

3 DSL (Data Science Lab.)

あらゆるサービスの利用状況を分析し 意思決定に役立てるためのデータ分析基盤を整備

技術開発部 担当部長 亀井 聡

NTTコミュニケーションズ（以下、NTT Com）技術開発部のデータサイエンスチーム“Data Science Lab.（以下、DSL）”は、インフラオペレーションカンパニーならではのデータ分析を行い、その結果を意思決定に役立てる取り組みを進めている。

意思決定に役立つ、NTT Com ならではのデータ分析を推進

回線の増設や、データセンターの設置など、設備投資はもちろん、インターネット上でのビジネスにおける意思決定を適切に行うためには、インターネットそのものの変化を正しく捉えた上での判断が重要だ。GAFaやBATのようなITジャイアントが主要通信の運び手として存在感を増しているが、NTT Comにはインフラオペレーションカンパニーとしての強みがある。お客さまのインフラ運用を任せられることも多く、ネットワーク（以下、NW）やサーバーの運用データを収集しやすい。NTT Com 技術開発部のDSLはこの強みを活かし、効果的な設備投資のほか、自動化、効率化、コスト削減などにつながるような、経営幹部やサービス部門の意思決定に

役立つデータ分析を行っている。

かつてはお客さまのデータ分析支援のような個別案件への対応が多く、スキルも属人的になりがちであったという反省から、ノウハウを収集し、再利用可能な形にまとめる devops 的な開発手法に加え Reproducible Analysis、Infrastructure As a Code といった取り組み、社内で不足しているデータサイエンティストの育成も行ってきた。

データ収集・分析の事例

トランジット比較

インターネット最上位の tier1 の通信を担う ISP は、tier1 の ISP 同士でトラフィックを直接融通し合うことでインターネット全体への接続性を確保している。他の ISP はこの ISP 群からトランジットを購入することでインターネットへの接続性を確保する。



亀井 聡

このトランジットの実態を調査するため、国内の主要トランジット業者である7社のISPと契約し、各ISPのNWにトラフィック計測を目的とした経路広告を行い経路の強度を継続的に観測している。

インターネット体感品質測定

インターネットサービスを利用するお客さまの体感品質には、NTT Comだけでなく、他社提供によるNWも関係する。お客さまの満足度を損なわない適切な設備投資の実現を目的に、日本全国でパートナー会社のスタッフ宅に PPPoE、IPoE、フレッツなど、さまざまな回線を引き、測定機器を約500台設置して実際の体感品質を測定した。

このほかにも2018年1月号で紹介した、「顧客接点データを活用した自社サービス状況の解析・可視化」など、さまざまなデータ収集・分析

インターネットの変化を捉えた意志決定	Infrastructure Operation Companyとして	新たなサービスを産み出す
<ul style="list-style-type: none"> 主要通信の運び手がキャリアからITジャイアントに。 サービス開発、設備投資を適切に行うことが困難に。 どこに回線を増設するか、データセンターを配置するか。 	<ul style="list-style-type: none"> 自動化やコスト削減品質向上等きめ細かなサポート。 自営インフラの効率化。 アウトソースされたエンタプライズNW・エッジNWの運用効率化。付加価値生成。 	<ul style="list-style-type: none"> CDN / DDoS / SeCaaS w/ VxNF基盤 / Data解析基盤 領域・組織横断で連携した意志決定を容易に、データドリブンで実施する。 解析基盤そのものの外販

図1 NTT Comにおけるデータ分析の主要な目的

- ・インターネット上の様々なコンテンツに対する利用者の体感品質を測定する
- ・さまざまな回線を、パートナー会社のスタッフ宅等に設置
- ・2013年11月から台数、測定エリア、回線、システムを更改しながら測定中
- ・現在国内300箇所。Probe x ISPは500。日々の計測数は20万程度。

