

# 電子商取引と人工知能技術

# 06

神場知成 (NECビッグロブ(株)) : kamba@cw.jp.nec.com

## 電子商取引とその種類

電子商取引は、大きくは B2B (Business to Business : 企業間の取引), B2C (Business to Consumer : 企業と個人との取引), C2C (Consumer to Consumer : 個人間の取引, いわゆるオークションなど) の 3 種類に分けられる。それぞれ取引のルールや参加者が異なるので、必要となる技術も若干異なるが、本稿では、この中で特にマーケティング領域に焦点を当てて紹介をする。これは、マーケティングという分野が本質的に、本特集のテーマである人工知能 (以下 AI) ときわめて関連が深いだけでなく、Google や Amazon といったキーワードに代表されるように<sup>1)</sup>、検索とそれに連動する広告や、個人の嗜好に基づいたマーケティングがビジネス的に非常に注目を集めており、技術的にも急速に発展している分野だからである。説明は、分かりやすさのために B2C 領域を例にとるが、B2B の領域でもほぼ同様である。なお、当然のことながら、電子商取引と AI との関連は、マーケティング分野に限定されるわけではない。これに関しては、スペースは限定されるが、最後にエージェント、セキュリティの分野に触れる。

## 市場規模

最初に、ここで紹介する技術が関連する市場の規模感を示しておく。経済産業省、電子商取引推進協議会、NTT データ経営研究所が 2005 年 6 月に発行した「平成 16 年度電子商取引に関する実態・市場規模調査」によると、国内の B2B 市場規模は 2004 年に 102 兆 6,990 億円で、2003 年調査の 77 兆 4,320 億円から 33% の拡大、B2C 市場規模は 2004 年に 5 兆 6,430 億円で、2003 年の 4 兆 4,240 億円から 28% 増となっている。

さらに、このような電子商取引市場の成長を支える分野としてのマーケティングという観点から、広告事業の市場を見ると、電通が 2006 年 3 月 1 日に報道発表した「2005 年 日本の広告費」によると (<http://www.dentsu.co.jp/marketing/adex/adex2005/index.html>),

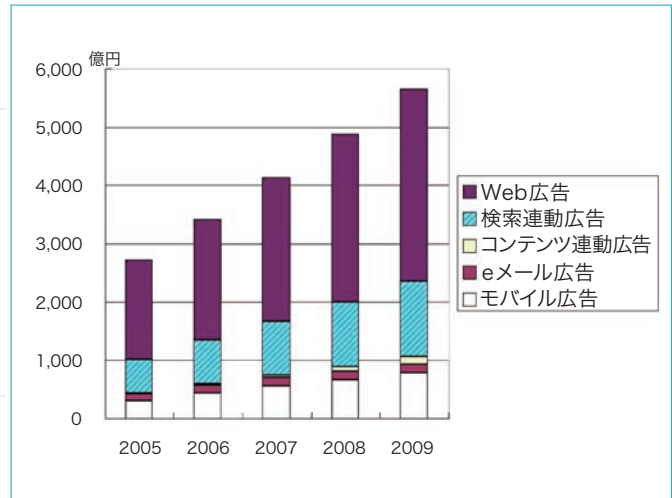


図-1 インターネット広告費の予測 (電通総研 2005 年 7 月 20 日発表数値)

2005 年の日本の広告費総額は 5 兆 9,625 億円で、内訳として、マスコミ 4 媒体 (テレビ, ラジオ, 新聞, 雑誌) 広告費が 3 兆 6,511 億円, セールスプロモーション費が 1 兆 9,819 億円, インターネット広告費が 2,808 億円となっている。この中でも注目すべきはインターネット広告費であり、マスコミ 4 媒体広告費が前年比 99.3% と伸び悩む中、前年比 154.8% という成長を示している。

ほとんどの広告には何らかのかたちで IT (information technology) 技術が関係し、その中でも特に、インターネット広告費の部分が、最も直接的に IT 技術の影響を受けている部分であることは間違いない。そして、現在のインターネット広告の急成長は、検索キーワードと連動して表示される、検索キーワード連動広告や、利用者の属性から効果的な広告を推定して提示するターゲティング広告に拠る部分が大きく (図-1)、これらはいずれも、IT 技術の中でも特に AI と深く関係する技術に支えられている。

