

NTT DATA

株式会社NTTデータ数理システム

優秀な技術者がコンサルまで行う データ分析のソフト開発企業

およそ30年間にわたって独立系ソフトハウスとしての独自の地位を固めた後、昨年NTTデータグループの一員となったNTTデータ数理システム。「数理科学を世のビジネスに活かし、役立てる」を創業の理念に掲げ、パッケージソフトウェアによるソリューションを提供している。グループ入りをしたことでさらに活躍の場を広げたいという同社の山本二郎社長にお話を伺った。



NTTデータ数理システム 山本二郎社長

独立系ソフトハウスとして創業 2012年にNTTデータグループ企業に

◆貴社の設立背景を教えてください。

1982年に先代の社長が独立系ソフトハウスとして当社を設立しました。先代は工学博士でもあり、「数理科学を世の中にどう役立てるか」ということを考えてビジネスを興しました。創業当初から、半導体ビジネスには力を入れていて、半導体設計のもっとも基礎的な部分で数理工学的知識が必須な領域となるプロセスシミュレーションやデバイスシミュレーションを手掛けてきました。半導体ビジネスが成長している時期でしたから、当社のビジネスも急速に拡大していき、国内では競争相手のいないトップクラスのレベルといわれるようになりました。設立当初からNTTの厚木研究所とは協力しながら仕事を進めていました。そのころからNTTグループとの縁があったわけです。しかし徐々に半導体産業が衰退していったため、半導体設計シミュレーション関連のビジネスもそれにつれて縮小してしまいました。

そこで半導体に代わるものとして力を入れ始めたのが、創業時から温めて来た数理計画を用いた最適化および1990年代から注目を浴び始めたデータマイニングです。その両方とも時流に乗っていましたから、先進的技術をビジネス化したといえます。

◆NTTデータグループ入りした経緯とは。

先代の社長が引退を考える時期となり、どのようなかたちで会社を次世代に引き継ぐかを考えた際、どこかのグ

ループに入って活躍するほうが会社として良いのではと判断したと聞いています。さまざまな候補の中から、それまでも付き合いがあったNTTデータを選ばれたようです。

NTTデータにとってももちろん大きなメリットがありました。データ分析ができる人は現在非常にニーズが高く、NTTデータがビッグデータに取り組むためにもそのような人材が必要で、それを容易に確保できるからです。

2012年2月に契約を結んで、NTTデータは当社の株式を80%取得しました。さらに2013年6月には100%を取得し、完全子会社化としました。これに伴って9月には社名をそれまでの「数理システム」から「NTTデータ数理システム」に変更しました。

数理計画とデータマイニングを使った ソリューションを提供

◆業務概要をお聞かせください。

メインは数理計画ソリューションとデータマイニングソリューションです。そのほかにシミュレーションソリューションと科学技術ソリューション、知識工学ソリューションがあります。

数理計画とは、さまざまな組み合わせの中から最適なものを見つけることです。そうするとトレードオフを考えながら計画を立てられます。例えば、配送について考えるときに、倉庫にどのぐらいのストックを置いて、何台のトラックを使うのがもっとも効率的かというのを数字で表すことができるのです。そのほかにシフトの組み方や、生産性が上がるような機械の配置など、さまざまな局面でこのソ

リユースが活かせます。当社ではこのために「NUOPT」という数理計画パッケージを開発・提供しています。

データマイニングとは、予測やリコメンデーションなどを行うためにデータを分析してパターンを見つけ、ビジネスに役立てることです。基になるデータが数値やカテゴリーの場合には「Visual Mining Studio」を、日本語や英語などの文章である場合には「Text Mining Studio」というツールをお使いいただけます。

◆どのような業種のクライアントがいますか。

当社のお客さまは多種多様で、あらゆる産業分野にまたがっています。その中でも最近ではWeb関係が増えています。その他、製造系の企業とのお付き合いも多く、最近ではコンサルティング企業や広告代理店なども増えてきました。

クライアントとのお付き合いは、都内にてほぼ毎日実施している、当社のパッケージソフトの説明セミナーや年1回のユーザコンファレンス、展示会、学会での口コミなどから始まることが多いです。お客さまが実際に抱えている問題について話し合ううちに、当社のソフトをご利用いただけるようになるのです。

◆NTTデータグループとの連携はありますか。

NTTデータグループとお付き合いのある企業の中にデータ分析ニーズをお持ちのお客さまがいる場合、私たちもプロジェクトに入ってお手伝いすることがあります。それは当社の全体の売上からするとまだ5%程度なのですが、これからはもっと増えていくでしょう。逆に当社からもお客さまを紹介して、相互に連携を強めていけたらと期待しています。

◆自社開発のパッケージソフトで 顧客の抱える課題を解決

◆貴社の強みを教えてください。

当社ではソフトを自社開発しています。この点が何よりも大きいですね。自分たちが開発したものですから、自分たちの判断で機能追加ができます。ソフトをつくっていますから、何か具体的な課題があればそれに対応するツールも容易につくれるのです。また、お客さまにとっては「パッケージ」として気軽にお使いいただけます。高額な費用を払わずとも、会社の抱えている課題が解決できるのです。