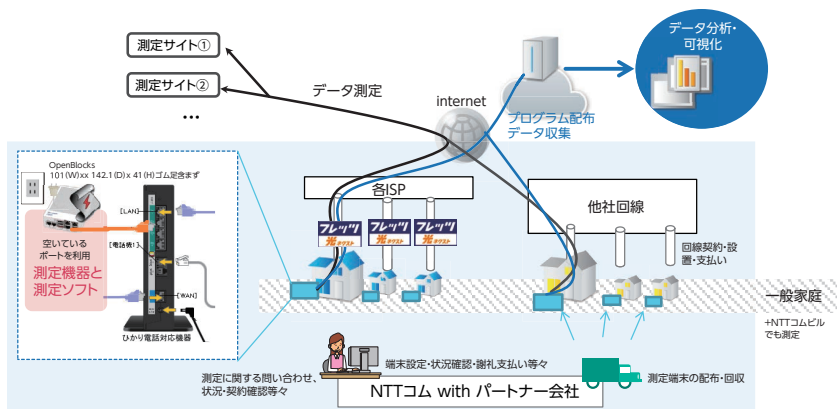


図2 データ分析の事例例：インターネット体感品質測定

の実績がある。グループ会社のデータ分析に協力するといった活動も行っている。

OSS を活用し低コストに 拡張可能なデータ分析基盤を実現

通信の状況を把握・分析しても、マネタイズにつなげることは難しい。継続して試行錯誤するためにも、分析そのもののコストや手間はなるべく抑える必要がある。また増大するデータへの対応も重要な課題だ。そこでHadoopをはじめとするOSSを最大限に活用し、Dockerのようなコンテナで展開可能な分散型のデータ分析の仕組みを作り上げてきた。

商用向けデータ分析基盤を整備 お客さまへの提供も視野に

NTT Com は次世代プラットフォーム推進室を立ち上げ、さまざまなサービスで共通的に利用できるプラットフォームの整備に取り組んでいる。その一翼を担うのが商用サービス向けデータ分析基盤の整備だ。

背景の1つに、NWでもあり、か

つサーバーでもあるNFV(Network Functions Virtualization)のような機能の増加がある。かつて明確に分かれていた領域・組織を横断して連携する必要性が高まっており、その難しさを軽減するためにも、データドリブンで意思決定を行えるようにしたい、という思いがある。

通常、サービス運用にはその運用観点で必要なデータしか蓄積されていないことも多い。DSLの取り組みを開始した当初は自社サービスのデータ収集も容易ではなく、サービスごとに主管組織の協力を得ながらデータを収集していた。今回、予めデータ分析基盤をサービス提供基盤に組み込むことで、今後開発するサービスに関してはサービス主管組織に負担をかけることもなく、データを効率よく収集・蓄積し分析することが可能になる。あわせて匿名化処理や統計化処理を一元化することで、お客さまデータを安全に扱うこともデータ利用者にとっては余計な負担をかけずに済む。

これまでのDSLの取り組みを

ベースにデータ分析基盤を開発し、サービスを横断してデータを分析、可視化する仕組みなど、基本的な機能の実装を終えている。近く商用サービスへの適用を開始する予定だ。まずは自社で使いこなした上で、このデータ分析基盤をお客さまに提供することで説得力のあるサービスとなる。設備投資に役立つサービス実態の把握といったメリットを訴求していきたい。

全トラフィックの把握

現在は「データを制する者がビジネスを制する」という状況にあると言って良い。NTT Com ならではのデータ活用を実現するという意味でも、今後、Flow分析によりNWの全トラフィックを把握することに挑戦したいと考えている。データ量が膨大であり容易ではないものの、技術的な目処はついている。

データ分析にはドメイン知識も 必要。人材育成にも注力

一方でデータに踊らされてはいけないという思いもある。データ分析をビジネスに活かすために必要と考えられている3つの要素のうち「統計や分析の知識」、「Hadoopなどを扱えるスキル」の獲得は技術開発部のメンバーにとって特に問題ではない。しかしもう1つの「各分野のドメイン知識」は不足しがちであった。サービスを知っている人の思いこそがカギを握ることも多いので、そうしたことを踏まえて人材育成を進めていく方針だ。