平成 29 年 4 月 21 日

各位

会社名　株式会社ソケッツ
代表者名　代表取締役社長　浦部　浩司
（証券コード：3634）

ソケッツ、歌詞と四字熟語の感情分析から「表現したい世界観」を可視化、
プロモーション支援に向け実用化へ

株式会社ソケッツ（本社：東京都渋谷区、代表取締役社長：浦部浩司、以下「ソケッツ」）は、自社の感情分析エンジンを活用して、歌詞の感情分析結果と、故事や仏典に基づく慣用句を筆頭に、相撲の昇進伝達式での口上や、企業経営・戦略に直結する示唆として多く用いられる四字熟語を関連付けることで「表現したい世界観」を可視化した分析レポートを公開いたしました。
本事例をはじめとした感情分析エンジンの研究開発の取り組みにおいては、アーティスト/クリエイターと企業や商品、そしてエリアをつなぐキャスティングやタイアップなどのプロモーション支援に向けた実用化を目指し、本年度、実証実験や準備を推進してまいります。

「表現したい世界観」「なりたいイメージ」が四字熟語でつながる？！

前回年末に行った、感情分析からの勝敗予測については、途中経過を覆し、最終的に見事的中しました！
（参照：http://www.sockets.co.jp/kansei/kansei_report01.html）
さて、今回は前回の予測でも行った感情分析に焦点を当てて、もう少し詳しくご紹介していきたいと思います。
まず前回も少しご説明しましたが、ソケッツの感情分析エンジンはエンターテイメントの感性メタを教師データとして感情を分析します。
エンターテイメント作品の感情を分析する、とは、言い換えれば、作品の「世界観」を見える化する作業ともいえます。ここでいう「世界観」とは、単なる知的な理解にとどまらず、より感情と意思、つまり心持などの情意的な評価を含むものと定義しておきます。
ということで、今回は歌詞データを感情スコア化、さらに感情スコア値の類似性から、歌詞の世界観に類似した情報を、歌詞とは直接的なつながりのない四字熟語に関連付けてみたいと思います。
ひと言で四字熟語といっても広義から狭義まで様々ですが、故事や仏典に基づく慣用句を筆頭に、相撲の昇進伝達式での口上や、企業経営・戦略に直結する示唆として、また企業・

経営理念などでも多く用いられ、いわば「真善美(※１)」の世界が四文字に凝縮されている四字熟語を媒体とすることで、エンターテイメント内での相関・つながりにとどまらない可能性をお届けしたいと思います。

（※１）真善美とは、人間の理想としての普遍妥当な価値のことを意味します。

今回のレポート条件

対象データ、概要は以下の通りとしました。

■1980 年以降にリリースされたシングル楽曲をソケッツの MSDB（※２）より抽出した 19,363 曲の歌詞データをもとに感情スコアを算出

■四字熟語は自然や季節などを表すものなどは除外し、気持ちや感情を含む熟語 848 つを対象に、四字熟語の説明テキストを感情スコア化

■今回は結果の精度判断をしやすいように、『おおよその歌詞を知っていて、かつ世界観がイメージできる曲』について、ソケッツ社内でアンケートを取り、その結果上位 10 曲の感情スコア値の類似性を解析、類似度の高い上位 2 つの四字熟語を関連付けることとします。

（※２）ＭＳＤＢ（メディアサービス･データベース）とは、ソケッツが開発した音楽、映像、書籍、放送、人物、施設、一般商品情報などを体系的かつ特徴情報を詳細に分類したデータベースであり、人の感性や感情を捕捉した「感性メタデータ」をキーとしたプロファイリング、アナリティクス、パーソナライズ、レコメンド、マーケティングサービスを実現します。

対象楽曲の感情スコアのサマリー

ここで改めて、感情分析エンジンでの感情スコア値の算出についてご説明しておくと、構文を考慮した自然言語解析で、歌詞のセンテンスごとに 10 種類の感情スコア値を算出、センテンスごとではなく楽曲単位での感情スコア化する場合には、検出された感情スコア値は単純に加算して評価しているのではなく、より人間の感覚に近い値が得られるというソ

