

ITソリューション フロンティア

IT Solutions Frontier

特集「適用領域を拡大するテキストマイニング」

07 | 2011 Vol.28 No.7
(通巻331号)



ジョギングに惹かれる訳	小粥泰樹	4
-------------	------	---

ソーシャルインテリジェンス構築のために —ソーシャルメディアが果たす役割と課題—	亀津 敦	6
---	------	---

ソーシャルメディアによる被災地ニーズの分析 —テキストマイニングを活用した「被災地の声 分析レポート」—	福島健吾	10
---	------	----

“顧客の声”分析の新しい手法 —評判情報を的確に抽出する“感性分析”—	牧 純一郎	14
--	-------	----

取引先審査に活かす情報収集の自動化 —企業情報モニタリングサービス「SmartSensor」—	大島 修	16
--	------	----

アジアから見た震災の影響 —災害に備えたサプライチェーン再構築の必要性—	澤井啓義	20
---	------	----

NRIグループと関連団体のWebサイト		22
---------------------	--	----

ジョギングに惹かれる訳

ここ数年、市民マラソン大会が各地で開催されるようになった。今春は震災の影響もあって多くの大会がキャンセルされたものの、その数はうなぎ上りに増えている。参加者もとりどりで、男性も女性も、また若者から中高年までと幅広い。マラソン人気を反映して、スポーツ用品店でもジョギングやランニングのコーナーに広いスペースが取られているようである。

筆者も、ご多聞に漏れず40代半ばを過ぎてメタボ検診の結果が気になりはじめたのをきっかけにジョギングを始めた。最近では毎月のペースで首都圏の市民マラソン大会に参加するようにしているが、人気の大会ともなると参加者があまりにも多く、エントリー自体ままならない状況である。

あらためてこのジョギング熱の理由を考えると、単に健康志向の高まりというだけでなく、人気スポーツに特有の条件が揃っていることに気づく。1つ目は一人でも手軽にできること、2つ目は熟練者でなくてもレベルに合わせて目標を設定できること、そして3つ目は精神修養につながる部分が多いことである。特に、これだけ長くブームが続くのは、ジョギングのメンタルな部分に魅力を感じる人が多いからではないかと考えられる。

レースで入賞するようなトップアスリートでなくとも、“走る作家”としても知られる村上春樹氏ぐらいのレベルになれば見える世

界は違う次元のものかもしれないが、ジョギング歴3年程度の筆者でも、走ることのメンタル面への影響はさまざまな形で感じることができる。

個人的には、走ることには大きく2つの効果があると思っている。

1つは、瞑想（めいそう）に近い効果を手軽に得られることである。“無の境地”に近いと言えば大げさかもしれないが、少なくとも雑念が払われるのは確かである。マラソンで5時間も走るという話をすると、「そんなに長い時間走って飽きませんか？」とか、「退屈じゃありませんか？」とよく聞かれる。ジョギングを始める前には筆者自身もそう思っていたから、人からそう聞かれることは不思議ではない。実際、筆者も初めの頃は、落語や英会話の録音でも聴きながら走ったら一石二鳥の時間活用にもなり、退屈でもなかろうと思って試してみたが、聴いてもほとんど頭に入らなかった。音楽でさえ疲れてくるとほとんど耳に入らない。だからといって退屈という感覚もほとんどない。走った後には、それまで思い悩んでいたこまごまとした事をすっかり忘れ、気持ちがすっきりしていることに気づく。雑念が払われるというのはこういうことなのだろう。

もう1つのメンタルな効用は、自分自身を客観的に見る機会になるということである。走ってみると、何よりもまず自分の限界を嫌というほど思い知らされる。レースともなる



と、それこそ無数のランナーに追い抜かれる。か弱そうな若い女性ランナーにも、二回り以上は歳が上と思われるランナーにも抜かれる。あまり抜かれるので、嫌でもそれが自分の実力なのだと思うようになる。抜かれるたびに悔しいなど感じていては切りがないので、そう納得するしかなくなるのだ。自分の限界が分かってくると、その限界を超えて頑張ることはしなくなる。無理して頑張ってみたところで長続きしないことが分かっているからである。

これを感覚的に言えば、自分の体を1つのプラントか制御対象のシステムと見たと、ゴールまで走るという目標のために制御スイッチを操っている気分である。特に走っている時の人間の体は、体力という限られたリソースの時間的配分が重要である。例えば「前半に思い切り頑張って、後半に疲れたら休む」などという戦略はまず失敗する。前半の頑張りが一定の限界を超えてしまうと、それが後半に致命的なダメージとなって現れ、完走も難しくなるのである。自分の限界を超えないぎりぎりのコントロールが最も高いパフォーマンスにつながるという意味で、やはり自分自身の力量を定量的にもよく把握していないと駄目なのである。

自分自身を客観的に見るという意味では、自分の性格についてさめた目で分析するよい機会になる。特に、自分の弱い部分は走る過

程で見事に露呈してしまう。筆者の場合、苦しくなってきた時に走るのをやめる言い訳を心の奥底で常に探している自分がある。弱い自分に気が付いているからこそ、走り始める時には、そのような言い訳をしなくてすむように万全の準備をして臨むのだが、苦しくなると巧妙な言い訳を見つけて走るのをやめさせようとするもう一人の自分が必ずといっていいほど現れる。もう一人の自分が、「今週は残業が多かったから無理しないほうがいい」とか、「足が痛いのによく頑張ったからそろそろ楽にしている」と言うのである。瞑想中のブッダを誘惑する夢魔のように、たとえばまたまた大げさに聞こえるかもしれないが、苦しい時には本当に甘いささやきに思えるのである。このようなもう一人の自分との駆け引きも、最近では楽しみの1つとなってきた。

最後にもう1つジョギングの楽しみをあげておこう。最近のマラソン大会は、たくさんのプロのカメラマンがデジカメで盛んに写真を撮っている。レース後に大会のホームページで自分のゼッケン番号を入力すると、自分が写っている写真の一覧を検索でき、気に入った写真を購入できるサービスが提供されているのである。レース中に自分の実力の無さを痛感した後で、何千人、何万人というランナーの中から、頑張った自分の証しを探し出す喜び、これもまた格別である。 ■

ソーシャルインテリジェンス構築のために —ソーシャルメディアが果たす役割と課題—

2011年3月11日の巨大地震と津波によって通信インフラが分断されるなか、Twitter（短文投稿サイト）などのソーシャルメディアが情報伝達手段として役立ったことが話題になっている。その一方で、デマの流布といった負の側面も見落とすことはできない。本稿では、ソーシャルメディアの特性や課題を検証し、これから社会全体としてどのように活用すべきかを考察する。

