

シリコンバレーに立つ

在学中に先輩と独自検索エンジンの会社を設立、事業が軌道に乗ったあと、シリコンバレーに新しい会社を起業。蓄積した技術で構築した分析プラットフォームでは、世界中から集まるデータが駆け抜けていく。



太田 一樹 Kazuki Ohta

Treasure Data, Inc. ファウンダー、最高技術責任者

太田 なつかしいな。僕は学生時代に会社を作ったけれど、オフィスがなかったから、夜になるとここによく集まっていたんです。

— スゴイですね。その後アメリカで起業と聞いて、ぜひお話を聞きたいと思っていました。

太田 僕も、2年前まで思っていなかったよ！

— 太田さんはACM ICPC国際大会の出場者でしたね。プログラミングはいつから？

太田 高校生だった2001年から。僕はネットゲームファンで、「ウルティマオンライン」や「ディアブロ」に熱中していたんです。高校3年生のときに、Javaプログラムが動作する携帯電話を買ってもらい、それでゲームを作って公開したら、なんと40万ダウンロード。

それからどんどんプログラミングにはまっていきました。それが、ふとしたきっかけでLinuxに触れて、これは性格かな、下のほう——基盤技術に向かって降りていきたくなったんです。

当時のLinuxは、うまく日本語を扱えない、ブラウザに日本語が出ない、というような不備がまだいろいろありました。それで、KDEというデスクトップ環境やUINという日本語入力環境を改良してコミッターになったり。

— 情報科学科に来ようと思ったのはいつですか？

太田 理科一類に入ったときから、情報科学科以外に進むつもりはありませんでした。(入学して)すぐに来たかったなあ。それで1、2年生のあいだは、週に1回フランス語と体育の授業がある日だけ登校して、あとはシステム開発のバイトばかりしていたんです。

— いまは、研究室の配属でどこに進むかを迷っています。太田さんはどうでしたか？

僕は昔から、OSやミドルウェアのようなシステムソフトウェアが好きでした。世の中を支えている感じがあるでしょう？ それで、研究室選考のときにいくつか書けるのに、1つしか

1985年生まれ。東京大学大学院情報理工学研究所修士課程修了。2008年、チームunknownのメンバーとしてACM ICPC国際大会に出場、13位を獲得。学部課程在学中の2006年、全文検索エンジンを事業の中核とする、株式会社プリファードインフラストラクチャ(PFI)の創設に参画するとともに、大規模データの分散処理ソフトウェアHadoopの普及とコミュニティを牽引。2011年に米国でTreasure Data, Inc.を設立、Hadoopを基盤とする、大量データの分散処理プラットフォームをサービス開始。大量データ処理の利用・導入をシンプルにしたことで、注目を集めている。

並列で素数の探索に熱中している人、マシン好きが高じてコンピュータ動物園の飼育係になった人、プログラム言語の美学を語ってやまない人たち、超成績がよいというより、いろいろな人がいる。そこが面白いね、情報科学科は。

書けなかったんです。ラッキーにも拾っていただけだけど、通らない可能性もあったから、そこは人生勝負ですね。所属していた石川研究室のテーマは、いかに速いコンピュータを作るか。そのために、並列計算やOSなどを研究しています。僕はそこで、ファイルシステムを勉強していました。

CPUとメモリはどんどん速くなるけれど、ディスクドライブのほうは実はぜんぜん速くなっていません。いちばん遅いところを速くしたいと思って注目したのが、分散・並列ファイルシステムだったんです。いかに多数のディスクを束ねて1つのファイルシステムに見せかけるか、読み書きを速くするか。そういうテーマを持って研究室に行っていました。

もっとも、3年生のときに先輩とPFIという会社を作って、会社をやりながら大学に通っていましたから、家で勝手にプログラミングしていることも多かったのですが。

世界に1つだけ解析用のデータベースがあるとしたら、それは自分たちのプラットフォームでありたい。

— そして、いまは独立して新しい会社を起業したんですね。そのころの話をお願いします。

太田 直接の契機は、PFIで海外との仕事で知り合った人と「アメリカで会社作らない？」という話になったこと。「いやあ、できたらいいですね」と言っていたら、本当にできちゃった。人生どうなるかわかりません(笑)。