ケッツの研究結果から、歌詞のセンテンスごとに算出した感情スコア値を元に、歌詞の流れを考慮したうえで、最終的なスコア値を算出しています。
つまり、楽曲単位での感情スコア値は最初から最後まで通して受ける感情を評価したスコアリングとなります。
まずは、ソケッツの MSDB から抽出した 1980 年以降にリリースされたシングル楽曲歌詞データの感情スコア値について、対象 19,363 曲のサマリーを簡単にご紹介します。

（1）感情スコア分布

19,363 曲の各楽曲で算出された感情スコアで、1 曲の中にどの感情がどの程度詰め込まれたストーリーとなっているかを感情スコアの分布から見てみたいと思います。まず曲のストーリー、世界観の要約、というところで、楽曲単位での平均値以上の感情数と各感情値での楽曲数分布をまとめてみました。

例）楽曲Aの感情スコア

幸福・恋愛	好き・喜び	安らぎ	昂ぶり	希望	不安・恐怖	怒り・苛立ち	哀しみ	嫌い・不愉快	退屈・うんざり
0.248377941	0.220629405	0.0469723	0	0.2933493	0.0387865	0	0.1235324	0.0283521	0

平均値 0.1

平均値以上の感情値　4個

平均値以上の感情スコア数（突出していた感情スコア）	対象全19,363曲分布	感情別分布										突出した感情スコアがなかったもの
		幸福・恋愛	好き・喜び	安らぎ	昂ぶり	希望	不安・恐怖	怒り・苛立ち	哀しみ	嫌い・不愉快	退屈・うんざり	平常
1個	81	9	2	0	25	10	0	0	32	3	0	0
2個	1,183	224	18	5	306	122	7	3	491	5	2	0
3個	5,682	1,034	85	25	1,483	751	56	9	2,209	15	3	12
4個	8,199	1,519	182	50	1,912	1,351	101	7	2,909	20	11	137
5個	3,663	9	114	36	748	535	51	2	1,143	14	7	398
6個	503	55	11	1	62	41	5	2	109	1	0	216
7個	15	0	0	0	0	0	0	0	1	0	0	14
8個	1	0	0	0	0	0	0	0	0	0	0	1
9個	0	9	0	0	0	0	0	0	0	0	0	0
10個	36	0	0	0	0	0	0	0	0	0	0	36

全体でみると、平均値以上の感情スコアが 4 種類以上の曲が 1 番多く、次いで 3 種類、5 種類となっています。

感情別分布でみると、
① 4 種類の感情スコアが平均値以上で「哀しみ」のスコアが 1 番高い楽曲　2,909 曲　対象楽曲数のうち 15.0%
② 3 種類の感情スコアが平均値以上で「哀しみ」のスコアが 1 番高い楽曲　2,209 曲　対象楽曲数のうち 11.4%
③ 4 種類の感情スコアが平均値以上で「昂ぶり」のスコアが 1 番高い楽曲　1,912 曲　対象楽曲数のうち 9.9%
④ 4 種類の感情スコアが平均値以上で「幸福・恋愛」のスコアが 1 番高い楽曲　1,519 曲　対象楽曲数のうち 7.8%

⑤ 3 種類の感情スコアが平均値以上で「昂ぶり」のスコアが 1 番高い楽曲　1,483 曲　対象楽曲数のうち 7.7%

平均値以上の感情スコアが 3，4 種類の「哀しみ」「昂ぶり」「幸福・恋愛」のいずれかが突出した楽曲は 10,032 曲で全体の 51.8%で対象の半数以上を占めました。

（2）高スコアの感情値掛け合わせ分布
次に、それぞれの楽曲で、どの感情値が高いのか、1 番高い感情値と 2 番目に高い感情値での分布を見てみました。

		対象全19,363曲分布	2番目に突出していた感情値										1つのみが突出していたもの
			幸福・恋愛	好き・喜び	安らぎ	昂ぶり	希望	不安・恐怖	怒り・苛立ち	哀しみ	嫌い・不愉快	退屈・うんざり	
1番突出していた感情値	幸福・恋愛	3,456		210	87	880	599	82	6	1,550	25	8	9
	好き・喜び	412	111		19	80	74	8	2	110	4	2	2
	安らぎ	117	26	11		23	14	2	1	40	0	0	0
	昂ぶり	4,536	1,143	190	113		1,223	178	19	1,599	35	11	25
	希望	2,810	581	112	31	925		65	5	1,061	10	10	10
	不安・恐怖	220	25	3	1	59	40		1	87	3	1	0
	怒り・苛立ち	23	4	0	0	5	3	2		7	1	1	0
	哀しみ	6,894	2,406	295	202	1,772	1,633	399	22		107	26	32
	嫌い・不愉快	58	6	7	1	9	4	3	0	25		0	3
	退屈・うんざり	23	5	2	0	4	2	2	0	8	0		0
	平常	814											