以下では、インターネット上での広告や、それを活用したマーケティング (以下、e マーケティングと呼ぶ) に関し、関連技術などを紹介する。



図-2 検索とキーワード連動広告

など)するので、マーケティングの観点からは、いったん顧客となった消費者に対して、再度 Attention (注目) 行動を喚起することが重要となる。

### 検索とキーワード連動広告

前章で述べたように、マーケティングにおける重要性が急速に成長しつつあるサービスが、「検索」であり、それがきわめて高度な技術をベースにしたサービスであることはいうまでもない。最終的に巨大なビジネスとなっているのは、検索と連動して表示される「検索キーワード連動広告」であるが、それはあくまで検索というサービスがあって初めて、その上に成立するビジネスモデルである。検索キーワードに対して、利用者が求めるような結果を適切に表示し、しかもそのときに、広告主の要望を満たすだけでなく、利用者から見ても納得のいくレベルの広告を表示することで、検索サービスに対する利用者の再利用が促進されるからである。

図-2は、Googleで「自動車保険」を検索した結果の表示画面であるが、画面右側には、保険会社各社のサイトへ誘導するリンクが列挙されている。

本章ではGoogleで用いられている検索技術、キーワード連動広告サービスを中心に述べる。まず、全体アーキテクチャに関しては、Webページを高速に収集してインデックスを作成し、検索キーワードが入力された際に、該当するものを高速に抽出することが必要になる<sup>3)</sup>。これは大規模なWeb検索エンジンにおいて必須の技術であるが、Googleの最も重要な特徴は、PageRankという概念により、ページのランク付けを行い、それを考慮することで検索結果の表示順序を決定していることである<sup>4)</sup>。通常のデータベース検索の分野では、検索結果においては精度と再現率とのバランスが最も重視される場合が多いが、Webページの検索では、ページの質に大きなばらつきがあり、しかも通常、検索結果が膨大な数になるため、検索結果の表示順序が最も重要である場合が多い。PageRankは、それをコントロールする重要なパラメータである。

### ⇒ GoogleのPageRank

GoogleのPageRankの基本的な考え方は、よく知られるように、「重要なページからたくさんリンクされているページは、重要なページである」という仮説に基づいている。計算の詳細は文献4)に述べられているが、きわめて単純化したモデルでは、図-3に例を示すように

- ランクが  $R(\alpha)$  であるページ  $\alpha$  から、 $N_\alpha$  個のページにリンクが張られているとき、それぞれの被リンクページに対して  $R(\alpha)/N_\alpha$  の重みを分け与える。

### 広告と商品の購買プロセス

まず、これから述べる技術が、eマーケティングのプロセスの中でどこに位置付けられるかを確認する。最近では、人々の購買プロセスに変化が起きつつあるといわれている。それは、「AIDMA (アイドマ) から AISAS (アイサス)へ」というものである<sup>2)</sup>。AIDMAは、マーケティングの分野でよく知られる、人々が物の購買に至るプロセスを示すもので、「Attention (注意喚起) → Interest (興味) → Desire (欲求) → Memory (記憶) → Action (購買行動)」の頭文字をとったものである。具体的に消費者の立場で考えてみれば、「テレビCMや雑誌広告を見て商品に注目し、興味を持つうちに欲しくなって、記憶に残ったものを店頭で見て購入する」というプロセスである。以前から、人々が新しいものを購入するときのプロセスは、典型的にはこのモデルに従うといわれていた。これに対し、最近の購入プロセス AISAS は「Attention → Interest → Search (検索) → Action → Share (共有)」というものになりつつあるといわれている。具体的には、「テレビCMや雑誌広告を見て商品に注目し、ネットで検索をして詳細を調べたり、比較検討してから購入し、その経験や感想などをブログに書き込むので、その情報は結果として他者と共有される」というものである。ポイントは、物を販売しようとする側にとって最も重要な、購買直前のプロセスに「検索」があってしかも、特にネット上で購買が完了する場合などはそこから購買までの距離がきわめて短いこと、さらに、利用者の購入経験は、ネットが存在しない時代には考えられないほどのスピードと範囲で人々に共有されるということである。

