

# Data Platform Modernization

## and its importance in Marketing Analytics at large-scale Media Agencies



### Problem Statement

The Customer manages a large portfolio of advertisement initiatives in various markets around the world. As part of the services, The Customer provides campaign performance, budget trending and conversion metrics, providing insights to its clients for maximum ROI from campaigns.

- The campaign performance and conversion metrics reports require data to be sourced from several data sources in the digital advertisement eco system
  - The volume, velocity and verity of data increased as the number of campaigns increased, and so the number of reports needed to get the required visibility
  - The data required wrangling and curation to make sure errors made in Ad Operations are corrected to avoid inaccurate reporting
  - Data processing and cleansing work was manually done by the data analysts who used Excel based reports
  - Data sources were managed in the analyst's desktops and network storage
- The whole process was error-prone and cumbersome to manage, and trending reports were not easy to produce
  - The manual processing of reports lacked:
    - » Ability to produce reports swiftly » Powerful business intelligence capabilities such as visual exploration of data » Easy to manage report authoring and distribution capabilities » Interactive dashboards » Ability to scale as the datasets grew
  - Without vital insights, The Customer and its clients were not able to dynamically plan and allocate media budget at the right time and hundreds of clients and partners asking for real-time reports on campaign performance made it extremely difficult to manage

## Customer Background

The case study was for one of the largest multi-national digital marketing communications company. Its principle services are communication strategy through digital creative execution, media planning and buying and marketing analytics. We will refer the agency as the Customer here after.

## Overview of Engagement with the Customer

- Softcrylic was originally asked by the Customer to create a reporting platform
- Goals of the project was to deliver campaign performance reports
- After initial platform was set up, it was quickly realized that the Customer needed more than just a reporting platform. They needed a way to modernize their data platform in order to support various analytics and reporting needs

## Challenges faced by the Customer in their Media Supply Chain

### Process Issues

- No standard approach to managing planned data and consequently no Single Source of Truth exists for what campaigns are running at any given time
- This means there is no centralized account / campaign list to validate a data repository against for data completeness
- Campaign activation process and Taxonomy Management is done manually by teams all over the world, causing a wealth of taxonomy and data issues that impact reporting downstream

## Reporting Issues

- Data and Reporting platforms are downstream systems and generally can report only on data that is provided to them. Therefore, any cleaning efforts are siloed, and other system would not receive that cleaned data.
  - Data used for reporting by multiple teams (Ex: Billing, Planning, Analytics) from different sources often created inconsistencies
  - Mostly these processes have been manual and different teams have different methods for pulling and cleaning data
- Campaign performance reporting delivered to clients are mostly descriptive in nature.
  - Analytics Team struggled with the fact that they have to spend most of their time putting together reports and fixing data issues, instead of providing insights on the data and optimization opportunities
  - Lack of advanced report delivery methods as most of the reports were done using Excel
- Analytics focused more on canned report delivery and was not able to generate in-depth insights due to lack of tools to explore data

## Solution

After assessing these challenges, Softcrylic moved away from a “report-first” approach and refocused on a “data-first” approach. The goal of this phase of the project was to bridge the gaps in media supply chain to improve quality in data delivery and to modernize the data platform to support advanced analytics needs.

- Turn the reporting platform into a data platform that will supply data for reporting, but also do much more:
  - Serve as a single source of truth to serve the needs of all parties involved in the supply chain (Planning, Activation, Optimization, Analytics, Data Delivery, Exploration and Research)
  - A platform that will help deliver both client reporting and capabilities for internal analytics teams to explore and infer insights on data
  - Improve the Customer capabilities to deliver advanced analytics (Dynamic Campaign Optimization, Data Modeling, Reporting) to clients
- Strengthen the campaign planning, activation, spend reconciliation process by providing quality data to downstream systems, so different teams are not using different data for their objectives

# Tools and Processes implemented to support a Modern Data Platform

From our experience at the Customer, we have learned that maintain a high-quality data platform, support tools and processes must be put into place.

