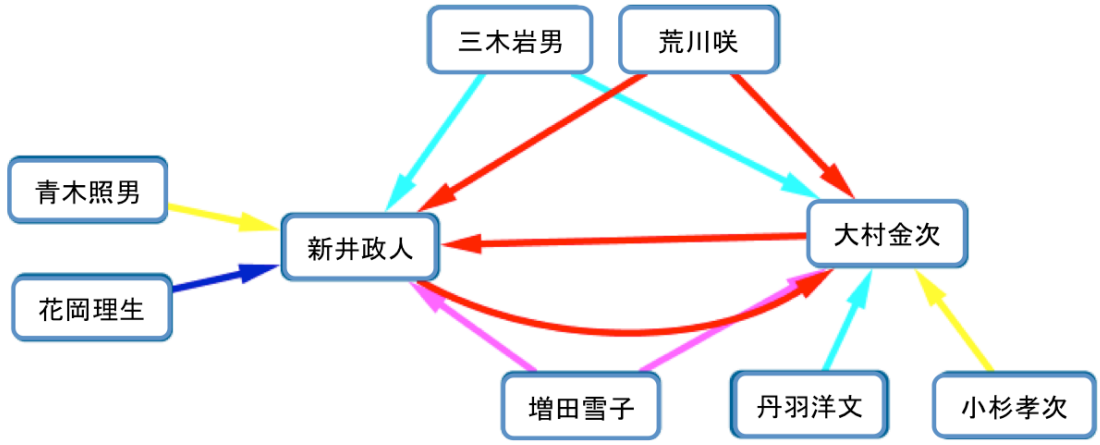
最後に資料 27 の松井さんの場合は田川さんとは逆で、同僚からの評価が低い。同僚と一緒に働く時に気後れせず、自分から引っ張っていく、などの行動が今後求められるであろう。

このように、上司向けのレポートにすることでその人が今後「誰に対して」「どのようにコミュニケーションしていけば良いか」までを示すことができ、職場のパフォーマンス向上に寄与することが可能ではないか。実際にこのレポートをデータ提供企業に見せたところ大変好評でわかりやすい、との評価を頂くことができた。360 度診断を実施しているディアグロー株式会社の内山社長からも高い評価を頂き、今後同じようなテキストでの評価方法を自社でも取り入れていきたい、という最大限の評価を頂くことができた。

また、この仕組みはもっと多くのデータを取り込むことができればよりスケールの大きな応用が可能になる。本研究ではサンプル数に限りがあったため応用まで踏み込むことは行っていないが、今後は部署内にとどまらず、社内でもっと多くのサンプルを取得することで、社内の人間関係をもっと広く見える化を行うことも可能であると

考える（資料 28 参照）。また、分析を深めていけば「仕事のマネジメントの数値が少し低い A さんと相性が良い人はリーダーシップの数値が高い B さんである」等の「相性」がある程度判断できるようになる可能性がある。このように、テキストから評価関係、ひいては「相性」を紐解くことで効率的な配属や異動が可能になると思われる。

資料 28：部署をまたいだ複数人の評価関係の例



7. まとめと今後の課題

本研究で得られたデータから検証を行った結果、テキストマイニングによる人事評価は既存の 360 度評価と完全に一致しているわけではなく、違う特色が出ていることが明確になった。また、そのテキストマイニングによる評価は「職場従業員間の評価関係」を表す際に、親和性が高そうであることも同時に判明した。

今後の課題について、少なくとも組織内評価をテキストマイニングで行っていく仕組みを作る、またその仕組みを多くの企業に使ってもらう、という未来を目指した場合、乗り越えるべき課題は大きく 3 点である。

①評価者の性格（甘口・辛口など）を考慮した採点ロジックの作成

人によって評価を全体的に甘めに書いたり、全体的に厳しめに書いたりする傾向は存在すると考えられ、このような回答者の傾向を補正すると、より精度が高まると考える。例えば、評価者が被評価者を必ず 5 人以上採点することとし、5 人の評価の平均点が高い人はある程度甘め、低い人は厳しめと定義して点数を調整しても良いであろう。しかし、本研究においては「同じ評価者が」「複数の被評価者を」評価することは無かったので、このような採点者の性格は考慮することができなかった。この部分を改善すれば精度は大きく変わると考えられ、今後の課題の中でも一番大きなもの