なお、上記のプロセスは1回で完了するわけではなく、消費者は同じ商品を定期的に何度も購入したり(例:歯磨き粉)、同じカテゴリの異なる商品を何度も購入したり(例:本、音楽CDなど)、1度購入したものに対する関連商品、周辺機器を多数購入したり(例:車、PC

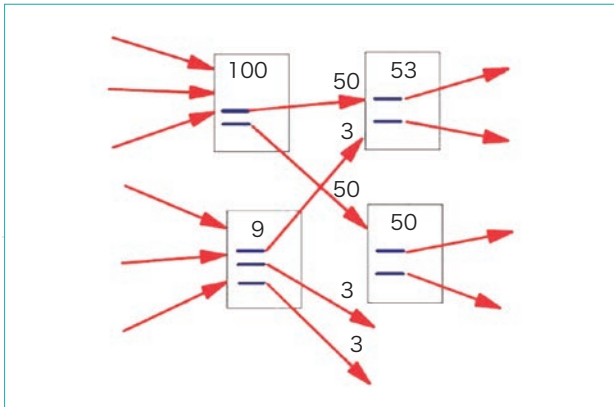


図-3 PageRank の考え方 (文献4) より引用)

- リンクを受ける側のページのランク  $R(u)$  は、そのページにリンクがあるすべてのページの集合を  $B_u$  とするとき、

$$R(u) = c \sum R(v) / N_v \quad (\text{ただし } v \in B_u) \quad (c \text{ は Web 全体での正規化のための定数})$$

となり、その計算は、多次元連立方程式を解くことに帰着し、限られた範囲のページであれば、一定の値が求められる。

なお、上記の計算方法は、文献4) に述べられる中で最も単純なかたちであるので、実際の PageRank にはさらに複雑な計算が必要になるし、検索結果の表示順序には、PageRank だけでなくページの更新頻度、ページ構成などさまざまな要素が影響を与えると考えられるが(その詳細なアルゴリズムは公開されておらず、また、ランクを計算する対象である Web ページの作成に用いられるテクニックが進歩するのに伴い、常に改良が継続されていると考えられる)、考え方の基本は分かるであろう。リンク元のページのランクが高ければ、被リンクページに分け与えられる値も大きいので、高ランクのページからたくさんリンクされているページは、ランクが高くなる。

ここで、上記の文献4) において、「このモデルは、直感的には、ネットをランダムにサーフィンする利用者が特定のページに行き着く確率をモデル化している」と述べられているのは興味深い。さらに正確な計算式においては、その利用者が途中でサーフィンを中断する確率も、考慮されている。余談であるが、筆者は AI というのはきわめて範囲が広く定義しにくい学問分野であるものの、その特徴は「シンプルな基本原理と、継続的に増大するアドホック処理とを柔軟に組み合わせた知的情報処理技術」と思っている。ここでの検索アルゴリズムに関していえば、各ページに対するランキングを定量的に行う「シンプルな基本原理」である PageRank は人間の行動モデルと整合するものであるし、その周辺で、検索

結果の品質向上のためには常にさまざまなアドホック処理が取り込まれている。

## ⇒ さまざまなメディアデータの検索

検索サービスは、単なる Web ページの検索だけでなく、あらゆるメディアデータを検索可能にする方向へと広がっている。前章では、検索の知的処理の基本原則が PageRank の概念にあるということ述べたが、メディアの種類によっては、それ以外のアドホックな処理や、ヒューリスティクスの方が重要なケースも多い。ここでは、いくつかのメディアデータによって、検索の際にどのような特徴が考慮されているか、例を紹介する。

### (1) Web ページ内の静止画、動画、音声など

Web ページ上に埋め込まれた静止画、動画、音声などは、そのままのかたちでは、検索の手がかりとなる「メタデータ」を付与されていないことが大半であるため、その処理の工夫が課題になる。たとえば静止画であれば、「その静止画の周辺に存在するテキスト」を解析してメタデータとする手法などが用いられている。たとえば、『大リーグ』や『イチロー』という単語が出現する文章の近くに埋め込まれている画像は、それらのキーワードと関係が深いであろう」というヒューリスティクスに基づいている。動画、音声などの場合も同様である。さらに、画像認識、音声認識などを用いてメタデータを抽出することも行われている。

