

ARTEFACT

DATA & AI MARKETING



ARTEFACT

AI IS ABOUT PEOPLE

WE ACCELERATE DATA AND AI ADOPTION
TO POSITIVELY IMPACT
PEOPLE AND ORGANIZATIONS.



21
COUNTRIES

1500
EMPLOYEES

+1000
CLIENTS

Artefact is a global leader in consulting services, specialized in data & AI transformation and data-driven digital marketing, from strategy to the deployment of AI solutions. We are offering a unique combination of innovation (Art) and cutting-edge AI technologies (Fact).



DATA READINESS | AI ACCELERATION | DATA & DIGITAL MARKETING | TECHNOLOGIES

Data & AI Marketing

4 Introduction by Vincent Luciani, co-founder & CEO of Artefact

DATA-DRIVEN MARKETING

8 Data-driven marketing: the rise of the customer data platform

10 Unleash the power of first-party data with a Customer Data Platform

13 **ORANGE BANK – Enhances its digital marketing with rapid deployment of a full funnel, cross-device activation strategy**

15 Interview of Florian Thiébaud “CDPs, platforms used for data centralisation and audience segmentation, become the norm”

17 **RED BY SFR – The value of proprietary data at Red by SFR**

19 Unlock revenue growth with data-driven digital marketing maturity

21 **RECKITT – Optimising media spend and increasing sales at Reckitt with the Artefact Audience Engine**

25 Building your own audience engine: turning the Google ecosystem into a cookieless Customer Data Platform (CDP)

29 **UBISOFT – How Artefact helped Ubisoft to build a powerful media campaign using data for its key video game launch of the year**

33 MROI | A new data-driven approach to measuring on and offline campaigns

35 **SANOFI CHC – Precision Marketing – Industrialising the deployment of data-driven campaigns**

39 CDP and the Business of Data in Sports and Entertainment

42 How AI can help brands discover consumer microtrends

DIGITAL COMMERCE

45 Becoming data-driven (and reaping immediate benefits) has never been easier in the retail sector

47 **CARREFOUR GROUP – How Data & AI can accelerate sustainable business transformation**

50 **CONFORAMA – AI-enabled personalization boosts Conforama CRM campaign revenues**

54 Retail Media: An indispensable asset for brands

56 **UNILEVER – How does Artefact support UNILEVER on Retail Media use cases to increase its sales?**

60 Data monetization opportunities for retailers: Retail Media within the CPG/Retailer data ecosystem

62 **CARREFOUR LINKS – With Artefact are helping brands to improve media targeting thanks to Carrefour's data**

65 **MATTEL x relevanC Advertising – How Artefact helped boost Mattel's online sales on Cdiscount (Relevant Advertising) retail media platform?**

67 **CARREFOUR GROUP – How to reduce food waste in the Bakery-Pastry department?**

70 Demand forecasting: Using machine learning to predict retail sales

72 Sales forecasting in retail: what we learned from the M5 competition

The outlook for data and AI transformation, today and tomorrow.

An interview with Vincent Luciani

The generative AI technology revolution has been a paradigm shift for all industries and sectors. Artefact sees AI as an incredible opportunity that, if used properly and ethically, will lead to economic, social, and democratic progress.



“At Artefact, we hold an optimistic vision, viewing AI as an incredible chance that, if used properly and ethically, will lead to economic, social, and democratic progress.”

How is generative AI profoundly transforming society and businesses?

We are at the beginning of a new era. The generative AI revolution is reshaping societal and economic landscapes. After an experimentation phase, generative AI will continue to change the game for the global community. It's a technology with the potential to improve the world in many ways, as long as solid checks and balances are in place to ensure its responsible and beneficial development.

- **Economically**, it offers undeniable productivity gains that will spur innovation and new business growth.
- **Socially**, generative AI will streamline administrative tasks, freeing up more valuable and creative time, which could lead to innovative job opportunities and the development of new skills.
- **Democratically**, the accessibility of GenAI to all will provide deep knowledge and solutions to address specific societal and educational inequalities and advance the cause of social justice.

How is Artefact leading the generative AI transformation for enterprises?

Since the availability of the first LLM models (Large Language Models), even before the official public launch of ChatGPT in November 2022, we at Artefact have been one of the key global pioneers using this powerful technology, designing and deploying many generative AI use cases with our clients throughout 2023.

As certified experts with major Clouds and open source GenAI, we've already acquired strong expertise and developed a solid ecosystem. In this context, we recently announced our official strategic collaboration with Mistral AI, the most powerful LLM platform for a European OpenAI.

Despite achieving notable reductions in development time and enhanced employee adoption, scalability of GenAI projects remains a challenge, emphasizing **the need for ethical and secure environments grounded in robust data foundations**.

For more than 10 years, Artefact has prioritized **the crucial role of data in AI success for enterprises**. Initiating data acceleration programs, we focus on elevated data quality, governance, and interconnected platforms, adhering to ethical and responsible guidelines.

Anticipating substantial growth thanks to these new LLM technologies, companies are urged to **embrace AI for a competitive edge**. This transformative year will necessitate new organizational models and widespread AI deployment across business value chains, with Artefact accompanying its clients from strategy to full operations.

However, the success of technology shifts depends on **fostering trust and enthusiasm among all employees**, requiring consultation and support from top to bottom, an area where hackathons and training can be instrumental.

What is Artefact's mission? What initiatives have you taken to anchor your AI adoption strategy to accelerate business growth and efficiency?

Artefact's primary mission is to accelerate the adoption of data and AI to positively impact people and organizations.

To translate this purpose, our motto is «AI is about people».

The companies that will endure are those that successfully foster a **data culture** with access to knowledge and data for all.

We've undertaken several initiatives in this area that are highly strategic to Artefact's positioning as a **major player in data democratization**, in order to fully realize its potential for positive transformation.

- The development of the **Artefact School of Data**, a key pillar in our strategy of providing clients with training adapted to the constantly evolving skills of the data industry. We are also developing «à la carte» e-learning platforms for clients to quickly share knowledge of data and AI with all of their employees. We've expanded our **Artefact School of Data** from France to Dubai and New York City, and soon to other cities to educate organizations about data and AI, while creating new job opportunities in this domain.
- We've also launched many **generative AI hackathons** at major companies to empower and inspire their employees with these new innovative technologies.
- For over five years, we've organized **large-scale conferences** such as **AI for Finance & Industry** and **AI for Health**. We also successfully launched the first edition of **AI for Luxury in NYC** and **AI for Life in Geneva**, bringing together top-level AI ecosystem stakeholders, including

major corporations, startups, and universities, to disseminate knowledge about data and AI.

Business data maturity has advanced rapidly over the past decade. How has Artefact evolved as a global leader in data and AI consulting services?

Companies have implemented **data governance policies**, which are a prerequisite for any transformation, but there are still sectors that lag far behind in terms of their data processing, with a real potential for efficiency.

We started to transform marketing departments by making them more profitable and relevant in their multi-channel media investments with pioneering targeting, measurement and personalisation solutions. For the past few years, we have also been deploying acceleration programs in all business areas (Sales, Supply Chain, Operations, Call Centers, HR and Finance, etc.).

We create value wherever there is data, and work with our clients to improve their processes and create customized business AI applications.

With significant growth expected thanks to generative AI technologies, we advise all organizations to embrace AI to gain a competitive advantage.

Can you give us concrete examples that show how Artefact designs AI solutions that improve business competitiveness?

Data is the key to understanding customers, developing better products and services, and streamlining internal operations to reduce costs and waste. Artefact supports over 1,000 clients worldwide, including 300 international brands in sectors from consumer goods, retail & e-commerce, and healthcare, to bancassurance, telecoms, industry, energy - and more.

For example, we've been working with the **Orange Telecommunications group** for over six years, and among the many use cases for leveraging the company's automation and AI potential, we deployed a solution with their teams to optimize their technicians' interventions on the fiber network. The solution is based on visual recognition technology that helps operators improve the quality of their installations or repairs. This application, available on a tablet, is currently used by more than 10,000 Orange technicians throughout the country - a resounding success!



ARTEFACT

This case perfectly illustrates Artefact's firm belief that to achieve true data maturity, companies have no choice but to make data accessible to everyone: not only to experts, but also to operational staff in the field. This will lead to new forms of augmented work, where applications and their interfaces put intelligent information in everyone's hands to work more efficiently and with more autonomy.

Artefact also helped the **Carrefour Group** in reducing the carbon impact of its e-commerce branch with a solution that can be implemented by the company and consumers. Carrefour's aim is to become the world leader in food system transformation for all by committing to four major objectives, including achieving carbon neutrality by 2030 for its e-commerce activities. The challenge for Artefact was to enable Carrefour to reliably measure all greenhouse gas emissions from data storage, transport and logistics activities, from first click to final delivery,

Our solution measured greenhouse gas emissions generated by e-commerce orders, then collected activity data to convert it into carbon emissions. All Carrefour business teams helped obtain the data – which is why the operation was a success, as it allowed all stakeholders to become ambassadors for **the group's "carbon neutrality 2030" objective**.

How is Artefact able to always be at the forefront of AI through core research and advanced technology?

At Artefact, we've implemented major projects to ensure that we always leverage **the best of data science and AI technologies** for our clients:

- **The launch of the Artefact Research Center**, which fosters a robust data and AI R&D ecosystem by connecting PhD talent at Artefact with esteemed professors from top universities (Polytechnique, Sorbonne University, and CentraleSupélec, University of

Paris-Saclay) and leading enterprises including Orange, Société Générale and Decathlon, with other companies joining us soon. Through the developments and publications of the Artefact Research Center, we aspire to shape a future where AI is not only a powerful tool but is also tailored to the needs of businesses with ethics and responsibility, thereby facilitating its adoption.

- **The creation of the SKAFF technology platform**, an open source developer portal that includes a central software components catalog supporting TechDocs and a scaffolder for automating engineering processes. This platform enhances efficiency by swiftly delivering high-quality outcomes through the consolidation of technical assets, convictions, and tutorials focused on our core technologies.

After a decade of exponential growth, what is Artefact's ambition for the coming years?

First of all, our gratitude goes to our clients for entrusting us, a cornerstone of our success.

I believe that our success also stems from our unique ability to transform data and AI into value for companies. We offer our 1000+ clients **a unique combination of innovation (Art) and data science (Fact)**.

By creating multidisciplinary teams and breaking down silos between business and technology departments, we generate real, immediate impact

for clients. Artefact has become one of the first and few consolidated pure data & AI players in the market, with the most comprehensive set of data-driven services and AI applications.

We offer data acceleration programs, industry specific AI solutions, and data-driven marketing services. Our engineers build tech agnostic solutions, combining custom code with open source and proprietary software, backed by strong partnerships with leading cloud providers, to create exactly what you need for your data and AI transformation.

Today, Artefact is present in 20 countries across Europe, Asia, the Americas (North & South), the Middle East and Africa, with 23 offices and 1,500 employees. And we have robust plans for geographical expansion as well as an ambitious M&A policy that will continue.

We're also continuously hiring new consulting Partners and Directors, experts in their respective fields, orchestrating collaboration across Artefact's regions. They provide dedicated support and industry-specific services. While strengthening our positions in CPG, Retail, and Luxury, we've also intensified our development in Financial Services, Healthcare & Pharmaceuticals, and Manufacturing, reinforcing human resources.

We're excited about the promising future that AI holds for individuals and organizations. The excellence of AI technology will be realized through the collective capabilities of human talent.



Data-Driven Marketing

- 8 | Data-driven marketing: the rise of the customer data platform
- 10 | Unleash the power of first-party data with a Customer Data Platform
- 13 | **ORANGE BANK – Enhances its digital marketing with rapid deployment of a full funnel, cross-device activation strategy**
- 15 | “CDPs, platforms used for data centralisation and audience segmentation, become the norm”
- 17 | **RED BY SFR – The value of proprietary data at Red by SFR**
- 19 | Unlock revenue growth with data-driven digital marketing maturity
- 21 | **RECKITT – Optimising media spend and increasing sales at Reckitt with the Artefact Audience Engine**
- 25 | Building your own audience engine: turning the Google ecosystem into a cookieless Customer Data Platform (CDP)
- 29 | **UBISOFT – How Artefact helped Ubisoft to build a powerful media campaign using data for its key video game launch of the year**
- 33 | MROI | A new data-driven approach to measuring on and offline campaigns
- 35 | **SANOFI CHC – Precision Marketing – Industrialising the deployment of data-driven campaigns**
- 39 | CDP and the Business of Data in Sports and Entertainment
- 42 | How AI can help brands discover consumer microtrends

Data-driven marketing: the rise of the **Customer Data Platform**



Florian Thiebaut
Managing Partner
Data Marketing Lead
ARTEFACT

A game-changing technical and legal environment

Following Safari's lead in 2016, the world's three main browsers eliminated (or will eliminate) the use of third-party cookies. On the mobile/tablet devices side, Apple's iOS 14 now requires explicit consent for any mobile ID collection.

As for regulation, GDPR laws in Europe have given consumers more control over their personal data, requiring them to give explicit consent for the use of cookies. This regulation represents a major shift in the world of data-driven marketing, as it has reduced the number of cookies placed on European devices by 30%.

This global trend restricting the use of IDs and advertising cookies sharply impacts the targeting capabilities of advertisers, who are often dependent on third party data. The vast majority of them use or have used retargeting and old generation DMPs that rely heavily on segments fed by third party data.

Everything seems to justify the current explosion of the Customer Data Platform (CDP) market. CDPs' main advantage over older generation Data Management Platforms (DMPs) is that they easily integrate identifiable first-party data (email, phone number) and aren't dependent on using third-party cookies or browsing data to refine customer and prospect knowledge.

CDPs are a true asset in a world that is becoming increasingly cookie- and ad ID-free. At a time when the pandemic is forcing brands to digitise at breakneck speed, and when the transformation of the technical and regulatory environment surrounding advertising trackers is forcing data marketers to revise their approaches, CDPs are here to optimise the customer experience.

Along with targeting, measurement must also be transformed. With more stringent consent collection requirements, it's more difficult to collect the consumer IDs needed to track impressions, clicks or views, and reconstruct complete customer journeys.

Four pillars for a sustainable data strategy

To maintain the same performance and differentiate themselves from the competition, advertisers must design a sustainable data strategy and exploit their customer and prospect data to its full potential.

This requires focus on four actions:

- **The CDP:** The first step is to establish a CDP environment based on a suite of tools that is both compliant and sustainable. This will enable data to be collected, stored, processed, visualised and activated, whatever the source. From this foundation, the focus must be on first-party data.
- **Data governance:** Brands need to rethink data governance and processes to enable secure and compliant end-to-end data collection.
- **Audience segmentation:** This data, centralised for a unified view of the

consumer, can then be used to create new audience segments and define new metrics for measuring campaign results.

- **Second-party partnerships:** In addition, it's becoming increasingly strategic to form so-called "second party" partnerships with other partner companies to exploit first-party data and create win-win situations.

This data completes a database that is incomplete at certain points in the consumer journey. Examples might be an agreement between an FMCG brand and a retailer, a mobile phone manufacturer with a telco or a hotel chain with an airline.

Three types of data to activate via a suite of tools

First- and second-party data are key to meeting the challenges of the post-cookie world. But what are they and what tools can be used to manage them?

- **PII or Personally Identifiable Information** is essentially CRM (customer relationship management) data. It can precisely identify an individual and is often an email address or a phone number for example. Once anonymised, it can be used via the APIs of media partners (e.g., Google Customer Match, Facebook Custom Audience/conversion API, Amazon, WeChat, etc.) to build audience segments, perform audience extensions, and reconstruct paths to measure the influence of digital campaigns on offline sales, etc.
- **Non-PII data** can be browsing data that cannot lead directly to the identification of an individual. It can be used to build more granular segments via analytics and audience creation solutions for measuring precision marketing actions without relying on third party data

- **Data that is purely media-related**, such as campaign impressions, video views and click rates, is more voluminous and less granular than the other two types of data. It is more difficult to use but there is a robust market of tools capable of treating it in a secure and compliant manner, such as Google Ads Data Hub, Facebook Advanced Analytics and Amazon Marketing Cloud.

These different data flows are injected into an ecosystem of interconnected tools, which are useful for a range of tasks – from data collection to performance measurement of the actions carried out – and can be activated on all channels, whether media, direct marketing or site personalisation.

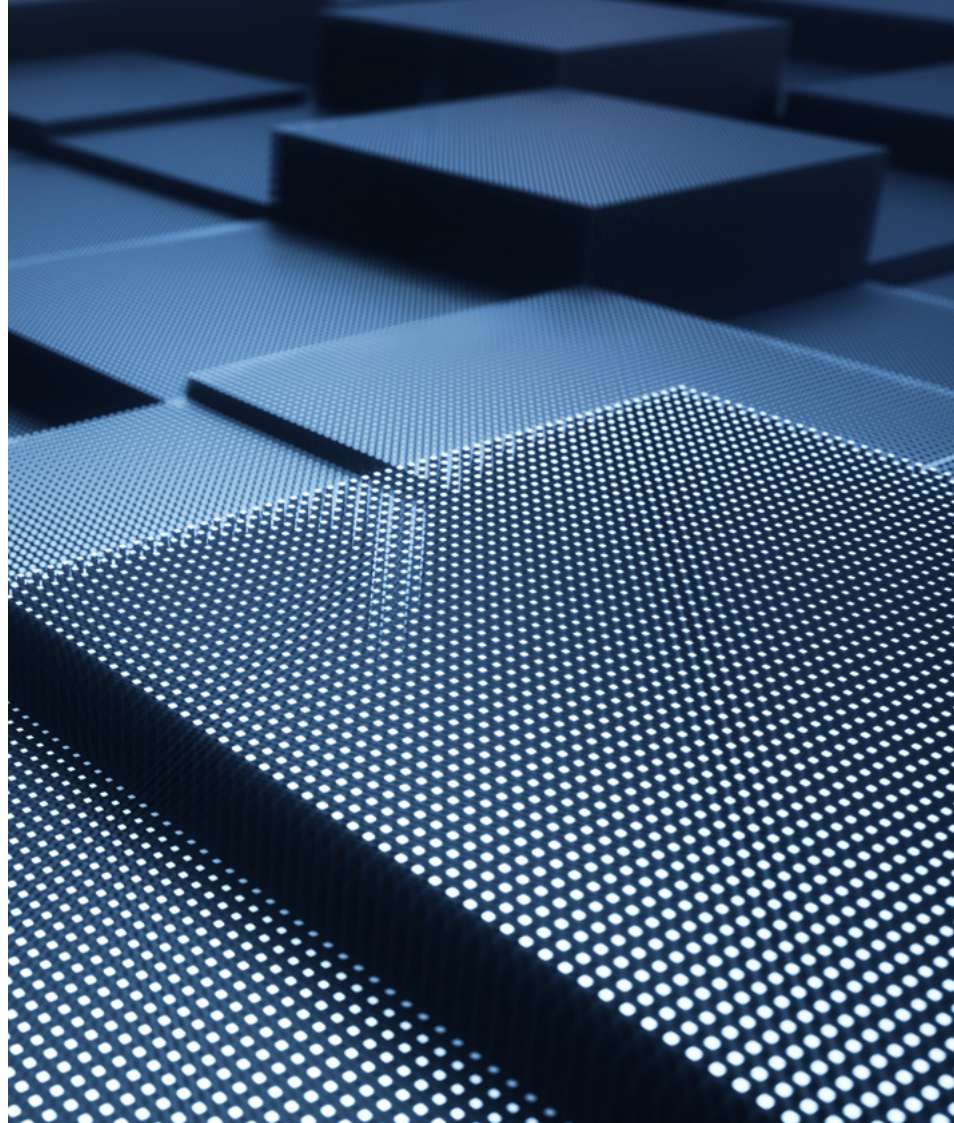
This entire ecosystem, the result of all the connections built between the different tools already used by the

company (also known as "full-stack" solutions), is what is called the CDP.

When it comes to the adoption of this way of working, the numbers don't lie. Fundraising for CDP providers is soaring, the tech giants are all positioned in the sector, and the number of users is exploding.

In fact, according to the Customer Data Institute, the market increased 30% from \$1 billion in 2019 to \$1.3 billion last year. Estimates see this figure reaching \$1.55 billion in 2021 as conditions are even more favourable for the adoption of CDPs.

As the data-driven world continues to evolve at a rapid pace, there seems little doubt in the business value of the CDP. Now is the time for organisations to consider deploying this future-facing technology.



Unleash the power of first-party data with a **Customer Data Platform**



Daniel Ferreira
Manager Digital & Data Consulting
 ARTEFACT

The demise of third-party cookies and evolving privacy regulations are impacting the way companies can gather and use consumer data. The digital media environment is becoming more technical, with crucial customer information siloed in walled gardens, making granular targeting and measurement difficult to achieve. Moreover, today's customers understand the value of their personal information and are choosy about what they consent to share.

Yet, with the expansion of digital channels, the number of customer touchpoints has multiplied, vastly increasing the number of data points marketers can collect and consume – and requiring a multiplication of tools (Meta, Google, Amazon...). Clearly, it's time for a new digital marketing and measurement approach.

But if marketers can no longer rely on third-party cookies for cross-site retargeting, audience buying from external marketing platforms, and digital marketing attribution, how can they keep their audiences happy? Because despite these issues, customers want value-added experiences that are more personalized than ever, and it's necessary to deliver them – or lose customer loyalty.