だと考えている。

②単語の採点ロジックの精度向上

そもそもの評価辞書が岩波辞書をベースに単語のポジティブ・ネガティブ具合を評価したものである。ただし、ビジネスの場ではポジティブ・ネガティブの強さは変わるかもしれないし、また、その企業内の言葉の使われ方、部署内の言葉の使われ方で単語の印象というのは少しずつ変わると考えられる。そのような意味で専用の辞書を整備すればさらなる改善が見込めると思われる。この課題については、Word2vec というテキストマイニングの新技術を応用することである程度改善されるかもしれない。Word2vec に関しては、日経ビックデータに以下のように記載されている。[4]

> Word2Vec は、その名前の表す通り、単語をベクトル化して表現するする定量化手法である。例えば日本人が日常的に使う語彙数は数万から数十万といわれるが、Word2Vec では各単語を 200 次元くらいの空間内におけるベクトルとして表現する。(中略)
> ・「男性－女性」を「おじ」や「王」に足すと、「おば」や「女王」になる。つまりこの差分は「男性と女性の違い」を表現している
> ・「フランス」－「パリ」＋「東京」＝「日本」などの結果も得られる。「フランスーパリ」は、「その都市を首都とする国」を表現している（中略）
> 実は Word2Vec は、非常に単純な前提をもとに構築されている。それは「同じ文脈の中にある単語はお互いに近い意味を持っている」というものである。「分散表現」とは、近年ディープラーニングなどの分野でも深く研究されている人工ニューラルネットワークの研究で提唱された考え方で、Word2Vec も人工ニューラルネットワークの研究から生まれた。
>
> 一方で、Word2Vec で使われている手法は、他の手法に比べその内部構造が非常に簡略化されている。ある単語が与えられたときに、近く（前後 5―10 単語くらい）に出現する他の単語を当てる、という問題の解を、与えられた文章中の単語全てに対して人工ニューラルネットワークに学習させる。似た意味の言葉はお互い近くにあらわれる可能性が高いため、学習を行っていく過程で、徐々に近しい方向のベクトルになっていく。(中略)
> この非常に面白い性質と、計算効率の高さから、この研究には昨年の発表以来非常に強い関心が寄せられ、実行プログラムも公開されている。C、Python、Java

[4] 日経ビックデータ「自然言語処理に新風を巻き起こした Word2Vec とは何か」（http://business.nikkeibp.co.jp/article/bigdata/20141110/273649/）（2014 年 11 月 13 日）

など、様々な言語での実装が開発され、多くの研究者に応用されている。

③単語の編集・辞書データ作成に手間がかかり、人手の介在が大きい

「自動化して簡単に多くの企業が使えるようにする」ということまで見据えると、現在の仕組みは多くの人手がかかっており、自動化するためには高いハードルを超える必要がある。具体的には単語の編集・辞書データ作成の作業は、精度を高めるために必要だが、人手での解釈が必要になる。この部分は、今後テキストの自動判別・辞書の自動作成などまで行えるテキストマイニング技術が必要である。現状では、google などの検索結果の精度を鑑みるに、一部の企業は既に前述の技術を実現していると思われるが、個人の力で実装するには難易度が高いので本研究では実施せず、今後の課題とした。

8. 終わりに

特に人事分野においては、まだ「大きな容量のデータ」「アンケート以外のデータ分析方法」は浸透していないように思える。しかし、筆者は最もデータ分析の活用の可能性がある分野も人事分野だと感じている。ビジネスの源泉は「人」である。その「人」のパフォーマンスを、データによって最大化できるのであれば、企業の業績に与える影響は大きいのではないだろうか。本研究がその 1 つの示唆となれば幸いである。

謝辞

最後に、分析に協力して頂いた方へのお礼の言葉をもって、当論文を締めくくりたい。まず、本研究について様々なアドバイスを頂き、統計の分野でも様々なことを教えて頂いた林教授に感謝したい。無理に教授のやり方に染めるのではなく、学生の自主性を重んじ、良い部分を伸ばすためのアドバイスを頂きながら筆者を指導してくださった。次に、本研究のきっかけを作って頂いた岩本先生、データを提供頂き、それ以外にも全面的にバックアップ頂いたディアグロー株式会社内山社長にも感謝する。また、一緒に切瑳琢磨した林研究室のメンバー（坂田さん・内田さん・萩原さん・山田さん）にも心から感謝している。

