ARTEFACT

DATA FOR HEALTHCARE

ARTEFACT

We transform data into value and business impact.

16 COUNTRIES

+1000 EMPLOYEES

+300 MAJOR BRANDS

Artefact is a global data-driven services company. Our offers sit at the intersection of consulting, marketing and data science, putting consumers at the heart of enterprises' digital transformation.

DATA CONSULTING | DATA & DIGITAL MARKETING | DIGITAL COMMERCE

TABLE OF CONTENTS

Data for Healthcare

Interview of Vincent Luciani, co-founder & CEO of Artefact

The data consulting specialist is gaining momentum and already has 1,000 employees in 15 countries, only 7 years after its creation.
In a highly fragmented market, Artefact, a consulting firm specializing in data, stands out by offering a particularly popular service around data and digital: consulting, development and integration of AI solutions applied to business, operational support for client teams, even training.

What are the services and solutions offered by Artefact?

Vincent Luciani: Artefact offers the full range of services needed to help a company exploit the potential of their data and create business impact.

Artefact was founded in 2014 by three partners, one with consulting expertise, one with digital marketing expertise and one with AI research expertise, with one goal: to bridge the gap between data and business. Artefact helps companies rethink their organization around the use of data through three types of services:

- **The broadest and most advanced data-driven marketing offering** on the market, based on our historical core business. The objective: to optimize marketing performance through the use of data. Because the digitalization of marketing took place long before that of other departments, we have taken a considerable lead in the development of personalization and performance measurement solutions. This has given us a unique experience in combining data science with digital marketing.
- **Offers around «data readiness»:** organization, data governance, creation of a Data Factory...which aim to transform the company in the long term
- **Business solutions based on advanced AI models,** semantic analysis or image recognition, for example, in order to predict sales volumes or to automate the processing of call center requests. The field of possibilities is very vast, and we study our clients' challenges carefully in order to respond to them in a specific way.

How can AI be used to improve performance?

V. L.: We have turned a corner in the last 2-3 years. Data, coupled with automation or AI models, has successfully demonstrated its profitability in all sectors. It has an important role to play in creating a competitive advantage by attracting more new clients or lowering costs. For example, Artefact has deployed a demand forecasting model with Carrefour for the Bakery and Pastry department, which gives a sales volume to be forecasted for the day, for each hypermarket, from multiple parameters (day, weather, school calendar...), allowing the department manager to adjust their daily production. This solution resulted in a significant reduction in food waste and an increase in client satisfaction as they get fresh bread all day long! Another significant example for L'Oréal is the use of Machine Learning algorithms analyzing millions of documents and images of influencers, in order to identify the weak signals of emerging trends in the field of beauty and cosmetics. Here again, AI proves its ability to be a real vector of product innovation anticipating consumer needs!

What is Artefact's added value?

V. L.: Artefact is Art...and Fact! It is very rare to gather all the expertise needed for these very complex projects under the same roof. Our clients are often forced to multiply the number of service providers: strategy consultants, developers, data scientists, product owners, etc...creating delays and risks of inconsistency. Artefact brings together three main types of professions in a single entity: consultants with business expertise, data experts (data scientists, data analysts, data engineers), and digital marketing experts, who are essential for mastering highly evolving media formats. Another asset: our vast expertise allows us to assemble and create standardized technological bricks tested in many different business contexts, which we make available free of charge to our clients and make available in open-source.

What types of organizations rely on Artefact?

V. L.: All functions and sectors of activity are affected. Our clients are mainly large international organizations that have challenges with organizational silos (data is present in several places in the company), and industrialization and scalability. Our clients include renowned international brands such as Danone, Samsung, Orange, Heineken and Unilever.

Any upcoming project?

V. L.: Our goal is to quadruple the size of the company in 4 years, both through organic growth (we are recruiting 500 people by 2022) and through strategic acquisitions. Hence, to gain flexibility, our recent exit from Euronext Paris and the arrival of two of the best LBO funds (Ardian and Cathay Capital) to help us accelerate our growth. We also want to support the structuring of careers around data: we have created the Artefact School of Data to address the huge shortage of talent in our business, by offering a professional opportunity to anyone wishing to transition to the data industry.

"Artefact bridges data to business to create value."

Guido Merighi Buitoni, *Global Head of Business Data & Analytics* - Sanofi,
Vincent Luciani *Co-founder & CEO* Artefact

How AI is improving the patient care journey

AI has the potential to transform the healthcare industry with digital solutions that streamline the way healthcare practitioners approach diagnosis and interact with patients. Vincent Luciani, CEO of Artefact, and Guido Merighi Buitoni, Global Head of Business Data & Analytics at Sanofi, talk about ways healthcare providers can use AI to improve the patient care journey – while handling data responsibly and securely.

Over the last few years, AI applications have appeared across the entire healthcare ecosystem, streamlining processes and improving patient outcomes. Advanced chatbots help emergency responders identify a heart attack in progress. AI-powered surgical robots perform minimally invasive procedures. And AI-based software platforms automate the healthcare industry's most repetitive tasks, saving precious time for busy administrators.

The exploding AI healthcare market (which includes software, hardware & services, algorithms, applications and end-users) is expected to reach $44.5 billion by 2026, with a projected average annual growth rate of over 46%.*

At Artefact, we've identified three key areas where AI can significantly improve the patient care journey: self-diagnosis, drug development and monitoring, and personalised health.

* MarketWatch report, 3 January 2022 "Healthcare Artificial Intelligence (AI) Market 2022 Industry Scenario, Strategies, Growth Factors and Forecast to 2026"

Digital accelerated self-diagnostics

We've all seen the wearable devices AI has given the world, like fitness trackers, smartwatches, and even background noise-diminishing hearing aids. They help keep people in shape, tally calories, and improve general well-being.