① 「哀しみ」「幸福・恋愛」　2,406 曲　対象楽曲数のうち 12.4%
② 「哀しみ」「昂ぶり」　1,772 曲　対象楽曲数のうち 9.2%
③ 「哀しみ」「昂ぶり」　1,633 曲　対象楽曲数のうち 8.4%
④ 「昂ぶり」「哀しみ」　1,599 曲　対象楽曲数のうち 8.3%
⑤ 「幸福・恋愛」「哀しみ」　1,550 曲　対象楽曲数のうち 8.0%

前回の紅白優勝予測でのロジスティック回帰分析の結果も、係数が 1 番大きかったのは「哀しみ」で影響度が高いとみなし予測モデルを構築しましたが、今回の分析対象楽曲中「哀しみ」が 1 番高スコアだった楽曲は 6,894 曲で 35.6%、邦楽は比較的「哀しみ」の要素が世界観の作品が多いようです。

（3）リリース時期ごとの高スコア感情分布
最後に、リリース時期ごとに見た、1 番突出していた感情値の分布は以下のようになっていました。

リリース時期	歌詞中で1番突出していた感情値分布										突出した感情スコアがなかったもの
	幸福・恋愛	好き・喜び	安らぎ	昂ぶり	希望	不安・恐怖	怒り・苛立ち	哀しみ	嫌い・不愉快	退屈・うんざり	平常
1980年代前半	36%	1%	1%	13%	5%	1%	0%	40%	0%	0%	3%
1980年代後半	26%	2%	0%	17%	8%	1%	0%	43%	0%	0%	3%
1990年代前半	24%	1%	1%	17%	11%	1%	0%	42%	0%	0%	3%
1990年代後半	21%	1%	1%	22%	13%	2%	0%	36%	0%	0%	4%
2000年代前半	18%	2%	1%	25%	14%	1%	0%	35%	0%	0%	4%
2000年代後半	16%	2%	1%	24%	15%	1%	0%	36%	0%	0%	4%
2010年代前半	15%	3%	0%	24%	17%	1%	0%	34%	0%	0%	5%
2010年代後半(～2017年2月)	13%	2%	1%	27%	18%	1%	0%	32%	0%	0%	6%

「昂ぶり」「希望」の割合は拡大傾向にあり、特に「昂ぶり」は「哀しみ」に迫る勢いとなっています。
J-POP の恋愛ソングの多さが話題になることが多い中、意外なことに「幸福・恋愛」の割合は減少しています。また、「希望」はわずかながらではありますが、徐々に拡大してきています。

歌詞と四字熟語の関連付け結果を見てみる

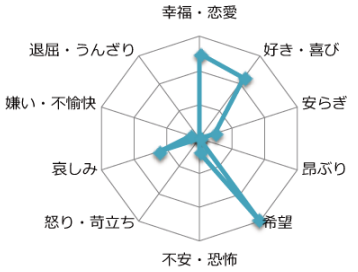
それではいよいよ『おおよその歌詞を知っていて、かつ世界観がイメージできる曲』（ソケッツ社内アンケート上位 10 曲）の楽曲をベースとして、感情スコア値の類似度の高い上位 2 つの四字熟語が、楽曲の世界観を表せているか、結果を見ていきたいと思います。

『おおよその歌詞を知っていて、かつ世界観がイメージできる曲』（ソケッツ社内アンケート上位 10 曲）

1．	世界に一つだけの花	SMAP
2．	ハナミズキ	一青窈
3．	何度でも	DREAMS COME TRUE
4．	One Love	嵐
5．	もう恋なんてしない	槇原敬之
6．	地上の星	中島みゆき
7．	CAN YOU CELEBRATE?	安室奈美恵
8．	春よ、来い	松任谷由実
9．	真夏の果実	サザンオールスターズ
10．	終わりなき旅	Mr.Children