注目されたソーシャルメディアの利点

3月11日の午後2時46分、地震が起きた時に筆者は東京駅の地下街にいた。地上に出ようとする人々の群れに混じって、それから行くことになっていた会議の相手と、自分のオフィスへの連絡を試みた。しかし、携帯電話も携帯メールも通じない。大きな地震の後などに電話や携帯メールが繋がりにくいことはこれまでの経験からも分かっていたが、交通がまひし被害状況も全く見えないなかで、コミュニケーション手段がないことがこれほど心細いものかと途方に暮れた。

程なくして、スマートフォン（多機能な携帯電話）に見入っていた同僚が「東北地方で大きな被害が出ているらしい」と言い出した。パケット通信は通常どおり利用可能だったのだ。被害を伝える“ツイート”（Twitter上の投稿）もすぐに流れはじめた。筆者もパケット通信に切り替えて何人かの安否を確認し、さらにGmail（Google社のメールサービス）で連絡を取り合うことができた。携帯電話の通話と携帯メールが利用可能になったのはその5～6時間後のことである。

その後、帰宅は諦めてオフィスで一晩を明

かした。貴重な情報源となるはずのテレビはオフィスにはなかった。その代わりに、NHKやフジテレビなどがインターネット上の映像配信サービスUstream上で例外的に番組を再送信していた。

テレビ放送とUstreamの両方を視聴している人が、自分の周りの状況や得た情報をTwitterに投稿すると、それが連携機能によってUstreamの画面に映し出される。「私鉄が動き出した」というニュースが伝わると、駅にいる人から「動き出したが非常に混んでいる」という投稿が飛び込んでくる。放送とネットが融合した“ソーシャルテレビ”が出現し、それが威力を発揮するのを目の当たりにすることになった。

このような現実を前にして筆者は、ソーシャルメディアが多くの人々の情報を結びつけることで、これまでのマスメディアを補完する役割を果たすことになるかと直感した。実際にニュースや新聞などでは、記者のレポートのほか、被災地からのTwitterの投稿が取り上げられることが特に震災直後は多かった。

あらためて浮き彫りになった課題

一方で、ソーシャルメディアの持つネガテ



イブな面もすぐに顕在化することになった。

千葉県にある製油所の石油タンクが地震の影響で爆発炎上したことから、「千葉のコンビナートから有毒の雲が飛来する」というデマが広がったのである。このデマはメールだけでなくTwitterでも急速に広がった。Twitterでは早くも3月11日の段階で伝播が始まり、翌12日に新聞報道などで否定されるまで“リツイート”（他のユーザーの投稿を再投稿する機能）によって急速に伝播している。この種のデマは、従来はチェーンメールで広まることが普通だったが、情報伝達速度の速いソーシャルメディアもデマの拡散に加わる形になった。

その後も、被災地からの情報として「どこそこの避難所で食料が足りず、乳幼児に餓死者が続出しています！」という、被災地の窮状を伝える投稿がネット上を駆け巡った。被害が広範囲に及び、自治体の機能も多くの地域で損なわれていたことを背景に、大勢の人がこれを信じ、「拡散希望」という言葉が添えられてさらにこの情報が広がっていった。

実際はどうだったのか。その後の新聞報道によれば、少なくともその避難所には支援物資が届いており、餓死者は出ていなかった。この種の話はその避難所についてだけでなく、他の避難所についてもあったようである。また、緊急の救援要請が発信されたものの、救助を要するような被害はなかったというケースもあったという。

誤まった情報が伝えられたとしても結果として餓死者や要救助者がいなかったからよかった、というわけにはいかない。このような誤った情報が広がれば、それを信じた善意の人々が公共機関やマスコミに支援要請をするかもしれない。その結果、支援物資が必要以上に集中したり、ただでさえ厳しい状況で動いている現地の自治体や救助隊に無駄な作業をさせたりすることにもなる。情報伝播の速さはソーシャルメディアの有用性を示しているが、真偽があやふやな情報も同様の速さで広がっていくリスクは十分に警戒しておくべきであろう。

また、ソーシャルメディア特有の“情報の錯綜”という現象も浮き彫りになった。前述のように、「千葉のコンビナートから有毒の雲が飛来する」というデマの事例では、デマがTwitter上で広まりはじめた3月11日の翌日の午後4時には、新聞やテレビで「そのような事実はない」という報道がされていた。それを見た多くのTwitterユーザーが、デマの流布を抑えようと「千葉から有毒の雲が来るという投稿はデマです」と投稿しはじめた。しかし、デマを否定する報道があった翌日ぐらいまでは元のデマ情報がまだ流れており、「いったいどっちが本当なのか」と困惑を訴える人がいたり、デマが完全には打ち消されない状況が続いた。

こうした情報の錯綜はこの事例に限ったものではない。東京電力が3月17日にTwitterの

公式アカウントを開設した際にも、それから数時間にわたって「なりすましアカウントではないか」という投稿と、「公式なアカウントであることを確認した」という投稿が錯綜する事態が生じた。

デマや情報の錯綜が生じる根本原因

デマの急速な伝播や情報の錯綜は地震の数日後から目立ってきたが、このようなソーシャルメディアの問題点は今回の震災で初めて明らかになったものではない。Twitterのユーザーの間では、ユーザー数が急拡大した2010年にはすでに指摘されていたことである。

ソーシャルメディアのマイナス面の原因をユーザーのリテラシーやモラルの低さに帰すだけでは問題の解決にはならない。注目しなければならないのは、ソーシャルメディアの利用スタイルやコミュニティ構造に特有の「中心性の欠如」と「非同期性」という2つの特徴である。

「中心性の欠如」は、Twitterに代表されるマイクロブログ（200字程度の短い文章を投稿・公開するブログ。ミニブログともいう）に顕著な特徴である。従来のブログや掲示板などは、情報が特定の場所に蓄積され、過去の情報を検索するためのポータルページがある。これに対して、マイクロブログにはこのような“中心”が存在しない。

Twitterのユーザーは、無名の市民であろうが企業の公式アカウントであろうが対等なア

カウントである。情報は、ユーザーが投稿を参照できるように設定した“フォロー”という関係の間でのみ伝達されていく。ここでは情報の正誤の判定は基本的に個人に任されているが、「世の中でその話題がどのように取り上げられているか」や「事実関係が時間とともにどのように推移しているか」などが分かるような“中心”が欠如しているため、情報が正しいか否かを判断することが難しい。そのため、誤った情報が訂正されないまま流れたり、相反する情報が同時に流れたりするリスクが従来よりも大きいのである。