実はこの業界の中でデータ分析にかかわるツールをこれだけ網羅的にそろえている会社は希少な存在です。しかも当社では自社開発のため、すべてがシームレスに連携できるように開発しており、そこが大きな強みとなっています。

さらに、開発技術者がお客さまに直接サポートやコンサルティングを行う点も当社ならではの強みです。

◆人材の採用・育成はどのように行われていますか。

中途採用はあまり行っておらず、新卒を育てています。ベースには数理科学の知識が必要となりますから、理学部の修士・博士出身の人が多くいますね。もちろん、学部卒や

他の学部の人もあります。「数理科学をビジネスに活かす」という当社創業時からの理念が学生の間にも共感を呼んでおり、入社希望者もとても多いです。今春には7名の新卒を採用しました。入社して直の3カ月は新入社員として技術的訓練を受けてもらいますが、後はお客さまとのやり取りの中で現場の技術を学んでもらいます。また、ビジネスマナーなどに関してはNTTデータグループの研修を受けています。



山本二朗社長

現在いる90名ほどの社員のうち、営業や総務に所属する数名を除いて、営業担当者も含め、すべて技術者。非常に優秀な人材がそろっています。例えば、いくつかの学会が毎年共催している「データ解析コンペティション」は今まで国内で19回行われている歴史あるコンペですが、昨年度は当社の若手が最優勝賞を獲得しました。

◆これから「とがった会社」として 世の中をリードしたい

◆今後の展望をお聞かせください。

先ほど申し上げたとおり、一通りのソリューションを持っているのが当社の強みですが、今後も世の中をリードする技術を持ち続けたいと思っています。「とがった会社」であり続けることが大事ですね。お客さまの抱えるいろいろな課題を解決する製品を生み出すことに加えて、さらにその製品にお客さまのニーズを組み込んで、より良い製品をつくっていただければ。そして、基礎的な研究も今までどおり継続していきたいです。

セミナーや国内の展示会、口コミをきっかけとしたお付き合いが主だったため、現在のお客さまは国内がメインです。しかし、グローバル展開を積極的に行うNTTデータグループに参加したことと、ビッグデータの時代を迎えたことにより、これからは国内だけではなく、欧米をはじめとする海外のビジネスにも取り組みたいです。

◆貴社の社風はどのようなものですか。

平均年齢は30代半ば、男性の社員が多いです。大多数が技術者ですので、自由な雰囲気ですね。服装も自由です。もちろんお客さまとの打ち合わせにはきちんとスーツを着ていきますが、外出前に着替えるために会社にスーツを置いておく人も多いです。

◆社員の方へのメッセージをお願いします。

いつも伝えていることですが、これまでもそうであったように、これからも技術の「とがった会社」であり続けましょう。また、お互いの知識やノウハウによる、さらなる切磋琢磨を期待します。

最先端アルゴリズムを使って 小さなメモリでも迅速にビッグデータを分析

数理計画部 研究員 白川 達也さん
数理計画部 研究員 二反田 篤史さん



左から二反田篤史研究員、白川達也研究員

新しい分野について学びながら ビッグデータの分析ツールを開発

◆開発メンバについてお聞かせください。

二反田さん：主に私たちを含めて4名で開発しています。私と白川は数理計画部に属しており、アルゴリズムの研究開発という、計算のコアな部分を担当しています。残り2名はデータマイニング部に所属しており、ビッグデータの分析・集計・加工といった前処理的な操作や手法・ツールの整備を行っています。つまり、当社の中でも部をまたがって技術を集約する混成チームで開発しています。

◆Big Data Moduleとはどのようなツールですか。

二反田さん：大規模データに対して、ルールやパターンを見つけるための機械学習やデータマイニングを行うツールで、2012年8月にリリースしました。例えばコンビニ店舗の売上が変動する際に、気候や周辺人口、イベントの有無などが影響を及ぼしている可能性があります。その規則を何かしらの手法で発見するものです。こういったものはデータが増えると、より確かな分析を行うことが可能になります。

◆開発時のご苦労はありましたか。

白川さん：2年少し前から準備を始めたのですが、そのころは私も二反田も完全な門外漢で、勉強も兼ねて頻りに打ち合わせを行うことが必要でした。ある程度知識が身につけてからも、こういった機械学習やデータマイニングの分野では新しい技術が日々生まれては陳腐化していくので、毎日それに追いつくだけでも大変でしたね。最新の論文をずいぶんたくさん読みました。逆にいうと、毎日新しいことが学べてもっとも楽しかった時期でもあります。

二反田さん：私は、いかに計算速度を速くするか、という点が大変でした。データ分析においては、データに意味のない情報やノイズが含まれているために、分析がうまくいかない場合が多くあります。その問題を解決するため、分析の過程においては試行錯誤を繰り返すのですが、大規模データを対象としていると、試行錯誤のたびの計算時間の増加が顕著なため、高速化が大きな課題でした。