Collect a maximum of first- and second-party data

To better understand customer behavior, preferences, and the ways they engage with messages across digital channels, follow Artefact's three-step plan:

- 1. Gather** as much first-party data as possible. The most advanced approach to collecting customer data is through developing product recommendation models, user scoring, propensity score models or CLV prediction models.
- 2. Explore** ways to leverage second-party data from large local publishers or marketplaces and combine it with first-party business-specific data.
- 3. Build** partnerships with strong media and tech partners such as Artefact's preferred media partners Google, Meta, and Amazon, as well as LiveRamp or InfoSum.

Remember: in exchange for sharing their personal information, today's consumers expect to receive something of value from the brand. In an article about a customer experience sentiment study on IT news-sharing website Spiceworks, an amazing 91% of consumers responded that they were likely to purchase from brands that heard their needs, and 76% of consumers said they would share their data in order to receive more personalization.

When offering incentives, ensure they have real benefit for customers. Examples of strong data value exchanges include:

- Retailers that offer customers a preview of new product lines;
- Automobile manufacturers that let prospective buyers vote on potential new features;
- Travel businesses that provide free upgrades to long-term high-value customers.

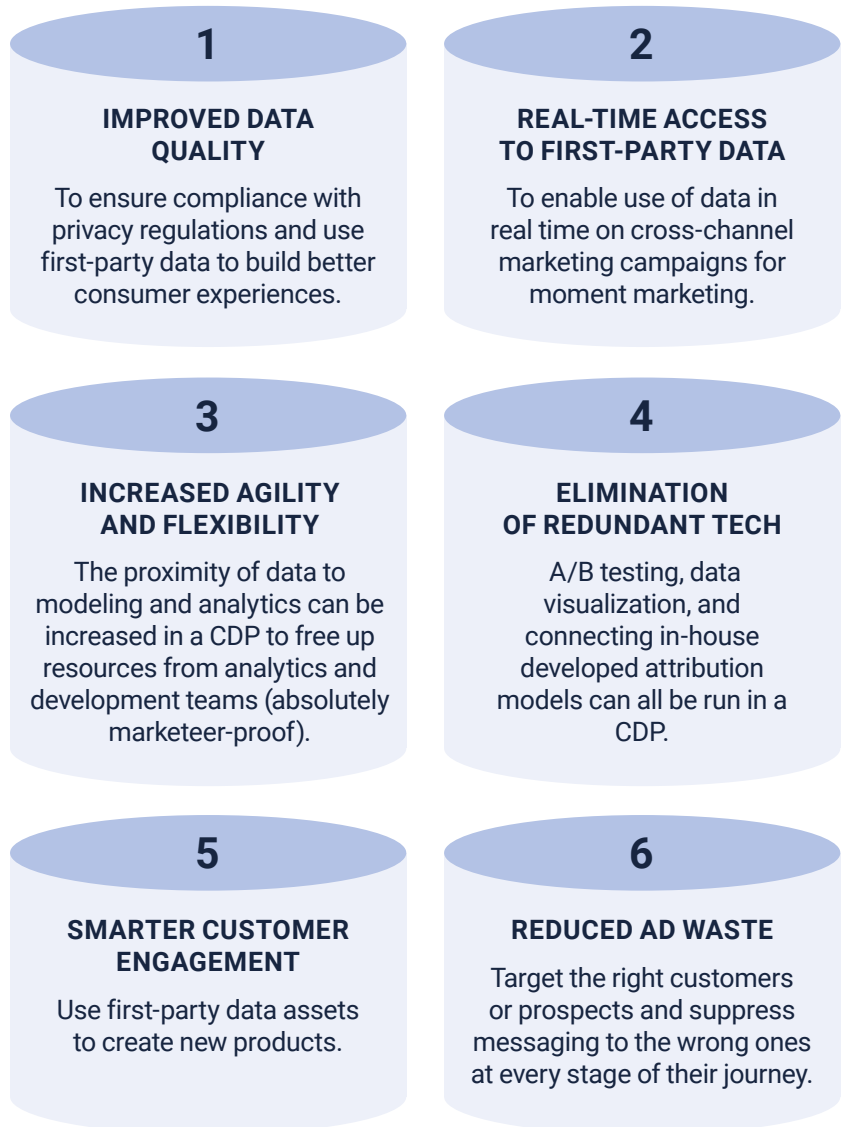
But once all this precious data has been amassed, how can it be used to benefit businesses?

Use a CDP to draw valuable insights from data

Walled garden platforms dictate audience buying, segmentation, attribution, and campaign analysis and optimization, leaving marketers with little control, ownership, transparency or efficacy.

A CDP breaks down these silos by offering marketers a unified view of their customers across different marketing platforms and channels, while assisting them with personalized marketing and targeted advertising.

The six major benefits of a CDP are:



DATA INGESTION

Data is integrated from different sources (e.g., brands, business units), potentially along with second-party data (publisher or marketplace) or third-party data (weather data or other business-specific data).

DATA TRANSFORMATION

A unique ID per customer (unified customer profile) is created to rapidly identify relevant segmentation, run customer analysis, and push that data into the data activation phase.

DATA ACTIVATION

Marketers use the transformed data to run smarter, more data-driven campaigns by sending audience data to various marketing platforms, such as Google ads, Dv360, Meta or others. In many CDPs, data for data visualization purposes can also be sent to a Tableau or Power BI dashboard, for example.

“Make and buy” to obtain the ideal CDP for each need

To assist clients as they move to CDP capability, Artefact assesses each situation and recommends the relevant mix of make and buy. For example, an organization considering building an audience engine in an environment such as Google Cloud Platform might also benefit from the countless out-of-the-box data connectors a CDP platform such as Treasure Data, Tealium or BlueConic may already have.

From an entirely custom-made CDP to a fully-integrated out-of-the-box CDP, or anything in between, the optimal solution is to find the sweet spot between make and buy that fits each company’s requirements.

Once a solution has been selected and integrated, it’s crucial to measure the success of CDP implementation. CDP ROI can be measured in two ways:

1. Through marketing-based KPIs, which are short-term, easy to measure, and show direct results:

- Increase in newsletter subscriptions.
- Incremental revenue from personalized creatives.
- Increased time on site.

2. Through operational KPIs, which are long-term, harder to measure, but demonstrate more durable results:

- Reducing time spent on manual tasks.
- Less reliance on IT & Development.
- Workflow and process optimization.

Conclusion: CDPs solve key business challenges

In a recent interview, Alexandra Mangear, Partner at Artefact, stated: “We have found that our clients face three main problems in data marketing:

- Measuring whether the impact of their data marketing strategy is aligned with their business objectives, through the building of Marketing ROI and CLTV (Customer LifeTime Value) measurement methodology and dashboards.

- Guiding their marketing team and external partners in adopting new technologies and data-driven best practices, which requires a change of habits/process and

- Adapting their paid media activation, mainly based on third-party cookie activation, to the future cookieless world.”

Artefact has strong expertise in CDP topics and first-party data strategy across multiple industries, as well as partnerships with trusted partners such as Google, Treasure Data, BlueConic, Tealium and Twilio Segment.





CASE STUDY

ORANGE BANK

Enhances its digital marketing with rapid deployment of a full funnel, cross-device activation strategy

CHALLENGES

Launched in November 2017, Orange Bank offers banking services designed natively around mobile uses so customers can autonomously perform all operations from their app.

The offer includes not only all basic banking services, but many innovative features and Premium offers as well, with a dedicated pack aimed at families, and a fintech designed for professionals. Orange Bank has a total of 1.2 million customers in Europe.

Orange Bank combines digital presence via mobile and the web with human contact through Orange boutiques. This model is an important differentiating factor compared to other neobanks. More than 2,000 qualified sales staff accompany customers in more than 300 Orange shops. Confirmation of account opening is subject to usage checks, which are carried out outside the application.

For this mobile bank, the marketing challenge was to enable subscription process management from start to finish on a mixed web and application path.

SOLUTION

Implementation

Working with Artefact, a GMP & GCP certified partner (Google Marketing Platform and Google Cloud Platform), Orange Bank achieved their steering objective in only one year, thanks to rapid deployment of the following projects:

Migration to Google Marketing Platform and Google Cloud Platform

- Redesign and enrichment of the data collected:
 - Redesign of the datalayer,
 - Definition of tracking plans, especially for compatibility with new Google Analytics,
 - Addition of variables and securing of reconciliation IDs.
- Consolidation of ad-centric and site-centric digital data,
- Consolidation of web and CRM data,
- Deployment of new Google Analytics to handle cross-device use-cases,
- Connection of media buying platforms to Google Analytics.

This deployment allowed Orange Bank to:

- Activate its app audiences in the web universe, thanks to Google Analytics 4 app and web audiences deduplication,
- Enable targeting of “similar users” with exclusion of existing customers,
- Improve bidding strategies,
- Revise steering KPIs to focus on high-value products,
- Refine attribution analyses,
- Profile prospects,
- Internalise SEA, the most profitable lever,
- And activate new marketing levers: affiliation, partnerships, and programmatic.

RESULTS

New Google Analytics and Google Analytics 4 properties

The new version of Google Analytics brings major innovations to marketers and analysts. It offers improved cross-device tracking capabilities, duplicating audiences between app and web environments, and has new features to automate certain activations.

New Google Analytics also includes machine learning capabilities to create audience segments based on user engagement, purchase intent, etc. — segments that can be targeted cross-device.

Attribution is also cross-device and cross-environment, allowing deduplicated reporting regardless of user device.

“CDPs, platforms used for data centralisation and audience segmentation, become the norm”



Florian Thiebaut
 Managing Partner
 Data Marketing Lead
 ARTEFACT

Customer Data Platforms have earned their stripes just several months after the end of third-party cookie collection in browsers.

DMP, CDP, CMP (Consent Management Platform)... Data is a sector that is particularly rich in acronyms, some more trendy than others. Since 2018, Data Management Platforms (DMPs), platforms that centralise data collected from cookies placed on websites, have been made obsolete by Customer Data Platforms (CDPs), DMPs... but with the focus on first party data, owned by the brand. In 2022, CDPs will definitively replace the on-site customisation tool. In a “cookieless” context, CDPs allow advertisers to reconcile online and offline user data within a single platform and then activate them via CRM channels – such as email or SMS – and online media, e.g. graphic and video ads.

This switch from one acronym to another, can be explained by three major reasons, including regulatory and technical developments (RGPD and ePrivacy most notably), shares Florian Thiébaut, Partner and Data Marketing Practice Lead at Artefact.

The result, mechanically, is a reduction in exploitable data... which will further amplify *“the scheduled end of third party cookie collection on all browsers by mid-2024 and users paying more attention to their personal data,”*

And a decrease in data volumes means a decrease in incremental revenue. Unthinkable for brands.

To continue to create precise and targeted audiences, based on their browsing habits, and to measure their marketing actions, advertisers must therefore adapt quickly. How about combining yesterday’s DMP with the CDP of today? Fabien Deriaz, Head of Data at “AGENCE 79”, says not: *“You have to think about DMP as a relevant working base, a first step, before moving on to the more recent tool, which is the CDP. Nevertheless, this transitional double run can be used for A/B testing in order to compare value creation in an environment with or without third-party cookies,”* the data specialist recommends.

Betting on first-party...

Instead, brands should deploy an ecosystem of CDP tools provided by several providers in order to leverage more of their proprietary CRM data, whether they are linked to Personally Identifiable Information (PII) such as an email, telephone number, or date of birth, or non-PII data, such as cookies or mobile IDs.

This is the strategy chosen by CMI, the media group comprising of around 15 paper magazines and their digital editions (notably *Elle*, *Marianne*, *Version Femina* and *Télé 7 jours*). In order to collect and centralise data, then activate audience segments and measure the results, the publisher has chosen an ecosystem of tools – CRM, big data, media, and CDP. CMI began this process six years ago and this was accelerated at the beginning of 2022 with a partnership with the data marketing platform 1plusX (TripleLift), with whom the group hopes to reconcile identities in a cookieless

ARTEFACT

context. As a media company, CMI needs to know its audience in order to increase its reader base. However, capturing their interests and buying intentions also gives CMI the means to provide advertisers with information about their prospects and better sell advertising space on its platforms. *"In a world where privacy is paramount, we have to, as a media company, continue to segment our audiences, which are often irregular and anonymous, and to offer advertisers the ability to target their future customers as accurately as possible, whether through anonymous or consent-less targeting methods, or by "next-gen" contextual targeting, as offered by 1plusX,"* explains Priya de Saint Olive, Chief Data & Digital Operations Officer at CMI. Thanks to 1plusX, CMI no longer offers advertisers targeting solely based on cookies, but on a mix of logged-in user identities, cookies, and contextual targeting. *"One option alone isn't enough anymore,"* comments Bastien Faletto, Senior Director of Sales EMEA at 1plusX, which is working to *"leverage publishers' and brands' first-party data assets,"* he adds, via CRM data modelling and real-time algorithmic analysis of the content and semantics.


The first use cases were deployed in just 6 weeks, with "satisfactory" results: +50–100% additional reach. *"It is possible to double the coverage of a campaign, with, at minimum, the same performance,"* adds Bastien Faletto. Priya de Saint Olive confirms: *"We can ensure better campaign delivery. Thanks to this combination of targeting strategies, the volume restraints – linked to the ability to reach a minimum number of people – have been removed."*

... and second-party data

"With a well-implemented CDP, marketers have access to nice user interfaces which allow them to carry out complicated data cross-referencing and build sophisticated

audiences," agrees Florian Thiebaut, who believes that *"nothing that is done with a CDP could be done without it..."* on the crucial condition that you master coding. Not everyone has this skill. There are two available options, then: build your own CDP internally, which *"requires financial and human resources and data maturity,"* Fabien Deriaz outlines, or make use of an existing solution, requiring, admittedly, *"sharing your data with a third party,"* but with the advantage of *"faster deployment"*. In this last case, brands have the choice between the promise of a *"full-stack" ecosystem,* from the big American players – Oracle, Adobe and Salesforce, from *"pure CDP players, which are technologically more mature,"* shares Florian Thiebaut, *"such as Treasure Data, Segment, Tealium or LiveRamp,"* or again, from *"smaller DMP players, agile enough to evolve their services, following the example of Weborama, Ysance, or Relay42."*

"Building CDP functionality with major players such as Google, Amazon or Microsoft is also feasible," he advises. *"They don't have CDP tools, but an*



"Due to the changes to regulations in Europe and around the world, in order to use personal data we must now have the explicit consent of the user, to whom we must describe the reasons for collecting data,"

overlap of functionalities, like the tool Looker from Google." Be aware though, when you call on these tech giants: "The CNIL [French National Commission on Informatics and Liberty] has singled out Google Analytics, but other American solutions that handle a lot of personal data are just as likely to be in it's crosshairs. This could lead some brands to build their own in-house CDP," warns 79's Head of Data.

With the ongoing phasing out of third-party cookies, the idea could be to find other sources of data. Already, second-party partnerships between businesses are multiplying, under the name of *"retail media,"* notably. *"To prepare for the end of third-party cookies, we are currently working on peer-to-peer dynamics with advertisers to expand the prospect pool,"* confides Priya de Saint Olive. *"If there are no technical impediments to doing so, we must resolve legal and business issues, such as how to share data – is it for analytics or for growing an audience?"* And for this, there's another technology that is already trending: the clean room.



CASE STUDY

RED BY SFR

The value of proprietary data at Red by SFR

OBJECTIVES

To implement their data marketing strategy, optimize the performance of their digital levers, and facilitate decision-making, RED by SFR deployed Google Analytics 360. This included:

- Building a solid technological foundation,
- Measuring the end-to-end user journey through the conversion tunnel,
- Understand the impact of each lever (or combination of channels) on revenue,
- Using real-time data to optimize advertising ROI (Return on Ad Spend (ROAS))
- Creating the reports and visualizations needed to manage the activity.

To accompany them in their technological and strategic choices, RED by SFR chose Artefact, a double GMP & GCP certified partner (Google Marketing Platform and Google Cloud Platform).

The objectives were achieved in only 5 months thanks to the deployment and configuration of Google Analytics 360, the implementation of data-driven attribution models and the development of dynamic dashboards in Google Data Studio.

PROJECTS

1. Tagging – the tagging of all RED by SFR assets as well as the tracking of all campaigns have been reviewed in order to collect data that better meets the needs of performance measurement, client knowledge and media activation.

NB: The configuration of the new Google Analytics (known as GA4) has been anticipated in the tagging plans, in order to be ready to deploy it and to integrate (in particular) the reconciliation of App and Web in the performance vision

2. Data Visualisation – a suite of dynamic dashboards has been implemented in Data Studio for better decision making. These consolidate all the measurement elements of the conversion tunnel, across all channels and according to new attribution models. They reconcile ad-centric and site-centric data via BigQuery.

3. Training – RED by SFR teams have been trained to ensure their operational autonomy in the use and maintenance of new tools such as Data Studio, Facebook Attribution, Google Analytics, Google Tag Manager and BigQuery.



Matthieu Myszak,
VP Consulting
ARTEFACT

"Deploying Google Analytics 360 was the perfect answer to RED by SFR's request: a reliable technological foundation for an intelligible, real-time and full funnel vision of media performance. Its ability to reconstruct client journeys, compare attribution models and refine the targeting of advertising campaigns has had a positive impact on the company's ROAS. The RED by SFR teams have reaped concrete benefits: time savings thanks to dynamic reports, optimization and acceleration of decision making, and autonomy thanks to training."

CO-CONSTRUCTION

In order to define the needs, the project team was made up of data and media experts from Artefact, as well as representatives from the IT, Acquisition, Analytics and Media teams on the RED by SFR side. The tagging plan, in particular, is the result of close collaboration between the two parties.

In addition, the three projects mentioned above were launched simultaneously. Thus, RED by SFR could have consolidated data and trained teams as soon as the tools were put into production.



Alexandre Millet
Multi-Lever Traffic
Acquisition and Data
Analysis Manager
RED by SFR

"We are delighted with our collaboration with Artefact, not only because of their technical expertise but also because of their strategic support in defining new use cases. The multi-disciplinary team mixing experts from our two companies allowed a quick deployment of the chosen technical foundation and delivered convincing results within a few months. By training our teams as the new solutions were deployed, we were able to be operational as soon as they went into production."

RESULTS

Optimized ROAS and
autonomous teams

**-36% reduction in
acquisition costs!**

On the operational side, the new technology foundation facilitates the deployment of new use cases and day-to-day decision making. The features offered by GA360 have a positive impact on media performance optimization. Good results have been obtained in particular on the "cashback and coupons" lever, on SEA (significant drop in spending per day, without a drop in sales) and in social paid.

8 hours a week!

The objective of empowering the RED by SFR teams has been achieved. Their mastery of the tools (GA360, Google Tag Manager (GTM), and Facebook Attribution) has enabled full internalization of media activities. Automated reporting has significantly reduced the time spent entering acquisition data.

As of this year, RED by SFR will be able to consolidate its dynamic and optimized marketing activation strategy according to each client/prospect profile.

What's on the horizon? Use cases based on more data sources for even finer marketing personalization, and therefore an even better ROAS.

Unlock revenue growth with digital marketing maturity



Lubin Arora
Senior Data Consultant

A luxury retailer with ambitious data goals partnered with Artefact to develop a customer-oriented data strategy. Artefact's approach included developing a comprehensive maturity assessment, identifying opportunities for advanced use cases, and recommending a holistic enterprise data roadmap with business opportunities. This partnership resulted in the qualification of 21 marketing operations use cases, unlocking €3 million in value over a 3-year roadmap through new client acquisition, improved customer lifetime value, and optimised operational costs. Furthermore, the engagement facilitated the design of a target architecture and data collection strategy to effectively support the identified use cases roadmap.

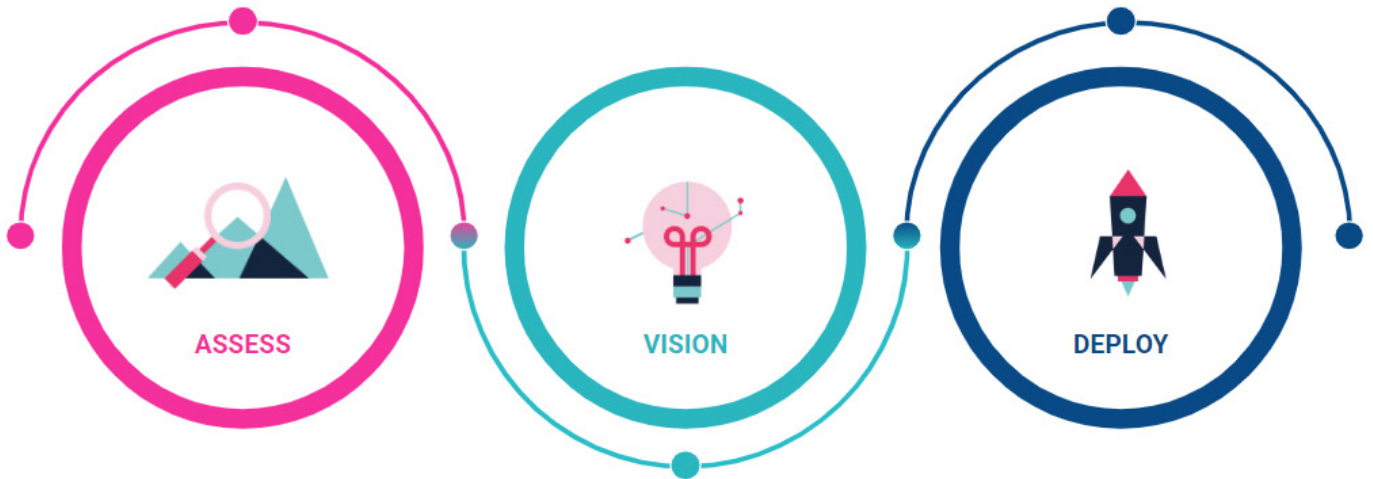
In today's hyperconnected world, where digital channels dominate the business landscape, does your business harness the full potential of digital marketing strategies? The stakes are high, but the rewards are significant – businesses like yours, with mature digital marketing functions, report 18% higher revenue growth and 29% higher cost savings.