もう1つの特徴である「非同期性」とは、ソーシャルメディア上に流れるメッセージを閲覧するタイミングが人によってそれぞれ異なるということである。電子メールなどが他者との緊密なつながりを前提にしたコミュニケーション手段であるのに対して、Twitterの“フォロー”や、mixi、FacebookといったSNS（ソーシャルネットワーキングサービス）における“友だち”は、それほど緊密な関係を表しているわけではない。必ずしも自分が発したすべてのメッセージが読まれるとは限らないのである。

例えば、友人から来た“ツイート”を“リツイート”によってそのまま拡散させたとする。後で友人が「あの“ツイート”は間違っていました」という訂正の“ツイート”を投稿しても、それが気づかれることなく誤った情報が流れ続けるということが起こる。

ソーシャルインテリジェンスへの期待

デマの流布や情報の錯綜という事態を前にして、ソーシャルメディアの中から積極的に情報の整理や誤情報の訂正などを行おうとするユーザーも少なからず現れた。

あるTwitterのユーザーは、震災関連のニュースや被災地の情報を確認し、デマを排除した上で実況し続けることで疑似的に中心の役割を果たした。また、情報の蓄積や双方向のディスカッションが可能なFacebook上で、被災地情報を集積する「助け合いジャパン」公式ファンページが有志によって開設され、内閣官房震災ボランティア連携室との連携が図られた例もある。

しかし、このような試みは有志の人々の多大な労力によるものであり、ソーシャルメディア上の膨大な情報を長期間にわたって整理・選別し続けるためには人の力を越えた仕組みが必要であろう。例えば、テキストマイニング技術やセマンティックWeb技術（Webページの内容を表すメタデータを付加することでコンピュータが情報を解釈できるようにする技術）は、人の言葉から有用な知を発見する。ソーシャルメディアを社会に役立つものにするためには、こうした技術を有効に活用することが必要である。

すでに述べたように、Twitterでは正しくない情報が急速に広がることもあれば、これを正す情報が直ちに現れて広まることもある。

これらの相反する大量の情報をテキストマイニング技術によって分析すると、話題になっていることが真実であるのか疑わしいのかを判定し数値化することができるようになる。また、その話題がどのように広まったのかをさかのぼって、信頼に足る情報かどうかを見極めることもできる。このような活用の仕方によって、現在は人の手で行っているソーシャルメディア上の情報の編集・加工や真偽の判定は、はるかに効率的に行えるだろう。

野村総合研究所（NRI）は今回の震災に際し、錯綜する情報の中から正しい情報を拾い出すため、Twitterの投稿をテキストマイニング技術により整理・分析して被災地のニーズをできるだけ正確に把握できるようにする「被災地の声 分析レポート」（<http://act311.nri.co.jp>）を提供している（P.10「ソーシャルメディアによる被災地ニーズの分析」参照）。

このような情報の整理をすることで、雑多な情報が錯綜するソーシャルメディアを知的データベース（＝ソーシャルインテリジェンス）として社会全体で活用できるようになるのではないだろうか。

今回の震災は、ソーシャルメディアが緊急時に“人々をつなぐ”有益なツールとなることをあらためて示した。今後の復旧・復興に向け、また緊急時の有益性をさらに高められるよう、課題を克服して社会的に有用なインフラに育てたいものである。 ■

ソーシャルメディアによる被災地ニーズの分析 —テキストマイニングを活用した「被災地の声 分析レポート」—

東日本大震災の被災地では、「いつ」「どこで」「何を」支援物資として必要としているかを的確に把握することが求められる。野村総合研究所（以下、NRI）はTwitter（140文字以下の短文を投稿するサイト）上の投稿を分析して被災地のニーズを正確に把握する取り組みを行っている。本稿では、テキストマイニング技術を核とするこの取り組みの概要を紹介する。

被災地のニーズを正しく把握するために

2011年3月11日に発生した巨大地震と津波は、東北から関東にかけての広範囲にわたる太平洋岸に甚大な被害をもたらした。この被害の大きさから、野村総合研究所（以下、NRI）は3月15日に社長直轄の「震災復興支援プロジェクト」を発足させ、グループを挙げた取り組みを開始した。

その取り組みの1つに、Twitter上の投稿である“ツイート（つぶやき）”を分析することによって被災地のニーズを市町村別に集計しインターネット上に公開する「被災地の声 分析レポート」（<http://act311.nri.co.jp>）がある。

地震の後しばらくの間は、被災地では避難所などへの生活物資の供給が満足できる状態ではなかった。善意の支援物資も、どこにどのタイミングで送るべきかという情報が錯綜し、受け入れ自治体側にも混乱が見られた。そこで「震災復興支援プロジェクト」では、以下の2つを目的に、広範囲な被災地のどこで何が求められているかを正しく把握するための「分析レポート」の提供を開始した。

①自治体やボランティア団体などが支援物資

を的確に発送したり振り分けたりできるようにすること。

②支援物資を送ろうとする人が何をどう送ったらいいか判断できるようにすること。

ニーズを整理・分析してレポートを公開

「被災地の声 分析レポート」は、Twitter上の膨大な数の投稿を収集しテキストマイニング技術を用いて分析することで、被災地のニーズを詳しく把握するものである。初回のレポートがNRIのホームページで公開されたのは2011年4月1日で、4月14日にはシステム化されて短時間での自動更新が可能になった。本稿の執筆時点では、2時間に一度の更新が行われ、最新の情報が無償で提供されている。

「被災地の声 分析レポート」は、特に被災地の状況を全体的に把握する立場の方からの反響が大きい。現地で震災対応の指揮をとっている方からも「大いに参考にしている」という声をいただいた。

以下では、「被災地の声 分析レポート」がどのように情報を分析しているか、その技術的なポイントや、同技術の他分野への応用について紹介する。

野村総合研究所
IT基盤インテグレーション事業本部
ビジネスインテリジェンス事業部
副主任データアナリスト
福島健吾（ふくしまけんご）



専門はテキスト分析を中心にしたBI、ソーシャルメディア分析、マーケティングサイエンス

テキストマイニングによる分析が不可欠

さまざまな情報が飛び交うTwitterの投稿から被災地のニーズを正しく把握するには、単にニーズを拾い出すだけでなく、信頼できる投稿を収集してその内容を的確に分析することが重要である。これにはテキストマイニング技術が不可欠である。その理由は次の2点から説明することができる。

(1) 状況を総合的・定量的に把握できる

Twitter上に被災地の生の声が集まっているとはいっても、その量は膨大であり、すべての投稿に目を通すことは事実上不可能である。たとえすべての投稿を読むことができたとしても、「いつ」「どこで」「何を」必要としているのかを定量的に把握することは極めて困難である。このことは、アンケート調査の自由記述欄の回答を全部読んだとしても、回答者の要望（「誰が」「何を」望んでいるか）を定量的に把握するのが困難であることと同じである。