But AI technology also offers self-diagnostic tools that can help patients avoid unnecessary trips to the doctor. Virtual nursing assistants can converse with patients empathically about their conditions, offer suggestions, and, if necessary, direct them to the most effective care unit and alert providers in case of problems.

Many pharmaceutical companies are forming partnerships with tech companies to develop new healthcare applications. In 2020, Sanofi teamed up with British company Babylon to use their online symptom checker to assist patients with digestive health conditions.

"The tool was able to give a Level 1 diagnostic, offering people fast recommendations and next steps on IBS (Irritable Bowel Syndrome), a condition that's often undiagnosed, under-addressed and improperly treated."

Guido Merighi Buitoni,
Global Head of Business Data & Analytics
Sanofi

Drug safety monitoring; drug discovery & development

AI can be used to monitor drug efficacy and safety and adverse effects. This practice, called pharmacovigilance (PV,) is an integral part of the role of pharmaceutical companies. Using AI-based insights, adverse effects can be analysed to rapidly identify emerging trends. These trends can then help identify which population segments would best benefit from (or avoid) a given treatment and help lead the way to new treatments and cures.

Pharmaceutical leader Bayer is working with its partners to explore the use of AI technology to support decentralised (virtual) clinical trials enabled by telemedicine, home delivery of clinical supplies, e-consent, the use of wearables, and live 24/7 data monitoring.

AI is also showing great promise for accelerating the drug discovery and development process: In April 2021, pharmatech company Exscientia and biotechnology experts Evotec announced a Phase I trial for the first AI-designed drug for the treatment of advanced tumours. The traditional discovery process would have taken up to five years to complete, but thanks to AI, the drug candidate was found in only eight months. Exscientia is now working with other pharmaceutical companies, including Sanofi, to develop oncology and immunology treatments.

Personalised health and wellness

The third area of research under study is how AI can improve human health behaviour, such as helping people improve their diets.

Using data-based diet planning, AI could personalise nutritional programmes and create meal plans for a person's specific metabolism and digestive system. While the first application that springs to mind is weight control, programmes like this could potentially save millions of lives by preventing diabetes, heart disease, and other conditions caused by poor nutrition.

Potential systems could integrate with smart wearables for more easily achievable results and sustainable healthy benefits.

With opportunity, risks and ethical considerations

While AI promises to improve the efficiency of healthcare delivery and quality of patient care, there is a need to minimise the ethical risks of AI implementation, which may include eventual threats to privacy and confidentiality, informed consent, and patient autonomy.

These are vital issues. But there is no single builder of the end product in AI and machine learning. An entire community of people is involved in

"Although AI has enormous potential to advance health care by improving the treatment and prevention of disease and expanding health equity, we must also be clear about its limits: nothing can replace the quality of advice you receive from your practitioner or doctor."

Vincent Luciani, *Co-founder & CEO*
Artefact

the life of an algorithm, from data scientists and others who make eventual improvements to it, to the software company that sells it and end-users. Consequently, there's an issue of accountability.

"While there's no simple solution, transparency is always the best policy. This means always obtaining consent for data use, anonymising personal information, and finding ways to ensure data confidentiality for patients", states Vincent.

"Even medical education would need to be reframed from a focus on knowledge recall to a focus on training future physicians to interact with and manage AI-driven machines", he adds. This reframing would also require diligent attention to the ethical and clinical complexities that might arise between patients, caregivers, and machines.

The critical role of data governance

While healthcare opens up a world of potential new uses for data and AI, there are also impediments to its success: gender bias during the analysis of use cases is one. The symptoms of a heart attack in a male are very different from those in a female, and unless those differences have been taken into account, any analysis will be flawed. Biases around race, ethnicity, or religion are also possible. To avoid them, the concept of "inclusive AI" should be a principle when developing any AI application.

Data governance is of paramount importance: guidelines must be established and best practices defined and consistently followed. In 2020, AstraZeneca engaged experts from inside and outside their company to assist them in developing a Data Governance Framework based on the principles of fairness, accountability, privacy & security, explainability & transparency. The Data Governance Framework will be part of their global Code of Ethics.

Other concerned companies around the world are undertaking new efforts to ensure fair and ethical use of AI. On 13 January 2022, Microsoft and leading public, private, educational and research organisations across the U.S. healthcare and life sciences industries announced the formation of the Artificial Intelligence Industry Innovation Coalition (AI3C) with the goal of maximising technology to provide recommendations, tooling and best practices for AI in healthcare.

CASE STUDY

SANOFI CONSUMER HEALTH CARE (CHC)

Driving Digital Transformation With Precision Marketing.

CHALLENGES

Sanofi is one of the world leaders in the pharmaceutical industry.

Its Sanofi CHC (Consumer Health Care) branch markets over-the-counter medications that respond to what people want: improving the quality of everyday life, at every stage of life. Until recently, the sector had been spared by digital transformation, mostly due to regulatory complexity.

Following the recruitment of a Global Director of Digital Transformation in 2017, the company identified a problem linked to their media mix. It was essentially focused on television, which prevented them from finding their growth points and effectively reaching their targets.

The media budget was under the umbrella of a major media group, but was operationally managed by numerous agencies; this generated massive process complexity and opacity. Rather than rely on media agencies to take on digital, Sanofi wanted the opposite: internal experts capable of challenging their agencies to seek out value through innovation. The client also wanted to abandon vertical logic in order to revitalise exchanges between global and local entities, while legitimising global expertise.

- **Business objective:** Modify the global media mix to increase impact and grow revenue while improving media efficiency
- **Organisational objective**: Integrate essential digital expertise into internal marketing teams to pilot the company's digital transformation

SOLUTIONS

A four year deployment

Phase 1 — Pilots

In pilot countries, Artefact tested pertinent use cases in programmatic advertising, passing from traditional media planning (based on reach, targeting large audiences) to agile programming (connected to business signals such as everyday life, user, and request spike times).