## Data Quality Automation Tool

- Data Availability
- Data Consistency
- Data Quality
  - Data correctness checks: Ensure vendors are delivering accurate data
  - Data Corrections: Help fix data downstream in supply chain
- For all data sources, a secondary of data pipeline is set up to sanity check the primary pipeline.
  - Ex. Facebook API data is the primary pipeline and a high-level UI report is the secondary pipeline.
  - Each day data will be brought in from both to validate against each other

## Taxonomy Management

- Tool Definition Tool: Fix process upstream
- Validation Tools:
  - Validate taxonomy downstream based on rules. It will help an agency create and properly traffic their advertiser's taxonomies
  - Will help an agency validate their taxonomies in multiple phases of the media supply chain and proactively find issues
- Taxonomy Exception Mapping Tool: Will help an agency clean up taxonomy downstream and create accurate reports for advertisers

## QA Process

- Go into detail about support team set up and how data warehouse is maintained
  - Data Availability
  - Data Mart Sync
  - Data Quality

## Data Standardization

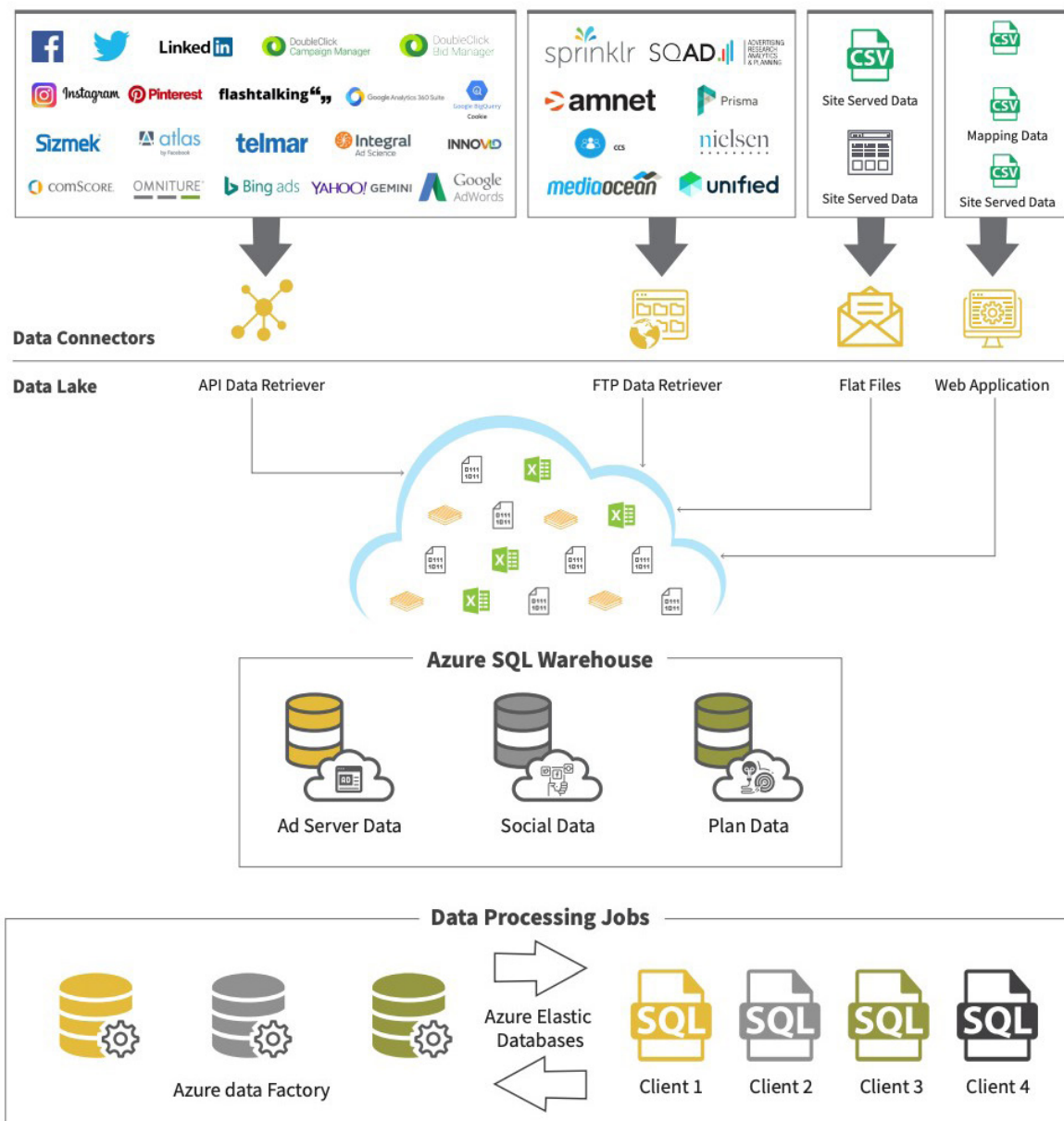
- Schema: Standardize ad-hoc data delivered publishers
- Ingestion Tool: Ingest inconsistent ad-hoc data provided by publishers in a standardized schema

Result of all of these tools and process is creating a single source of truth that will feed all parts of the media supply chain, not just the end reporting needs.