インターネット上の書き込みやアンケート調査などの自由記述テキストを分析するためには、NRIのテキストマイニングツール「TRUE TELLER」が有効である。「TRUE TELLER」は、今回の震災における被災者のニーズを把握するためにも活用することができる。具体的には、Twitter上の投稿のテキストから、「地名」、「必要とされている物資・サービス」を抜き出し、時系列、市町村



図1 「被災地の声分析レポート」の画面

別、物資・サービス別に集計を行う（図1参照）。このようにすることで、「いつ」「どこで」「何を」必要としているかを理解することができる。例えば、2011年4月3日に宮城県では他の市町村と比べて「南三陸町」を含む投稿が急増したこと、「食料」や「ガソリン」が不足しているという投稿が複数あることを瞬時に確認することができた。

(2) 「ハッシュタグ」の問題点を回避できる

Twitterには、ユーザーが特定のトピックに関係ある投稿を一覧表示できるようにする「ハッシュタグ」という機能（キーワードの前に「#」を付けて投稿する機能）があり、広く使われている。このハッシュタグを使えば投稿を整理・分類することが容易であるため、地震の直後は震災に関するハッシュタグを整理して表示させ、情報提供を行うサイトが複数見られた。

しかし、ハッシュタグは以下の3つの理由で、被災地のニーズを的確に把握するのに十

分とはいえない。

まず一点目に、ハッシュタグが“自然発生”することがあげられる。地震直後、Twitterの公式ブログでは、震災に関する話題を特定するための複数のハッシュタグが提唱された。しかし、これには拘束力がなく、時間が経過するに従ってニーズが細分化されたことでハッシュタグも多数発生することになった。そのため、初めに提唱されたハッシュタグをモニタリングするだけでは、すべてのニーズを把握することはできない。

二点目に、ハッシュタグは被災地からの投稿にだけ付けられているわけではないことがあげられる。震災に関連するハッシュタグは被災地の情報を伝える投稿だけでなく、被災地の投稿に対する返信のほか、震災について意見を述べる投稿にも多数付けられていたのである。

三点目は、被災地の状況を投稿する人がハッシュタグを付けずに投稿するケースが多いことである。震災をきっかけにTwitterを始めたような場合、ハッシュタグの使い方が分からないということも多いはずである。

注意が必要な“リツイート”

Twitterには、他のユーザーの投稿を再投稿する“リツイート”と呼ばれる機能がある。マーケティングや広告であれば、多くの人に情報を伝達できることからその効果が期待される。しかし、被災地の投稿からニーズを把

握するという視点では“リツイート”を分析に含めるべきではない。なぜなら、10回の“リツイート”が10倍のニーズを表しているわけではないからである。

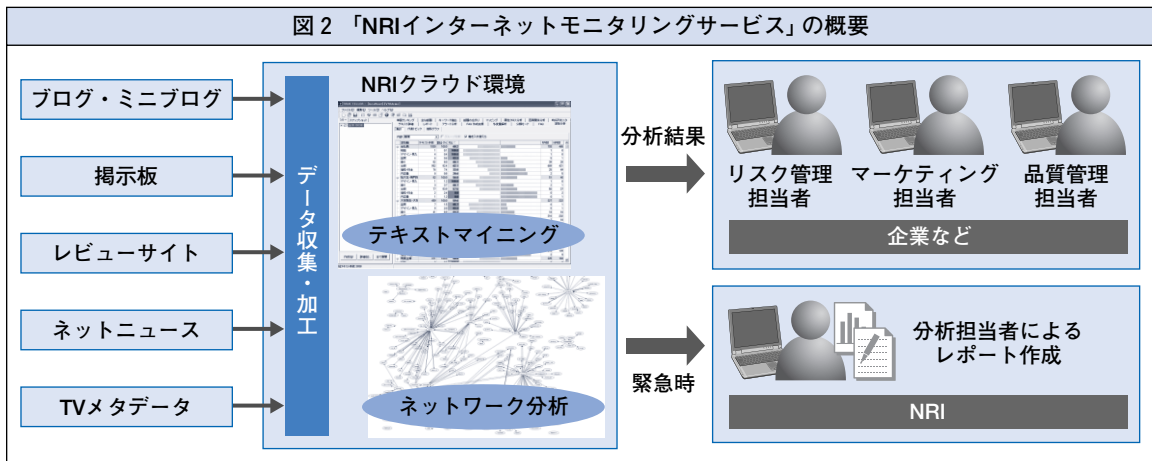
また、“リツイート”された情報は、過去には正しかったかもしれないが、いまでは不要な情報かもしれない。被災地では対応が急速に進められているため、誰かが気づいて“リツイート”した時にはすでに解決済みだったという現象も見られた。今回、解決しても1カ月近く“リツイート”されていた情報も存在した。

とはいえ、単純に“リツイート”を分析対象から外すだけでは有用な分析にならない。“リツイート”も含め、パターンマッチングなどを応用しながら重複している情報やデマと考えられる情報を省いていくことで、より精度の高い分析が可能となる。

平時からネット上を監視する

今回の震災では、Twitterのようなマイクロブログ、mixiやFacebookのようなSNS（ソーシャルネットワーキングサービス）をはじめ、ソーシャルメディアが災害時の重要な情報インフラとなることが実証された。しかし、ソーシャルメディアはその利用者が急増しているように、災害時に限らず社会的に大きな影響力を持つ情報インフラとなったことは間違いない。

一般に、ソーシャルメディア上の情報は瞬



時に拡散する。このため、企業は平時からソーシャルメディアの動向を監視し、問題となり得る情報にはますます短時間で対応することが必要になっている。

一方で、ソーシャルメディアの情報拡散力に注目して、ソーシャルメディア上でさまざまなマーケティング施策を実施する企業も多い。マーケティング施策の効果を検証するためには、ソーシャルメディア上の情報の分析だけでなく、関連する報道記事などとの関係も理解することが必要不可欠である。

このようなニーズに応えるため、NRIはテキストマイニング技術を核とした「NRIインターネットモニタリングサービス」を提供している。このサービスは、企業などを対象にソーシャルメディア上の書き込みやインターネット上の報道情報などを収集・分析するものである（図2参照）。

同サービスでは「被災地の声 分析レポート」と同様に、ユーザーは最新の分析結果を

Webブラウザを通じて閲覧することができる。インターネット上で対象企業に関する話題が急増した時にアラート（警報）を発することも可能なので、企業は即座に問題に気づき、対策を講じることができるようになる。このような緊急時にはNRIの分析担当者が緊急レポートを発行しており、いま何が起きているかを詳細に知ることが可能である。