本研究を始め、大学院生活の 2 年目を実りあるものにして頂いているのは皆様のお陰です。本当にありがとうございました。

参考資料

高橋 潔 (2001)「多面評価法(360 度フィードバック法)に関する多特性多評価者行列分析」 『経営行動科学 14(2)』 pp.67-85

日経ビジネスオンライン (2007)「イヤな上司を排除するのにも使える「不幸」な制度職場の雰囲気を乱す「360 度評価」の危険性」(http://business.nikkeibp.co.jp/article/skillup/20070612/127148/?P=2) (2014 年 6 月現在)

日経ビックデータ(2014)「自然言語処理に新風を巻き起こした Word2Vec とは何か」(http://business.nikkeibp.co.jp/article/bigdata/20141110/273649/) (2014 年 11 月現在)

プレジテントオンライン (2004)「360 度評価が「裏目に出る」のを防ぐには」(http://www.president.co.jp/pre/backnumber/2004/20040329/804/) (2014 年 6 月現在)

金 明哲 (2007)「R によるデータサイエンス -データ解析の基礎から最新手法まで」森北出版

石田 基広 (2008)「R によるテキストマイニング入門」森北出版

石田 基広・小林雄一郎 (2013)「R で学ぶ日本語テキストマイニング」ひつじ書房

R ユーザ会 (2011)「R ではじめる Twitter 解析」(http://www.slideshare.net/abicky/rtwitter) (2014 年 6 月現在)

高村大也「単語感情極性対応表」(http://www.lr.pi.titech.ac.jp/~takamura/pndic_ja.html) (2014 年 6 月現在)

高村大也, 乾孝司, 奥村学 (2006)「スピンモデルによる単語の感情極性抽出」 *情報処理学会論文誌ジャーナル*, Vol.47 No.02 pp.627-637.

廣瀬紳一 (2012)「日本におけるアセスメント・センター運営の妥当性検証と今後の社会的要請に関する考察」 慶應義塾大学大学院経営管理研究科,博士論文

林 洋一郎 (2011)「アセスメント・センターの概要と心理測定上の課題」 法政大学キャリアデザイン学部紀要 8, pp211-228.

Nishtha Langer Indian School of Business Sandra A. Slaughter Scheller School of Business(2014) “Using Text Mining to Measure the Practical Intelligence of Project Managers” STATISTICAL CHALLENGES IN eCOMMERCE RESEARCH

付録１：多特性多評価者行列による妥当性検証

本研究で算出したテキスト評価はデータとして妥当性があるのかどうか、という問いを心理学で使用される MTMR 行列という分析手法を使って検証した。MTMR 行列とは、「多特性多評価者**行列**(multitrait-multirater matrix: MTMR; Conway, 1996)」という意味で、高橋(2001)は自身の論文でMTMR行列のことを以下のように説明している。

> 異なる評価領域を同一のフォーマットで評価しても正しい違いが見られるか,もしくは評価者が違っても同じ概念を正しく評価できるかといった妥当性を計るために用いられる。(中略)
>
> 評価対象者の資質が様々に立場の異なる人々から多面的に評価される状況では,多面評価法で問題にされる構成概念の妥当性は,2 つのより個別な妥当性概念一収束的妥当性(convergent validity)と弁別的妥当性(discriminant validity)―によって検討される。前者の収束的妥当性とは,評価者が同じであっても異なっていたとしても,同一の特性(構成概念)を評価している場合には,その評価が一致していることを意味している。また,後者の弁別的妥当性は,異なる特性(構成概念)に関しては,同一の評価者や異なる評価者による評価 が相違していること,逆にいえば,評価者による影響が出ないことを意味している。より詳しくは a)同一の評価者が下した同一構成概念の評価結果が相関 していること(収束的妥当性) b)異なる評価者が下した同一構成概念の評価結果が相関していること(収束的妥当性) c)同一の評価者が下した異なる構 成概念の評価結果が相関していないこと(弁別的妥当性) d)異なる評価者が下した異なる構成概念の評価結果が相関していないこと(弁別的妥当性) 4 者 によって検討されるのである。

もう少し具体的には、多特性多評価者行列では、以下のような方法で妥当性を検証する。①それぞれの評価方法で同じ被評価者を評価した場合の相関係数、②同じ評価方法で異なる被評価者を評価した場合の相関係数、③それぞれの評価方法で異なる被評価者を評価した場合の相関係数を算出し、上記の相関係数の関係が①＞②＞③となれば、そのデータは「妥当性が有る」と見なす。

本研究において、多種類の評価方法は「上司・同僚・後輩」と言い換えられ、多特性の対象は「10 人の被評価者」と言い換えられる。MTMR 行列の考え方に基づくと、評価者（上司・同僚・後輩）がそれぞれ被評価者を正しい違いが出るように評価しているかを検証するために、それぞれの相関係数を算出し、関係性を検証する。作成し

