

Google Cloud による放送アクセスログ分析で、 放送業界全体が利益を創出できる仕組みを目指す



コンテンツビジネス局 ネットデジタル事業部 三浦 一樹 氏



北海道テレビ放送株式会社

<https://www.htb.co.jp/>
〒060-8406 札幌市中央区北 1-1-6

1968 年に北海道初の UHF 局として誕生した民間放送局。コールサインは JOHH-DTV で、リモコンキー ID は 6。ANN / テレビ朝日ネットワーク系列に属しています。2018 年に開局 50 周年を迎えたのを機に、さっぽろ創世スクエアに本社を移転。「HTB は夢見る力を応援する広場です」という企業理念に基づいて、テレビジョン放送、コンテンツビジネス、各種イベントなどの事業を展開。開局 50 周年ドラマ「チャンネルはそのまま！」が、日本民間放送連盟賞グランプリを受賞しています。

数多くの受賞歴を持つニュース・ドキュメンタリーやバラエティ番組、ドラマなど、さまざまな分野の番組制作や放送を通じ、「ユメミル、チカラ」を応援し、地域の未来への貢献を目指す北海道テレビ放送株式会社（以下、HTB）。放送サービス向上の一環として、1 秒間に 100 万回のアクセスにも耐える放送アクセスログ分析システムを構築。HTB のネットデジタル事業担当者に、今回のシステム構築について話を伺いました。

（利用している Google Cloud サービス）

[Cloud Functions](#)、[Cloud Pub/Sub](#)、[Cloud Dataflow](#)、[BigQuery](#)、[Cloud Datalab](#)、[Google データポータル](#)

BigQuery にアクセスログ データを 蓄積しておけば後は何とでもなる

「現在の放送業界では、視聴率という指標により、ビジネス スキームが確立されています。一方で、インターネット接続型のテレビの普及により、アクセスデータログが取得できるようになりました。とはいえ、北海道内の約 50 万台分のデータ放送のアクセスログを収集し、分析して、新しい指標を作っても、現状では放送業界全体の指標として認知されません。また、スポンサーにその指標を認めてもらわなければ価値もありません。そこで、放送アクセスログ分析の標準化が重要になります」と話すのは、コンテンツビジネス局 ネットデジタル事業部の三浦一樹さんです。

現在、放送アクセスログ分析では、主にチャンネルの「オン / オフ」の時刻を取得しています。三浦さんは、「データ放送のアクセスログを分析したデータと番組表のデータを組み合わせ、どれだけの価値を生み出せるかが今後の課題です。最終的には、放送業界全体で利益を創出できる仕組みを目指しています」と話します。

放送アクセスログ分析を、放送業界全体に浸透させるためには、日本全国の放送局が同じ指標を使う必要があります。しかし地方の放送局は、技術職が 1～2 人しかいないことも多く、そこで利用してもらうためには、簡単かつ低コストで運用できる仕組みでなければ受け入れてもらうことは困難です。そこでまずは、必要なデータを安価に収集できる方法を模索します。

「テレビは 24 時間放送されているので、せっかくのデータが捨てられてしまうのはもったいないと思っていました。そこで、BigQuery にアクセスログ データを蓄積しておけば、後は何とでもなると考えました。蓄積したデータは、すぐには分析に使えないかもしれませんが、放送業界全体でデータの標準化ができれば、短期間で対応することができそうです。すでに系列の地方局から、話を聞きたいという要望もあります」（三浦さん）

Cloud Functions のベータ版がリリースされたことで GCP の採用を決定

放送アクセスログ分析を開始した背景を三浦さんは、次のように話します。「2015 年にデータ放送による情報提供を担当するようになり、その年の統一地方選挙の番組用に、当選、落選のデータ管理の仕組みを作成しました。このときのアクセスログにより、8 時に開票開始から約 4 時間で 4,000 ～ 5,000 人がデータ放送を見ていることが分かりました。そこで、オンプレミスのサーバーで、2016 年より放送アクセスログ分析を開始しました」

当初は、オープンソース ソフトウェア (OSS) の大規模データ分散処理ソフトウェア フレームワークである Apache Hadoop やクラスタコンピューティング フレームワークである Apache Spark を使うことを検討していました。しかし Hadoop や Spark は、かなりのハードウェアリソースが必要で、また専任の管理者も必要のために断念。クラウドサービスを使うことを前提に、2016 年より再検討を開始、2017 年、ちょうど当時 Cloud Functions のベータ版がリリースされたことを知り、Google Cloud Platform (GCP) の採用を決定、2018 年には放送アクセスログ分析をスタートします。