NRIでは、今後も被災地のニーズを把握するための分析を続けていくことにしている。被災地のニーズも、当初は食料をはじめとする生活物資が中心であったが、いまは復興に必要な物資の比重が高まっているため、こうした変化を反映させていくことも必要である。また、今回の震災が被災地のみならず日本全体や世界にさまざまな影響を及ぼしていることから、人々にどのような意識の変化が起きているかを広く調査・分析することも検討している。 ■

“顧客の声”分析の新しい手法

— 評判情報を的確に抽出する“感性分析” —

ソーシャルメディアの普及によって、企業の評判を左右する“顧客の声”が、企業の関知しない場所で広がっていくケースが増えている。企業にとって、幅広くかつ素早く“顧客の声”を探知し分析することは重要な課題である。本稿では、テキスト情報を解析するための“感性分析”の有効性と、野村総合研究所（以下、NRI）の感性分析エンジンの特徴を紹介する。

ソーシャルメディアにあふれる“顧客の声”

企業には、コールセンターへの問い合わせやアンケート調査への回答などを通じてさまざまな“顧客の声”が寄せられている。最近では、個人のブログをはじめ、SNS（ソーシャルネットワークワーキングサービス）のようなネット上のコミュニティなど、いわゆるソーシャルメディアが普及したことにより、企業の評判が思わぬ場所で発信・形成され、口コミとして消費者の間に急速に広まるケースも増えてきている。そのため、企業はこれまで以上に“顧客の声”を広く丹念に拾い上げる確に把握する必要に迫られている。

通常、“顧客の声”の分析にはテキストマイニングツールが用いられる。テキストマイニングとは、構文解析技術によってテキストを要素に分解したり係り受け構造を抽出したりすることで、テキストの意味や話題の傾向を客観的に把握する手法である。膨大なテキストデータの全体像を定量的かつ速やかに把握できるため、“顧客の声”の分析だけでなく特許情報の分析やマーケティング調査などさまざまな場面で活用されている。ソーシャルメディア上の“顧客の声”を分析するためにも、

もちろんテキストマイニングは有効である。

この目的でテキストマイニングを活用する際にポイントとなるのは、テキスト中の評判に関係する情報をいかに的確に抽出するかという点である。従来から一般に用いられてきたのは、評価語（ポジティブまたはネガティブな評価を意味する語句）を設定して、その語がどのように現れているかを抽出・分析する手法である。しかし、次々に新しい情報が飛び交うソーシャルメディア上のテキストから適切な評価語を選定することは難しいため、この作業を省力化し、分析を高度化する手法に対するニーズが高まっていた。これに応える形で登場したのが“感性分析”である。

ソーシャルメディアに適した“感性分析”

感性分析は次の3つの技術から構成される。

- ①整備された評価語の辞書（感性辞書）
- ②評判抽出に特化した解析エンジン
- ③顧客の評価（ポジティブかネガティブかなど）の数値化

感性分析を使うと、複雑な設定をしなくても評判情報を抽出でき、ポジティブ・ネガティブ比率などの集計を簡単に行うことができる。また、感性辞書が用意されているため、

野村総合研究所
IT基盤インテグレーション事業本部
プロダクト開発部
副主任テクニカルエンジニア
牧 純一郎（まきじゅんいちろう）
専門は自然言語処理技術、テキストマイニングツールの製品開発



これまで分析者が行っていた評価語設定などの作業負担を軽減するだけでなく、分析者の主観に左右されない客観的な分析結果が得られることもメリットとしてあげられる。

感性分析はすでに実用化が進んでおり、欧米ではソーシャルメディア分析における必須技術として定着しはじめています。

NRIの感性分析エンジンの特徴

NRIが開発したテキストマイニングツール「TRUE TELLERテキストマイニング」も、この感性分析の技術を標準搭載するツールの1つである。NRIは10年以上に及ぶテキストマイニング開発の知見を活かし、他の感性分析エンジンにはない独自の改良を加えることで分析の質を向上させている。以下に「TRUE TELLERテキストマイニング」の感性分析エンジンの特徴を紹介する。

(1) 評価語の正確な抽出

評価語は必ずしも単語であるとは限らない。例えば「生産性が高い」のように、2つの単語が文を成したときにはじめて評価語となるケースも多い。NRIの感性分析は、単語だけでなく構文レベルのマッチングも行うよう設計されているため、複雑な評価語でも高精度に抽出することが可能である。

(2) 文法規則による補正

例えば「良いとは思わない」という文では、「良い」というポジティブな評価語に「思わない」という打ち消し表現が続いている。この

場合、「良い」という意味を反転させてネガティブと判定する必要がある。このような文法規則による補正ロジックは、判定精度を向上させるために欠くことのできない技術である。

(3) 評価軸に基づく自動カテゴライズ

NRIの感性分析エンジンの最も大きな特徴といえるのが、評価軸という観点に基づく評価語のカテゴライズである。例えば洗顔料に関する口コミ分析を例にとると、「メイクは落ちるが匂いが苦手」という文は、「洗浄力：ポジティブ」「香り：ネガティブ」というように自動でカテゴリーを付与した状態で出力される。複数の評価語が混在する文章でも、評価の対象（上記の例では「洗浄力」や「香り」）を明確に区分して抽出できるため、企業は自社製品の評判の良い側面と評判の悪い側面を具体的に把握することが可能となる。

NRIはこの評価軸を企業利用に最適化し、13業種54製品（サービス）に対してそれぞれの評価軸のカテゴリーを整備している。企業は自社の製品に必要な評価軸を自由に選択して分析に使用することが可能である。

感性分析はソーシャルメディア上の評判の的確な把握を支援する優れた技術である。ソーシャルメディアの影響力が無視できない規模となったいま、企業における感性分析の活用はますます広がっていくことは間違いない。NRIは感性分析技術の精度向上に向け、さらなる技術開発に努めていく考えである。 ■

取引先審査に活かす情報収集の自動化

—企業情報モニタリングサービス「SmartSensor」—

企業の取引先審査においては、取引のリスクを的確に把握するための外部の情報源（企業データベース、新聞、雑誌、Webなど）からの情報収集が欠かせないが、人の手による網羅的な情報収集には膨大な手間がかかる。本稿では、野村総合研究所（以下、NRI）が提供する企業情報モニタリングサービスと、そこで用いられている自然言語処理技術の概要を紹介する。

取引先企業のリスク審査の重要性

企業にとって取引先企業の動向が自社のビジネスを左右するような重要な意味をもつことは少なくない。そのため、相手企業と取引関係を築くに当たっては取引先審査を実施するのが普通である。