Sanofi CHC consolidated this experience into a Playbook* : this media strategy bible contains guidelines for creating, configuring and optimising campaigns, etc. This document centralises all important information and forms the framework for upskilling Sanofi CHC teams. This work was carried out hand-in-hand with their media agency on a global level to organise deployment and ensure that the new methods will be sustainable locally.

In 2018, ten pilot markets demonstrated that this new advertising model generated value by lowering costs while increasing sales. These POC were translated by twenty pilot campaigns in one year, supported by the most digitally mature teams to test the use cases. The highly collaborative operation united several functions within each project group. Artefact coordinated all marketing, media and creative agencies in every country, along with local, global, CMI, and legal teams. These teams were trained at the highest levels to become internal champions of data-driven marketing.

Phase 2 — Scaling

Ten countries were added to accomplish this phase. Several new pilots had to be made in the additional regions to adjust to local markets. Support by local Artefact sites in Europe, Asia, and America fostered this acceleration of the system. Data visualisation tools and dashboards to streamline and automate reporting were created to facilitate communication of results to all stakeholders.

From the start of this transformation, an internalisation model had been defined whose objective was to make the advertiser autonomous over time. The Precision Marketing Center of Excellence was launched to develop the technical competences of internal teams and recruit profiles able to sustain the transformation. The teams are responsible for establishing training and upskilling schedules in different markets to transfer knowledge. The CoE also updates the Playbook according to new use cases and orchestrates the ecosystem of direct partnerships with the technological platforms (GAFAM). It guarantees business continuity: the sustained transfer of competences and knowledge.

Key results

- Double the impact of campaigns on sales.
- Up to +27% on measured awareness.
- Up to +19% on consideration.

Phase 3 — Industrialisation and Innovation

Now that the foundations of the new organisation are in place, Sanofi and Artefact are initiating an important innovation and pilot phase of process and new driver automation, aimed at increasing customer centricity. Sanofi CHC wants to be positioned as a key player in health and well-being via connected services – a goal which requires new digital transformation considerations in which AI comes into play.

The company also wants to develop their e-commerce, which wasn't a core part of their model until now, to improve their presence in markets such as Germany or China.

2020 is also the year when Precision Marketing is being launched in other business units, such as Sanofi Pasteur, with high stakes on vaccines. Accompanying these teams in their transformation represents a major challenge, as it's the first time that a vaccine-producing laboratory is communicating directly to consumers: until today, such communications were published by public institutions and NGOs.

RESULTS

The first result observed is transformation, defined by a key KPI: share of Precision Marketing in the total media mix, which substantially increased thanks to the project as a whole.

On a performance basis, these actions made it possible to double the impact of campaigns on sales and to reinforce awareness, with up to +27% on measured awareness, and up to +19% on consideration.

These results show how Precision Marketing impacts the entire funnel: from brand image to product consideration to purchase. Even the employer brand benefits: internally, thanks to changes in practices, and externally, as the existence of a Center of Excellence attracts sought-after digital profiles.

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

Florian Thiebaut
Partner – Data-Marketing Practice Lead

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.

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.

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.

CASE STUDY

SANOFI CONSUMER HEALTH CARE (CHC)

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	
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.

AI Requires a Holistic Framework and Scalable Projects

Ghadi Hobeika
CEO Artefact US

Artificial intelligence and digital transformation projects have a low success rate, but best practices help.AI has the potential to drive change in almost every industry. In other words, there are big incentives for organizations to start their AI journey now; there is also the risk that if they don't, playing catch up will be difficult, if not impossible, in a discipline that will become increasingly critical the more widely it is adopted. So, it's no surprise that AI is causing so much interest and excitement. However, a lot of AI projects fail.

Ever since I can remember, artificial intelligence has been the holy grail. Films have portrayed it, from BladeRunner to the more recent Her. In the meantime, business leaders promised it would revolutionize the workplace. In both cases, we've been presented with scenarios in which AI transforms the daily grind.

Indeed, AI has been talked about as a scientific discipline since 1956. And although the math-based fundamentals have existed for more than 70 years, the computing power required has only recently been a reality, with the cloud being the ultimate AI catalyst.

Significant progress has been made — and the sector is no longer in its infancy. According to McKinsey's The state of AI in 2020 survey, 50% of respondents said their companies had adopted AI in at least one business function.

AI has the potential to drive change in almost every industry. In other words, there are big incentives for organizations to start their AI journey now; there is also the risk that if they don't, playing catch up will be difficult, if not impossible, in a discipline that will become increasingly critical the more widely it is adopted. So, it's no surprise that AI is causing so much interest and excitement.

However, a lot of AI projects fail.

POCs Should Be Designed for Long-Term Success

Many proofs of concepts (POCs) are not designed to scale. They do no more than prove that something can be done. Then, they are left to fester because it wasn't determined in advance whether the concept in question was relevant and required by the entire organization, or whether an enterprise-wide roll-out was technically feasible.

Moreover, the cost structure of projects of this nature: Getting to this point is likely to have devoured 70% of the overall budget, without the result ever seeing the light of day. That is bad business on every level. So, what is the alternative?

In short, scale must be an integral part of the POC, and reflected in the metrics that determine if it was successful.

There are some straightforward tactics for achieving this. A good option is running the POC in two regions and requiring both streams to deliver on pre-determined goals before it can be signed off as a success. It is also important to identify parallels and variations between the projects. This approach develops process and structure as part of the initial venture, and it underpins adoption in the wider environment if the project moves ahead.

Skillsets Demolish Silos

Organizational silos, rooted in the traditional business structure, are still commonplace. They are a constant thorn in the side of smooth-running operations, and they can be the death knell for scalable AI implementations. Addressing this means building the right skills into every part of the project.