Cloud Functions のテスト期間は約 1 年。その間、Google Cloud にさまざまなリクエストを出しています。三浦さんは、「特に、スパイクに耐えられるかが重要でした。ベータ版では 1 秒間で 1 万アクセス程度に

しか対応できなかったのですが、テストが終わるまでには正規版がリリースされるだろうと思っていました。正規版では、100 万アクセス程度まで機能拡張されていたので、われわれのリクエストのおかげだと思っています」と笑います。

北海道テレビ放送で、もっともスパイクする番組は、北海道日本ハムファイターズの試合です。三浦さんは、「人気選手が登場したときとか、得点が入ったときなどに一気に増え、ピンチになったり、点を取られたりすると一気に減るといったことが顕著です。それでも、Cloud Functions は、1 秒間に数万アクセスが普通に処理でき、エラーも発生しないので、この機能をこの価格で実現できる Google Cloud はすごいと感じました」と話します。

システム構成は、Cloud Functions でデータを受け、Cloud Pub/Sub で Dataflow とメッセージングを行い、BigQuery に蓄積していく仕組みになっています。BigQuery に蓄積されたデータは、Cloud Datalab で分析しています。三浦さんは、「もともとは、BigQuery を使いたいと思いついたプロジェクトですが、Cloud Functions や Cloud Datalab、データポータルのようなサービスの充実も、本当に便利です」と話します。

「SQL 文も BigQuery で習得したのですが、やりたいことが簡単に、安く、速くできます。R 言語や Python を使って従来の環境で分析するときには、1 つのクエリ処理に 1 晩かかることもあったのですが、クエリの書き方にもよりますが、BigQuery では同じ処理が 5～10 秒程度で終わります。また、サーバレスの構成が可能で、コストメリットも大きいです。」(三浦さん)。

また、アップデートで機能追加できるのも便利だと語ります。「日に日に、機能がよくなっている機械学習は、今後の活用を期待しています。機能追加は Google Cloud に任せ、放送アクセスログ分析により価値のある放送サービスの構築に集中できます。」

「可視化に関しては、最終的にはデータポータルを使いたいと思っています。ただしデータポータルは、現状では何時何分という単位での可視化しかできないので、早い時期にアップデートして、何時何分何秒までの単位で可視化ができるようになることを願っています」(三浦さん)。

今後、Video Intelligence API で画面や音声から知見を収集

今後、HTB では、OSS のダッシュボードツールである Re:dash でダッシュボードを作り、現場の人たちに使ってもらう計画です。現在は、何時何分何秒に何人の人が視聴したかだけが分かるのですが、ユーザー属性が想定できる仕組みを検討しています。「たとえば、朝テレビを見て、その後夜までテレビを見ていなければ、単身のサラリーマンかもしれないとか、昼間もついていたら専業主婦がいるのかも、などが想定できます」(三浦さん) 放送アクセスログ分析により、今後さらに詳細な分析も期待できます。三浦さんは、「まだまだあいまいなデータもありますが、今後も試行錯

誤しながら、より正しく納得性のある情報が出せる段階に近づいていきたいと思っています」と話します。

さらに三浦さんは、「今後は、Video Intelligence API を使い、画面に映っているものや音声などに関する知見を収集したいと思っています。現在、番組が何時からはじまり、CM と本編がこの時間だったというデータは取れるようになっています。画面からさらに有意義な情報を引き出す仕組みを実現し、詳細な分析に繋げていきたいと思っています」と、今後の期待を語っています。

Google Cloud を活用することで、ビジネスの将来に注力できるようになります。インフラストラクチャの管理やサーバーのプロビジョニング、ネットワークの構成などに起因する負担を軽減することができます。つまり、イノベーターもプログラマーも、自分の本来の仕事に集中することができます。

お問い合わせはこちらから
<https://goo.gl/CCZL78>



Google Cloud の詳細については、右記 URL もしくは QR コードからアクセスしていただくか、同ページ「お問い合わせ」よりお問い合わせください。
© Copyright 2020 Google
Google は、Google LLC の商標です。その他すべての社名および製品名は、それぞれ該当する企業の商標である可能性があります。