### (2) ブログ、携帯電話用ページなど

Web ページの中で重要なページを見つけるには前述の被リンク分析が有効に機能しているが、ブログ検索結果のランキングでは、従来とは異なるヒューリスティクスも必要だといわれている。理由は、Web ページにおいては、あるページ A を作成するときに、すでに存在するページ B にリンクを張ることは一方向的であり、ページ A の被リンク数に対してはまったく影響を与えないため、ページ作成者が自ページへの誘導をコントロールすることはできないのに対し、ブログ C の作成者が、そこから他者のブログ D にトラックバック(あるブログへリンクを張ると、リンク先の相手に対してリンクを張ったことを通知する仕組み)をする際には、トラックバックを受けたブログ D 側から、ブログ C へのリンク(トラックバックへの参照)が自動的に作成される場合が多いことに起因している。これを悪用し、内容とは無関係な多数のブログに対してトラックバックを行い、それらのブログから、強引に自ブログへの誘導を行う「トラックバックスパム」も知られている。

また、これとは異なるが、携帯電話用 Web ページのランキングも、従来の Web ページとやや異なるテクニク

クが必要となるといわれている。これは、一般に携帯電話用のページが、PCを想定したページほど柔軟に相互リンクをせず、それぞれのサイト内で閉じたかたちで階層的に作成されることが多いからである。

通常の Web ページ、ブログ、携帯電話用 Web ページなどそれぞれ、仕組みや、それを作成する制作者の側の意図の相違により、ランキングには異なるテクニックが必要となっている。

### ⇒ キーワード連動広告

さて、このような検索サービスとそのランキング方式が非常に注目を集める技術分野となっているのは、前述のように、そのビジネス規模が急速に成長しつつあるからである。つまり、利用者が入力したキーワードに対する検索結果を表示する際に、それと並列して、そのキーワードと関連の深い広告を表示する手法である。広告費については、いろいろなかたちがあるが、よく知られているのは、オークション方式である。利用者が、表示された広告をクリックすると、そのクリック回数に応じて広告主に対して費用が請求されるが、その際に、1クリック当たり、広告主がいくらの費用を支払うかがオークション方式によって自動的に決定される方法である。キーワードによっては、1クリック数千円というような単価になっていることもあるが<sup>1)</sup>、広告主から見れば、具体的に意図を持って情報を探している人にターゲティングした広告を表示できるわけで高い効果が期待できるといわれている。

### レコメンデーション

ここまで、前述の AISAS というプロセスの中で、特に真ん中の S、つまり Search (検索) を中心に述べてきたが、Attention (注意喚起) の側面でも、AI 技術は非常に重要な役割を果たす。それは、レコメンデーション技術である。

ここで、技術の説明の前に、いったん、実際の e マーケティング業務において、どのようなことが行われているか、BIGLOBE における事例をもとに説明する。図-4 は BIGLOBE におけるマーケティング業務のプロセスを、図-5 はマーケティングを実行する eCRM 基盤 (CRM: Customer Relationship Management) を示す。位置付けを明確にするために背景を述べておくと、BIGLOBE は、もともとは NEC が提供するインターネット接続サービスからスタートしているが、現在は、それに加えて、保有するポータルサイト (<http://www.biglobe.ne.jp/>)