We need mathematical expertise, IT skills and a coding specialist delivered (respectively) by data scientists, solutions architects, and machine learning (ML) engineers. The business perspective, provided by product managers and owners, is also an essential part of the mix. This multidisciplinary team should have an open and collaborative way of working, with good communication channels and a deep level of trust throughout the lifecycle of the project so they can collectively lay the groundwork, roll out the implementation and, finally, train the people that will run the application on a day-to-day basis once the POC is completed.

Technology Is Important, Too

Cloud computing has made AI projects a reality for many businesses. It does away with the need for big, costly IT implementations, relying instead on agile tools and technologies that are customizable and available on an on-demand basis.

As with the hybrid-team approach, the tech toolbox should comprise the applications and software specific to the project in question. And it goes without saying that it should be scalable.

The AI Risk Paradox

AI presents organizations with a dilemma: Implemented badly, it is likely to fail, creating business risk. However, not implementing AI at all risks falling behind more future-facing competitors as they reap the rewards of exploring this next-generation technology.

The key is to view any AI project in terms of its role in the long-term direction and success of the overall enterprise and its operations. This approach will inform the technical and people-based framework that is essential for successful implementation and a holistic AI vision.

CASE STUDY

THE CARNOT CALYM INSTITUTE

Fighting against lymphoma with AI

Artefact, Microsoft and the Carnot CALYM Institute combined their expertise to build a Lymphoma Data Hub enabling researchers to leverage AI for accelerated early-stage diagnosis and therapeutic innovation.

A technology partnership to improve lymphoma diagnosis and treatment

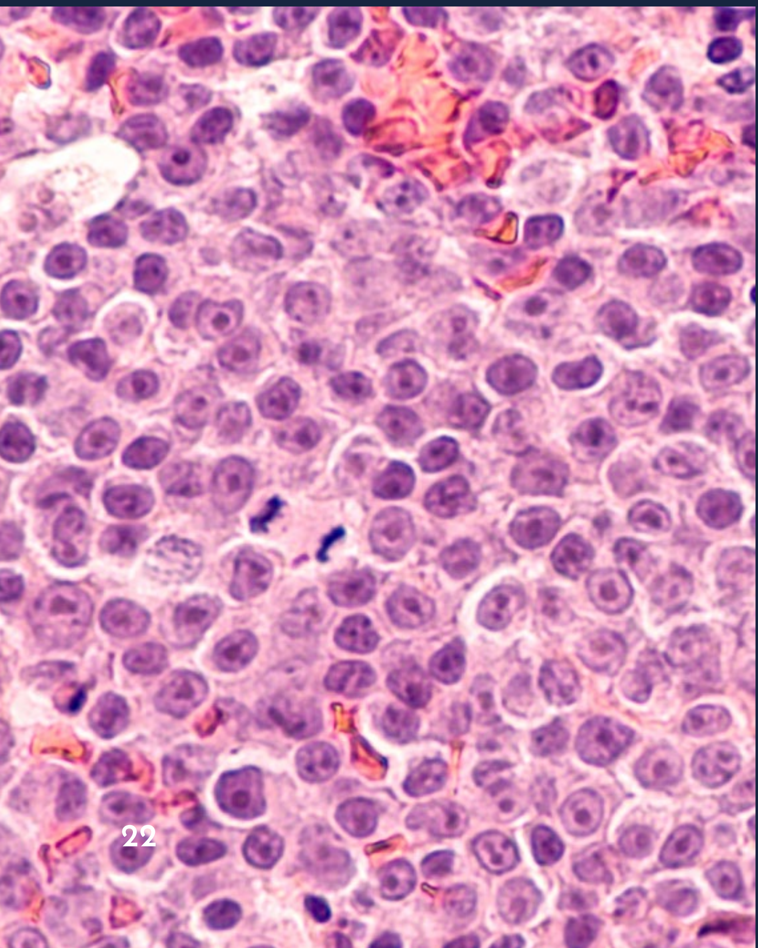
Artefact transforms data and AI into value and impact not only for clients, but for humanitarian causes as well. The Artefact 4 Good programme gives our employees the chance to contribute to meaningful projects. Microsoft's Share AI skills sponsorship programme allows employees to do the same. The Carnot CALYM Institute, an international academic research consortium, constantly seeks new ways to fight against lymphoma. Members from all three organisations joined together to see how AI could be used to identify new avenues of research.

"By breaking down existing silos of data, we open new perspectives in terms of research questions."

Delphine Sondaz, Project Leader - Carnot CALYM Institute

"AI can safely accelerate product development and time to market for new medical treatments which can traditionally take years or even decades."

Vincent Luciani, CEO - Artefact



CHALLENGES

What lymphoma is and how AI can help

Lymphoma is a blood cancer that begins in infection-fighting cells (lymphocytes) of the immune system. It is the most common blood cancer and the 6th most frequent cancer worldwide, with some 80 sub-types. Due to their heterogeneity, lymphomas are difficult to diagnose and require different therapeutic strategies.

Because AI can analyse and integrate massive amounts of data, CALYM believes it can play a major role all along the innovation chain in R&D in Health, from disease characterisation and diagnosis to drug development and personalised therapeutic strategies.

For the Lymphoma Data Hub project, CALYM already had assets to facilitate AI integration:

- A wide network of international experts in research and care;
- Massive quantities of data generated by over 30 years of clinical trials and 25,000 patients;
- The added benefit of annual data generation increases.
- But there were also obstacles:
- Data collected around lymphoma are diverse (clinical, pathology, imaging, omics) and complex; patient data is sensitive and personal.
- Data are sometimes fragmented or of poor quality, in incompatible formats, and governed by disparate and complicated rules.
- Above all, CALYM lacked the tools to ethically and securely exploit their data.