取引先審査では、財務諸表に基づいて与信審査を行い、経営状況や倒産リスクなどを計量して信用力を評価する。また、コンプライアンス（法令遵守）に関するリスクのように、財務諸表に表れないリスクも審査の観点に盛り込む必要がある。代表的なコンプライアンスリスクには刑事事件、民事訴訟、情報漏えいなどに代表される事件・事故や、それに伴う行政処分などがある。取引先審査においては、審査対象の企業でそうした事例が過去になかったか、現在、係争中の訴訟・紛争がないかなどを検証し、今後の取引におけるリスクの有無を判断することが必要になる。

こうしたビジネス上のさまざまなリスクの検証は、新たに取引を始める企業についてだけ行えばよいわけではない。すでに取引関係にある企業でも、事件・事故が突然に発生する可能性があるため、企業の動向を常にモニ

タリングし、重大な出来事が発生した場合に迅速に対応できるようにしておかなければならない。

例えば業務委託先の企業で情報漏えい事件が発生した場合、委託元に対してはすぐに報告されても、委託元以外には情報漏えいがあったという報告が行われなくてもいいかもしれない。本来であれば、漏えいした情報とは関係のない取引先にも、速やかに情報漏えいの有無や是正対策を報告するなどの対応が求められるが、委託先で情報漏えい事件があったことを後で新聞報道などによってはじめて知るということも起こり得る。

情報収集を自動化する「SmartSensor」

事件・事故の情報を収集するための情報源は、新聞、インターネット上のニュースサイト、官公庁・地方自治体が発表する行政指導情報などを網羅する必要がある。

膨大な情報の中から事件・事故の情報を収集し、さらにそこから取引先企業が含まれる情報を拾い出すのは大変な手間と時間のかかる作業となる。取引相手の企業だけでなく、親会社や子会社などのグループ企業にまで調査対象を広げると、作業規模がさらに大きく

野村総合研究所
IT基盤インテグレーション事業本部
ビジネスインテリジェンス事業部
主任テクニカルエンジニア
大島 修（おおしまおさむ）
専門は自然言語処理応用システムの開発



なってしまう。

そこでNRIは、取引先審査業務の効率化を支援するための企業情報モニタリングサービス「SmartSensor」を2011年5月に提供開始した。SaaS（Software as a Service：ソフトウェアの機能をインターネット上のサービスとして利用する仕組み）型のサービスである「SmartSensor」は、以下の3点を大きな特徴としている。

(1) 豊富なデータソース

過去25年分の記事アーカイブを含む新聞記事データ、官公庁・地方自治体のホームページで発表された行政指導情報、さらにブログや掲示板などソーシャルメディアのデータを検索できる（図1参照）。これらのデータは日々更新されており、企業に関する事件・事故の情報を最新のものも含めて網羅的に検索することが可能となる。ソーシャルメディアのデータも対象とすることで、ネット上での評判という観点で情報を検索することも可能である。

組み込まれている企業データベースには、全国の上場企業4,000社、中堅企業160万社を含む約350万社のデータが格納されている。対象の企業だけでなく、親会社・子会社などのグループ企業も含めたシームレスな情報の検索が可能である。

(2) 企業情報の自動整理

同一の出来事に関する複数の新聞記事やブログ記事を自動的にまとめ上げたり、記事中



図1 「SmartSensor」の検索画面

の企業名やリスクキーワードに関する箇所を自動的に色付け表示するなど、企業に関する情報を整理された形で参照できる仕組みになっている。

記事をデータベースに格納する際は、カテゴリ情報が自動的に付与される。カテゴリには、リスク種別（違法行為、訴訟、行政処分、システム障害、情報漏えい、リコール・製品不具合、労務トラブル、事故、災害など）と企業活動種別（生産、販売、サービス提供、研究開発、建設、業務提携、資本提携など）がある。これにより、ある企業に関して「刑事事件」「訴訟」関連の事象に絞った事例検索なども容易であり、一般的なインターネットの検索エンジンを使う場合のように、膨大な検索結果から必要な情報を探す手間を大幅に省くことができる。

このような情報の自動整理は、収集した新聞記事などのテキストデータに対して自然言語処理を施すことによって実現している。

(3) モニタリング業務の省力化

自社とすでに取引のある企業をあらかじめ登録しておくことで、それらの企業に関する情報を常時モニタリングすることが可能である。何らかの事件・事故があった際に自動的にメールで通知する機能を備えているため、取引先企業がいくら多くてもモニタリングの手間がかからず、何らかの問題が発生した場合にも迅速な対応が可能となる。

自然言語処理技術を活用

「SmartSensor」にはNRIが開発した情報抽出エンジン「SmartCollector」が組み込まれている。「SmartCollector」は非構造のテキストデータの自動解析を行い、企業名や企業活動、リスク種別などの情報を抽出する。

「SmartCollector」の主な機能は以下のとおりである。

(1) エンティティの抽出

膨大なデータソースのテキストの中から、企業名やリスク種別、企業活動、商品名、地域などの情報が自動的に抽出される。このような処理は自然言語処理の分野でエンティティ（ひとまとまりのデータの実体）抽出と呼ばれる。

企業活動に関するエンティティを正確に抽出するため、エンジンには企業名辞書のほか、企業活動、リスクキーワードなどのシソーラス（類義語）辞書を内蔵して表記の揺れを吸収している。例えば記事中の「野村総合研究

所」も「野村総研」も「NRI」も同じ企業と認識することができるほか、「情報漏洩」も「情報漏えい」も「機密情報の漏洩」も同一の概念として認識することができる。

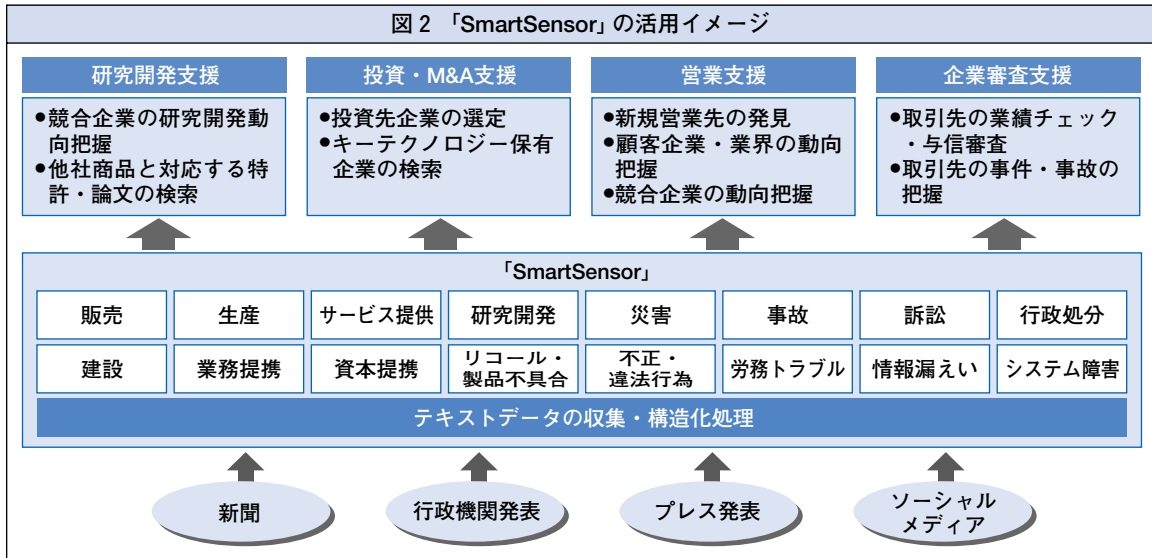
このようにして抽出したエンティティの種類から、記事カテゴリーを自動的に判別することが可能になる。