Roadblocks to your digital marketing maturity journey

The post-COVID era has witnessed remarkable growth in digital marketing investments, with consistent annual growth of over 10%. As companies strive to advance their digital marketing efforts and bridge existing gaps, significant investments are expected to continue. However, many organisations lack a coherent approach, leading to various pitfalls that slow down progress. This is similar to challenges faced by companies to achieve digital transformation objectives where around 70% of companies fail.

While companies are investing to transition to a cookieless world, implement Customer Data Platforms (CDPs), or employ hyper-targeting activation strategies. However, despite these efforts, many companies are falling short of their expected ROI and struggling to scale these initiatives.

Navigating the complexities of the digital marketing ecosystem requires assessing its completeness and relevance, and examining whether your current digital operating model effectively addresses your business challenges.



What is our actual maturity level?

How mature are we on each dimension compared with the market?

(channels performance, tools, use cases, measurement, op-model, etc.)

Where can we be in 3 years from now?

Where can we realistically be in 3 years from now?

Which global ROI can be expected ?

What should be our short-term priorities?

Where should we start?

What use cases can we utilize to prove value of data driven use case?

How can test our operating model?

Kickstart your digital marketing maturity journey with Artefact

Do you know where you stand on your digital marketing maturity journey? And where do you want to be? At Artefact, we help organisations embark on their transformational journey, where first assess the current state, determine the desired future state, and create a roadmap that aligns with realistic expectations.

At Artefact, we help organisations embark on their transformational journey, where first assess the current state, determine the desired future state, and create a roadmap that aligns with realistic expectation

ASSESS: WHAT IS OUR ACTUAL MATURITY LEVEL?

To determine your digital marketing maturity level, we conduct a thorough assessment across various dimensions, leveraging Artefact’s unique seven-dimension framework. Our team of

experts helps you identify gaps and prioritise areas for improvement, providing a solid foundation to initiate your digital transformation.

VISION: WHERE CAN WE BE IN 3 YEARS FROM NOW?

Once the current maturity level is established, we work with you to outline your digital marketing maturity goals for the next three years, ensuring they align with your internal resources, market dynamics, and industry benchmarks. By setting achievable goals, we help ensure a steady progression towards digital marketing maturity.

DEPLOY: WHERE SHOULD WE START?

To kickstart your transformation journey, we help prioritise areas that deliver quick wins and incremental value. We help identify the top five priorities to launch within the first few months, focusing on initiatives that have a significant impact on your digital marketing performance. By prioritising the right initiatives and

focusing on quick wins, we help you drive incremental value and launch your digital marketing efforts towards long-term success.

Artefact’s Expertise in Action

Artefact has collaborated with clients from diverse industries such as retail, luxury, pharma, consumer goods, transport and logistics etc. to help them assess maturity levels, define. Artefact’s project teams bring in expertise from media advisors, data strategy consultants, and media experts to evaluate and benchmark maturity levels.

Your journey to digital marketing maturity begins here. With Artefact, leverage our expertise in data, analytics, and digital marketing channels. We collaborate with clients to optimise their digital marketing efforts, capitalise on data-driven opportunities, and drive growth in the digital landscape.



CASE STUDY

RECKITT – Optimising media spend and increasing sales at Reckitt with the Artefact Audience Engine

CHALLENGES

Transform into a digital-first business and use data effectively to increase sales

Reckitt is a multinational consumer packaged goods (CPG) company, with offices all over the world. Its purpose is to create a cleaner and healthier world and its products include household favourites such as Dettol, Durex, Harpic, Lysol, Nurofen and Vanish.

Like many businesses today, Reckitt recognised that it relied on traditional methods to understand its customers – consumer panels that delivered audience insight based on surveys, brand knowledge, demographics, consumption and market statistics; shoppers were predominantly reached through TV advertising.

Reckitt's mission was to transform itself into a data-first business, a move that would enable it to exploit the value of its own data to strengthen its understanding of its consumers, and from there make its digital advertising more effective in order to increase sales.



"We know the world has changed; we want to use our data more effectively to drive increased sales."

Stanley Wang
Audience Engine Product Owner
RECKITT

SOLUTION

The Artefact Audience Engine: machine learning and AI models turn data into actionable insight

The Artefact Audience Engine takes a scalable and artificial intelligence (AI)-driven approach to first-party data.

Looking at the user journey (on its websites) provides Reckitt with insight about the user's intentions, and their propensity to purchase. Artefact's custom built machine learning models use this data to create hyper-targeted audiences for digital marketing campaigns. These audiences are built and sent directly to a demand side platform (DSP) of choice through API connections and are used as seed segments for DSP lookalike models to find scale. The Audience Engine allows the brands to answer their specific business questions, better understand their customers, and target them according to their needs.

One of the very few Google partners to be certified on both the Google Cloud Platform (GCP) and Google Marketing Platform (GMP), Artefact used both platforms to build the Reckitt Audience Engine.

The Audience Engine helps Reckitt's brand marketers deliver campaigns that are more effective at reaching core consumers. But the benefits also extend across the business: the media team is able to create more efficient media campaigns; communication strategists can better understand audience performance and use this to prioritise future business objectives and activities; while the CRM team is able to build stronger programmes and content.



"Leveraging the GCP and the GMP at the same time allows us to move data into an environment where we can apply our machine learning and artificial intelligence models and then push it directly back onto the GMP platforms for insights and activations."

Derek Li – Senior Data Consultant, ARTEFACT



DEPLOYMENT



Multi-disciplinary teams roll out and scale the Audience Engine across Reckitt's global business

Artefact worked with Reckitt's transformation office to deploy the Audience Engine at a global level. Data scientists developed the product in terms of its capabilities and algorithms; consultants rolled it out to all the markets and brands; and data engineers worked to industrialise the tool, building a platform so that anyone from Reckitt can run the Audience Engine for any brands automatically. This took place in four streams, with a robust project management methodology implemented by Artefact to deliver a programme of this scale:

- The rollout team worked directly with the martech leads, the media managers and the brands to deploy the tool.
- The product innovation team, made up of data scientists and consultants, listened to the needs of the business and developed new capabilities for the Audience Engine that enabled these to be resolved.
- The product industrialization team worked to scale the product, developing a tool so that the Audience Engine runs automatically.
- The adoption team demonstrated the value of the Audience Engine to the broader business, consolidating results and talking to different stakeholders about how it could help them in their own day-to-day tasks.

"We needed a very strong mix of data scientists, data engineers and consultants to deliver what Reckitt was expecting from the Audience Engine. The roll out has demonstrated increases in both efficiency and effectiveness across all markets, and we are now focusing on product innovation with the team at Reckitt."

**Manuela Mesa - Director
ARTEFACT**

RESULTS

Test and learn campaigns show an average of 21% increase in media spend ROI

Reckitt has big ambitions for its Audience Engine, wanting to ensure it will be used across all key brands and markets. To achieve this, it adopted a methodical approach, first creating digital centres of excellence in specific regions to cover key markets, and then using agile teams to drive development and innovation.

In Mexico, sexual health brand Sico believed its core consumers to be young males however the Audience Engine debunked this long standing idea and showed that young females were equally as important and that the message to consumers shouldn't only align to male interests. The Audience Engine also unearthed an even more exciting piece of new insight – it found that those who concentrated their time on articles related to safe sex and STI prevention on the website, the core audience was dominantly young females.



ARTEFACT

This led to the realisation that in Mexico, there was an appetite for young women to understand how Sico would help them with safe sex practices and there needed to be a purposeful effort to reach young women differently.

The second example is in the US, where media planning and buying is based on the Oracle Data Platform's pre-determined purchase data segments, which look at the specific products or categories of products that people buy. But trials showed that modelling raw data using the Audience Engine machine learning algorithms could increase ROI on media spend by 30%. By taking control of what model to use on a given dataset to drive a specific objective, the Audience Engine has demonstrated to be significant more effectiveness in impacting ROI while also driving greater media buying efficiency. By lowering the cost of media by 20% in this test & learn, we were able to reach more relevant people that were more likely to impact our sales – the Audience Engine makes media budget work smarter and harder.

The Audience Engine's purpose is to enable Reckitt to use data in a highly sophisticated manner to be more effective than conventional digital targeting. The test and learn campaigns set out to prove the value of Audience Engine has solidified the product's credibility and inspires confidence to drive adoption to Reckitt's benefit. Test and learn campaigns such as these inspired confidence by showing the Audience Engine to be highly credible; having quantified the benefits, Reckitt could then start to drive adoption so that it is used in as many campaigns as possible.

FUTURE PLANS

Using the Audience Engine to continue to evolve as a digital and data led company

Reckitt's objective is that the Audience Engine continues to deliver incremental value, making its business more effective and efficient; this will see it support both large and small brands across all its markets by enabling them to gain an in-depth understanding of who their customers are. As technology develops, it will look for new ways to leverage data, adapting the Audience Engine to fit as the company becomes more digital and data led.

"The Audience Engine has changed how our business is able to operate and we will continue to evolve as it enables us to make better and more effective use of our data."

Stanley Wang
Audience Engine Product Owner
RECKITT



"The strategic mindset of Artefact's multi-disciplinary teams has been instrumental in allowing us to leverage the different types of consumer data that we have, as well as build a sophisticated product that can scale globally and is now used every day by our teams around the world."

Stanley Wang - Audience Engine Product Owner, RECKITT

Building your own **audience engine**: turning the Google ecosystem into a cookieless **Customer Data Platform** (CDP)



Pascal Coggia
CEO
ARTEFACT UK



Manuela Mesa
Director
ARTEFACT



Natacha Zouein
Senior Data Consultant
ARTEFACT

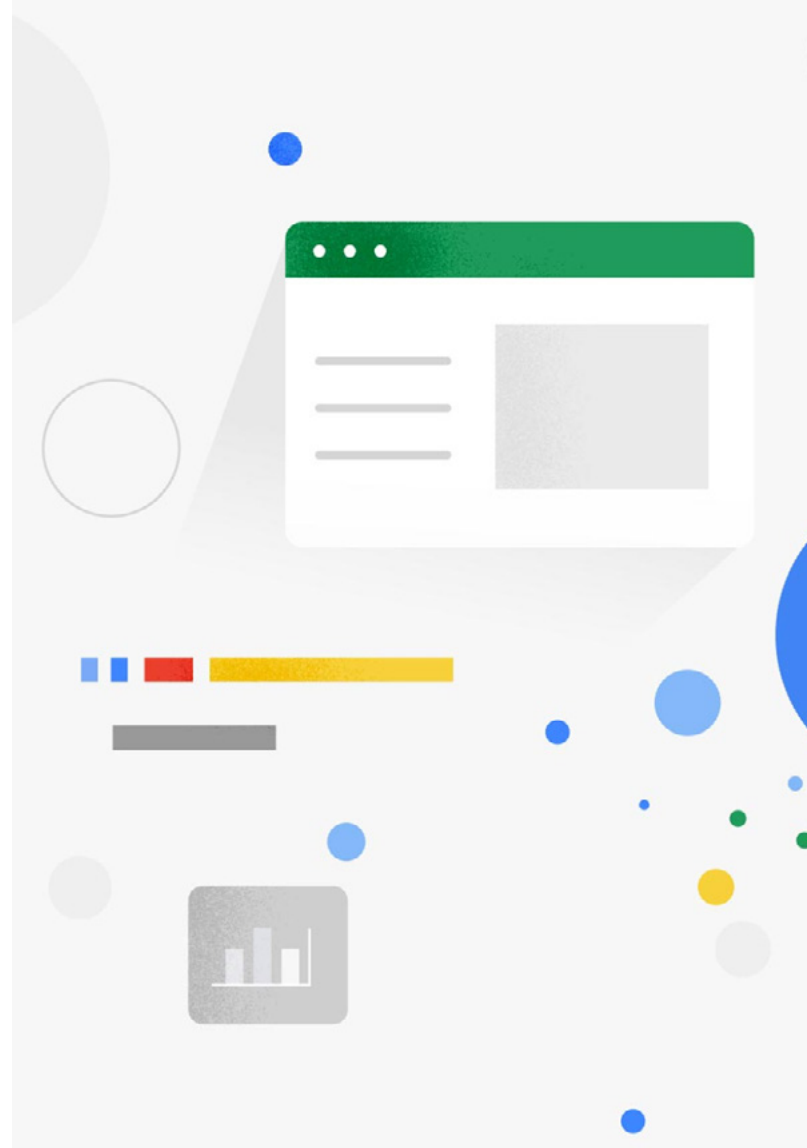
The race for more privacy, yet greater personalization

The dichotomy between the increasing awareness and concerns about customers' data privacy and the rising need for personalized content and ads is forcing companies to rethink their personalization and segmentation strategies.

Research provides strong arguments for personalization (with 74% of customers feeling frustrated when website content is not personalized for example, according to Instapage). This is mirrored by the feedback we receive from clients:

- More than half, across a variety of industries reported sales uplifts of more than 20% in 2020 as a result of personalization;

“The dichotomy between the increasing awareness and concerns about customers’ data privacy and the rising need for personalized content and ads is forcing companies to rethink their personalization and segmentation strategies.”



- Hypertargeting in online media investments represents around \$20bn of additional revenue for consumer packaged goods (CPG) companies;
- The cost of media can be reduced by 5-15% on Facebook and 20-25% on programmatic display.
- Search impression share can be increased across all keywords by around 15% while retaining the same budget.

First-party data ownership is becoming an increasingly significant source of competitive advantage

This situation puts companies under pressure; they have no choice but to enrich their data assets and focus on building a strong first-party data

strategy and customer database.

Our data shows us that global players have realized that:

- Category leaders have set objectives requiring them to collect first-party data for millions - and in some cases billions - of customers;
- CPG brands have digital assets that have seen an increase in traffic of more than 20% in 2021, as a result of which a third of their global customer base engaged online with the brands;
- Companies should be collecting at least 10% of first-party data from their most valuable customers and deploying collection tactics as a priority.

Companies need to equip themselves with tools dedicated to the collection

and processing of first-party data in compliance with the GDPR regulations, as well as strengthen their audience segmentation capabilities. *But how will the imminent removal of third-party cookies impact user identification?*

CDP platforms are flourishing as off-the-shelf solutions for companies to quickly lift and shift their first-party data maturity

The move towards a cookieless world should be seen as an opportunity for new technologies to emerge. To consolidate and properly manage that data, as well as drive holistic insights and make them actionable across marketing channels, Customer Data Platforms (CDPs) are regarded as the most attractive solution. By using this data-driven technology, marketers

can leverage first-party data to extract insights that can help them deliver one-to-one personalization and thus enrich the customer experience.

- First-party data is the first choice among marketers for customer insight and has the greatest impact on customer lifetime value (CLV);
- By 2025, 20% of direct-to-consumer revenue will come from recurring customer relationships, leading to a tenfold increase in first-party data collection.

Limitations of relying solely on IDs, would they be cookies, emails, IPs or other forms of personal identification

CDPs have traditionally been used to consolidate customers' data through a combination of IDs, but IDs generally raise difficulties:

- Loss of an average 32% in ID matching when sharing them media buying platform, and it ranges from 0 to 95% depending on many factors like browser, OS, user consent, browser settings and add-ons, the type of IDs and the limitations of the media buying platform itself. We also expect this percentage to worsen as regulations become stricter, wall gardens increase privacy rules and browsers add new technical barriers to ID collection and matching.
- We estimate that only 57% of media reach can be targeted with IDs and look-alikes with meaningful accuracy, which leaves an immense opportunity and need for non ID based targeting to extend reach with high relevance.
- Fast growing media formats like YouTube trueview and Bumper, Pinterest,

Without the right tools to translate this first-party data into media and marketing strategies, CDPs can be too much of a blank canvas. Indeed, adopting a CDP approach is only one piece of the puzzle; it is crucial to go

beyond IDs and replace more data with more intelligence.

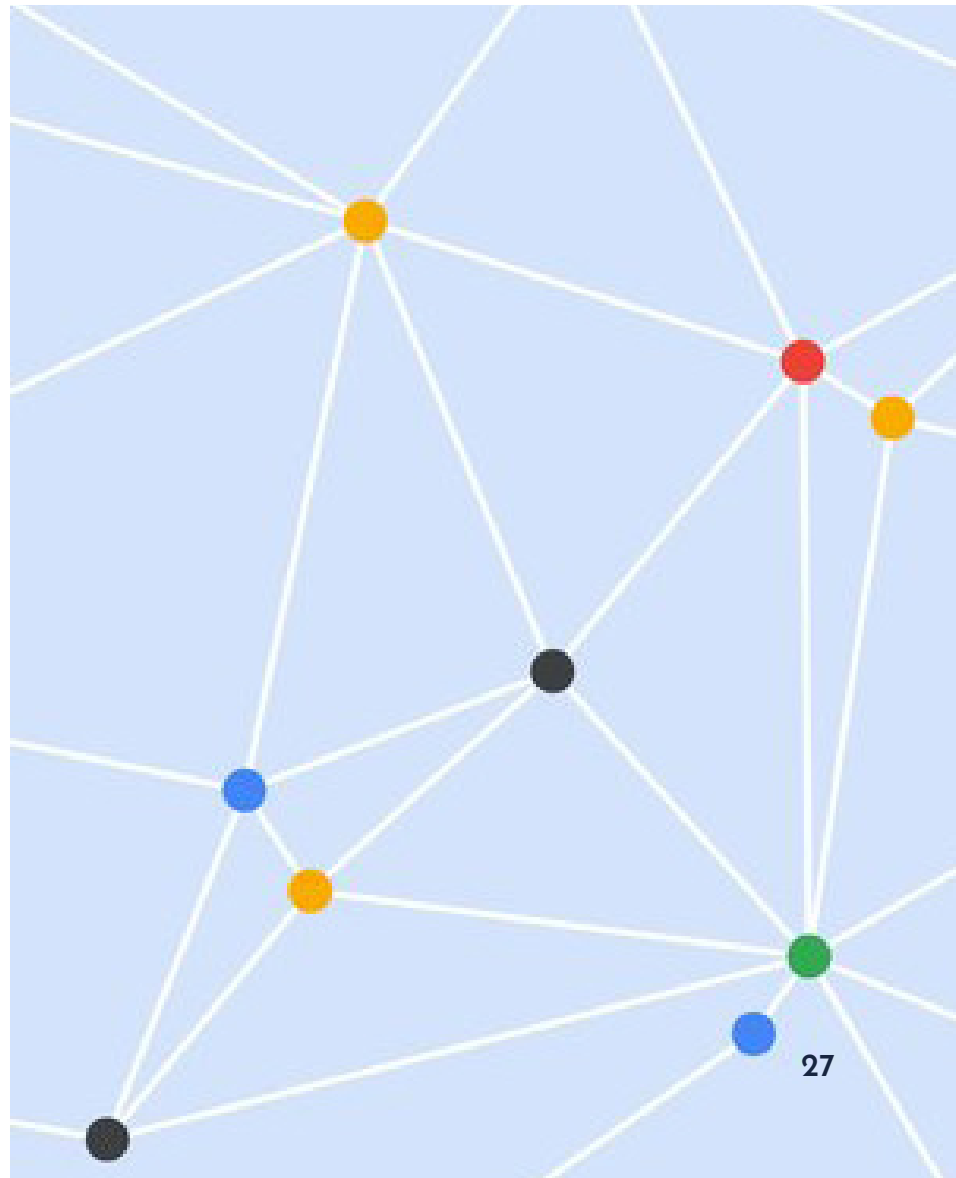
New, rising, and yet big opportunity for brands to leverage first-party data without using names and cookies.

Artefact and cloud providers such as Google help brands to navigate a world without cookies - and Personally Identifiable Information (PII)

The Audience Engine, one of Artefact's bespoke tools for clients, is a CDP that powers personalized digital marketing campaigns by building, sharing, measuring and optimizing audiences from first-party, second-party and third-

party consumer data. It was created based on attributes, not IDs, with the aim of building a scalable and AI-driven approach to first-party data. This entailed adding a layer of intelligence on top of the client's data to ensure an enhanced gathering of insights and learnings about their customers.