SOLUTION

Development of a collaborative Lymphoma Data Hub

To improve and facilitate the ethical exploitation of massive quantities of health data, CALYM asked Artefact to develop an environment allowing collaboration, data sharing and state of the art data learning in a real-time and secured way, in order to increase the number and quality of research partnerships for the benefit of lymphoma patients.

CALYM had five requirements. Their solution had to be traceable, modular, collaborative even for remote work, agile, and above all, secure.

Our solution was a Lymphoma Data Hub: a platform enabling the collection, processing and sharing of data, functioning like an AI Factory with raw materials, where all data available for research projects is concentrated in a single environment.

The success of this environment is ensured by two essential pillars:

1 — Data governance

Governance guidelines for the data platform must be created to define and maintain roles and responsibilities as well as processes to follow. These include clear and managed data supervision, data security according to current legislation, data monitoring, data quality, and data accessibility.

2 — Architecture

The security, quality and confidentiality of data must be ensured by defining the ideal infrastructure for the collection, treatment, storage and sharing of data. Microsoft's Azure cloud solution was selected for deployment of the Lymphoma Data Hub. Azure is the first cloud provider with HDS* certification in France, and enables clear definition of roles, multiple security controls, performance optimisation, scalability, traceability and data organisation.

This hub for data sharing and algorithm development is also an evolutive catalogue of data available for research collaborations with partners who share our vision. Its use will strengthen academic and partnership research to improve the understanding, diagnosis and treatment of lymphoma.

**Hébergeurs de Données de Santé (Health data hosting)*

RESULTS

Measurable improvements in cancer detection accuracy

Prior to the launch of the Lymphoma Data Hub, a diagnostic pilot was carried out using computer vision to develop a deep-learning algorithm to assist pathologists in diagnosing follicular lymphoma, one of the disease's many subtypes.

Artefact's experiment was successful in detecting the presence of cancer with an accuracy of between 90% and 98%.

"Because patient information is private, its collection, storage and use raise both legal and ethical concerns. The importance of data security cannot be overstated."

Vincent Luciani, CEO -Artefact

Learn more about the LDH: Lymphoma Data Hub

Eight other research projects for the near future have also been identified, ranging from earlier disease detection to enhanced monitoring of its evolution in response to medical care.

Artificial intelligence represents a new opportunity to improve care for patients with lymphoma. As the Lymphoma Data Hub evolves and grows with input from members of the consortium, the disease will be better understood, allowing new avenues for the development of treatments and drugs to be explored; diagnosis will become faster and more accurate, enhancing patient comfort; and treatment responses will be better anticipated, leading to more personalised therapeutic strategies.

«Our strategic vision for the consortium has always been based on innovative research projects [...] resulting from collective reflection and breaking down the barriers between data from different medical specialties."

Emmanuel Gomez, Director of R&D - Carnot CALYM Institute

How did we use computer vision to help medical experts diagnose Follicular Lymphoma?

Yague Thiam
Senior Data Scientist

With the introduction of opt-in permissions for apps, iOS 14 will make it harder for brands to target consumers on an individual level and to measure results of marketing activities. Bobby Gray, Head of Analytics and Data Marketing at Artefact, considers the impact and explains how brands can respond using first-party data.

Introduction

This project is part of Artefact's contribution in Tech for Good. The project has been conducted in collaboration with Institut Carnot CALYM, a consortium dedicated to partnership research on lymphoma, and Microsoft.

In autumn 2019, the Institut Carnot CALYM launched a structuring programme aimed at setting up a roadmap to optimise the valorisation and exploitation of data from the clinical, translational and preclinical research conducted by the members of the consortium for more than 20 years. This project, proposed by Pr Camille Laurent (LYSA, IUCT, CHU Toulouse, France) and Pr Christiane Copie (LYSARC, Pierre-Bénite, France), both members of Institut Carnot CALYM, is part of this structuring programme.

The primary objective of this research project is to develop a deep-learning algorithm to assist pathologists in diagnosing Follicular Lymphoma. A secondary objective is to identify informative criteria that could help medical experts understanding the morphological differences between Follicular Lymphoma and Follicular Hyperplasia which will be referred below as FL and FH.

What is Follicular Lymphoma? What are the challenges in its diagnosis?

FL is a subtype of Lymphoma, the most frequent blood cancer in the world. There are more than 80 types of Lymphoma and this diversity makes its diagnosis difficult, even for experts. Moreover, FL is very similar to FH which is not cancerous, adding challenges to its diagnosis.

In this article, we will describe our approach in building a classifier for FL and FH using only labelled whole-slide images. Whole slide images are high resolution digital files of scanned microscope slides. In our case they contain extract of lymph nodes.

How could deep learning help in its detection?

Using whole-slide images of FL and FH, we trained a binary classifier through a patch-based approach. Our model architecture is a simple Resnet-18 trained on a few epochs (~10).

After predicting the class of an observation with the classifier, we extract the last activation layer to build a heatmap on top of the input image to highlight parts that have prompted the model in defining a given class.

Why did we use a patch-based classification?

Patch-based classification is a classification technique where the class of a given observation is built based on the aggregation of the predictions of its components (patches). In our case it is used because the images are way too large to be used directly on the model.

In fact, whole-slide images are very large (~10^5 pixel square). Their size makes training a deep learning model almost impossible with common tools. To solve this issue, we divided them into patches of the same size following two important criteria:

The patches must be big enough so that the follicles remain visible in them.

The patches should be small enough so that training a model can be done in a reasonable amount of time.

In patch-based classification, the model output can be interpreted as that of a classical classification except that the first layer of computation is at the whole-slide level. For example, when predicting the class of a slide of FL, a score of 98% would mean that 98 % of the patches it is composed of have been predicted to be FL.

At the dataset level, this slide will be predicted with a score of 0.98 for the FL class.

PS: We made the hypothesis of dividing the images into patches based on medical experts' conclusions stating that in a whole-slide of FL, the follicles are expected to be present everywhere.