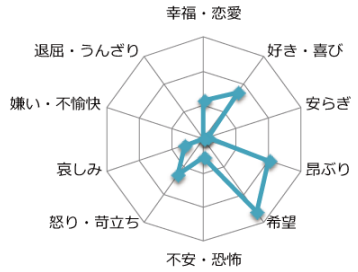
NO.1

楽曲/アーティスト	♪世界に一つだけの花　/　SMAP
リリース年	2002年
平均値以上の感情スコア数	4
1番高い感情スコア	希望
2番目に高い感情スコア	幸福・恋愛
サビ歌詞一部	世界に一つだけの花　一人一人違う種を持つ　その花を咲かせることだけに　一生懸命になればいい
類似度の高くマッチングした四字熟語	管鮑之交 … 終生変わることのない親密な交際 海内無双 … この世に並ぶものがないほど優れていること



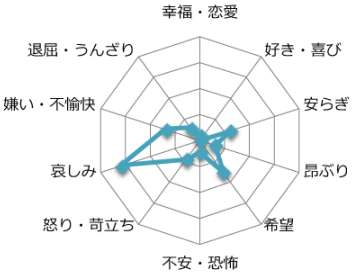
NO.2

楽曲/アーティスト	♪ハナミズキ　/　一青窈
リリース年	2004年
平均値以上の感情スコア数	5
1番高い感情スコア	希望
2番目に高い感情スコア	昂ぶり
サビ歌詞一部	薄紅色の可愛い君のね 果てない夢がちゃんと 終わりますように 君と好きな人が 100年続きますように
類似度の高くマッチングした四字熟語	一致団結 … 多くの人々が一つの目的に向かって一緒に行動すること 和光同塵 … 優れた才能や知性を目立たないようにして、世の中に交わること



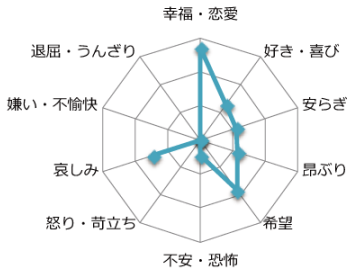
NO.3

楽曲/アーティスト	♪何度でも　/　DREAMS COME TRUE
リリース年	2005年
平均値以上の感情スコア数	4
1番高い感情スコア	哀しみ
2番目に高い感情スコア	希望
サビ歌詞一部	何度でも何度でも何度でも 立ち上がり呼ぶよ きみの名前 声が涸れるまで
類似度の高くマッチングした四字熟語	堅忍不抜 … 意志が強く、辛いことや困難なことがあっても耐えて、信念を貫くこと 一喜一憂 … 物事の状況が変わるたびごとに、喜んだり心配したりすること



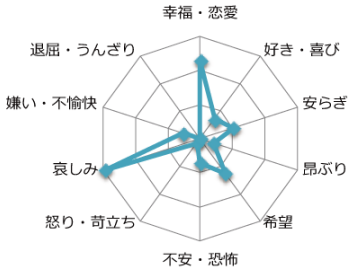
NO.4

楽曲/アーティスト	♪One Love　/　嵐
リリース年	2008年
平均値以上の感情スコア数	6
1番高い感情スコア	幸福・恋愛
2番目に高い感情スコア	希望
サビ歌詞一部	100年先も 愛を誓うよ 君は僕の全てさ 信じている ただ信じてる 同じ時間(とき)を刻む人へ
類似度の高くマッチングした四字熟語	撥乱反正 … 世の乱れを治め、平和な世の中に戻すこと 金石之交 … 金属や石のように堅固で、永久に変わらない交わり どんなことがあっても破れることのない厚い友情



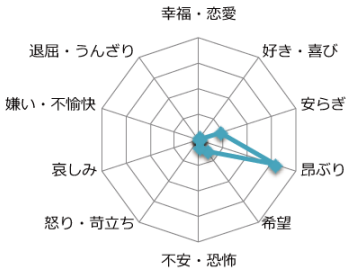
NO.5

楽曲/アーティスト	♪もう恋なんてしない　/　槇原敬之
リリース年	1992年
平均値以上の感情スコア数	4
1番高い感情スコア	哀しみ
2番目に高い感情スコア	幸福・恋愛
サビ歌詞一部	もし君に 1 つだけ 強がりを言えるのなら もう恋なんてしないなんて 言わないよ絶対
類似度の高くマッチングした四字熟語	堅忍不抜 … 意志が強く、辛いことや困難なことがあっても耐えて、信念を貫くこと 一日千秋 … 一日会わないと千年ものあいだ会わないように感じること 待ち焦がれる気持ち