# Data Platform

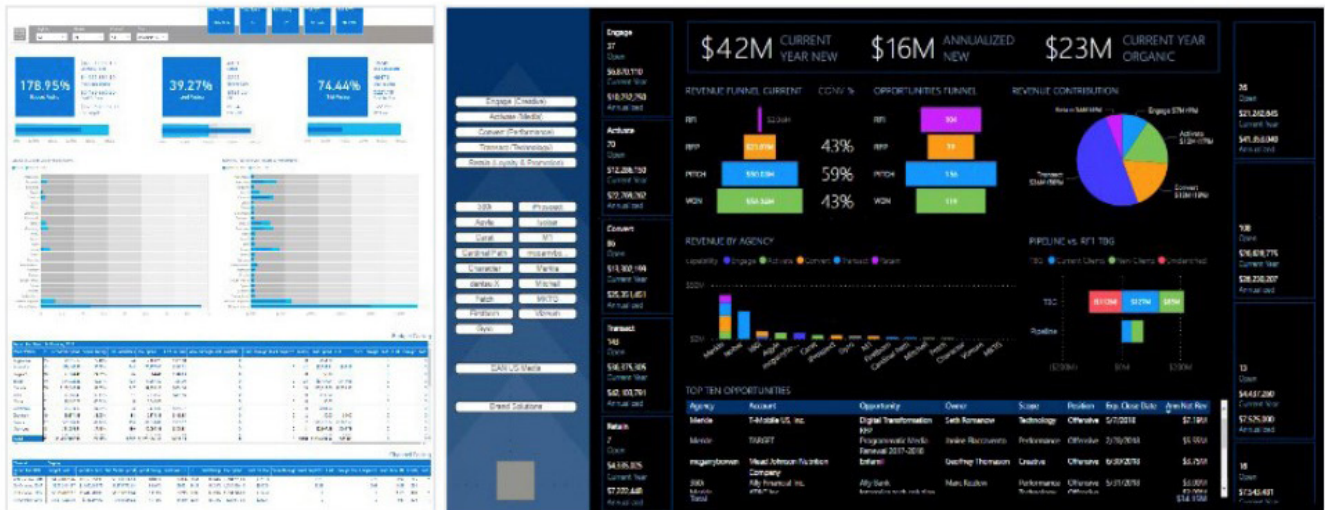
After careful analysis of the requirements and goals, Softcrylic came up with a proposal to create 'single-source of truth' (data warehouse) for all their reporting needs by automating data collection and curation using **Microsoft Azure Cloud** platform and technologies. Softcrylic designed and developed a number of Data Connectors for retrieving various data sources using C# and node.js and deployed them using Azure Functions which are instantly scalable. These connectors pushed data into a **Data-Lake** (BLOB storage). Data Connectors were designed in such a way that new data connector development would be plug-and-play.

Azure SQL Warehouse was chosen to be the cloud-based datastore that can be scaled on-demand for both storage and performance needs. The data stored in Azure Data-Lake was processed using Azure Data Factory (ADF) which provided rich features, simple to use configuration and faster processing. It helped us to easily handle the enormous need for data curation.



# Visualization

Microsoft PowerBI was selected to deliver visualization. Reporting datasets were developed in modular data marts on Azure SQL instances, enabling re-useable and scalable reporting across individual requirements. Publishing through Microsoft's PowerBI service allowed for high-performance scheduled extracts and granular security for internal customers as well as external clients.



# Customer ROI

- The platform offers a single-source of truth for all data processing and reporting needs
- Secure and scalable data platform
- Data available via SQL, as well as both Excel and PowerBI through the same logical architecture
- Powerful and interactive visualization features using PowerBI
- Automation of repetitive analytics tasks frees internal resources to focus on customer benefit
- Timely and accurate delivery of up-to-date metrics transforms the Customer's internal culture and client relationships.
- The platform becomes integral to pitches to new business and servicing existing clients
- The platform was a key factor in retaining one of the Customer's largest accounts totaling over \$1B in annual advertising budget

## About Softcrylic, LLC.

Softcrylic helps organizations navigate and execute the path of Digital Transformation through IT solutions and services in a variety of technical disciplines including Software Development, Test Engineering, Data and Analytics. Since 2000, Softcrylic has worked with both start-ups and Fortune 500 organizations to help make their company goals a reality.

## Get in Touch with us

Email: [info@softcrylic.com](mailto:info@softcrylic.com)

Phone: +1 609 452 7638

For more information,  
visit [www.softcrylic.com](http://www.softcrylic.com)



---

© 2019 Softcrylic. All rights reserved.

Softcrylic is a registered trademark of Softcrylic, LLC.