Google has been promoting this vision over the past few years as it moves away from cookie-based third-party data and behavioral targeting towards the ability to work with emerging aggregates; initial proposals were based on its Federated Learning of Cohorts (FLoC), with this being replaced in early 2022 by Topics, a new proposal for interest-based targeting. *What does this mean for Artefact clients?*



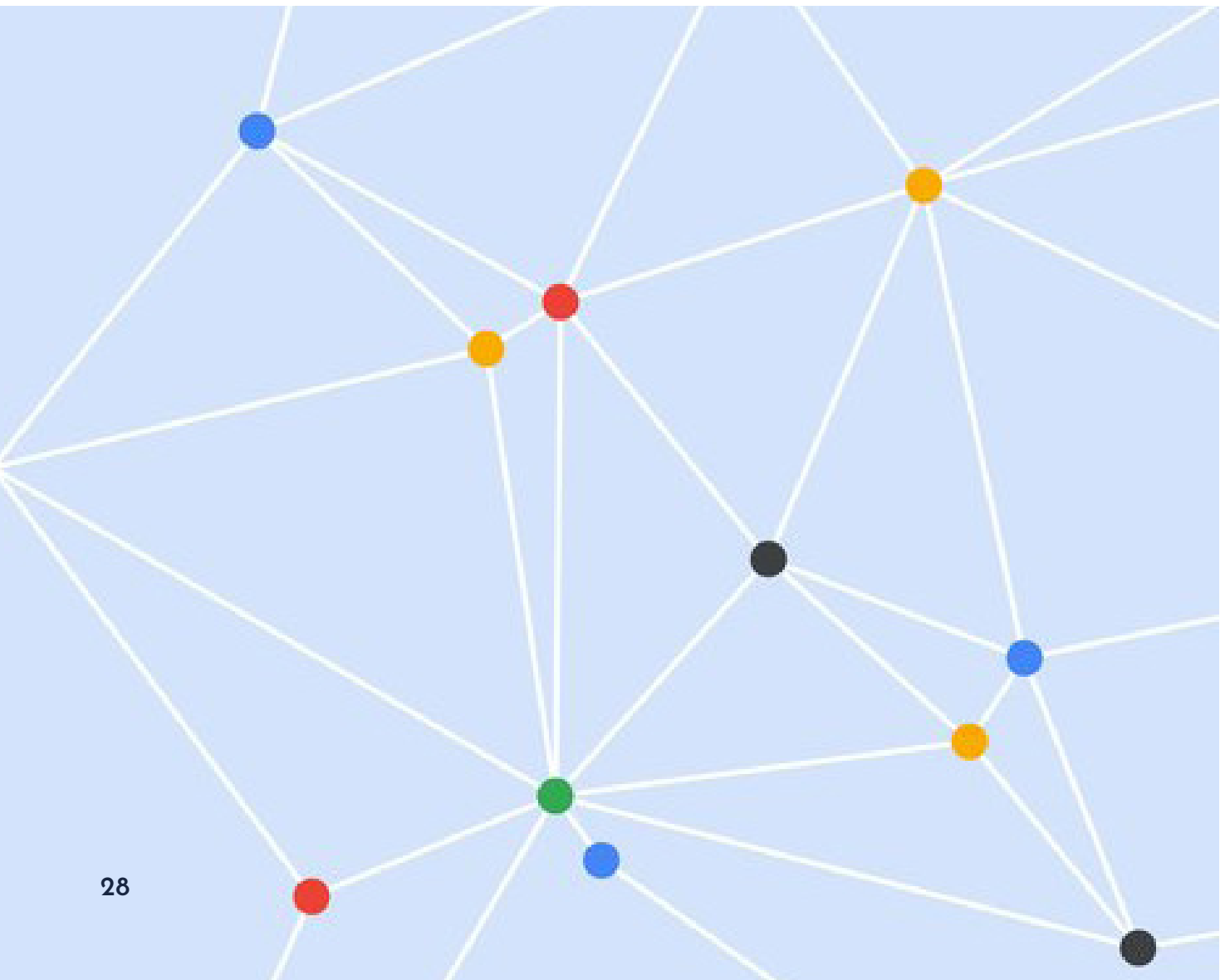
“The deployment of cookieless targeting across various of our clients is showing that it can be as efficient as ID targeting, if not more so.”

A cookieless and PII-less approach offers large advertisers a new gold mine, delivering high levels of performance with a greater reach

The deployment of cookieless targeting across various of our clients is showing that it can be as efficient as ID targeting, if not more so. With enough data, which is gathered from a combination of sources including market and affinity segments, keywords, online behavior patterns, location and consumption reports, we can create audience segments with the equivalent personalization that meet clients’ specific business needs. This type of targeting is delivering effective reach, offering precise and customized personalization on a larger scale.

Data science is only one part of data strategy: companies need to switch from data-driven rules to business-driven machine learning

Personalization is about much more than having a customer’s name and email; what the organization understands about them is also key. Rather than simply focusing on identifiable data, it is essential to use machine learning to add more precise insight to the data collected. CDPs and solutions like the Audience Engine enhance the collection and interpretation of data, and are showing how the switch from data to intelligence can help companies’ personalization and data privacy strategies reach their full potential.





CASE STUDY

UBISOFT

How Artefact helped Ubisoft to build a powerful media campaign using data for its key video game launch of the year

The video games industry is one of the most competitive markets online with many big actors and titles that need to capture gamers' attention all year long and not just during launch. Game publishers are fully aware that the games business is now more globalized, complex, and saturated than ever before, making robust strategic processes throughout development and publishing even more crucial.

Ubisoft was set to launch the new title of its successful franchise, *Far Cry 6*, in 2021.

The 18-month long media campaign was built around different milestones during the year: announcement of the game in June 2020, E3 conference in June 2021, start of the preorders in June 2021, previews of the game in September 2021 and official launch of the game in October 2021.

In this type of competitive environment, data is key to bring more intelligence to digital strategies and enable a virtuous circle leading to a more effective, informed and data-driven media buy.

Three steps are necessary to achieve Ubisoft's performance objectives:

- 1 – Targeting - precisely define marketing segments and associated audiences
- 2 – Messaging - acutely orchestrate campaigns and messages
- 3 – Measuring - accurately measure the performance of campaigns

For a game publisher, the main challenge when marketing a new game is that it speaks to different types of gamers and one common message can't realistically address all these disparate audiences.

“The challenge is to find the targets, orchestrate and optimise campaigns according to our KPIs.”

Natacha Kocupyr

There are 2.7 billion gamers on the planet, and people across all ages and demographics are engaging with games with all varying levels of interest, loyalty and advocacy. No two gamers are alike and broad labels such as “casual gamers” and “core gamers” are no longer useful on a granular level.

“The evolution of profiles of gamers creates a richness of targets that will enable us to identify the players that are the most receptive to what we want to offer.”


Clément Prevosto
Director Media & CRM, Ubisoft EMEA

Using its proprietary data marketing model and Ubisoft's strong target knowledge, Artefact and Ubisoft came up with different targets, each representing a unique segment of gamers, mutually exclusive, that would be interested in the game to boost performance, spot growth opportunities and maximise potential.

“For each game, we start by defining the targets that we want to focus our investments on that are most likely to convert.”

Clément Prevosto

Relying on a data-driven methodology helped Artefact's team of experts translate these marketing segments into relevant media audiences. To achieve this step, Artefact involved media planners and media experts to track and understand potential target groups in different countries. Together, they identified the right advertising partners and formats capable of reaching these audiences and to capitalise on untapped gamer groups.



“What makes a good partner is their ability to be agile and to react quickly”

Clément Prevosto

Once the media audiences were defined, the team managed to evaluate the number of people that each audience represents and the volume of impressions needed to achieve the required cover rate. The demographics data allowed experts to break down the core demographics of the franchise into different key variables of a player base.

This data, supplemented with easily accessible consumer engagement data ranging across media tools and Ubisoft's own ecosystem, was key to serve the audiences by reaching out through the right channels.

“It is about relying both on Ubisoft customer knowledge and data from advertising tools offered by platforms.”

Natacha Kocupyr

Working on a 18-month long media campaign helped the team optimise, streamline and forecast the operation even further, identify patterns and adjust the marketing plan across different markets and territories to get the most of each euro spent.

Insights gained from Artefact experts have been integral to the marketing strategy and Far Cry's brand growth.

“This digital media strategy centers on Far Cry 6 but it is used on all of our games”

Clément Prevosto

Compared to the control group, the audiences targeted by the media campaign generated +7% video views and +40% clickthrough rate.

“We measure the performance around business KPIs (such as the volume preorders before the launch of the game), media KPIs and impact KPIs on campaigns and audiences.”

Natacha Kocupyr
VP Digital Marketing, Artefact

+7%
video views

+40%
clickthrough
rate

Ubisoft relies on three types of KPIs to evaluate the performance of its media campaigns:

- Business KPIs (preorders, beta downloads)
- Marketing KPIs (awareness, organic views, search levels, sentiment)
- Media KPIs (brand lift, search lift)

During the duration of the campaign, brand lift surveys were a particularly useful tool for the management of the media strategy. More than 70 of them were conducted in the span of 18 months.

“We absolutely have to optimise very quickly and achieve performance in a short duration of time when launching games.”

Clément Prevosto
Director Media & CRM, Ubisoft EMEA

To conclude, Artefact and Ubisoft, using a variety of tools, defined the most relevant segments, **selected the right audiences and estimated their numbers to create a continuously optimised successful media campaign** leveraging data that delivered increased engagement in a competitive market environment.

The gaming audience is much more diverse than most media plans assume – and failing to account for these nuances only leads to lower view rates and a higher cost-per-view. **Make no assumptions about your audience and strive to get a holistic understanding of the market and individual games** that will allow you to segment gaming enthusiasts, based on playing, viewing, owning, and social behavior, to reflect the games market’s changing engagement.

Fundamentals of marketing: targeting the right person, at the right time with the right message remain the same, but **data algorithms make the consumer insights much more accurate and by consequence audience targeting much more powerful** and allocation of media budgets increasingly profitable. The golden triangle of marketing deserves its bright color nowadays thanks to data !



MROI A new data-driven approach to measuring on and offline campaigns



Florian Thiébaud
Managing partner
Data Marketing Lead
 ARTEFACT

Florian Thiebaut, spoke to the podcast “Les 5 minutes data” created by Intuiti. In this episode, Florian Thiebaut is interviewed about the innovative solutions for measuring the performance of online advertising campaigns available to Artefact clients.

Indeed, techniques such as MMM, attribution, or geographic experiments are no longer sufficient to meet the needs of advertisers in a rapidly changing context. For more granular, more frequent, and more activable insights, it is necessary to bring a new dimension to these solutions by relying on data. This new approach, called MROI at Artefact, allows advertisers to better manage their investments by answering questions such as: “How can I optimise my media mix to meet my brand awareness objective?” or “What incremental value do the new formats we are testing bring compared to traditional campaigns?”

Transforming performance measurement is essential in a changing economic, technological and legal context

Marketing performance measurement is becoming an increasingly complex topic for advertisers. Advertisers must rethink traditional methods to meet today’s challenges, under pressure from 4 factors:

- 1. Digital channels and formats have multiplied** in recent years. Social networks are diversifying and display formats are evolving, especially those offered by Google such as Discovery campaigns, Performance Max campaigns or local campaigns on Google Maps for example.
- 2. The regulations related to the processing of personal data** have been strengthened. The collection of consent prior to the deposit of cookies has repercussions on the measurement of campaigns since it was historically based on this type of tracer.
- 3. The technical environment is changing.** Third-party cookies and certain personal identifiers, such as mobile IDs for example, are about to disappear. These are decisions made by web and mobile browsers, not regulations.
- 4. The uncertain economic context** (pandemic, inflation, hyper-competition...) impacts advertisers’ budgets and the way they manage their investments.



Historical measurement tools are insufficient to manage advertising investments in this new context

Until now, digital marketing professionals have mainly used 3 types of measurement:

1. Marketing Mix Modelling, usually performed twice a year or quarterly. MMM provides insights and recommendations across all channels (including offline). However, these are not granular enough, not easily actionable, and infrequent. Moreover, these analyses are very often outsourced to third-party players or technologies.

2. Attribution, obtained for example via Google Analytics, provides good granularity (at the level of campaigns, audience, and format) and frequency (almost real-time) of insights. On the other hand, this information remains limited to the performance of online advertisements on sales and does not really make it possible to make the link between the various channels, particularly offline.

3. Geo-targeting experiments, which make it possible to test incrementality on “twins” and to understand the impact of a change in strategy on a campaign. However, this requires frequent testing and therefore requires large budgets to obtain actionable results and recommendations.

These measurement techniques are no longer sufficient: they communicate little with each other, are only partially consolidatable and therefore do not provide an exhaustive view of performance measurement. It is therefore difficult for marketers to adapt their budgets quickly and regularly on the various levers.

MROI, a new dimension for traditional measurement techniques

At Artefact, we are implementing a complementary approach to the three techniques mentioned above, which our longstanding clients have already taken advantage of. This approach, which we call MROI, consists of internalising and automating MMM analyses to go beyond frequency, granularity and actionability of insights.

To achieve this, Artefact’s MROI solution leverages the computational capabilities of cloud technologies such as GCP, data visualisation technologies such as Looker, and the highly granular cross-channel data provided by the Google GMP suite of tools.

In this way, data processing is automated to obtain insights that can be visualised dynamically via Looker. This makes it possible to help marketers answer the questions they ask themselves on a daily basis, such as:

- What is the optimal media mix for a given audience and period, such as Christmas, for example?
- How much budget should I allocate to YouTube, TV, and other video channels to meet my brand awareness objective?
- What value and increment do the new types of campaigns and ad formats offered by Google bring to my media mix?
- To what extent do these new campaigns and formats allow me to create incremental value on the performance and branding KPIs of my campaigns? How do they complement my traditional campaigns, and with what synergy effects?

Finally, it should be borne in mind that the automatic ingestion of offline data remains a challenge. The transformation of cross-channel campaign measurement is therefore still in progress. On the other hand, these new approaches are perfectly feasible for online channels and bring a new dimension to the management of digital advertising investments.



CASE STUDY

SANOFI CONSUMER HEALTH CARE (CHC)

Precision Marketing - Industrialising the deployment of data-driven campaigns

CHALLENGES

Scaling advanced Precision Marketing across 30+ markets.

Sanofi is one of the world leaders in the pharmaceutical industry. In the past 3 years, Artefact has helped the Sanofi CHC (Consumer Health Care) business unit market its over-the-counter medications via digital-first tactics and enablers to reach the right consumers at the right time with the right message, across more than 30 markets.

For its seasonal products category, Sanofi CHC has developed a forecasting-based approach to adjust digital media spend according to predicted demand peaks. Through multiple pilot campaigns, the Global Digital Transformation team was able to

prove the added value of this approach with an ROAS multiplied by 2 to 4 according to geographies.

However, setting up a new campaign remained time-consuming: data scientists had to go through a series of manual, repetitive and error-prone tasks, preventing them from focusing on other innovative projects. In order to scale its innovative ML pipelines, the Sanofi data science team defined their needs to industrialise the use case and called for the support of Artefact to jointly design and implement a robust solution.

SOLUTIONS

A co-designed industrialisation process based on 6 key solutions.

Through a close collaboration between Artefact and Sanofi's data and business teams, a comprehensive industrialisation process leveraging the unified Databricks platform was designed. Our joint objectives were to:

- **Simplify the end-to-end setup of a new seasonal campaign**
- **Automate data ingestion and processing tasks**
- **Make the solution more robust to prevent errors and manual maintenance**
- **Improve project maintainability and scaling**

Following a swift 1-week audit to map out the current process and technical pain points, the team aligned on implementing a future-proof infrastructure based on 6 key solutions:

1 Separation of concerns
Having a separate ETL pipeline from the forecasting model process makes it easier to maintain and scale. This allowed us to implement automated checks alongside a monitoring system that sends detailed reports to the relevant teams about the ingestion status.

2 Use of Delta Lake as a data golden source
In DS teams where infrastructure can be a pain to obtain/maintain, Delta Lake combines the key features of data warehouse and data lakes solutions, thereby removing the complexity of SQL database admin. It also has versioning capabilities – important for ML reproducibility – and will serve as the unique source of truth for data.

3 Packaging as much code as possible into a Python library to simplify processes

Part of the initial code was scattered among several notebooks within Databricks, complexifying management of dependencies and code reusability. Notebook-based development is relevant for prototyping but can create challenges for ML projects industrialisation. Having clearly defined Python libraries implemented on the notebook and keeping only Databricks as entry point for Compute made it easier to generalise notebooks and organise incoming campaigns.

4 Leveraging Spark and Databricks
Training the model using hyperparameter search methods can be time-consuming and demanding. This is where the autoscaling infrastructure of Databricks and the managed ML runtime with Spark and HyperOpt come in handy. Using memory computations in a distributed manner over a set of workers speeds up performance and considerably improves training time.

5 Use of ML Flow tracking
With ML Flow tracking in place, Sanofi now has a User Interface where Data Scientists can compare model runs and keep track of all parameters used (Data version and model parameters) and results obtained.

6 Simplified new ML model testing and implementation
A generic model factory framework was set up, making it easier to implement new machine learning models, and to try them during a Precision Marketing campaign with very little effort.

RESULTS

A setup time divided by four for data ingestion and configuration.

Thanks to this project, Sanofi CHC was able to greatly simplify its data pipeline and accelerate the scaling of its forecasting-based Precision Marketing use case.

Reduction of setup time for new campaigns

- Setup time for data ingestion and configuration reduced by up to a fourth.
- Number of tasks performed by data scientists to set up a new campaign reduced by up to a third.

Simplification of the creation of new forecasting models

- Accessible platform to easily test, manage and visualise models.
- Generic process to include new data sources.
- Automated data pipeline.

Key Achievements

Key Achievements	
Reduce the setup time for Data Scientists	Facilitate the creation of new triggers based on new data sources or new ML models to improve accuracy
The setup time of the data and configuration has been divided by 4.	Create a platform to easily test, manage and visualize models
	Define a generic process to include new data sources
The number of tasks done by a data scientist to set up a campaign have been approx. Divided by three from 15 to 4	Migrate the Python lib to an automated pipeline to simplify processes

“The key to the success of the project was the close collaboration between Sanofi business experts and Sanofi data scientists with the Artefact team.”

– Albert Pla Planas, Data Science Team Lead, Sanofi

The project also allowed the teams to generate 4 important learnings for future ML-driven projects:

1**Integrate data engineering in ML projects**

Involve Data Engineers from the beginning of a project to accelerate industrialisation of the pipeline, and clearly decouple the different stages of the pipeline (all data handling, transformation and curation must happen before jumping into the ML stages).

2**Tap into pre-packaged tools**

The use of Databricks with Delta Lake and ML Flow was crucial to industrialisation success, ensuring an easy self-service infrastructure without the need for DevOps.

3**Deep collaboration between Business and Data teams**

Possibly the most important success factor was the close collaboration between Sanofi business experts and data scientists, who ideated and drove the project, with the Artefact team, who brought additional industrialisation experience and know-how.

4**Use agile methodologies to industrialise**

The agile methodology (sprints, and quick iterations followed by feedback & alignment weeks) was very efficient to identify and address all Sanofi's pain points and ensure value delivery for Sanofi teams.

CDPs and the Business of Data in Sports & Entertainment



Wayne Campbell
Partner
ARTEFACT

In the entertainment industry, like any other, data is collected from thousands of different consumer touchpoints 24 hours a day, every day. To draw actionable insights from this information and enable data-driven decisions that deliver concrete benefits to fans and consumers, companies need powerful analytical tools. Customer data platforms (CDPs) have become the preferred solution for many brands.

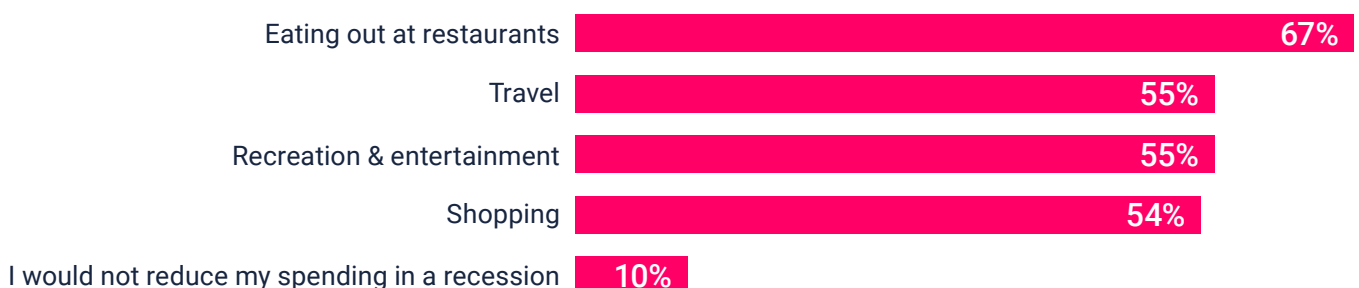
Vast quantities of data on production trends, viewing habits and programming preferences are collected by production studios, streaming services and media companies. By analyzing this data, they can surface precious insights that deliver concrete benefits to fans and consumers, from forecasting and personalization to content recommendations.

Sports and entertainment have been with us for over 3000 years; the first recorded Olympic Games took place in 776 BCE and have been played every four years since (with a “minor” pause of 1500 years after Emperor Theodosius banned them due to their pagan influences). Sports are played and watched and enjoyed by people of all cultures around the world and will always feature largely in our lives.