TRAINING SET

Our training set is composed of 58k randomly selected patches (*1024 pixel square*) of FL and FH extracted from a set of 30 whole-slide images in each of the 2 classes.

VALIDATION SET

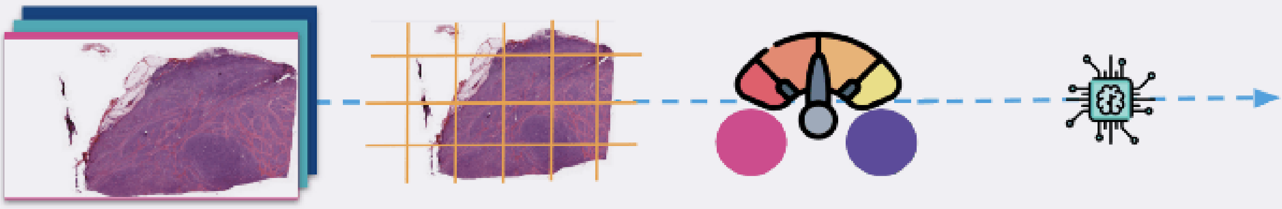
20% of the patches was sampled for validating the model performance at training time.

TESTING SET

Our testing set is composed of 15 whole-slide images, each divided into patches. This reference set has been used to compare the results of different training approaches that we will precise below.

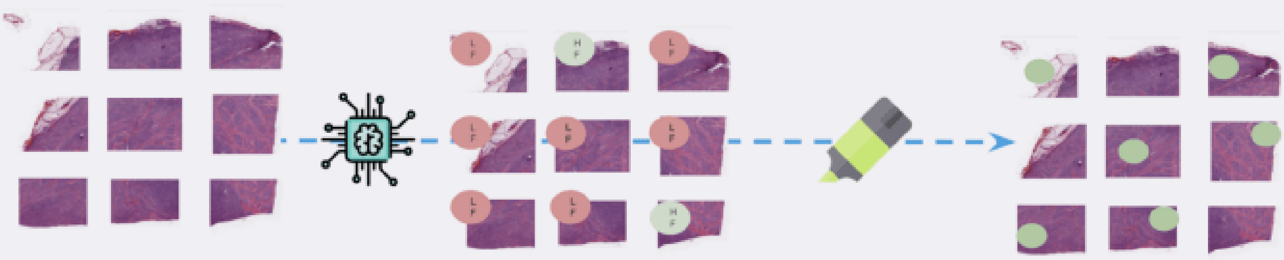
MODELLING

Before training the deep learning classifier: Image preparation and processing



The images are first divided into patches, then normalised before they are fed to the model for training.

After training: Inference and interpretation



At inference time, new whole-slide are divided into patches before the model predicts a class foreach one of them. Parts of images responsible for predicting FL class are highlighted to help monitoringthe results.

Data preparation and processing

1 — TILING

As stated earlier, whole-slide images are very large and cannot directly be fed to a classification model unless you are using a super galactic hardware. We used the library openslide to read the slides and its deepzoom support to divide the images into relatively small tiles of size 1024 pixel square. After breaking them into tiles we ran them into a basic cleaner that dropped all tiles that were not at the center of the tissue (borders, holes etc).

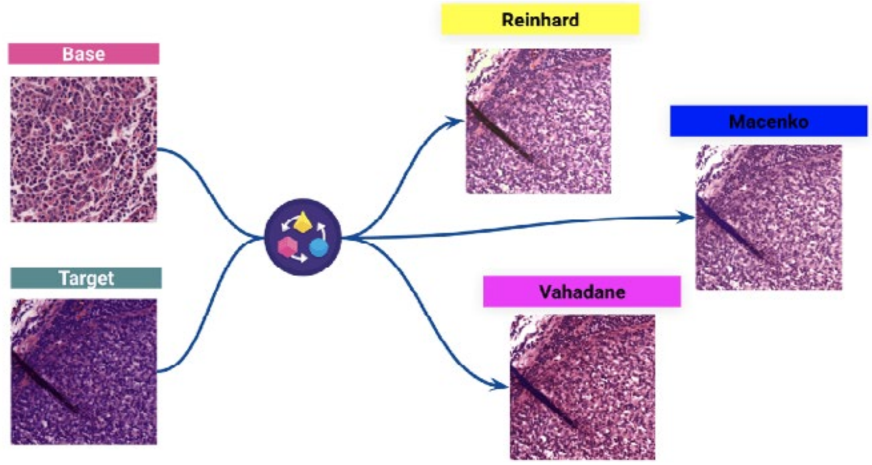
2 — STAIN NORMALISATION

The second step of our data processing, which is also the most important step, is the stain color normalisation. Staining is the process of highlighting important features on slides and enhancing the contrast between them. The staining system used is the common H&E (Hematoxylin and Eosin).

However, since the images are coming from many different laboratories, we have observed variations in the colouring of the slides. They mainly come from differences in the dying process from one laboratory to another. These differences can affect the model's performance a lot.

We used classical techniques to normalise the coloration of the dataset before training the model.

We picked the Reinhard technique to see the impact on the model.



Results of three different stain normalisation : a target image colouring is normalised to a base image colour distribution.

Training a Resnet-18 classifier

After processing the whole-slide images, the training went smoothly (dropout, weight decay, etc..). Nothing fancy except from adding mixup in the data augmentation. We used a Resnet18 trained from scratch since pre-trained models were not significantly improving our results. We also preferred the Resnet-18 since the Resnet-34 and Resnet-56 were not improving our performances. After ~10 epochs, our model was ready for testing.

We used the very practical Fastai library to build our models with few efforts.