た相関行列は、資料 29 のようになった。

資料 29：多特性多評価者行列（MTMR 行列）

	方法	上司	上司	上司	上司	上司	上司	上司	上司	上司	上司	同僚	同僚	同僚	同僚	同僚	同僚	同僚	同僚	同僚	同僚	部下	部下	部下	部下	部下	部下	部下	部下	部下	部下
方法	特性(被評価者No)	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
上司	1	1.00	−0.28	−0.01	0.49	0.16	0.19	0.71	−0.12	−0.16	−0.02	0.38	−0.12	0.34	−0.48	0.19	0.21	0.68	−0.57	−0.83	0.51	0.41	0.23	0.50	0.62	0.27	0.14	0.34	−0.75	0.40	−0.73
上司	2		1.00	0.53	−0.07	−0.41	0.65	−0.04	0.01	0.02	0.27	−0.27	0.60	−0.33	0.72	0.60	−0.08	−0.26	−0.26	0.30	−0.69	0.20	0.25	0.07	−0.06	0.11	0.33	0.07	0.03	0.23	0.62
上司	3			1.00	0.34	0.07	0.91	0.50	0.75	0.45	0.17	0.61	0.11	−0.05	0.66	0.45	0.46	−0.06	−0.58	−0.14	−0.30	0.22	0.61	0.56	0.29	0.53	0.09	0.19	−0.02	0.11	0.35
上司	4				1.00	−0.42	0.33	0.25	0.40	−0.37	−0.38	0.37	0.07	0.16	−0.05	−0.22	−0.23	0.81	−0.19	−0.64	−0.20	0.41	0.13	0.83	0.84	0.35	−0.06	0.25	−0.26	0.04	−0.17
上司	5					1.00	−0.07	0.48	0.19	0.75	0.18	0.58	−0.54	0.40	−0.32	0.01	0.84	−0.25	−0.39	−0.04	0.75	−0.20	0.31	−0.19	−0.29	0.08	0.04	0.29	−0.18	0.27	−0.23
上司	6						1.00	0.53	0.55	0.35	0.38	0.48	0.38	−0.17	0.60	0.71	0.35	0.02	−0.59	−0.25	−0.36	0.42	0.71	0.55	0.44	0.47	0.02	0.13	−0.17	0.11	0.25
上司	7							1.00	0.22	0.29	0.20	0.69	−0.40	0.42	0.06	0.46	0.76	0.15	−0.88	−0.56	0.55	0.05	0.48	0.39	0.26	0.42	0.35	0.19	−0.67	0.34	−0.55
上司	8								1.00	0.46	−0.13	0.70	0.00	−0.21	0.28	0.00	0.28	0.10	−0.21	−0.28	−0.15	0.31	0.40	0.63	0.40	0.60	−0.37	0.20	0.39	−0.11	0.27
上司	9									1.00	0.56	0.65	−0.06	0.18	0.10	0.23	0.76	−0.43	−0.20	0.24	0.15	−0.05	0.73	−0.15	−0.16	−0.04	−0.28	0.16	0.02	−0.04	0.31
上司	10										1.00	0.20	0.28	0.11	0.27	0.61	0.39	−0.50	0.02	0.37	−0.14	−0.12	0.74	−0.42	−0.12	−0.48	−0.27	−0.38	−0.29	−0.38	0.14
同僚	1											1.00	−0.33	0.41	−0.02	0.07	0.73	0.17	−0.47	−0.39	0.33	0.10	0.70	0.45	0.40	0.30	−0.21	0.23	−0.29	−0.01	−0.17
同僚	2												1.00	−0.58	0.12	0.37	−0.52	0.15	0.25	0.05	−0.74	0.69	0.27	0.17	0.40	−0.03	−0.40	0.12	0.22	−0.03	0.54
同僚	3													1.00	−0.24	−0.36	0.50	0.13	−0.21	0.06	0.40	−0.48	0.24	−0.17	−0.09	−0.44	0.38	0.19	−0.75	0.16	−0.25
同僚	4														1.00	0.49	0.12	−0.55	−0.17	0.40	−0.55	−0.27	0.19	−0.01	−0.25	0.14	0.25	−0.42	0.20	−0.26	0.45
同僚	5															1.00	0.31	−0.33	−0.51	−0.10	−0.10	0.25	0.47	0.07	0.04	0.27	0.09	−0.20	−0.19	0.06	−0.04
同僚	6																1.00	−0.34	−0.64	−0.02	0.55	−0.29	0.57	−0.06	−0.22	0.13	0.24	0.14	−0.40	0.20	−0.18
同僚	7																	1.00	−0.15	−0.77	0.05	0.63	−0.04	0.73	0.84	0.30	−0.07	0.56	−0.32	0.36	−0.31
同僚	8																		1.00	0.51	−0.40	−0.19	−0.31	−0.47	−0.17	−0.62	−0.58	−0.44	0.52	−0.67	0.32
同僚	9																			1.00	−0.39	−0.63	−0.04	−0.81	−0.76	−0.68	0.05	−0.38	0.29	−0.35	0.63
同僚	10																				1.00	−0.23	−0.13	−0.09	−0.17	0.16	0.21	0.13	−0.36	0.28	−0.78
部下	1																					1.00	0.25	0.69	0.78	0.51	−0.37	0.54	0.08	0.34	0.08
部下	2																						1.00	0.18	0.36	−0.06	−0.30	0.13	−0.37	−0.08	0.21
部下	3																							1.00	0.84	0.78	−0.06	0.47	−0.06	0.30	−0.12
部下	4																								1.00	0.39	−0.38	0.34	−0.21	0.07	−0.17
部下	5																									1.00	0.12	0.35	0.22	0.41	−0.18
部下	6																										1.00	0.23	−0.46	0.63	−0.14
部下	7																											1.00	−0.21	0.85	0.17
部下	8																												1.00	−0.35	0.47
部下	9																													1.00	−0.05
部下	10																														1.00