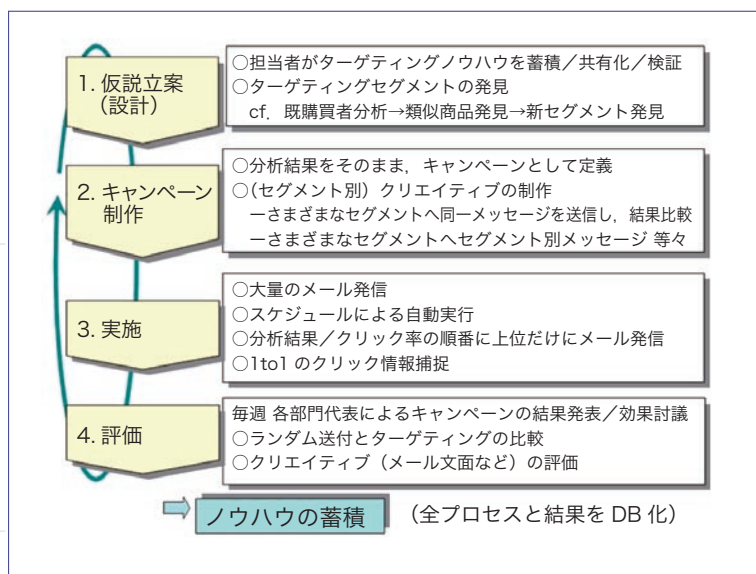


図-4 マーケティング業務プロセス

等) を中心として、広告、有料コンテンツ販売、オンラインショッピング等のブロードバンドメディア事業を行うとともに、そこで培われるプラットフォームシステム、マーケティングノウハウを活用した、企業向けのソリューション提供ビジネスを行っている。eCRM 基盤は、そのようなソリューション提供を行う際の基盤の一部である (詳細については文献 5) など。

eCRM 基盤上では、数百万人規模の顧客に対し、毎日発生する大規模な購買データ、クリックデータをもとにマーケティングを行うことができる。このような eCRM 基盤を用いて実際に行ったサービスプロモーションの事例として、あるアイドルに関するクリスマスギフトの例を挙げておく。図-6 は、その年の 10 月 9 日から 11 月 5 日までの、ギフトの予約状況を示している。図中にあるように、売り上げの山が 3 回あるが、1 回目は BIGLOBE トップページに掲載をして誘導したもので、2 回目と 3 回目は利用者に対してメール配信でプロモーションを行ったものである。ここで 2 回目のメール配信の方が 1 回目のメール配信よりも明らかに効果 (購入予約数) が大きい。実は、メール配信数は、第 1 回目の方がはるかに大きい。これは、1 回目の配信で、利用者の各セグメントでの反応を分析した結果、アイドルギフトであるにもかかわらず、意外にも 40 歳以上の男女の反応がよかったため、マーケティング担当者が、「お子様へのプレゼントなどの目的で購入しているのではないか」という仮説を立て、2 通目の配信では配信対象を該当セグメントに絞り込むとともに、メール文面の中でプレゼントとしての魅力を訴求するなどの工夫を行ったためである。これにより、2 回目のメール配信では、1 回目の実に 9 倍というクリック率があり、多数の購入予