TESTING

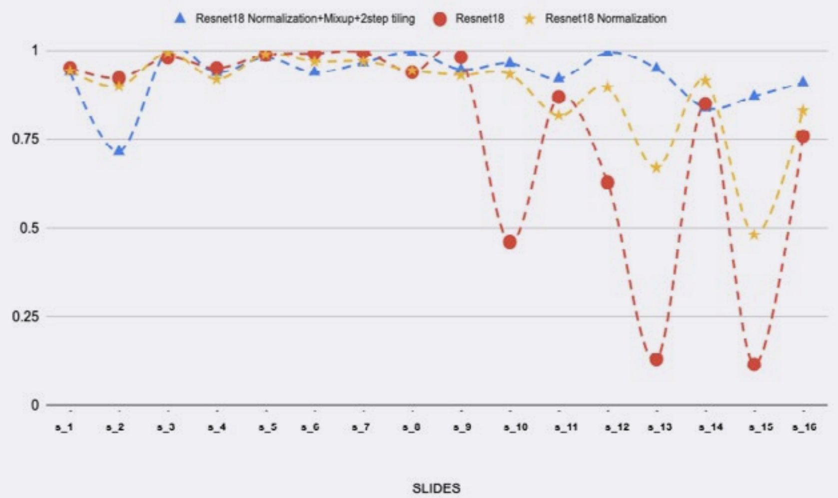
The results of 3 experimentation are worth being mentioned:

A simple resnet-18 as baseline

A resnet-18 + stain normalisation on the dataset

A resnet-18 + stain normalisation on the dataset + mixup as data augmentation

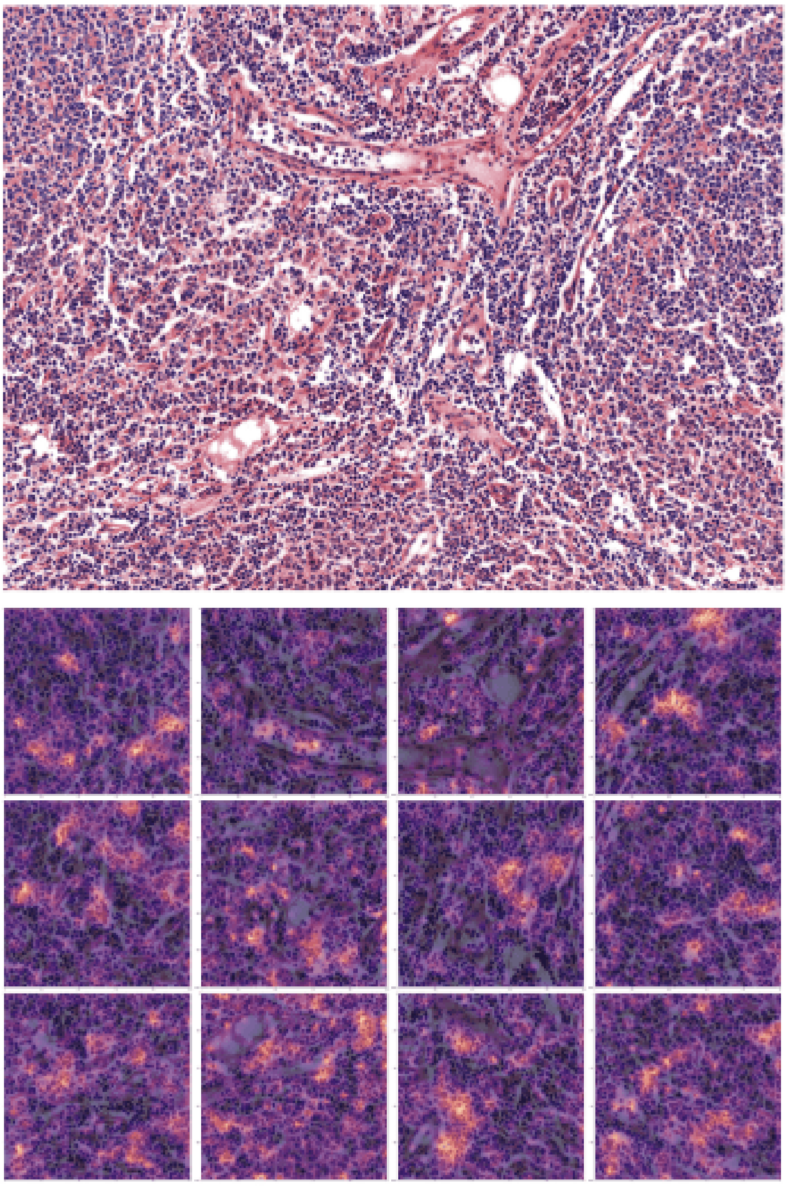
The results on the test set for these 3 experimentations.



The results of 3 different models on the 16 selected slides of Follicular Lymphoma. We can see the effect of stain normalisation and mixup on performance.

Stain normalisation is by far the most important step in our modelling approach. We were experiencing generalisation problems (red line) but it definitely help in solving the issue. Adding mixup and a 2-step tiling makes it even better.

MixUp is a data augmentation technique which consists of creating new observations by linearly interpolating many samples.



Parts of the image that has most contributed to the prediction of the class Follicular Lymphoma are highlighted on the bottom image — 12 patches

Interpreting the results of a computer vision classifier

In order to easily communicate the results to medical experts, we provided images with heatmaps to highlight where the model's focus was when predicting a given label. We did that by extracting the last activation layer of the convolutional network and by linearly extrapolating it on the image we were predicting onto.

Interpreting the model's output with heatmaps has been very useful in adjusting the modelling approach as it gives experts ways to analyse what the model is actually doing. Through our exchanges with experts, we (data scientists) were able to adjust how we to handle better the dataset and make the model more robust (i.e able to adapt to different types of inputs). And also to make sure it serves its purpose. It was in fact how we realised the need to normalise the staining of the images.