(2) ナレッジデータベースの検索

文書構造解析によって、抽出したエンティティの関連付けを行う。例えば、「A社は顧客情報を格納したUSBメモリーを紛失したと発表した」という記事から、「A社」→「情報漏えい」という形で企業名と事象を関連付けて抽出する。文書の構造や使われているキーワードの揺れを吸収できるため、「A社は顧客情報の流出があったと発表した」といった表現の記事であっても、全く同一の事象として抽出することができる。

一方、「A社は情報漏えい対策ソリューションの販売を開始する」といった記事であった場合、「A社」→「販売活動」→「情報漏えい対策ソリューション」という関連付けを行い、情報漏えい事故とは全く別の企業活動を抽出することができる。

このようにして、企業名とリスク種別や企業活動を関連付けたナレッジデータベースが構築される。このデータベースを利用することで、ユーザーは特定の企業に関して探したい事象のみに絞ったノイズの少ない検索を行うことが可能となる。



さまざまな業務への応用が可能

「SmartSensor」は、取引先審査以外にもさまざまな用途に利用することができる（図2参照）。

例えば、銀行が「オペレーショナルリスク」へ対応するために利用することも可能である。日本では2007年3月から適用された「バーゼルII」（新しい自己資本比率規制）は、事務事故、システム障害、不正行為などで損失が生じるリスクをオペレーショナルリスクと定義し、これを計量化することを求めている。オペレーショナルリスクは、過去に自行で発生した損失実績のほか、他の金融機関の事例にも基づいて発生頻度や損失金額を推定して計量化する。このように、他の金融機関で過去に発生した事件・事故情報をモニタリングし、自行に当てはめた場合のリスクの有無を調査

する目的で「SmartSensor」を活用することができる。

リスクへの備えだけでなく、ビジネスチャンスを発見する目的での活用も期待できる。「SmartSensor」は幅広い企業活動（販売、生産、研究開発など）の情報を収集・蓄積しているため、特定の技術を保有する企業の検索や、投資先企業の選定、顧客企業・業界動向の把握、新規営業先の発見などの用途でも活用することができる。

取引先審査のための情報収集のように、特定の業務領域に特化した、よりスマートな業務支援を実現する自然言語処理技術の重要性が高まりつつある。NRIは「TRUE TELLER」をはじめとするテキストマイニングソリューションの開発で培ってきた自然言語処理技術を活用して、企業のさまざまな業務の高度化を支援する取り組みを進めていく。 ■

アジアから見た震災の影響

—災害に備えたサプライチェーン再構築の必要性—

東日本大震災により半導体や電子部品などハイテク部品の工場が被害を受けた結果、日本製部品の供給不足が全世界の製造業の生産に深刻な影響を与えている状況が明らかになってきた。本稿では、アジアに進出する日系製造業を中心に、今後必要になる取り組みをサプライチェーンの再構築とIT活用の側面から考察する。

分断された部品サプライチェーン

3月11日の地震発生から本稿の執筆時点で約2カ月が経過したが、ここにきて世界の大手自動車メーカーや電機メーカーの海外工場が日本製部品の調達難から生産を縮小したり、ラインを一時停止したりする事態が明らかになってきた。

地震の後しばらくの間は、工場在庫や流通在庫で生産を継続できたが、いよいよ追加の部品が入らなくなり、その原因を遡っていくと日本製の部品に行き当たったということのようである。このことは、日本製部品のサプライチェーンが想像以上にグローバルにかつ密接な連携をもって構築されていたことを世間一般に知らしめる結果となった。すなわちグローバルなサプライチェーンに歪みがあることが明らかになったのである。

サプライチェーンの歪みは、想定外の需給変動が起きた時に顕著に現れる。2008年9月のリーマンショックの際には急速な需要の減少に直面した製造業は川上から川下に向かって一斉に大幅な生産調整を行った。大手メーカーの生産調整の情報は連鎖的に川下の取引企業へと伝わって行くが、川下に行くほどタ

イムリーな生産計画の変更ができずに大きな損失が生じたという。

今回は、情報の流れはこれとは逆である。川下である部品メーカーからの供給が滞ってはじめて、さらにその部品メーカーのサプライヤーが震災の影響を受けていたことが分かったのである。裾野の広い川下で起きた異変は実態の把握が難しく、影響度の見極めに時間がかかるのではないかと考えられる。

製造業では、価格競争などを背景として、部品を外部から調達して自社ではなるべく在庫を持たないようにするための水平分業体制がグローバルに構築されてきた。その結果として、自社が発注するサプライヤーのその先は全く見えないという事態が起きている。また最近では、グローバルに分散していた生産拠点を効率性を重視して1カ所に集中させる傾向もあり、これも有事の際のサプライチェーンの分断をもたらしている。

アジアに進出する日系製造業に現れた影響

NRI香港では、アジアに進出している50社100拠点以上の日系製造業・商社に対しサプライチェーンマネジメントを支援するシステムソリューションを10年以上にわたり提供して

NRI香港

社長

澤井啓義（さわいひろよし）

専門はアジアにおけるITソリューション



いる。具体的には、生産管理や販売・購買・在庫管理・会計といった企業の基幹業務を支援するERP（統合基幹業務システム）のクラウドサービスと、これと連携して社内外を含めた販売先・調達先との間の情報共有を支援する「NRI Intercom」という仕組みである。