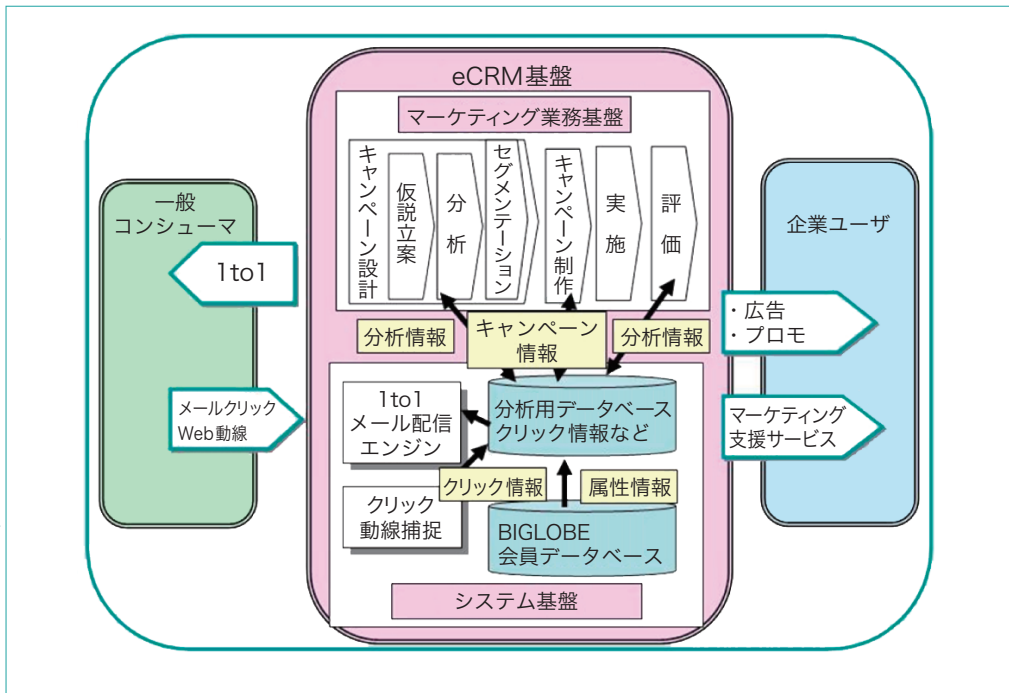


図-5 BIGLOBE の eCRM 基盤

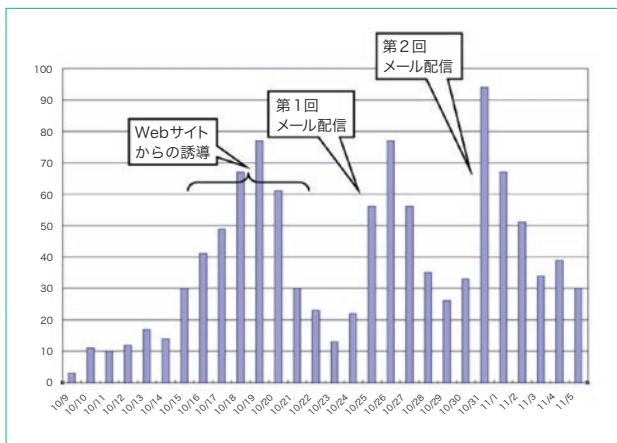


図-6 クリスマスギフト予約状況の事例

約に結びついている。

ここではメールによる e マーケティングの事例を挙げたが、きわめて知的な判断が行われていることが分かるであろう。レコメンデーション技術というのは、このような処理を部分的に自動化していこうとする試みである。現在のところ、レコメンデーション技術は、オンラインショッピングのサイトにおいて、購入ページあるいは購入直後のページにおける「お薦め機能」として実現されることが多い。利用者の属性や過去の購買履歴から、その興味にあっているであろうと推定される商品を表示することで、利用者の、次の行動を促すのである。レコメンデーションの方法としてよく知られるものは、次の2つである。

#### (1) 内容に基づくレコメンデーション

#### (2) ソーシャル・レコメンデーション

本や音楽 CD の購入を例にとると、(1) は、たとえば利用者が過去にモーツアルトの交響曲 CD を購入していれば、「その他のモーツアルト CD」「同時代の古典派作曲家の交響曲の CD」「モーツアルトに関連する本」などを推薦する方法である。このケースをシステム側から見れば、「モーツアルト」というキーワードを「古典派」「交響曲」などのシソーラス（同義語辞書）の中で管理して、それを手がかりとして関連商品を推薦することになるので「内容に基づく」方法といわれている。

これに対して、「ソーシャル・レコメンデーション」というのは、商品購入パターンから、利用者間の類似関係を学習し、それに基づいて推薦をする方法である。たとえば、前述のようにモーツアルトの交響曲 CD を買った人がいれば、同じ CD を買っている他の人たちが、他によく買っている商品を推薦することになるので、場合によっては、モーツアルトや交響曲とはまったく異なる、小説を推薦するようなこともある。システム側から見れば、購買パターンに基づき、購入者あるいは商品を自動的に分類し、利用者に対して商品を推薦したときの購買率を予測するアルゴリズムが重要となる。

これらのレコメンデーション技術は、オンラインショッピングサイトなどで利用されており、たとえば Amazon (<http://www.amazon.co.jp/>) のサイトに顧客がログインすると、そのトップページにお薦め商品が表示されるのと並んで、「おすすめの理由は？」というリンクがあり、それを見ると、推薦の根拠となった購入履歴や、同じ購入履歴を持つ人がよく購入している商品