In the sports and entertainment industries, as in many others today, fan and consumer attention is increasingly solicited, but disposable income is increasingly limited: with inflation at more than four-decade highs, consumers are reducing their discretionary spending. According to a survey conducted by Variety magazine, 55% of respondents said they have begun making changes to spending on activities such as travel, recreation and entertainment.

Discretionary Spending Reductions in a Recession

Question: Which categories would you pull back on in the event of a recession?



Source: variety/morning consult
Note: survey conducted among 2,200 u.s. Adults july 6-7, 2022

ARTEFACT

Meanwhile, fan and consumer expectations are higher than ever. They not only want brands to present them with personalized offers for sports and entertainment-related events and merchandise – they demand it. But due to ongoing changes across the regulatory landscape, especially with regard to the demise of third-party cookies which prevent the collection of consumer browsing data, it's more difficult to precisely identify and target audiences, especially given that fans and consumers transcend geographical, cultural and time boundaries.

How data is changing the game in the sports industry

Intelligent use of customer data can benefit all areas of sports, from partnership and sponsorship activation to sales of merchandise, food and beverage (F&B), and tickets.

- **Partnership and sponsorship activation:** Their participation facilitates identification of specific audiences (fans and consumers) for any given event, online or in the physical world. Other new entrants such as short-form fan-made content, e-gaming, betting, web 3.0, metaverse and NFTs represent engagement paradigms that should be added to the sports/entertainment portfolio as possible partnership and sponsorship assets.
- **Merchandise:** By leveraging consumer data to power retail strategies such as data-driven demand forecasting, inventory management, product selection, product mix, and omni-channel availability and pricing optimization, brands can increase sales, conversion and price competitiveness.
- **F&B:** As with merchandise, consumer data allows companies and brands to understand what fans like, forecast demand, manage inventory, employ dynamic pricing,

optimize menu options and bundles, inform staffing levels, and measure what worked well, what worked poorly, and everything in between.

- **Ticket sales:** Data can forecast ticket demand for any given event (physical and digital), optimize ticket pricing, and create event hospitality bundles (tickets, F&B, parking, accommodation, experience), all of which can be extremely valuable in driving sales and securing revenue.

What data is doing to light up the entertainment world

The entertainment industry relies heavily on data analysis today. Production studios, streaming services and media companies all collect vast quantities of data on production trends, viewing habits and programming preferences, and they need data analysis to make data-driven decisions. "You do not make a \$100 million investment these days

"Fan and consumer expectations are higher than ever. They not only want brands to present them with personalized offers for sports and entertainment-related events and merchandise - they demand it."





How CDPs can enhance fan and consumer engagement

By maximizing the value of their customer data, brands and companies in sports and entertainment could differentiate themselves by creating revenue-generating ideas for boosting attendance, driving engagement and brand affinity with unique fan and consumer experiences, and growing sponsorship opportunities.

Ten years ago, Data Management Platforms (DMPs) facilitated data-driven strategies, but today's regulatory and technological landscape makes them obsolete as they were designed around third-party data (cookies). To make a difference today and in the future, data assets must be enriched and the trust of fans, consumers and audiences must be won. To meet these challenges, tools dedicated to the collection and processing of first-party data in compliance with regulations must be acquired. Customer Data Platforms (CDP) have become the preferred solution for many brands.

In the last nine months alone, Artefact has delivered more than 25 CDP projects. Brands are investing more resources into becoming consumer- and fan-centric because they see the continuously changing landscape driven by tightening data privacy regulations, the shift from CRM to CDP, omni-channel fan and consumer journeys, heightened consumer expectations, and consumer fatigue. Today, brands want to build longer-term relationships with their consumers that are more tailored, personalized and rewarding.

As the data-driven world continues to rapidly evolve, there seems to be little doubt about the business value of CDP. It's time for organizations to consider deploying this forward-looking technology.

without an awful lot of analytics," as Dave Hastings, Netflix's director of product analytics, famously commented.

Insights from data enable companies to deliver concrete benefits to fans and consumers, including forecasting, personalization, and content recommendations by:

- **Adapting content**, artists, or entertainers based on analysis of consumer data including demand, tastes, trends and customer sentiment;
- **Estimating demand** for specific content to inform ticket and/or streaming pricing, and deciding when to release and deliver content;
- **Analyzing box office performance** by using consumer and market insights to reveal why certain movies were flops or hits and predict which new ones will be successful;
- **Offering a content delivery channel mix** (social media, connected TV, mobile, in-venue, virtual, etc.) based on consumer preferences drawn from collected data.



How AI can help brands discover consumer microtrends



Cryil Fekete
Partner
ARTEFACT

Widespread lockdowns have changed shopping habits and strengthened the influence of online prosumers - brand advocates who also create their own products. These microbrands can capitalise on emerging trends faster than the competition. To keep up, consumer brands should leverage AI tools to help them predict the next big craze.

We've all heard how COVID-19 has pulled us into an online-only market, yet these consumption habits have been evolving for years now. With the rise of social networks and user-generated content, online prosumers - those who consume but also produce their own products - have become increasingly influential as microbrand competitors, setting trends and steering consumer conversations at speed.

This has created a new challenge for big brands who often lack the agility to keep up. It typically takes between 15 and 22 months for a big brand to move a consumer product from an idea to being product-ready. Yet in the modern world, these smaller and more agile 'microbrands' are able to capitalise on niches faster.

Unlike big brands, they operate at just the right level to spot a gap in their market first. They also don't have the same, slow-moving decision-making or product development structures the big brands do.

Glossier is a good example. Starting out as prosumer beauty blog Into the Gloss, founder Emily Weiss was able to tap into the frustrations of millennials and rapidly meet them with targeted beauty product ranges. It took only five years for Glossier to generate \$150m in sales, and it's now a niche beauty brand valued at \$1.2b.

Microbrands, like Swedish haircare brand Maria Nila and Skin Laundry skin treatment, may not become the next market leaders but, collectively, they're cannibalising market share and profits for the big players. In fact, between 2016 and 2017, small companies comprised only 19% of FMCG sales but generated 53% of US industry growth.

So, how should big brands respond to stay relevant?

Seeing the future with AI

Beyond shortening the process for new product development, brands should focus on capitalising on the sweet spot as a trend reaches its height, ahead of the competition. That means forecasting trends before they happen.

This is no easy task, but the deluge of online, publicly available data out there gives brands a window into evolving consumer wants – and with the right tools, it can even tell you how tastes will develop in the future.

Change begins small, but, in business, a ‘microtrend’ can hint at a much larger megatrend that sweeps the market only months later. To have your product ready in time, the key is being aware of emerging microtrends long before they turn macro.

Crucially, this kind of trend detection is not necessarily about looking into the future; it’s about noticing what’s happening to consumers right now and being able to predict accurately how those trends will develop.

Traditional market research won’t cut it though – it moves too slowly for the online world. Instead, a new generation of AI-driven predictive and analysis tools makes the trend detection process efficient, repeatable and highly accurate.

Market intelligence departments have long had an overabundance of search and social data to draw from – but what they lack is the time, resources and mental computing power to properly understand it. AI solves that problem. It can crunch huge amounts of public data quickly and repeatedly, giving marketers invaluable, real-time insight into market trends.

Getting started with AI trend detection

Online influencers tend to be the source or a reflection of emerging microtrends. The first step for brands, then, should be to find out who are the most influential opinion leaders in their target niche. It might be tempting to simply monitor those with the largest followings – but they are rarely where microtrends originate.

The closer you can get to the source, the more time you have to develop your product before the microtrend really hits the market. Unsupervised AI algorithms – unbiased and uninfluenced by existing market knowledge – can pick out those early opinion formers whose content regularly goes on to start profitable microtrends.

Once you have that core group of influencers, you can monitor them for the next big trends. Of course, even at this stage, there will be more information than you can handle. Deciding which microtrend to focus on and invest in can be difficult when there are thousands of potentials to consider.

Fortunately, powerful predictive intelligence can anticipate how trends will interact and evolve once they’ve made contact with the wider market – helping brands see which trends have legs and what are likely to be most profitable. Making a decision becomes much easier.

With AI computing power and predictive intelligence, a company can confidently predict the next big thing. A large beauty brand, for example, could use AI to identify its most powerful influencers and communities. It can analyse the potential profitability of the topics they mention on social media as it matches with former trends’ expansion pattern. With foresight of what’s going to be popular down the line, the brand can start development, acquire a new company or revamp an existing product, just in time to capitalise on the craze.

Ultimately, when used correctly, AI can help a big brand become as agile as a micro one. It won’t make the development process any shorter, but it will help brands start earlier and finish right on time – beating the microbrands in the process.



Digital Commerce

- 45 | Becoming data-driven (and reaping immediate benefits) has never been easier in the retail sector
- 47 | **CARREFOUR GROUP** – How Data & AI can accelerate sustainable business transformation
- 50 | **CONFORAMA** AI-enabled personalization boosts Conforama CRM campaign revenues
- 54 | Retail Media: An indispensable asset for brands
- 56 | **UNILEVER** – How does Artefact support UNILEVER on Retail Media use cases to increase its sales?
- 60 | Data monetization opportunities for retailers: Retail Media within the CPG/Retailer data ecosystem
- 62 | **CARREFOUR LINKS** – With Artefact are helping brands to improve media targeting thanks to Carrefour's data
- 65 | **MATTEL × RelevanC Advertising** – How Artefact helped boost Mattel's online sales on Cdiscount (RelevanC Advertising) retail media platform?
- 67 | **CARREFOUR GROUP** – How to reduce food waste in the Bakery-Pastry department?
- 70 | Demand forecasting: Using machine learning to predict retail sales
- 72 | Sales forecasting in retail: what we learned from the M5 competition

Becoming **data-driven** (and reaping immediate benefits) has never been easier in the retail sector



Jérôme Petit
Managing Partner
Retail & eCommerce
 ARTEFACT

Data technologies are now at the head of the aisle in the great supermarket of retail opportunities. Serve yourself: take advantage of these unbeatable offers today!

Particularly turbulent market conditions in 2022 disrupted the retail sector's pricing and supply chain policies. While price optimization has always been a daily concern for this industry, the 5.2% increase in consumer prices over the whole of 2022 (compared to 1.6% in 2021) and its acceleration in the first quarter of 2023 have intensified the need to react. More frequent price adjustment decisions must be made every day to absorb cost increases or react to competitor repositioning, and be immediately deployed in store networks.

To make matters worse, a shadow still hangs over supply chains, inherited from pandemic events. Many retailers are no longer sure they can offer all products due to lack of stock.

To address both of these challenges, players who are able to rely heavily on their data and process it instantly and at scale through machine learning stand out. Of course, you can't go from data-rich to data-driven overnight, but contrary to common fears, the maturity of the technologies is such that the first building blocks, which can rapidly bear fruit, can be deployed in mere weeks.

Machine learning to optimize prices...

Inflation has focused retailer attention on the daily challenge of price optimization: how to identify products that are most sensitive to price variations (elasticity)? Which products shape the price image of a brand (known value items)? How to react in real time to competitor repositioning? Since promotions represent a growing share of sales (more than ever, consumers are looking for bargains that fit a tight budget), how can they be optimized without impacting profitability? All the answers are in the data! Customer data and proprietary sales history can be combined with external data such as competitor price and promotion records, seasonality, calendar events, and, yes, even the weather (!) to make the right decisions. Retailer pricers and category management teams usually do all this. Unfortunately, they often need to work in Excel but have limited time to devote to it, due to their other responsibilities.

So, what's new? Why, it's the availability today of proven off-the-shelf algorithms that automate the simplest decision-making



ARTEFACT



“In a business where margins are so tight that operational excellence is a necessity, the notion of a data-driven company is far from new. What’s changing today is the ease of access and use of technological platforms.”

processes. It’s the seamless and massive availability of third-party data in the market. It’s the ease of use of technology stacks that allow millions of transactions to be processed in a few milliseconds. Today, it only takes three months to build a data platform that combines all transaction data, promotions, stock, product hierarchy, store hierarchy, customer data, etc. And a technology partner can manage the infrastructure, resource deployment and network dimensions in the cloud through its managed services. Today, like Monsieur Jourdain, if you know Excel and PowerPoint, you’re a data analyst without realizing it: in a matter of days, you can take control of data in BigQuery (Google), Synapse (Microsoft), or Snowflake and build interactive dashboards in Looker, Power BI or Tableau.

...and better manage inventory

In recent years, the health and geopolitical context has also challenged supply chains. Today, the supplier service rate varies from week to week and delivery times can be very uncertain. Distribution channels have also grown highly complex: not only do stores need to be stocked, but home deliveries, click-and-collect, and partnerships

need to be served as well. Once again, data science comes to the rescue of retailers by allowing them to better control inventory management. By leveraging machine learning, retailers can now analyze receipts in real time to immediately detect out-of-stock items, calculate the spread of uncertainties across all links in the chain to better size buffer stocks, or improve stock allocation under an infinite number of constraints (to optimize costs, shorten delivery times, or reduce carbon footprint).

Democratizing data use throughout the company: data is about people

In a business where margins are so tight that operational excellence is a necessity, the notion of a data-driven company is far from new. What’s changing today is the ease of access and use of technological platforms.

If technology is no longer a barrier, the challenge is still to make these solutions available to the widest possible audience. To democratize their use, simple solutions need to be deployed on a massive scale, employee training programs need to be multiplied, whether on-demand or more intensive, and events (e.g., hackathons) need to be organized to engage the managers who are driving

the transformation. Data is about people. This is Artefact’s slogan, and rightly so.

Directly monetizable data... But perhaps not immediately for everyone

The icing on the cake is that data itself is a goldmine, thanks to retail media and data sharing. As digital signals become more difficult to capture, the billions of transactions and customer interactions that retailers generate have become a critical strategic advantage for them. This data, which provides in-depth understanding of consumer expectations, has great potential for monetization. But it represents a profound, existential transformation of the retailer business model: moving from a self-financed model (with negative working capital) but with very narrow margins, to a model where initial investments are substantial but margins are high. A Copernican revolution, perhaps not the easiest to undertake for all players.

In the great supermarket of retail value creation opportunities, data technologies are now at the top of the shelf, in self-service. Retailers, why wait to share these unbeatable offers with your partners?

The image shows the Carrefour logo in large, blue, 3D letters mounted on a building facade. The building has a green and yellow geometric pattern. The sky is blue with light clouds.

CASE STUDY

CARREFOUR GROUP

How Data & AI can accelerate sustainable business transformation

CHALLENGES

How Artefact is helping Carrefour achieve carbon neutrality for its e-commerce activity?

Carrefour Group is the leading European retailer and the world's second largest retailer, and is present in more than 30 countries. Carrefour's international profile raises a number of ecological challenges and a desire to offer its customers, regardless of their level of awareness, quality food and services accessible to all.

The Group's aim is to become the world leader in food system transformation for all by committing to four major objectives, including achieving carbon neutrality by 2030 for its e-commerce activities.

To achieve carbon neutrality, three main levers of action have been identified:

Measuring the ecological impact of a delivery in order to manage the strategy;

Reducing the carbon emissions of its logistic and digital infrastructures;

Engaging customers to become participants in ecological transformation.

This aim also provides a triple opportunity for the Group's e-commerce activity: reducing its operating costs, improving its NPS score (customer satisfaction indicator) and anticipating legal changes.

To seize these opportunities and take concrete action on each of these levers, Carrefour must be able to measure all greenhouse gas emissions by drawing on real data that compiles all data storage, transport and logistics activities, from first click to final delivery, whether to the home or by store pickup.

SOLUTION



Defining a reliable, actionable, transparent carbon measure for Carrefour and its customers

The first step for the Artefact and Carrefour teams was to agree on the scope of action for measuring this carbon footprint. They decided to limit themselves to measuring greenhouse gas emissions generated by e-commerce orders in 2021

“The challenge we gave Artefact was to calculate the CO₂ emission of an online order. How much CO₂ will a customer produce if their order is delivered or if it’s picked up at the store?”

Bertrand Swiderski
Chief Sustainability Officer
CARREFOUR

The second step was to collect activity data in order to convert it into carbon emissions. As this data wasn’t already present and documented in Carrefour’s data platform, the business teams (logistics, warehouses, e-commerce) had to be brought together to obtain it. This step proved to be crucial to the operation’s success, as it allowed all stakeholders to become ambassadors for the group’s “carbon neutrality 2030” objective.

Carrefour’s strategy for measuring its carbon footprint was based on a systemic, unifying, long-term, iterative approach. The strategy was successful thanks to the participation of over 30 employees and the involvement of Carrefour customers via their “Engaged Consumers Clubs”.

“Today, we recognize that consumers are becoming experts on these topics. They want to understand how things are done and want to challenge companies. Thanks to them, the project has matured.”

Léonard Cahon
Consulting Manager - ARTEFACT

Encouraged by these initial results, Carrefour will continue its commitment by publishing the carbon footprint of each of its orders on its e-commerce site in the near future.

“Soon, customers will clearly see the number of kilograms per CO₂ on their orders, thanks to the insights gained from our carbon assessment.”

Manuel Chatain
E-commerce CSR Manager - CARREFOUR





RESULTS

Opening a wider field of possibilities and ecological alternatives

By analyzing its carbon footprint and implementing this first measure, Carrefour now has a way to pilot its e-commerce carbon emissions reduction strategy. The group can now encourage its customers to review their consumption patterns in order to be more responsible, encourage its service providers to reduce their emissions, and also promote internal awareness by proposing several possible levers of action:

- Act on the choice of delivery slots in order to optimize truck loading, routes and schedules;
- Increase the number of clean vehicles (biogas, electric or hydrogen) by 2030;
- Reduce the amount of packaging used.

And to ensure the sustainability of this measurement and its easy use by the teams, Artefact teams worked on three elements:

- A dashboard to run trajectory simulations by combining forecasted activity data;
- Training to teach how to use and modify the dashboard;
- Detailed documentation to enable employees and clients to understand and reproduce the measurement process from start to finish.

Using data as a key lever to help businesses achieve their environmental objectives

Carrefour's e-commerce, supply chain and logistics platform teams worked together to meet this challenge, supported by the collaboration and expertise of the Artefact, Aktio and Google teams. The project is part of Artefact's "Data for Sustainability" solutions, which aim to create a positive impact on the environment through data by accelerating the ecological transformation of businesses.

"We expected a very clear vision of what each basket order would emit. Our request was complex, but Artefact responded to it with great dynamism and agility."

Bertrand Swiderski
Chief Sustainability Officer - CARREFOUR

"At Artefact, we believe data will play a major role in helping companies achieve their carbon neutrality goals."

Vincent Blaclard
Partner - ARTEFACT



CASE STUDY

CONFORAMA

AI-enabled personalization boosts Conforama CRM campaign revenues

CHALLENGES

Conforama is the second largest home furnishings retailer in France and is present in seven countries, with 300 stores, including 200 in France. The company sells furniture and decorative items in kit form and posted sales of 1.7 billion euros in 2022.

As a gateway brand, Conforama's goal is to "Make what people want most accessible at the best price." It's an ambition backed by a transformation plan to deliver an omnichannel experience through data and AI. An initial audit and data marketing vision with Artefact identified and prioritized 12 use cases and 25 technical and organizational enablers. The first use case was to integrate a personalized product recommendation into the company's weekly emails.

Several challenges needed to be addressed through this use case:

- How to understand the needs of three million customers and recommend the most relevant products from a catalog with 42,000 references?
- How to propose only products currently in stock, on promotion, and not already suggested to customers?
- How to easily operate and maintain the technical solution?

Conforama

"Time savings, yes, but above all a business benefit for our CRM teams. Because thanks to this personalization, customers click more and therefore buy more. We've gained 15% of the click rate following the personalization of these emails, which represents several million in incremental sales."

**Mélodie Charles, Marketing Director
CONFORAMA**



SOLUTION

Saving consumers time, improving business productivity

By using machine learning algorithms to analyze user data, such as preferences, purchase history and online behavior, artificial intelligence-based product recommendation suggests products relevant to consumers in a personalized way. This allows companies to better understand their customers' needs and recommend products that match their interests, resulting in increased sales and customer retention.

One of the main benefits of this solution is that it saves customers time. Rather than scrolling through countless product pages to find what they're looking for, customers can quickly access a selection of recommended products that specifically meet their needs. AI-based product recommendation can enhance the online shopping experience and encourage customers to return for more purchases. A strategic advantage, given that 72% of consumers only interact with marketing messages that are personalized and tailored to their interests.

In addition, AI-based product recommendation can boost business productivity: machine learning algorithms can analyze large amounts of data in real time, allowing companies to continuously monitor customer trends and buying behavior. This can help organizations better understand customer desires and quickly adapt their product offerings accordingly. It can also enable companies to optimize their inventory by offering products that are more likely to sell, which can lower costs and maximize profits.

Lastly, AI-based product recommendation can offer significant business benefits. By suggesting relevant and personalized products to customers, companies can improve their conversion rate, increase sales and strengthen their brand image. From a market perspective, AI-based product recommendation has been shown to deliver +2.5% incremental growth.

A first use case focused on email campaign personalization

Conforama

Prior to this project, all Conforama customers received emails featuring the same eight products selected each week by the marketing teams. This was a labor-intensive task, as it required identifying the eight products most likely to interest three million customers, each of whom had unique interests. All this time spent analyzing data could have been spent on more strategic activities, such as creating editorial content for those emails.