今回の震災を受けて、日系企業のアジア拠点からは以下のような声が上がっている。

- ①リーマンショックの時は、顧客の発注が突然キャンセルになって“お手上げ状態”だったが、今回は日本側で一部の部品や原料の調達に影響が出ていると聞いている。それがいつどういう形で海外拠点に影響するのか分からない。
- ②日本で生産された部品や製品をアジアで販売しているが、日本の工場の生産調整で部品や製品の調達ができなくなると、自社のシェアをアジアや欧米の競合他社に取られるのではないか。
- ③これまで現地のサプライヤーからの調達比率を上げてきたが、そのサプライヤーが震災の影響を受けた日本の部品メーカーの部品を使っていたため、調達に困る事態になっている。

サプライチェーンの再構築に向けて

これらの問題に対しては、以下のような施策が必要と考えられる。すなわち、震災による日本からの部品供給の分断が世界の製造拠点の生産へどのような影響を及ぼすか、その

影響範囲を本社が把握し、最適なグローバル生産計画を策定することである。そのためには生産計画のシミュレーションが必須であり、海外拠点の持つデータを本社に集約しなければならない。具体的には、各拠点の代替部品を含む部品構成表、受注、顧客からの内示、発注などのデータである。NRI香港では、これらのデータを抽出して本社に送るといったサポートを、サービスの一環として可能な限り無償で行っている。

アジアに進出する企業は、これまで各拠点ベースの最適化を主な目的にシステムの構築を行ってきた。拠点単位の需要予測に基づいた生産と発注計画の策定である。しかし今回の震災を機に、先のリーマンショックの際に迅速な対応を取れずに後手に回ったことの反省を踏まえ、グローバル最適の視点でサプライチェーンマネジメントを見直すべきである。例えば、一部の地域で起きた需給バランスの変化に対するグローバルな影響の迅速な把握と、それに対応する最適な生産活動のシミュレーションができる仕組みを本社主導で作る機運が高まるのではないかと考えられる。

今後、サプライチェーンの分断は徐々に修復されていくと思われるが、電力供給事情などにより、生産量がすぐに元の水準に戻ることは難しいため、生産の一部を海外に移転する動きも出てくるだろう。NRI香港は、このような動きに対してもITの側面から可能な限り支援を継続していく。 ■

NRI Web Site

- 『ITソリューション フロンティア』本誌記事およびバックナンバーは、野村総合研究所(以下、NRI)ホームページで閲覧できます。
URL : <http://www.nri.co.jp>
- 『ITソリューション フロンティア』に関するご意見、ご要望などは、氏名・住所・連絡先を明記の上、下記あてにお送りください。
E-mail : it-solution@nri.co.jp

NRIグループと関連団体のWebサイト

野村総合研究所 http://www.nri.co.jp	NRIネットコム http://www.nri-net.com	NRIセキュアテクノロジーズ http://www.nri-secure.co.jp
	NRIサイバーパテント http://www.patent.ne.jp	NRIデータテック http://www.n-itech.com
	NRI社会情報システム http://www.nri-social.co.jp	ユビークリンク http://www.ubiquiink.co.jp
	NRIパシフィック http://www.nri.com	野村総合研究所(北京)有限公司 上海支店 http://beijing.nri.com.cn http://shanghai.nri.com.cn
	野村総合研究所(上海)有限公司 http://consulting.nri.com.cn	野村総合研究所(香港)有限公司 http://www.nrihk.com
	NRIシンガポール http://www.nrisg.com	NRIソウル支店 http://www.nri-seoul.co.kr
	NRI台北支店 http://www.nri.com.tw	(財)野村マネジメント・スクール http://www.nsam.or.jp

マッチング・ポータルサービス

B2Bポータルサイト「BizMart」 http://www.bizmart.ne.jp	情報収集、情報交換、商取引などの企業活動を総合的に支援する企業間ネットワークサービス
--	--

ナレッジ・ポータルサービス

NRIサイバーパテントデスク http://www.patent.ne.jp	国内外の特許情報や主要企業の技術雑誌(技報)の検索・閲覧サービス
情報技術本部サイト http://www.nri-aitd.com	最先端のITに取り組む技術集団である情報技術本部の活動内容や研究開発を紹介
日本企業台湾進出支援「ジャパンデスク」 http://www.japandesk.com.tw	台湾経済部と共同で、日本企業の台湾進出を支援

ソリューション・サービス

オブジェクトワークス http://works.nri.co.jp	MVCモデルに基づくWebアプリケーション開発のためのJ2EE準拠開発フレームワークの紹介
BESTWAY http://www.bestway.nri.co.jp	金融リテール投信ビジネスの“De-facto”スタンダードシステム。100社を超える金融機関が利用中
TRUE TELLER (トールテラー) http://www.trueteller.net	コールセンターからマーケティング部門まで、様々なビジネスシーンで活用可能なテキストマイニングツール
統合運用管理ソリューション (Senju Family) http://senjufamily.nri.co.jp	NRIが培ったノウハウを結集した統合運用管理製品群。企業の「ITサービスマネージメント」の最適化を実現
PCLifecycleSuite http://www.pcls.jp	企業内のPC運用コスト削減と品質向上を同時に実現する、PC運用管理の再構築サービス

インターネットリサーチ

TRUENAVI http://truenavi.net	NRIが戦略策定等のコンサルティングに際して独自に開発したインターネットリサーチを企業向けに提供
---	--

ナビゲーションサービス

携帯電話の総合ナビサービス「全力案内!」(ユビークリンク) http://www.z-an.com	携帯総合ナビサービス。世界初の携帯プロブ交通情報で道案内も。NTTドコモ、au、ソフトバンクから提供中
--	---

編集長 野村武司
編集委員(あいうえお順) 安藤研一 五十嵐 卓 井上泰一
岡田充弘 尾上孝男 佐々木 崇
鈴木昌人 田井公一 武富康人
鳥谷部 史 野口智彦 広瀬安彦
三浦 滋 見原信博 八木晃二
吉川 明 若井昌明
編集担当 小沼 靖 墨屋宏明

IT^{ソリューション}フロンティア

2011年7月号 Vol.28 No.7 (通巻331号)

2011年6月20日 発行

発行人 嶋本 正
発行所 株式会社野村総合研究所 コーポレートコミュニケーション部
〒100-0005 東京都千代田区丸の内1-6-5 丸の内北口ビル
ホームページ <http://www.nri.co.jp>

発 送 **NRIワークプレイスサービス株式会社** ビジネスサービスグループ
〒240-0005 横浜市保土ヶ谷区神戸町134
電話 (045) 336-7331/直通 Fax. (045) 336-1408

本誌に登場する会社名、商品名、製品名などは一般に関係各社の商標または登録商標です。本誌では®、「TM」は割愛させていただきます。

本誌記事の無断転載・複写を禁じます。

Copyright © 2011 Nomura Research Institute, Ltd. All rights reserved.