図-7 BIGLOBE VIDEO STORE

のリストなどが表示される。同様に、BIGLOBEの有料動画ポータルサイト「ViDO STORE (<http://broadband.biglobe.ne.jp/vstore/>)」でも、サイト利用者のアクセス履歴や購入履歴などの利用動向の分析データから、そのコンテンツの利用者に人気のある他のおすすめコンテンツを表示するレコメンド機能を用いている(図-7)。

内容に基づくレコメンデーションも、ソーシャル・レコメンデーションも、従来からさまざまなアルゴリズムが研究されているが、すべての場合に最適なアルゴリズムというものがあるわけではなく、実際にオンラインショッピング等に適用する場合は複数の方式を組み合わせるパラメータのチューニングを行うことになる。これは、レコメンデーションというのが、いわば通常の店舗で、店員が顧客に対して商品のお勧めをするのに相当していることを考えれば当たり前である。商品カテゴリ(自動車か歯磨き粉か)によっても販売テクニックは異なるであろうし、同じ商品(たとえば自動車)であっても、ターゲット層(50歳のエグゼクティブか、20代前半の若者か)によって最適な売り文句はまったく異なるであろう。AIとの関係でいえば、レコメンデーション技術というのは、いわば、「優秀な店員を模倣するエキスパートシステムを作る」という課題に対する、AI分野における伝統的なエキスパートシステムとは異なるアプローチである。

### その他の研究分野

本稿ではここまで、AIと関連する電子商取引の分野として、主にB2CやB2Bの分野で重要になっている、eマーケティングを中心に述べてきたが、C2C、つまりオークションの分野に目を広げると、その他のAI関連の課題も多い<sup>6)</sup>。主なものは次の2つである。

#### (1) 取引エージェント

##### (2) 不正防止、セキュリティ

(1)の取引エージェントとは、価格交渉や購買などを自動的に行うエージェントである。最も単純な形態は、オークションサイトにおいて、買い手の側があらかじめ上限価格を設定しておく、その範囲内で、他者の入札価格を若干上回る価格を常に設定するような自動入札機能として、すでに多くのサイトで導入されている。さらに、複数のエージェントが存在する中での市場形成のメカニズムや、その中で目的を設定して最適な行動を選択する方法などについては、マルチエージェントの分野で研究が行われている。

(2)の不正防止、セキュリティも、AIと関連する話題の多い分野である。B2BやB2Cなど、販売する側が企業であるケースでは、認証方式の充実などで主要な問題が解決する場合もあるが、個人間売買が行われるC2Cの分野では、それだけでは手に負えない問題も多い。(1)とも関連するが、不正を防ぐためのオークションルール(オークションプロトコルと呼ばれる)を検討したり、あるいは、実際に発生している取引の状態から、データマイニングによって不自然なパターンを見つけることで、不正の発見を行う研究などが行われている。

### おわりに

本稿では、eマーケティングを中心に、AIと関連する技術が電子商取引の分野でどのように活用されているかを述べてきた。事例でも見たように、これは人間のきわめて知的な処理を自動化することでビジネスを拡大しようとするものであり、技術的にも、ビジネス的にも非常にチャレンジングな分野である。文中で述べたGoogleが、初期のAI分野でもきわめて有名なStanford大学のWinograd教授の研究室から生まれたことは偶然ではないと思う。AI分野から、ますます大きなビジネスにつながる技術が今後生まれることを期待している。

#### 参考文献

- 1) 日経ビジネス、Googleとアマゾン ネット消費の真の支配者、日経ビジネス2005年10月31日号、日経BP、pp.26-39(2005)。
- 2) 秋山、杉山：ホリスティック・コミュニケーション、宣伝会議、p.163(2004)。
- 3) Brin, S. and Page, L.: The Anatomy of a Large-Scale Hypertextual Web Search Engine. WWW7 / Computer Networks 30 (1-7), pp.107-117(1998)。
- 4) Page, L., Brin, S., Motwani, R. and Winograd, T.: The PageRank Citation Ranking: Bringing Order to the Web (<http://dbpubs.stanford.edu:8090/pub/1999-66>) (1998)。
- 5) NEC 技報 BIGLOBE 特集, Vol.57, No.6 (2004)。
- 6) 人工知能学会編：人工知能学事典、共立出版、996p。(2005)。

(平成18年6月5日受付)