Today, an email is sent to every Conforama customer each Tuesday containing eight product recommendations. But these recommendations are personalized according to purchase history, and filtered exclusively for products that are on sale, are available in stores, and that haven't been featured in previous activations.

The implemented AI solution includes 4 main data processing steps:

- Collection of transaction histories, customer and product references, then data preparation;

- Building the "Collaborative Filtering" model to calculate customer appetite for the product catalog;
- Product filtering based on available inventory, commercial news (sales, promotions, etc.), past activations and purchases;
- Product data enrichment (photos, prices, descriptions, etc.) for activation.

This solution is based on 16 data tables, 25 transformation and modeling steps, and 40 automated quality tests. Dozens of iterations of the model made it possible to choose the most efficient approach based on transaction history. Thanks to this solution, Conforama now generates several million recommendations each week in 45 minutes at a cost of 50 euros per week.

In other words, if you count development and operation costs, as well as incremental sales, the project break-even point is reached in one week, with an automated and reliable solution.





Conforama

RESULTS

A smooth transition to AI: lessons from Conforama's success story

For many players, there are three challenges linked to their level of maturity:

LEVEL 1

Personalizing a currently rule-based touchpoint using an AI algorithmic approach;

LEVEL 2

Extending AI-based personalized recommendation across the entire customer journey (similar products / complementary products / suggestion based on purchase history);

LEVEL 3

Optimizing the orchestration of recommendations across channels to ensure an omnichannel experience.

Level 1 is often the most difficult, as it requires laying the foundations for four separate dimensions: target vision, user experience and priorities; data sources; technological tools; project team and work method.

The Conforama example offers valuable lessons about these four dimensions:

- Select a first use case and functionalities that can be quickly implemented and measured to put the organization on the road to success. For example, this initial victory means Conforama can now plan the deployment of product recommendations in stores or the improvement of their algorithm thanks to browsing data.
- Ensure the data is reliable. Good data modeling relies first and foremost on good quality data. For Conforama, exploratory analyses were performed on more than 50 tables to select data sources in areas such as customer knowledge, product repositories and transactions.
- Use technologies that allow teams to deploy a technical solution quickly and collaboratively. Conforama selected the most appropriate tools for this type of workflow: DBT, BigQuery ML and Vertex AI for their performance, modularity and portability.
- Build a dedicated team capable of dealing with all potential problems, and adopt a test and learn approach. To do this, a multidisciplinary IT / Conforama business team was formed, and a 2-week sprint approach was adopted.

Retail Media: An indispensable asset for brands

While Retail Media represented only 9% of digital media investments for brands in 2019, it will soar to 43% of these investments in 2023 and is expected to double in value by 2024 to reach €100bn. Vincent Cailliot, Director of Data Consulting and Sidney Zeder, Senior Consulting Manager – Data Marketing, both of Artefact, explore the opportunities of retail media for Consumer Packaged Goods (CPG) brands.



Vincent Cailliot
Director Data Consulting
ARTEFACT



Sidney Zeder
Senior Consulting Manager
ARTEFACT

Retail Media investment driven by regulatory and tech developments

The increased importance of Retail Media in the digital strategies of brands can be explained by the evolution of the availability of consumer data, which is at the heart of any media personalization strategy.

In the past, third-party cookies were mainly used to drive this strategy. Regulatory developments to better protect consumer data, including the GDPR in Europe, have led to new technological constraints, making cookies obsolete: This process began as early as 2016, when Safari removed cookies from its platform. It continues today with Google Chrome, the most popular browser, announcing the removal of third-party cookies from its platform next year.

As a result, brands need to find new data sources to build their digital activation strategy. One solution is for brands to better leverage their first-party (1P) data assets, i.e., their proprietary data, by collecting more and better data from their customers. But to create digital activation strategies with long-term value, proprietary data is no longer sufficient: it needs to be enriched with data – especially transactional data – from retail partners.

Facilitated data sharing between brands and retailers allows brands to deploy ever more relevant digital marketing strategies with a high level of autonomy. This increased collaboration and data sharing between retailers and brands is possible thanks to “data clean rooms” such as Amazon Marketing Cloud, LiveRamp or Decentriq, which allow

the sharing of personally identifiable consumer data at the individual level in an anonymous way.

A rapidly evolving ecosystem

The ecosystem of technology partners around Retail Media is highly fragmented and constantly evolving, with partners that are more or less specialized depending on major Retail Media activities: first-party, second-party or third-party cookie data collection tools, data processing and audience creation, activation or analysis, etc. The challenge for brands will be to identify which combination of technology partners will best meet their needs, depending on their current technical ecosystem and their own business challenges.

Retailers are also an essential part of this ecosystem, providing access

to transactional data to build their Retail Media strategy. While the majority of retailers in the US have launched Retail Media offerings, most retailers in Europe are still in the experimentation and use case-testing phases; few have yet industrialized use cases with brands.

Valuable use cases beyond Retail Media

LeRetail Media allows brands to address marketing use cases from consumer insight generation to digital campaign activation and marketing performance measurement. The availability of transactional data (previously unavailable to CPG B2B2C brands) at the “individual” level enables the construction of insights and activation plans that are all the more impactful. The same data can be used to measure their effect on

sales and ROI, enabling effective optimization of activation plans.

Retail Media is just the next step towards more collaboration between retailers and brands. In a long-term partnership perspective, collaboration and data sharing can enable the implementation of more advanced category management and supply chain use cases, such as the analysis of the long-term value of existing promotions or the prediction of in-store product demand and thus optimize supply chain operations.

Which Go-To-Market strategy to launch?

For retailers, it’s important to define a new offer to monetize their data. This can range from monetizing their owned media inventory (website), to sharing data “as a service” in a clean room, to offering services (campaign management or reporting as proposed by Amazon for example). These new offers can be marketed internally or via partners. The internal or external development strategy will determine the associated costs, in terms of salaried resources (commercial and technical profiles to be recruited) and technical resources (clean room tools, technical infrastructure to be set up).

For brands, the approach will be more traditional, from defining their business needs by identifying and prioritizing use cases, to setting up the partnership with their key business partners. Likewise, the implementation of pilots with a retailer to test the value of use cases can be carried out by a third-party partner.





CASE STUDY

UNILEVER

How does Artefact support Unilever on Retail Media use cases to increase its sales?

CHALLENGES

Thanks to Retail Media, Unilever identifies new growth opportunities and increases sales of its priority product categories.

The global Unilever group has a portfolio of 400 brands that are anchored in the daily lives of its 5 billion consumers. Present in France for 125 years, Unilever is the leader in several market categories including ice cream, condiments and deodorants.

The consumer packaged goods (CPG) sector has the particularity of being intermediated. Since the distribution of their products is carried out by different retailers, there is little direct relationship between CPG brands and their clients. Yet, consumer knowledge is key to optimize media and promotional strategies, product assortment in the territory or to identify new growth levers.

It is from this challenge that the so-called Retail Media

offers were born, i.e. 2nd party data partnerships. A retailer's data is made available to a brand to enrich its own data assets in a win/win approach. This movement was initiated in 2012/14 by pure players such as Amazon and Alibaba, and traditional players such as Casino, Walmart and Carrefour have gradually followed.

Accompanied by Artefact since 2019 on various data issues, the Unilever Group seized this opportunity to identify new growth levers, develop a common consumer and product approach with a leading retailer and increase sales of certain priority product categories. This is the case, in particular, for the Magnum brand's mini ice cream bars range.



“Retail Media is a win/win strategy for brands and retailers. Retailers’ data allows us to enrich the shopper’s knowledge and accurately measure our activities on all channels, throughout the transformation tunnel. For their part, retailers find a new source of revenue and differentiation from their competitors. In addition, it is a way to better satisfy their clients with more personalized offers and a better anticipation of stock shortages.”

Sarah Baqa
Head of Performance Marketing - Unilever

Retail Media in a win/win partnership with Trade Marketing

Advertising expenditure (media and traditional marketing) in the consumer goods sector amounts to nearly 680 billion dollars worldwide and 16 billion in France. The investments allocated to Trade Marketing are just as massive.

While the advertising levers have been highly digitized over the last twenty years, Trade Marketing is still very underdeveloped. However, driven by e-commerce, it represents half of the budgets of CPG advertisers (600 billion dollars worldwide, 16 billion in France). Trade marketing is strategic for these brands because it consists in carrying out actions in partnership with mass retailers to better meet consumers’ expectations: prospectus, merchandising, point of sale animations but also sales forecast and supply chain optimization. Thus, the digitalization of trade marketing represents a real growth opportunity for advertisers and brands.

Retail Media is based on this concept of partnership in the promotion of brands’ products. The sharing of 2nd party data favors client knowledge and allows a better segmentation and therefore a better targeting of advertisements or promotions. This targeting can be activated in both proprietary brand environments and in external audience extension environments.

Retail Media infuses Unilever’s media, promotional and supply chain strategy

Unilever, together with Artefact, has identified 6 strategic axes to exploit the full potential of Retail Media:

- **Media activation**
- **Measuring Customer Lifetime Value**
- **Coupon customization**
- **Optimization of store assortments**
- **Short-term sales forecast**
- **Supply of shelves**

Since the CPG group initiated these retail partnerships, each of these axes has been worked on through concrete use cases. For this, Unilever has benefited from Artefact’s end-to-end support. This concerns data strategy, the launch of pilots, the construction of new data products, but also the provision of human resources (data analysts, data scientists and data engineers), as well as the training of Unilever employees.



SOLUTION

An iterative approach

The Retail Media project was envisaged through an iterative approach in 4 steps:

- **Identify and select the right distribution partners** based on their pre-existing relationship with Unilever and their technology infrastructure.
- **Laying the foundations of the partnership**, i.e. the use cases that can be implemented as well as the necessary legal and technical requirements. The modus operandi is also built to facilitate collaboration between all stakeholders (agencies, internal and external teams of Unilever and the brand).
- **Implement pilots for each use case.** They allow for testing and optimization of devices on key Unilever brands (Magnum and Sun) before scaling up.
- **Deploy devices on a greater number of brands,** use cases, markets and partner brands.

The vision of this data partnership is therefore based on a virtuous circle: the measurement of actions carried out makes it possible to collect insights which in turn feed the next activations.

“Our data strategy is really ambitious, so we wanted to be accompanied by a team of experts. We chose Artefact, which already has experience with similar products and works in a very advanced way on data and Retail Media solutions. In addition, they have the technical capabilities and resources to take use cases to scale. This is the key to this long-term partnership.”

Sarah Baqa



“Retail Media allows us to activate audiences and precisely measure the link between digital actions (advertising, promotions) and sales. The data from the retailer’s loyalty cards also has the advantage of tracking consumers over time (Customer Lifetime Value) to see if they have increased their purchase frequency, if they have increased the value of their basket or if they have switched brands, etc.”

Florian Thiebaut,
Managing Partner – Data Marketing Lead - ARTEFACT

RESULTS

A use case conducted on Magnum’s mini bars segment

For example, retailers’ data was activated in the media to meet a challenge for the Magnum brand: to confirm its leadership during the summer period, which is key for this product line, and to acquire new clients in the Confectionery segment (mini bars).

To achieve this, the first step was to use shopper data to identify the different consumer segments. We showed them the same banners during a defined period. This allowed us to measure the increase in on and offline sales, but also to track the buyers recruited to retain them post-campaign. Through this first wave, Unilever was able to identify the best performing and most profitable audiences. The ultimate goal is to use these results at scale to target these audiences with more personalized ad creative and promotional offers.

This pilot campaign significantly increased the brand’s revenue.

Unilever’s data and Retail Media strategy will continue in 2022 alongside Artefact. In addition to the two «Precision Media» levers, the exploitation of the 4 other identified work axes is also planned.



Data monetization opportunities for retailers: Retail Media within the CPG/Retailer data ecosystem

Retail media has been around for quite a while, but thanks to the evolution of new uses of consumer data, its potential is gaining attention. Sidney Zeder, Senior Manager and Gaéтан Bélan, Senior Data Consultant & Product Owner, both of Artefact, explore the opportunities of data monetization for retailers.



Gaéтан Bélan
Senior Data Consultant
& Product Owner
ARTEFACT

Retail media has been on the rise on digital platforms for the last six years, most notably on Amazon. The Covid-19 crisis accelerated this trend for traditional retailers. Retail media, in simple terms, is the means for retailers to sell media inventory on their e-commerce platforms. Because the Covid-19 pandemic fueled the shift to digital ways of buying, such as e-commerce or click-and-collect, even for grocery shopping, retailers had no choice but to go with the flow.

82% of CPG companies surveyed said they were already investing in at least one retail media platform. This represents approximately 17% of digital budgets already allocated to retail media.



Sidney Zeder
Senior Consulting Manager
ARTEFACT

In fact, between 2019 and 2020, CPG e-commerce penetration increased by five points, from 10 to 15 percent. For retailers, the downside is that margins are lower in e-commerce than in brick-and-mortar. The upside is that by selling online, they collect a lot of consumer data that can be monetized or used to create new services. In a Goldman Sachs study,

Media investment has effectively shifted down the marketing funnel. Although many brands' search investments still flow into the "Google family," we're seeing brands diversify their digital spend into e-commerce platforms to capitalize on the "search destination" status they hold. When you're on Amazon as a consumer, you're very close to the "moment of truth": you're in a purchasing mindset. Therefore, when you're on Amazon as a brand or product, the closer you can get to that funnel, the better. Goldman Sachs expects this trend to translate into a 6-8 percent increase in total CPG e-commerce sales through retail media over the next four years.

Retailer data monetization opportunities with CPG brands

This close-to-the-funnel media investment trend has created opportunities for retailers around three types of data monetization with CPG brands:

1. Inventory monetization: traditional retail media consisting of selling media inventory on proprietary assets. This can be offline inventory – retailers have long monetized their customers to offer coupons or specific promotions to brands in their stores – but also their online inventory on their own platforms, such as their e-commerce website, where brands can display banner ads, emails or even shopping mobile applications to deliver personalized promotions to their customers.

2. Data monetization: retailers are monetizing existing consumer data across CPG brands to support their customer centricity. 1P data shared by retailers originates from their loyalty program. The cardholder data they share can be socio-demographic (e.g. the age

of their consumers), transactional (e.g. what did they buy), behavioral (e.g. what did they look at), loyalty data (e.g. did they buy again), etc. This data is shared “as-a-service” in a data clean room where brands can access the retailer’s data in a secure environment to carry out specific use cases defined by the two partners.

Carrefour, for example, has created a consumer intelligence service called Carrefour Links, based on the LiveRamp clean room, where partners can access their cardholder data. This is a self-service platform that allows users to perform basic activities such as reconciling retailer and brand databases on individual customers to build a more complete view of the consumer and thus improve their experience. It also provides analytics and measurement capabilities that Carrefour can bill to its partners.

Access to this data can unlock three types of use cases for brands:

- **Marketing:** the data shared by retailers allows brands to gain insights about their consumers, activate them with media, or measure marketing performance

through transactional data. For example, an ice cream brand partnered with a retailer to build advanced audiences for a digital marketing campaign. Using the retailer’s transactional data, the brand was able to build and activate two audiences: the brand’s current ice cream buyers and ice cream buyers from competitor brands. As a result, the brand was able to increase the uplift of its campaigns by targeting the two relevant audiences with adapted messaging.

- **Trade:** the data shared by retailers allows brands to perform revenue growth management use cases by better optimizing promotions or assortment... It also unlocks store optimization use cases through enhanced in-store experience or sales force optimization. For example, one brand worked with a retailer to analyze the short- and long-term impact of promotions on incremental margin. This enabled them to identify certain types of promotions that were margin-destroying for both the brand and the retailer, as opposed to those that generated a positive long-term business impact.



- **Operations:** the data shared by retailers allows brands to optimize their supply chain through demand forecasting and demand management use cases. It can also fuel sustainability as well as production and innovation use cases.

Brands usually start implementing marketing use cases to deliver short-term value with a simple set-up, while long-term partnerships can then address very valuable use cases for Trade and operations that benefit both partners. Across all three categories, data monetization unlocks better measurement capabilities: optimized media performances, customer lifetime value calculation, dynamic budget allocation and global ROI optimization.

3. Service monetization: along with inventory and data monetization, retailers can unlock additional revenue by providing different levels of services to brands.

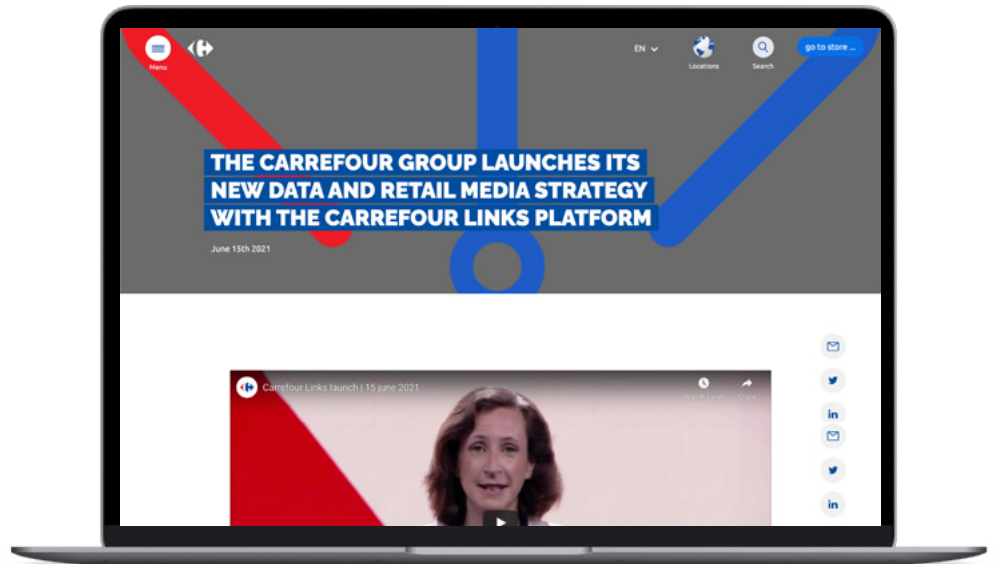
In the most developed additional service offering, retailers can propose managed services to brands, based on shared SLA and KPIs. In a long-term partnership approach, retailers optimize their revenue potential through best-in-class services for brands. Amazon offers its key clients advanced services such as category management studies and dashboards or MMM (Media Mix Modeling) studies to help them improve their strategy in relation to this retailer. Des cas d'usages à valeur, au delà du retail media

Data monetization: maturity in the Retail Media market

Artefact has benchmarked over 20 retailers in the US Retail Media market and analyzed their maturity with respect to Retail Media based on their value proposition and the comprehensiveness of their capabilities. We found that most retailers in the US have launched a Retail Media offering but are still in a nascent stage, mainly offering inventory to brands, while more mature retailers are focusing on data monetization or even service monetization for the best-in-class retailers.

Data monetization offerings require setting up technical capabilities, such as a technical infrastructure to collect, store and process the first-party data to be shared, as well as a clean room to share data with brands, or even partnerships with DSPs to enable brands to directly activate audiences created in clean rooms. But the business opportunity is worth it: selling data as-a-service delivers margins often in excess of 80%, compared to inventory monetization where margins typically only reach 40%, as inventory assets are limited and therefore less scalable.





CASE STUDY

CARREFOUR LINKS

With Artefact are helping brands to improve media targeting thanks to Carrefour's data

CHALLENGES

Retail ecosystem is currently going through a revolution

The old rules of retail have been deeply impacted by two new factors: digital assets (digital media) and shopper data, that helps leverage knowledge of customer behavior and granular activation of digital assets.

Retail media mixes digital assets and shopper data. For many years, retail media has been helping brands and manufacturers become more efficient in four different ways:

- **customer insights:** a deeper understanding of your customers thanks to data from identified members of customer loyalty programs
- **precise targeting:** moving from mass marketing towards 1:1 marketing
- **activation:** omnichannel activations that can be synced through several touchpoints enhanced by shopper data
- **measuring:** measure of the return on ad spend and impact on sales

SOLUTION

Carrefour's vision on their Retail Media system

Carrefour provides manufacturers with cutting-edge solutions to conduct retail media campaigns and to develop their sales and display presence both in store and on their ecommerce website.

According to Nicolas Trolé, Carrefour Media has the unique opportunity to be able to leverage assets from the whole entire division, such as 5000 physical stores, a digital platform that reaches 10 millions customers every month, and the key element in their omnichannel offering, their unique data.

In 2021, Carrefour took the decision to develop a new full stack platform on carrefour.fr, with Criteo as partner, to offer advertisers new activation solutions and better measuring and targeting.