— それはどんな会社なんですか？

太田 世界中にあふれている大量のデータを、僕らが開発したクラウドプラットフォームにどんどん送ってもらい、並列処理によって短時間で集計・分析し、結果を返す、そういうサービスです。

大規模なWebサイトやオンラインゲームのログ、コンビニで何か買ったときの記録、セン

サーが出し続けるデータ。それから携帯電話も、実は近くの基地局と通信してログを出しています。そういう膨大なデータが相手です。これまでできなかったような大量のデータ処理が可能になれば、新しいサービスや運用の効率化につなげることができます。

— その会社で、普段はどんな仕事を？

太田 いまはプログラムを書くというより、プロダクトの方向付けや、その実現に必要な技術——エンジニアを集めることなどを行っています。それから、お客さんのところに行って、プロダクトに足りないものを聞いたり、投資家にプレゼンしたり。社内のLANをセットアップするような雑用もしますよ(笑)。ひと言でいえば、コンピュータサイエンスのバックグラウンドがあるビジネスの人ですね。

— 卒論や修論でやっていたファイルシステムは、直接は関係がないのですか？

太田 僕が卒論や修論にしたテーマは、スーパーコンピュータが対象だったから、直接活きているわけじゃない。でも、論文を書く過程で得たいろんな知識や、それが体系化されたものは、すごく活かしています。

どうということかという、論文を書くときにはかなりの量の論文を読んで、いろんな分野を調べます。すると、この分野にはこの人が考えたこういう技術A、それとは別の技術Bがあり、そこから技術Cが派生して、これからこういう方向性で進化するだろう、というように自分の中でマップを作れます。知識を体系立てるといって、ちょっとかっこ良さすぎるかな。

僕はファイルシステムの論文を調べていくなかで、世界にどういう課題があり、何がボトルネックになっているか、世界にどんなデータベースや解析システムがあり、それはどういう技術をベースにしているか、ひととおり頭の中かに整理できました。だから、卒論や修論で読んでいた有名な国際学会の論文は、いまだに読んでいます。

— 会社でしているプログラミング以外のことは、どこで身に付いたのですか？

太田 やっぱ最初創設にかかわったPFI。最初6、7人で始めて、徐々に売上げやメンバーが増えて30人以上に成長してくると、組織をつくる、束ねる、ということが必要になります。だから、会計の本や経営の本は意識的に読んでいま

した。異分野の本も面白いですよ。ライバルの会社をいかに自分のテリトリーからたたき出すか、そんなことが書いてあったり。ホントかな？(笑)。

— なぜ日本ではなくアメリカで会社を興したのでしょか？

シリコンバレーからは新しいソフトウェアの会社がたくさん出ています。いろいろある理由のひとつは、ソフトウェアで世界にインパクトを与えた人たちが投資家になっていること。あとから出てきた会社にお金を出すだけじゃなく、アドバイスをしながらいっしょに会社を育てていく文化があります。会社を作るのなら、そこかなと思いました。

それまでの海外経験も大きいですね。3年生だった2005年、まだすごく小さくて急激に成長しているさなかのGoogleキャンパスに遊びに行きました。そこでは、OSカーネル、GCC、PhthonからYouTubeの開発者までが、一堂に会していた。それを見て、シリコンバレーへの憧れのようなものができました。

もうひとつ、これは指導の先生にとっても感謝していることなんですけれど、海外の研究所に送ってもらえたこと。それで向こうの雰囲気をつかみ、それまでぜんぜん話せなかった英語にも慣れました。

コンピュータは世界を変えられる。ほんとは、コンピュータが人より得意なことを、いかに現実世界に応用するか、それはいちばん楽しい仕事です。

— 最後に、いま情報科学科にいる人、これから情報科学科に来ようとしている人へのメッセージをお願いしますか。

太田 困ったな……ちょっと大きく言うと、コンピュータは世界を変えられると思っています。ほんとに変えられるよ。個人でも大量の人に同じ物をネットワーク経由で配信できるのは、とてつもない魅力です。個人でチョコレートで40万人に売るビジネスはできないものね。人より得意なことを、いかに現実世界に応用するか、それはいちばん楽しい仕事。

いま3年生？ ちょうどCPU実験のころ？ 大学時代のなかでも、君たちくらいのころはすごく楽しかったな、僕は。ぜひ、学生生活を楽しんでください。

聞き手：生田 拓人+島田 工