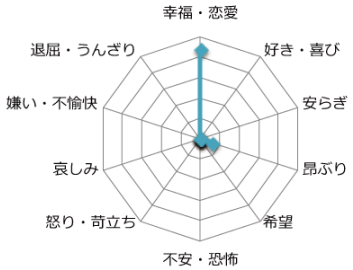
NO.6

楽曲/アーティスト	♪地上の星　/　中島みゆき
リリース年	2000年
平均値以上の感情スコア数	3
1番高い感情スコア	昂ぶり
2番目に高い感情スコア	安らぎ
サビ歌詞一部	つばめよ高い空から教えてよ 地上の星を つばめよ地上の星は今 何処にあるのだろう
類似度の高くマッチングした四字熟語	三三九度 … 新郎新婦が三つの杯を用い、一つの杯で酒を三度ずつ三回飲み合うこと 鼓舞激励 … 人を奮い立たせ元気付けて励ますこと



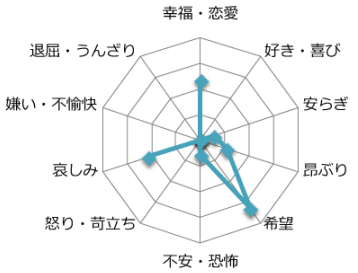
NO. 7

楽曲/アーティスト	♪CAN YOU CELEBRATE? / 安室奈美恵
リリース年	1997年
平均値以上の感情スコア数	2
1番高い感情スコア	幸福・恋愛
2番目に高い感情スコア	昂ぶり
サビ歌詞一部	Can you celebrate? Can you kiss me tonight? We will love long long time 永遠ていう言葉なんて 知らなかったよね
類似度の高くマッチングした四字熟語	比翼連理 … 男女の愛が深いこと 武陵桃源 … 俗世間から離れた、幸福な別天地



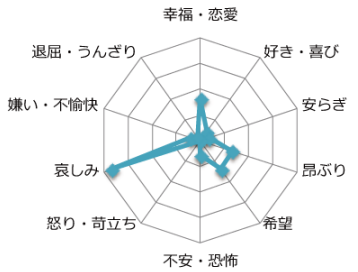
NO.8

楽曲/アーティスト	♪春よ、来い / 松任谷由実
リリース年	1994年
平均値以上の感情スコア数	4
1番高い感情スコア	希望
2番目に高い感情スコア	幸福・恋愛
サビ歌詞一部	春よ 遠き春よ 瞼閉じればそこに 愛をくれし君の なつかしき声がする
類似度の高くマッチングした四字熟語	自問自答 … 自分で自分に問いかけ、自分で答えること 大器晩成 … 偉大な人物は、普通の人より遅くに大成する



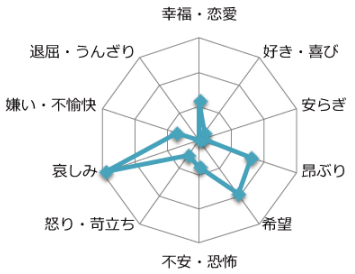
NO.9

楽曲/アーティスト	♪真夏の果実 / サザンオールスターズ
リリース年	1990年
平均値以上の感情スコア数	4
1番高い感情スコア	哀しみ
2番目に高い感情スコア	幸福・恋愛
サビ歌詞一部	四六時中も好きと言って 夢の中へ連れて行って 忘れられない Heart & Soul 声にならない
類似度の高くマッチングした四字熟語	落花狼藉 … 花が散り乱れていること 一日千秋 … 一日会わないと千年ものあいだ会わないように感じること 待ち焦がれる気持ち