“When we talk about data at Carrefour, we are talking about data from our customer loyalty program, which represents 14 million cardholders, 7 million customers registered online and offline, and 5 million digital-only customers”

Nicolas Trolé,
Chief Revenue Officer - Carrefour Links

“Retail media is an ecosystem that needs a certain number of mandatories on the retailer's and manufacturer's sides”

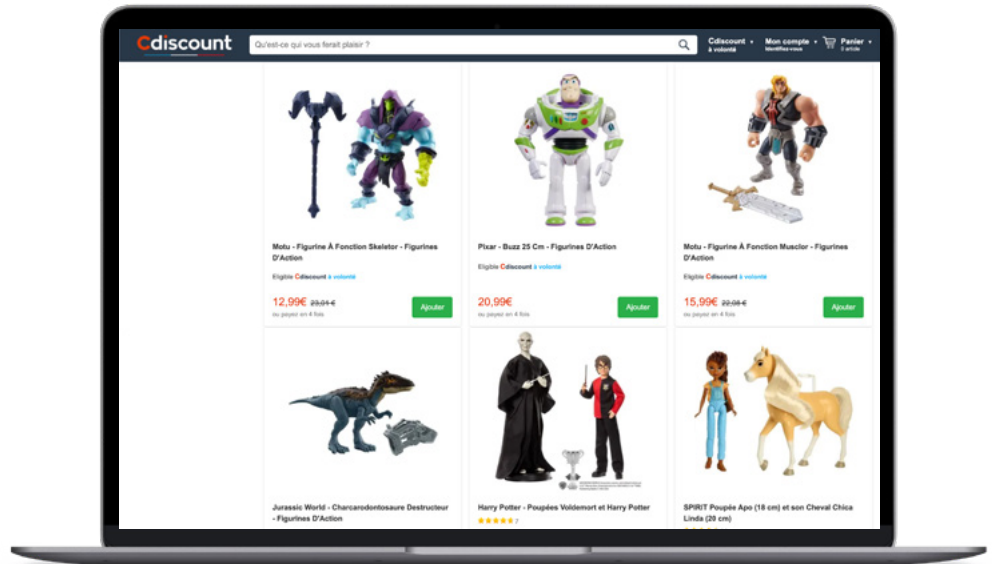
Thomas Faure,
Senior Consulting Manager
E-Retail Lead - Artefact

RESULTS

Artefact as a trusted partner to maximise the value of Carrefour's data

Thomas Faure reviews the four pillars of the partnership between Carrefour Links and Artefact:

- **Business consultancy:** Manufacturers want to understand where growth comes from. “Carrefour's data helps brands identify business opportunities for them to accelerate”
- **Profiling and customer analysis:** “Who is my customer?” One of the main challenges for manufacturers is to define who their customer is. Carrefour's data helps brands represent the personas of their brand buyers
- **Strategy, operation and activation:** helping brands create media activation strategies (budget, KPIs)
- **Advanced data use cases:** projects that take into account the whole retail value proposition



CASE STUDY

How Artefact helped boost Mattel's online sales on Cdiscount (Relevance Advertising) retail media platform?

CHALLENGES

Increase presence and visibility for the brands of the Mattel Group on e-retailers media ecosystems by leveraging retail media.

E-retail is a strategic distribution channel for the Mattel Group as **32% of transactions** in the gaming category were conducted online during the 1st semester of 2020.

Retail media is used in an objective of performance enabled by leveraging shopper data. Retail media is especially helpful for the conversion part of the sales funnel.

Retail media lives in different types of environments such as:

- **in-platform**, products that maximise the presence of products directly on the eCommerce website such as sponsored brands, sponsored products, merchandising banners
- **off-platform**, solutions that generate additional traffic redirected to the product pages on the eCommerce website such as search engine ads, social ads, display & video ads

SOLUTION

High touch handling of Mattel's retail media investments in partnership with relevanC Advertising.

In 2020, Cdiscount cemented its position as the number one French e-retailer, generating 25 million unique visitors and 10 million customers (adding a million new buyers during the year).

Cdiscount's vision of its retail media ecosystem covers the entire sales funnel and encompasses different steps (brand awareness, consideration, traffic, acquisition, and insights)

In February 2020, relevanC introduced a self-serve platform, relevanC Advertising Platform, built entirely in-house, that offers retail media solutions through search and display on Cdiscount.

"relevanC Advertising is a best-in-class tool to operate and manage retail media campaigns,"

says Maïana Darmendrail – *Digital Manager & E-Retail Manager, Mattel France*

"CDiscount possesses very mature retail media solutions in search, display and video,"

states Thomas Faure – *E-Retail Lead, Artefact*

Artefact's made the decision to perform an all-year round online campaign aimed towards buyer intent. Artefact defined custom audiences out of precise shopper data items (purchase history, buyer intent, search history, browsing history, income level ..).

The key to increase performance of media investments was to conduct continuous optimizations on both retail and media KPIs.

"We have the intimate conviction that continuous optimization brings performance improvements,"

explains Cédric Chamoux – *Directeur Retail Media, relevanC Advertising*

- Optimizations were made on **several factors: channels** (onsite, offsite, search, display), **segments** (audience, keywords), **creatives, formats**.
- Optimizations were based on **several metrics: media metrics, business metrics, retail metrics** (stock level, promotions, organic positioning)
- Optimizations happened on different solutions: **search advertising**, daily improvements on **keyword selection, display advertising**, bi-weekly improvements on **impact measurement**

In addition to leveraging the e-retailer ecosystem, Artefact kept looking for ways to innovate and test new solutions to build more expertise and scale projects.

In that way, Artefact conducted off-platform campaigns through Shopping for Partners solutions (*Google for Retail*) that were activated to support milestones, such as product launches.

relevanC
ADVERTISING



RESULTS

Ad campaigns six times more effective!

Mattel achieved extremely positive results through retail media activations in 2020:

- **600% increase** on return on ad spend for media investments using retail media data (compared to traditional campaigns)
- **from 4 up to 10 euros** earned when spending one euro on display and search ad campaigns

"relevanC Advertising has the most qualitative data that you can find on the market. The daily optimization by Artefact really moved the needle,"

mentions Maïana Darmendrail – *Digital Manager & E-Retail Manager, Mattel France*

Brands are able to find success when mixing business metrics with media expertise, i.e. sync retail signals (stock level, promotion level) with media KPIs to better optimize advertising campaigns.



CASE STUDY



CARREFOUR GROUP

How to reduce food waste in the bakery-pastry department?

CHALLENGES

How to reduce food waste in the Bakery-Pastry department?

Predicting demand as accurately as possible is one of the foundations of the retail business. However, this challenge is becoming more complex as consumer habits evolve. From now on, it is necessary to take into account the combined use of different sales channels (from e-commerce, to local stores, to supermarkets) and the increasing demand for social and environmental responsibility. At the top of the list of CSR issues linked to mass retailing: food waste.

In France, bread is the third most wasted product after fruits and vegetables. Indeed, bread, pastries and cakes are fresh products that have a very limited shelf life. But, they are also items that can generate a

lot of frustration among customers when they are out of stock at the end of the day. This is why Carrefour Group, together with Artefact, sought to use machine learning and data science to optimize the prediction of sales in the bakery-pastry department of its supermarkets.

This project meets a double objective: produce enough to meet the demand, while reducing “scrapage”, i.e. the volume of unsold products.

This use case’s success relies, in particular, on the teams involving both the “jobs” concerned (department managers) and the various technical profiles of Artefact and Carrefour Group.

SOLUTION

A new algorithm to prevent “scrappage” in the bakery-pastry department

In the retail world, every day is different. Sales are highly dependent on the context: holidays, weather, current promotions, merchandising highlights, etc. To take into account all the variables that have implications on demand, we need to be able to analyze the petabytes of data generated by the billions of transactions carried out every year at Carrefour, to which we need to add external data that influences consumption. These calculations are only possible with Artificial Intelligence.

The Carrefour Group and Artefact teams thus started with data from the sales receipts generated by more than 200 supermarkets in France. Every day, this data is collected, cleaned and enriched with external sources – such as calendar data, for example – to build a sales history over several years. This represents thousands of configurations for a single day, depending on the assortment, product prices, promotions, etc.

This data is used to train supervised machine learning models, built on the basis of decision trees, which determine the relationships between the target variable (future sales) for each product and the explanatory variables (promotion, cannibalization, etc.).

Close collaboration with the “field” teams

From an organizational point of view, the project was led by multidisciplinary teams. Two teams on the Artefact and Carrefour sides combined technical and business profiles. The operational skills of Carrefour’s retail professionals played a crucial role. They were able to explain their business, their needs, and bring their vision, in order to guarantee the success and adoption of the solution “in real life”.

For example, the algorithm’s recommendations were first tested in pilot stores, only on pastries, to get feedback from the field teams. Their comments were used to improve the models, before the solution was deployed across all bakery and pastry products in the supermarkets.

Easy-to-access and interoperable tools

All of Carrefour’s data-driven use cases are fed by a centralized platform in the cloud, which makes the data accessible, formatted and documented. The results of the processing are then fed back into Carrefour’s information systems, shared by a wide range of employees. Several teams, made up of data scientists, data engineers, but also data translators (pivotal profiles, acting as a link between business and data), are likely to consult, transform and process them for specific uses.



RESULTS

Concrete results on the number of discarded products...and on the brand image

A few months after the implementation of this new prediction model, the results are very positive.

In fact, over the last five months of 2021, approximately 100 tons of pastries were saved. At the same time, sales have increased due to fewer shortages at the end of the day.

Finally, the Net Promoter Score, a performance indicator closely monitored by Carrefour, has evolved very positively.

Carrefour Group is multiplying AI use cases to improve customer experience

For the group, future experiments follow the same model: responding to business needs, working jointly with operational teams, to feed the customer experience. This acceleration of Carrefour's digital transformation was made possible by the creation of complete and expert data teams within the company, and the deployment of data platforms in all countries where the group operates.

The volume and wealth of data collected by Carrefour provides a unique opportunity to explore the major challenges facing the retail sector: omnichannel, e-commerce, anticipation of consumer habits, etc. Carrefour recently unveiled other examples of how data can be used to improve the customer experience: five-minute shopping on Carrefour.fr, the implementation of personalized assortments for local stores, and the personalization of promotions.



In fact, over the last five months of 2021, approximately 100 tons of pastries were saved. At the same time, sales have increased due to fewer shortages at the end of the day.

Demand forecasting: Using machine learning to predict retail sales



Jérôme Petit
Managing Partner
Retail & eCommerce
ARTEFACT FRANCE



Pascal Coggia
CEO
ARTEFACT UK

All industries aim to manufacture just the right number of products at the right time, but for retailers this issue is particularly critical as they also need to manage perishable inventory efficiently. Too many items and too few items are both scenarios that are bad for business. (Estimates suggest that poor inventory management costs US retailers close to two billion dollars per year.)

Looking beyond past sales to accurately predict future sales

Massive incremental profit can be unlocked by retailers managing orders and inventory effectively. But as this requires the processing of data for a huge number of stock keeping units (SKUs), which often include perishable goods and items that are ordered daily, it is also a significant challenge.

Retailers used to rely solely on the data from previous years to predict future sales (and therefore manage their inventory), but this method is only useful up to a point. However, machine learning has now evolved to the stage that it can provide accurate predictive models using different signals based on how they influence purchases.

Predicting sales is complex because, in any given period, purchases are affected by many factors: weather, shopping trends, regulation, new products, buying behaviours, a pandemic... And predictions based on previously recorded data don't factor in specific events, making monthly sales appear evenly distributed when this is unlikely to be the case.

For example, an item that is often out of stock might cause a slowdown in the sales of that particular product or category, but it won't show in the monthly reports. Even worse, poor figures are often regarded as a mark of buyers' disinterest, when the opposite is true; consumers' over-purchase of an item has caused it to sell out.



Or a product missing from the store might actually be in stock – just not yet out on the shelves. Big box retailers often struggle to restock in real time, so an instantly popular item might disappear from the shelves very quickly, and thus not perform as well as expected, despite it being available in inventory. This calls for technology that can help retailers seamlessly align supply and demand.

Using machine learning and multiple signals to assess inventory levels

Machine learning provides a solution to these challenges. Predictive models can forecast sales months in advance by using a number of the signals that affect them (seasonality, consumption trends, price levels, etc). To be as accurate as possible, it's important that the models use more indicators than the standard day, product and store that it is usual to factor in.

To illustrate this, a retailer might analyse the seasonality to predict sales for the forthcoming period. However, the data will be skewed because using dates is not 100% accurate; a certain date can be a weekday one year, but the weekend the following year, causing sales to vary greatly. Other factors, such as whether that date falls on a holiday (Christmas, Easter, etc) or a major sporting event, also influence consumer buying patterns.

It is a similar story with price level signals. Promotions at store level can markedly affect the sales of a product from a given category or even make the store as a whole more attractive.

Both these examples illustrate why it is necessary to take many different signals and indicators into account to accurately forecast sales: a task that was a pain in the neck before machine learning and advanced artificial intelligence models made it achievable.



Adopting machine learning for inventory management

The technology is there, but for retailers to use it effectively and make accurate predictions, they need to collect and analyse huge amounts of data. Much of this is in different data sources and it can be complex to try to process multiple Excel and PDF files that contain previous reports and media plans. Big data tools are needed to process this information into the clean and readable format required to create predictive models that can prevent inventory issues.

Past sales data for a given store may be 'inaccurate' due to one-off events (promotions, adverse weather, traffic congestion, etc). To remove this bias, predictive models combine past sales numbers with those of similar stores.

The other big challenge is preventing items being unavailable on the shelves while they are in stock (caused by it being almost impossible for employees to monitor shelves in real time and restock them immediately).

Technology solutions using surveillance cameras and weight sensors do exist but are a huge investment. However, readily available

information such as real-time sales at SKU level may be leveraged to detect 'empty shelf' situations. Models can analyse the usual flow of sales of an item, so the normal time between two sales of a product in a given store is known. Human intervention can be used to review and resolve statistical anomalies.

Predictive analysis is but one of the many ways that traditional retailers can benefit from machine learning. They have a lot to gain from relying on advanced technology for better inventory management to increase store revenue. Processing vast amounts of data can also help them to optimise the assortment, offer more attractive and profitable promotions and set prices more efficiently.

Well-devised tools can undertake complex and time-consuming tasks and quickly deliver accurate reports. This is the real value creation lever of artificial intelligence in retail: freeing managers from tedious multi-sources comparisons analyses and allowing them to focus on the continued improvement of the customer experience.

Sales forecasting in retail: what we learned from the M5 competition.



Maxime Lutel
Senior Data Scientist
ARTEFACT

In this article, Data Scientist Maxime Lutel sums up his learnings from the M5 sales forecasting competition, which consisted in predicting future sales in several Walmart stores. He will walk you through our solution and discuss what machine learning model worked the best for this task.

Using machine learning to solve retailers' business challenges

Accurate sales forecasting is critical for retail companies to produce the required quantity at the right time. But even if avoiding waste and shortage is one of their main concerns, retailers still have a lot of room for improvement. At least, that's what people working at Walmart think, as they launched an open data science challenge in March 2020 – the M5 competition – to see how they could enhance their forecasting models.

The competition aimed at predicting future sales at the product level, based on historical data. More than 5000 teams of data lovers and forecasting experts have discussed for months about the methods, features and models that would work best to address this well-known machine learning problem. These debates highlighted some recurring issues encountered in almost all forecasting projects. And even more importantly, they brought out a wide variety of approaches to tackle them.

This article aspires to summarize some key insights that emerged from the challenge. At Artefact, we believe in learning by doing, so we decided to take a shot and code our own solution to illustrate it. Now let's go through the whole forecasting pipeline and stop along the way to understand what worked and what failed.



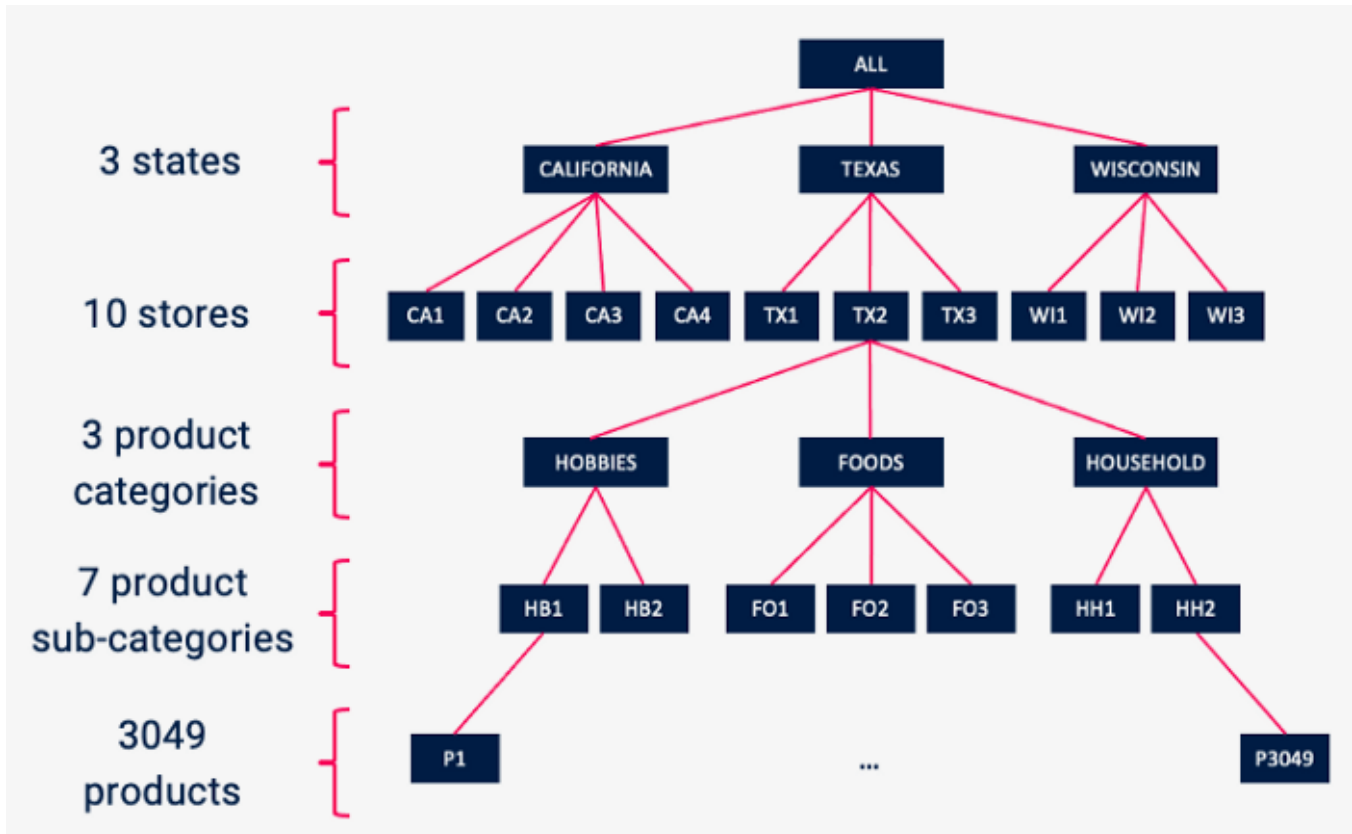


Image 1

Problem statement: Hierarchical times series forecasting

The dataset contains 5-year historical sales, from 2011 to 2016, for various products and stores. Some additional information is provided, such as sell prices and calendar events. Data is hierarchically organized: stores are divided into 3 states, and products are grouped by categories and sub-categories. (cf Image 1)

Our task is to predict sales for all products in each store, on the days right after the available dataset. It means that 30 490 forecasts have to be made for each day in the prediction horizon

This hierarchy will guide our modeling choices, because interactions within product categories or stores contain very useful information for prediction

purposes. Indeed, items in the same stores and categories might have similar sales evolution, or on the contrary they could cannibalize each other. Therefore, we are going to describe each sample by features that capture these interactions, and prioritize machine learning based approaches over traditional forecasting ones, to consider this information when training the model.

Two main challenges: intermittent values and an extended prediction horizon

At this stage, you might think that it is a really common forecasting problem. You're right and that's why it is interesting: it can relate to a wide range of other projects, even if each industry has its own characteristics. However, this challenge has 2 important specificities that will make the task more difficult than expected.

The first one is that the time series we are working with have a lot of intermittent values, i.e. long periods of consecutive days with no sales, as illustrated on the plot below. This could be due to stock-outs or limited shelves' area in stores. In any case, this complicates the task, since the error will skyrocket if sales are predicted at a regular level while the product is out of shelves. (cf image 2)

The second one comes from the task itself, and more precisely from the size of the prediction horizon. Competitors are required to generate forecasts not only for the next week, but for a 4-week period. Would you rather rely on the weather forecast for the next day or for 1 month from now? The same goes for sales forecasting: an extended prediction horizon makes the problem more complex as uncertainty increases with time.

Feature engineering – Modeling sales' driving factors

Now that we have understood the task at hand, we can start to compute features modeling all phenomena that might affect sales evolution. The objective here is to describe each triplet Day x Product x Store by a set of indicators that capture the effects of factors such as seasonality, trends or pricing.

SEASONALITY

Rather than using the sales date directly as a predictor, it is usually more relevant to decompose it into several features to characterize seasonality: year, month, week number, day of the week... The latter is particularly insightful because the problem has a strong weekly periodicity: sales volumes are bigger on the weekends, when people spend more time in supermarkets.