上図の見方について説明を加える。まず相関係数が何を示しているかであるが、例えば「方法：上司　被評価者：1」と「方法：上司　被評価者：2」の場合、被評価者 1 の上司が評価した 8 種類のコンピテンシーのスコア」と「被評価者 2 の上司が評価した 8 種類のコンピテンシーのスコア」の相関係数を計測している。上司・同僚・部下が複数人存在している場合、算術平均のスコアを各要素のスコアとして採用している。背景ピンクのセルは異方法・同特性の評価の相関係数（①を検証できる）を示し、背景ブルーのセルは同方法・異特性の評価の相関係数（②を検証できる）を示し、背景白のセルは異方法・異特性の評価の相関係数（③を検証できる）を示す。

この表をまとめると、①を示すピンクのセルの相関係数の算術平均は 0.15　②を示すブルーのセルの相関係数の算術平均は 0.11　③を示す白のセルの相関係数の算術平均は 0.07　という結果が見られ、①＞②＞③の関係が成立していることがわかる。よって、「上司・同僚・後輩が、被評価者間の違いを正しく評価している」ことが示唆される。

付録 2：本研究で作成した R のコード

```r
#①：評価辞書に載っていない単語をMDSによって点数化するためのコード
#距離行列の算出
scorefile <- file("/Users/makoto/Desktop/human/deargrow/data/data/wordscore.csv", encoding='SJIS')
scoredata <- read.table(scorefile,header=T,row.names=1,sep=',')
scalescore <- scale(scoredata)
score.dist <- dist(scalescore)

#固有ベクトルを取る
library(stats)
cmddist <- cmdscale(score.dist,k=1) #kは次元数
cmddist #以降はこの距離データをExcelで編集して、単語を得点化する。

#出した距離が最初の距離をうまく表現できているか検証
dhat <- dist(cmddist[,1])
cor(score.dist,dhat)^2
#最初のデータから作成した距離行列と、cmddistで出した距離データから作成した距離行列との相関係数。これが高いと最初の行列をうまく表現することになる。


#②：①の処理で作った「評価辞書に掲載されてない単語」も含めて、各評価コメントを点数化するためのコード
#RMeCabライブラリの読み込み(R上でワークするためには、先にMeCabをPCにインストールする必要があることに注意)
library(RMeCab)

#データの読み込み
master <- file("/Users/makoto/Desktop/human/deargrow/data/data/master_2013.csv", encoding='SJIS')
masterdata <- read.csv(master,header=T,sep=',')

#単語感情極性対照表(評価辞書) の読み込み。
#取り込む辞書データ(pn_ja_merit_0620.dic)は、KBS学生にアンケートした単語の点数も含めた辞書データ
#まずは「強み」のコメントのみ抜粋するため、評価辞書のデータ(pn_ja_merit_0620.dic)は事前にポジティブな単語のみ抜粋している。
pndic_m <- read.table("/Users/makoto/Desktop/human/deargrow/data/data/dic/pn_ja_merit_0620.dic",
                      sep = ":",
                      col.names = c("term", "kana", "pos", "value"),
                      colClasses = c("character", "character", "factor", "numeric"),
                      fileEncoding = "Shift_JIS")

pndic_m[1:10,1:4]  #例
```

```r
#名詞＋品詞で複数の候補がある場合に平均値を採用するための処理
#例えば「アイス」の読みとして「アイス」、「アイスホッケー」が割り当てられている例がある
pndic2_m <- aggregate(value ~ term + pos, pndic_m, mean)

#「強み」のコメントデータから、上記pndicに登録されている品詞の単語を全て抽出
pos_m <- unique(pndic2_m$pos)
merit_DF <- docDF(masterdata, col = 7, type = 1, pos = pos_m) #masterdataの7列目「強みコメント」を形態素解析
merit_DF[16,1:10]  #例

#抽出した単語の中で、pndicにも登録されている単語(つまり、評価辞書による採点が可能な単語) のみ抽出
meritwords <- subset(merit_DF, TERM %in% pndic2_m$term)

#単語感情極性スコアを付与