ミドルクラスのデータ分析に最適 分割すればどんなに大きなものも扱える

◆Big Data Moduleの特長を教えてください。

白川さん：世の中のビッグデータの分析は集計程度のレベルにとどまっている例が多く見受けられます。集計はもちろん重要なのですが、私たちのモジュールが目指すのはそれを超えた深い分析です。ビッグデータというとテラバイトやペタバイトオーダのデータというイメージがあるかもしれませんが、ある論文によると、Hadoopのような大規模分散環境で実行される分析タスクでも実際には10ギガぐらいのデータ量が平均的であるといわれています。分析が必要なデータに整形するとミドルデータといっていい程度のサイズになるんですね。しかし、ミドルデータ程度でも、多くの分析アルゴリズムは計算限界を超えてしまうので、適切なアルゴリズムを採用することが必要となってきます。

二反田さん：私たちのBig Data Moduleはそのミドルクラスである数ギガ～数百ギガを主なターゲットとしています。そして、大規模な分散環境を用意しなくても、1台のマシンで高速に分析を行うことができる点が大きな特長といえるでしょう。もちろん、その上のサイズとなると、1台のマシンでは取り扱えませんので、その基盤としてHadoopとの連携も視野に入れていました。こちらは私たちとは別の2名が主に担当しています。

◆最先端のアルゴリズムを採用されたそうですね。

白川さん：オンラインマイニングアルゴリズムを採用しました。これはデータ1つを与えると微小な変化をモデルに加えるという操作を繰り返すものです。それによってどんなサイズのデータも処理が可能になります。

二反田さん：メモリにデータを載せて分析すると扱えるサイズが限られてしまいますが、オンラインマイニングアルゴリズムはデータを分割しながら少しずつメモリにロードして処理するので、全体のデータのサイズを気にしなくて済みます。メモリ領域自体を大きくするHadoopとはそこが異なる点ですね。

◆現在抱えている課題はありますか。

白川さん：ビッグデータのイメージが先行していて、ユーザがすぐにHadoopに飛びついてしまうことがもどかしいですね。Hadoopはオーバスペックになりがちで、10ギガ、100ギガ程度のデータをHadoopにかけると非効率です。そのような現実的なデータサイズであれば、当社のオンラインマイニングアルゴリズムのほうが効率的なので、これから普及させていきたいです。

教師なし学習を用いた
全自動分析ツールの開発が夢

◆今後の夢や目標を教えてください。

白川さん：実現は難しい夢物語ではありますが、ボタンを1つ押すと欲しい結果が出るような全自動分析ツールをつくれたらいいですね。その第一歩としては、教師なし学習に基づく分析を強化していきたいです。機械学習には教師あり学習と教師なし学習という2つの手法があるのですが、教師ありのほうは、機械がルールを発見するために人間があらかじめ正解データを用意しないとイケないものです。例えばスパムメールを検知する際には、「こういう文面はスパ

ム」という知識を前もってインプットする必要があります。教師なしでは、入力データを受け取ったら人間が正解データを用意せずとも、アルゴリズムが自動的にルールを発見して分析結果を出してくれます。全自動分析を行う場合もそうですが、ビッグデータを扱うときにも、正解データを前もって用意することが難しい場面が多いのです。ですから、十分にビッグデータを活用しようとする、教師なし学習が必要になると考えています。

二反田さん：私も教師なし学習に興味がありますが、直近の目標としてはBig Data Moduleのストリームデータ対応を考えています。センサデータ等のリアルタイムに送られ続けるデータを効率的に、分析をできればと思っています。

◆NTTデータグループとの連携に期待することは。

白川さん：NTTデータとの共同研究にも大いに期待しています。私たちはアルゴリズムを持っていますので、実際の業務システムとの連携など、NTTデータの持っている実データへの適用を多くこなしていきたいです。

二反田さん：NTTデータの所有する大規模な計算環境もぜひ使ってみたいですね。（インタビュー：村上百合）

NTTデータ数理システム ア・ラ・カルト

■和気あいあいと楽しむ社員旅行

2年に1回行われる社員旅行。1泊2日で、毎回50名ほどが参加するのだとか（写真1）。2013年6月7～8日に行われた湯河原旅行のテーマは「近場でゴージャス」！社員たちは宿の素晴らしさや食事の美味しさに大満足したそうです（写真2）。1日目は水族館かアーチェリー、2日目はお皿の絵付けが蒲鉾づくりと、好きなアクティビティが選べるのも好評の秘訣。部署の垣根を越えた交流をたっぷりと楽しみました。

■他部署の活動が分かる成果発表会

毎年5月半ばに催される成果発表会では、各部署から1名ずつの代表者が出て、40分ほどの発表を行います。技術開発や研究に関する成果など、他部署でどんなことをしているのかを知ることができるチャンス。それに関する質疑応答も熱心に行われます。成果発表会がある日の夜は新人歓迎会で盛り上がり、公私両面でかわわりを深められる良い1日になるそうです。

■テニスや映画鑑賞などのサークル活動も

多趣味な社員の多いNTTデータ数理システム。テニスや映画鑑賞、スキー、茶道、華道といったさまざまなサークル活動が営まれています（写真3、4）。活動イベントが決まると回覧板やメールで告知され、参加者を募るといったかたちで運営されているそうです。



写真1



写真2



写真3



写真4