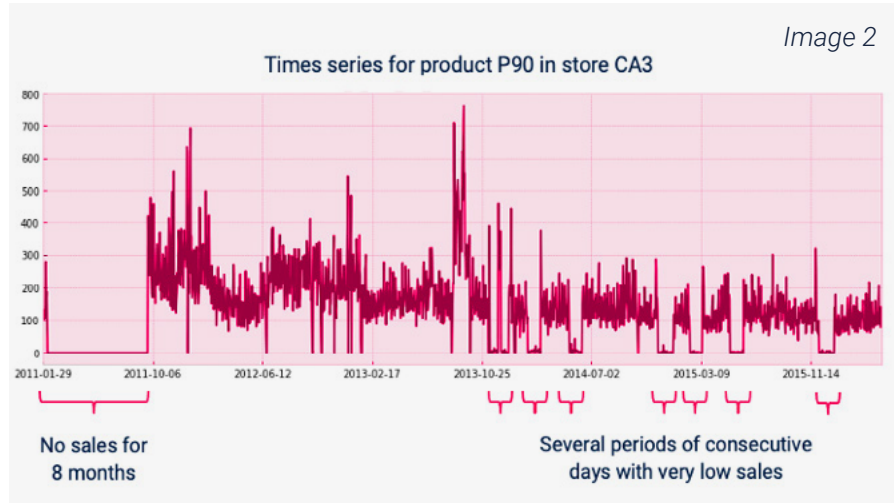
Calendar events such as holidays or NBA finals also have a strong seasonal impact. One feature has been created for each event, with the following values:

- Negative values for the 15 days before the event (-15 to -1)
- 0 on the D-day
- Positive values for the 15 days following the event (1 to 15)
- No value on periods more than 15 days away from the event

The idea is to model the seasonal impact not only on the D-day, but also before and after. For example, a product that will be offered a lot as a Christmas present will experience a sales peak on the days before and a drop right after.

TRENDS

Recent trends also provide useful information on future sales and are modeled thanks to lag features. A lag is the value of the target variable shifted by a certain period. For any specific item in a given store, the 1-week lag value would be the sales made one week ago for this particular



item and store. Different shift values can be considered, and the average of several lags is computed as well, to get more robust predictors. Lags can also be calculated on aggregated sales to capture more global trends, for example at the store level or at the product category level.

PRICING

A product's price can change from one store to another, and even from one week to another within the same store. These variations strongly influence sales and should therefore be described by some features. Rather than absolute prices, relative price differences between relevant products are more likely to explain sales evolutions. That's why the following predictors have been computed:

- Relative difference between the current price of an item and its historical average price, to highlight promotional offers' impact.
- Price relative difference with the same item sold in other stores, to understand whether or not the store has an attractive price.
- Price relative difference with other items sold in the same store and same product category, to capture some cannibalization effects.

Categorical variables encoding

Categorical variables such as the state, the store, the product name or its category also hold a significant predictive power. This information has to be encoded into features to help the model leveraging the dataset hierarchy. One-hot encoding is not an option here because some of these categorical variables have a very high cardinality (3049 distinct products). Instead, we have used an ordered target encoding, which means that each observation is encoded by the average sales of past observations having the same categorical value. The dataset is ordered by time for this task to avoid data leakage.

All categorical variables and some of their combinations have been encoded with this method. This results in very informative features, the best one being the encoding of product and store combination. If you wish to experiment other encoders, you can find a wide range of methods here.

Tweedie loss to handle intermittent values

Different possible strategies can be used to deal with the intermittent values issue. Some participants decided to create 2 separate models:

one to predict whether or not the product will be available on a specific day, and a second one to forecast sales. Like many others, we have chosen another option, which is to rely on an objective function adapted to the problem: the tweedie loss.

Without going into the mathematical details, let's try to understand why this loss function is appropriate for our problem, by comparing sales distribution in the training data and the tweedie distribution (cf image 3).

They look quite similar and both have values concentrated around 0. Setting the tweedie loss as an objective function will basically force the model to maximize the likelihood of that distribution and thus predict the right amount of 0s. Besides, this loss function comes with a parameter – whose values are ranging from 1 to 2 – that can be tuned to fit the distribution of the problem at hand (cf image 4)

Based on our dataset distribution, we can expect the optimal value to be between 1 and 1.5, but to be more precise we will tune that parameter later with cross-validation. This objective function is also available for other gradient boosting models such as XGBoost or CatBoost, so it's definitely worth trying if you're dealing with intermittent values.

How to forecast 28 days in advance?: Making the most out of lag features

As explained above, lag features are sales shifted by a given period of time. Thus, their values depend on where you stand in the forecasting horizon.

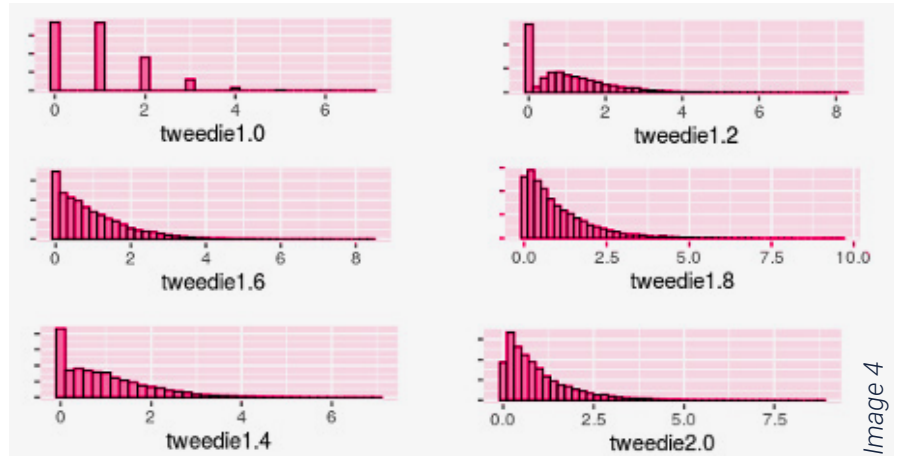


Image 4

The sales made on a particular day D can be considered as a 1-day lag if you're predicting one day ahead, or as a 28-day lag if you're predicting 28 days ahead. The following diagram illustrates this point (cf image 5)

This concept is important to understand what features will be available at prediction time. Here, we are on day D and we would like to forecast sales for the next 28 days. If we want to use the same model – and thus the same features – to make predictions for the whole forecasting horizon, we can only use lags that are available to predict all days between D+1 and D+28. This means that if we use the 1-day lag feature to train the model, that variable will also have to be filled for predictions at D+2, D+3, ... and D+28, whereas it refers to dates in the future.

Still, lags are probably the features with the biggest predictive power, so it's important to find a way to make the most out of this information. We have considered 3 options to get

around this problem, let's see how they performed.

OPTION 1: ONE MODEL FOR ALL WEEKS

The first option is the most obvious one. It consists in using the same model to make predictions for all weeks in the forecasting horizon. As we just explained, it comes with a huge constraint: only features available for predicting at D+28 can be used. Therefore, we have to get rid of all the information given by the 27 most recent lags. It is a shame as the most recent lags are also the most informative ones, so we have considered another option.

OPTION 2: WEEKLY MODELS

This alternative consists in training a different LightGBM model for each week. On the diagram above, every model is learning from the most recent possible lags with respect to the constraint imposed by its prediction horizon. Following the same logic as the previous option, it means that each model can leverage all lags except those that are newer than the farthest day to predict.

This method allows us to better capitalize on lag information for the first 3 weeks and thus improved our solution's forecast accuracy. It was worth it because it was a Kaggle competition, but for an industrialized

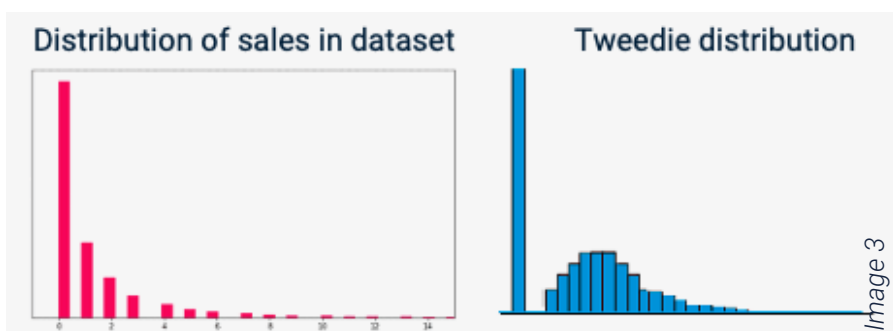


Image 3

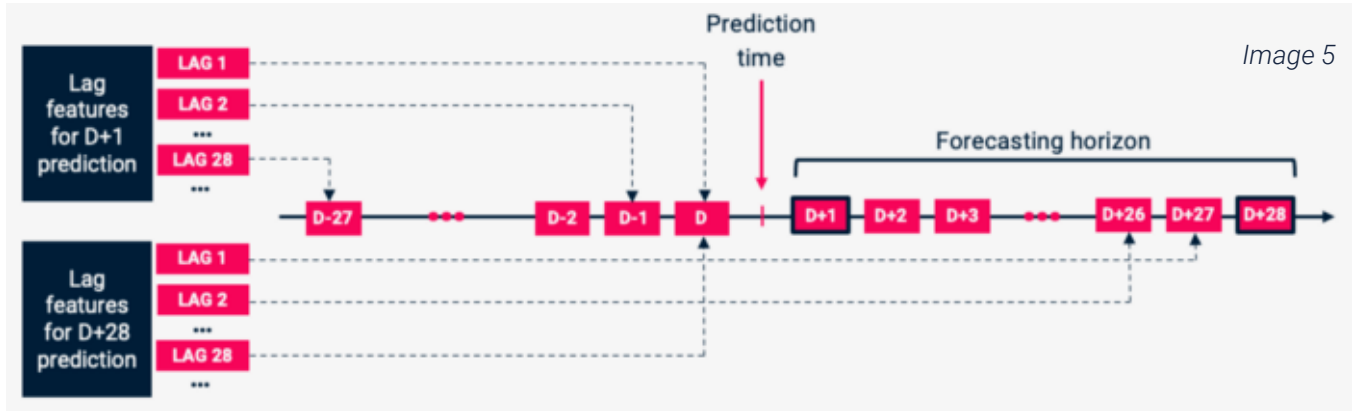


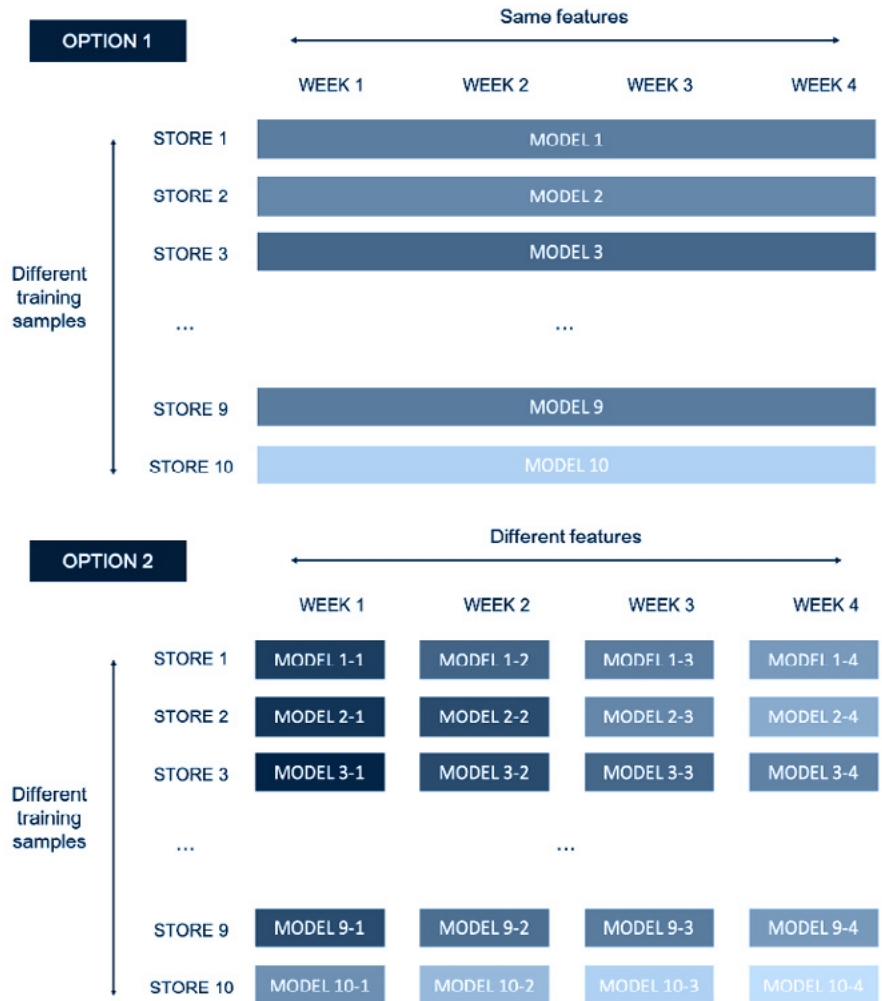
Image 5

project, questions of complexity, maintenance and interpretability should also come into consideration. Indeed, this option could be computationally expensive and if we are aiming at a rollout on a whole country scale, it would require maintaining hundreds of models in live. In that case, it would be necessary to evaluate if the performance increment is large enough to justify this more complex implementation.

OPTION 3: RECURSIVE MODELING

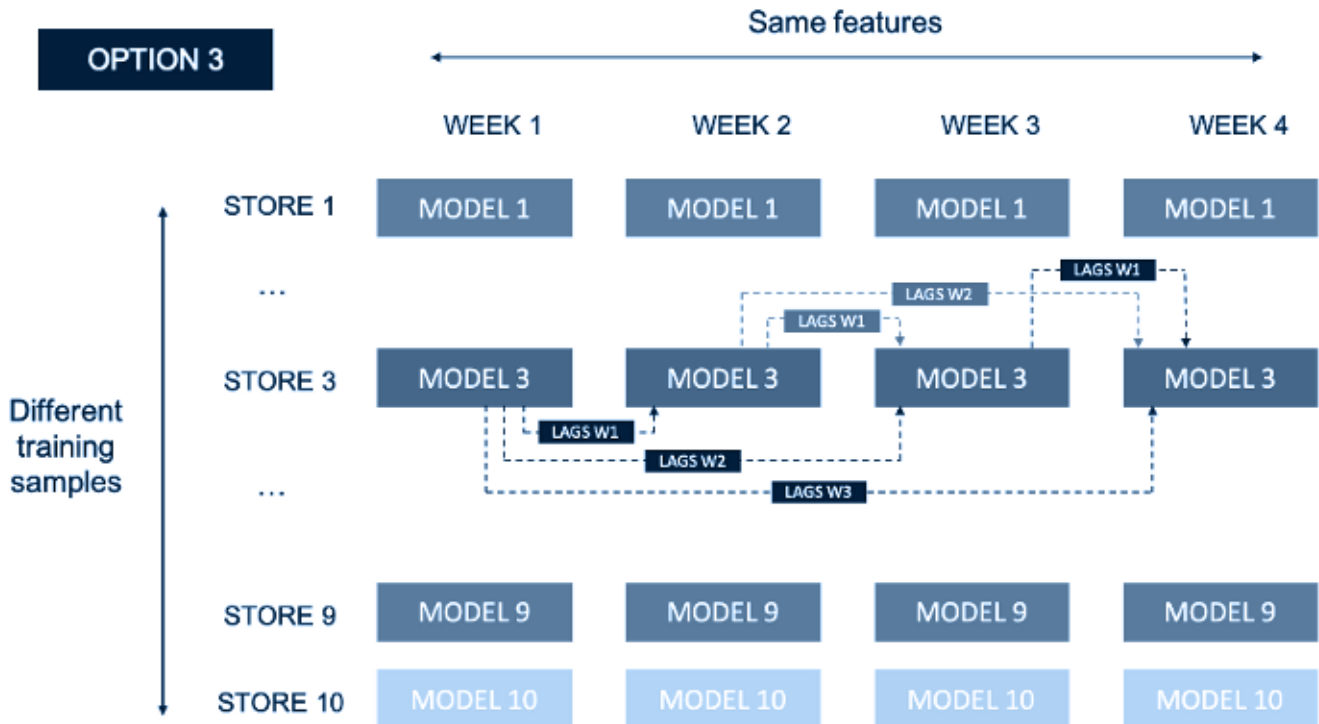
The last option also uses weekly models, but this time in a recursive way. Recursive modeling means that predictions generated for a given week will be used as lag features for the following weeks. This happens sequentially: we first make forecasts for the first week by using all lags except the 6 most recent ones. Then, we predict week 2 by using our previous predictions as 1-week lags, instead of excluding more lags like in option 2. By repeating the same process, we always get recent lags available, even for weeks 3 and 4, which allows us to leverage this information to train the models.

This method is worth testing, but keep in mind that it is quite unstable as errors spread from week to week. If the first week model makes important errors, these errors will be taken as



MORE PRECISELY:

- Model 1 makes forecasts for days 1–7, relying on all lags except the 6 most recent ones.
- Model 2 makes forecasts for days 8–14, relying on all lags except the 13 most recent ones.
- Model 3 makes forecasts for days 15–21, relying on all lags except the 20 most recent ones.
- Model 4 makes forecasts for days 22–28, relying on all lags except the 27 most recent ones just like in option 1.



the truth by the next model, which will then inevitably be poorly performing, and so on. That's why we decided to stick with option 2, that seems to be more reliable.

Ensuring model robustness with an appropriate cross-validation: Why cross-validation is critical for time series

In any machine learning project, adopting an appropriate cross-validation strategy is critical to simulate correctly out-of-sample accuracy, select hyper-parameters thoroughly and avoid over-fitting. When it comes to forecasting, this has to be done carefully because there is a temporal dependency between observations that must be preserved. In other words, we want to prevent the model from looking into the future when we train it.

The validation period during which the model is tested also has a greater importance when dealing with time series. Model performance and the

optimal set of hyper-parameters can vary a lot depending on the period over which the model is trained and tested. Therefore, our objective is to find which parameters are the most likely to maximize performance not over a random period, but over the period that we want to forecast, i.e. the next 4 weeks.

Adapting the validation process to the problem at hand

To achieve that goal, we have selected 5 validation sets that were relevant to the prediction period. The diagram below shows how they are distributed over time. For each cross-validation fold, the model is trained with various combinations of parameters on the training set and evaluated on the validation set using the root mean squared error. (cf Image 9)

Folds 1, 2 and 3 aim at identifying parameters that would have maximized performance over recent periods, basically over the last 3 months. The problem is that

these 3 months might have different specificities than the upcoming period that we are willing to forecast. For example, let's imagine that stores launched a huge promotional season over the last few months, and that it just stopped today.

These promotions would probably impact the model's behavior, but it would be risky to rely only on these recent periods to tune it because this is not representative of what is going to happen next.

To mitigate this risk, we have also included folds 4 and 5, which correspond to the forecasting period respectively shifted by 1 and 2 years. These periods are likely to be similar because the problem has a strong yearly seasonality, which is often true in retail. In case we had a different periodicity, we could choose any cross-validation strategy that has more business sense. In the end, we have selected the hyper-parameters' combination with the lowest error over the 5 folds to train the final model.



Image 9

STEPS	INCREMENTAL FORECAST ACCURACY
Feature Engineering (seasonality, trends, pricing)	+ 3 pts
One model per store	+ 0.3 pts
Tweedie Loss	+ 0.5 pts
Weekly modeling	+ 0.7 pts
Cross-validated hyper-parameters optimization	+ 1 pts

Results

The different techniques mentioned above allowed us to reach a 0.59 weighted RMSSE – the metric used on Kaggle – which is equivalent to a weighted forecast accuracy of 82.8%. The chart below sums up the incremental performance generated by each step.

These figures are indicative: the incremental accuracy also depends on the order in which each step is implemented.

Key takeaways

We have learned a lot from this challenge thanks to participants' shared insights and we hope it gave you food for thoughts as well. Here are our key takeaways:

- Work on a small but representative subset of data to iterate quickly.
- Be super careful about data leakage in the feature engineering process: make sure that all the features you compute will be available at prediction time.

- Select a model architecture that allows you to leverage lags as much as possible, but also keep in mind complexity considerations if you're willing to go to production.
- Set-up a cross-validation strategy adapted to your business problem to evaluate correctly your experiments' performance.

DATA READINESS

- ^ Data Strategy & Organization
- ^ Data Governance
- ^ Data Platform
- ^ Data for Sustainability
- ^ Data & AI Adoption
- ^ The Artefact School of Data

AI ACCELERATION

- ^ AI Strategy
- ^ AI & Analytics Factory
- ^ Generative AI Factory
- ^ AI for Operations
- ^ AI for Customer Care
- ^ A Supply Chain 2.0

DATA & DIGITAL MARKETING

- ^ CX & Creative Studio
- ^ Customer 360 & CDP
- ^ MROI & Measurement
- ^ Digital Analytics
- ^ Media Buying & Certifications
- ^ Retail Media

TECHNOLOGIES

- ^ The Artefact Research Center
- ^ Skaff platform: Tech Accelerators
- ^ AI Partners
- ^ Cloud Certifications
- ^ Data & AI technologies (Fact)
- ^ MLOps



ARTEFACT

AI IS ABOUT PEOPLE

CONTACT

hello@artefact.com
artefact.com/contact-us

ARTEFACT HEADQUARTERS

19, rue Richer
75009 – Paris
France

artefact.com