NO.10

楽曲/アーティスト	♪終わりなき旅　/　Mr.Children
リリース年	1998年
平均値以上の感情スコア数	4
1番高い感情スコア	哀しみ
2番目に高い感情スコア	希望
サビ歌詞一部	閉ざされたドアの向こうに 新しい何かが待っていて きっときっとって 僕を動かしてる
類似度の高くマッチングした四字熟語	自縄自縛 … 自分の言ったことで自分を束縛して、自由に振舞えず、苦しめるようになる 自問自答 … 自分で自分に問いかけ、自分で答えること



いかがでしょうか。四字熟語の意味と曲の世界観、なかなか相関性の高い関連付けとなっているのではないでしょうか。
感情分析エンジンは、単純なキーワード抽出処理ではなく、取り扱いデータにより、話し言葉、擬音等の混入など、それぞれでコーパスを作り精度を上げ、構文解析によって難しい文脈や言い回しにも対応し、高精度な感情スコア化を実現していますので、NO.5『もう恋なんてしない』に登場する、「もう　恋なんて　しない　なんて　言わないよ　絶対」という人が最後まで聞かないと理解できない意味の解釈も正確に行い、スコアリングしています。

その結果として、NO.３『何度でも』と NO.5『もう恋なんてしない』、曲のテーマ、印象が全く異なる両曲でともに関連付けられている「堅忍不抜（意志が強く、辛いことや困難なことがあっても耐えて、信念を貫くこと）」は、両曲の歌詞を表す四字熟語として違和感ない結果を得ています。

まとめ

このようにソケッツの感情分析エンジンは、従来のキーワードマッチングによる方法ではなく、「気持ち」や「感情」の類似した情報を見つけ、歌詞データにとどまらず、人の感情に基づく他のデータの取り扱いが可能で、さらに関連付けることができます。
たとえば、“辛いことや困難なことがあっても耐えて、信念を貫く”という世界観やイメージを持たせたい企業や商品があったとして、今回の結果からいえば、NO.３『何度でも』をタイアップソングやテーマソングとするなど、「なりたいイメージ」と「表現したい世界観」の楽曲が、四字熟語を媒体としてつながるわけです。

近年、自然言語処理で潜在意味解析という文脈で発展してきたトピックモデルは、簡単にいえば、データの背後にある隠れた「トピック」を推定するということですが、今回の四字熟語はまさに“隠れたトピック≒潜在意味”と同等であると位置づけられるかもしれません。
今後、四字熟語は前述した相撲の口上や企業経営視点以外に、ドラマや CM などのキャスティングやプロモーション・タイアップにおいても、四字熟語を媒体として、アーティスト/クリエイターと企業や商品がつながる、なんていうことが当たり前になるかもしれません！？

（※３）トピックモデルとは、文書中の各単語およびそれらの単語が属するトピックが確率的に生成されていると仮定するモデルで、文書から「何らかの話題(＝トピック)」を発見するための統計的なモデルです。

「ソケッツレポート」の記事のご利用、また共同研究などに関するご相談はこちらのお問い合わせフォームよりご連絡ください。
http://www.sockets.co.jp/inquiry/contact/policy.html

●公開日

平成 29 年４月 26 日（水）

●ソケッツレポート

「表現したい世界観」「なりたいイメージ」が四字熟語でつながる？！

http://www.sockets.co.jp/kansei/kansei_report02.html

株式会社ソケッツ：(http://www.sockets.co.jp/)
株式会社ソケッツは、平成 12 年６月に設立された、「人の気持ちをつなぐ」という事業目的を持ったモバイルインターネット分野を中心としたサービス、アプリケーション、データベースの開発会社です。現在、ＫＤＤＩ株式会社や株式会社ＮＴＴドコモなどへの音楽や映像、電子書籍などの検索サービスやストリーミングサービスに加え、カルチュア・コンビニエンス・クラブ株式会社や楽天株式会社への作品おすすめ（レコメンド）機能の提供、国内最大級の楽曲データベースを活かした自社開発パーソナライズドサービス「LIFE’s radio」の提供などを行っております。

本リリースに関するお問い合わせ先：
株式会社ソケッツ
〒151-0051 東京都渋谷区千駄ヶ谷４-23-５JPR 千駄ヶ谷ビル３F
担当：コーポレート本部　小田嶋
Tel：03-5785-5518　 Fax：03-5785-5517
E-mail：ir@sockets.co.jp

以上