meritwords_score_list <- merge(meritwords, pndic2_m, by.x = c("TERM"), by.y = c("term"))
meritwords_score_list[1:20,1:5]#例
merit.score <- colSums(meritwords_score_list[4:(ncol(meritwords_score_list) - 2)] * meritwords_score_list$value)

#出てきた点数を標準化して、csvに出力
merit.score_data <- as.data.frame(merit.score)
fixed.merit.score <- scale(merit.score_data)
write.csv(fixed.merit.score, "meritscore.csv")
#強みコメントの点数の出力まで完了。以降はexcelにてデータ編集、cytoscapeにて見える化。

#次に「弱み」のコメントのみ抜粋する。評価辞書のデータ(pn_ja_demerit_0620.dic)は事前にポジティブな単語のみ抜粋している。
pndic_d <- read.table("/Users/makoto/Desktop/human/deargrow/data/data/dic/pn_ja_demerit_0620.dic",
                      sep = ":",
                      col.names = c("term", "kana", "pos", "value"),
                      colClasses = c("character", "character", "factor", "numeric"),
                      fileEncoding = "Shift_JIS")

pndic_d[1:10,1:4]  #例

#名詞＋品詞で複数の候補がある場合に平均値を採用するための処理
#例えば「アイス」の読みとして「アイス」、「アイスホッケー」が割り当てられている例がある
pndic2_d <- aggregate(value ~ term + pos, pndic_d, mean)

```

```r
#「弱み」のコメントデータから、上記pndicに登録されている品詞の単語を全て抽出

pos_d <- unique(pndic2_d$pos)
demerit_DF <- docDF(masterdata, col = 8, type = 1) #masterdataの8列目「弱みコメント」を形態素解析
demerit_DF[16,1:10]#例

#抽出した単語の中で、pndicにも登録されている単語(つまり、評価辞書による採点が可能な単語）のみ抽出
demeritwords <- subset(demerit_DF, TERM %in% pndic2_d$term)

#単語感情極性スコアを付与
demeritwords_score_list <- merge(demeritwords, pndic2_d, by.x = c("TERM"), by.y = c("term"))
demeritwords_score_list[1:20,1:5]#例
demerit.score <- colSums(demeritwords_score_list[4:(ncol(demeritwords_score_list) - 2)] * demeritwords_score_list$value)

#出てきた点数を標準化して、csvに出力
demerit.score_data <- as.data.frame(demerit.score)
fixed.demerit.score <- scale(demerit.score_data)
write.csv(fixed.demerit.score, "demeritscore.csv")

#弱みコメントの点数の出力まで完了。以降はexcelにてデータ編集、cytoscapeにて見える化。



#③：各コメントのテキスト評価点と、360度診断の点数の相関係数(上司・同僚・部下に分けて実施)
library(psych)
text_data <- file("/Users/makoto/Desktop/human/deargrow/data/data/syokusyu.csv", encoding='SJIS')
text_score <- read.csv(text_data,header=T,sep=',')
score_boss <- subset(text_score,text_score$被評価者との関係 == "上司")
score_colleague <- subset(text_score,text_score$被評価者との関係 == "同僚")
score_subordinate <- subset(text_score,text_score$被評価者との関係 == "部下/後輩")
corr.test(score_boss[,-1])
corr.test(score_colleague[,-1])
corr.test(score_subordinate[,-1])
```