Conclusion and Key learnings

The goal of this study was to explore the process of creating a good deep learning base classifier for differentiating Follicular Lymphoma and Follicular Hyperplasia. Our keys learnings are listed below:

- **The high importance of color normalisation when training a model with this type of dataset**
- **Usage of advanced data augmentation technique such as mixup can help increase performances**
- **The tight collaboration with medical-expert to challenge models at each iteration**

Including ethics best practices in your Data Science project from day one

Karim Si Larbi
Senior Data Scientist

The use of machine learning as a means for decision making has now become ubiquitous. Many of the outputs of services we use every day are the result of a decision made by machine learning. As a consequence, we are seeing a gradual reduction in human intervention in areas that affect every aspect of our daily life and where any failure in the algorithmic model's judgment could have adverse implications. It is therefore essential to set proper guidelines to build trustworthy and responsible machine learning solutions, taking into consideration ethics as a core pillar.

In recent years, ethics in machine learning has seen a significant surge in academic research, with major conferences such as FACCT and AIES, as well as in large tech companies that are putting together fast-growing teams to tackle the ethical challenges.

Ethical AI is a broad subject that covers many topics such as privacy, data governance, societal and environmental well-being, algorithmic accountability, etc. In this article we will mainly focus on the following components of ethics in machine learning: fairness, explainability and traceability. We'll first discuss what is at stake and why paying attention to ethics is mandatory, then we'll explore how to frame and develop your machine learning project having ethics in mind and how to follow up on ethics once deployed into production.

Why we should pay attention to ethics

With machine learning algorithms and the set of abstractions and hypotheses underlying them becoming more and more complex, it has become challenging to fully grasp and understand all the possible consequences of the whole system.

There have been several high-profile real-world examples of unfair machine learning algorithms resulting in suboptimal and discriminating outcomes. Upon which, the well-known example of COMPAS. COMPAS was a widely used commercial software that measures the risk of a person to recommit another crime, that was compared to normal human judgment in a study and was later discovered to be biased against African-Americans: COMPAS was more likely to assign a higher risk score to African-American offenders than to Caucasians with the same profile.

In the field of NLP, gender biais was detected in early versions of Google Translate that was addressed in 2018 and more recently.

In the field of credit attribution, Goldman Sachs was being investigated for using an AI algorithm that allegedly discriminated against women by granting larger credit limits to men than women on their Apple cards.

In the field of healthcare, a risk-prediction algorithm used on more than 200 million people in the U.S. demonstrated racial bias.

With no clearly defined framework on how to analyze, identify and mitigate biases the risks of falling into ethical pitfalls may be fairly high. It is thus increasingly important to set proper guidelines in order to build models that produce results that are appropriate and fair, particularly in domains involving people. Building trustworthy AI makes end users feel safe when they use it, and it allows companies to exert more control over its use in order to increase efficiency while avoiding any harm. For your AI to be trustworthy, you actually need to start thinking ethics even before processing data and developing algorithms.

How to think ethics even before your project starts

Ethics must be considered from the beginning of a new project, particularly at the problem framing phase. You should have in mind the targeted end-users as well as the goal of the proposed solution to establish the right analysis and risk management framework to identify the direct or indirect harms that may be induced by the solution. You should ask yourself, in these conditions, could my solution lead to decisions that could be skewed toward a particular sub-group of end-users ?

It's thus critical to build KPIs to track the methods that carry out your risk management strategy's efficacy. A robust framework could also incorporate, when possible, ethical risk reduction mechanism.

When dealing with a sensitive subject that has a high risk potential, it is necessary to extend the time allocated to the exploration and build phase in order to inject thorough ethical assessment analysis and bias mitigation strategies.

You must as well establish mechanisms that facilitate the AI system's auditability and reproducibility . A logic trace must be available for inspection so that any issues can be reviewed or investigated further. This is done by enforcing a good level of traceability through documentation, logging, tracking and versioning.

Each data source and data transformation must also be documented to make the choices made to process the data transparent and traceable. This makes it possible to pinpoint the steps that may have injected or reinforced a bias.

How to include ethics when developing your data project

To include ethics when developing your data project, it is important to include at least three components: fairness, explainability and traceability.

FAIRNESS

The first step in most machine learning projects is usually data collection. Whether going through the data collection process or using an existing dataset, knowledge of how the collection was performed is crucial. Usually, it is not feasible to include the entire target population so features and labels could be sampled from a subset, filtered on some criteria or aggregated. All these steps can introduce statistical bias that may have ethical consequences.

REPRESENTATION BIAS

arises from the way we define and sample a population. For example, the lack of geographical diversity in datasets such as ImageNet has demonstrated a bias towards Western countries. As a result of sampling bias, trends estimated for one population may not generalize to data collected from a new population.

Hence there is a need to define appropriate data collection protocols and to analyse the diversity of the data received and report to the team any gaps or risks detected. You need to collect data as objectively as possible. For example, by ensuring, through some statistical analysis, that the sample is representative of the population or group you are studying and, as much as possible, by combining inputs from multiple sources to ensure data diversity.

Documenting the findings and the whole data collection process is mandatory.

There are in fact many possible sources of bias that can exist in many forms, some of which can lead to unfairness in different downstream learning tasks.

Since the core of supervised machine learning algorithms is the training data, models can learn their behavior from data that can suffer from the inclusion of unintended historical or statistical biases. Historical bias can seep into the data generation process even given a perfect sampling and feature selection. The persistence of these biases could lead to unintended discrimination against certain groups or individuals, which can exacerbate prejudice and marginalization.

Not all the sources of bias are rooted in data, the full machine learning pipeline involves a series of choices and practices along the way, from data pre-processing to model deployment.

It is not straightforward to identify from the start if and how problems might arise. A thorough analysis is needed to pinpoint emergent issues. Depending on the use case, the type of data and the task goal, different methods will apply. In this section, we will explore some techniques to identify and mitigate ethical biais through an illustrative use case. We'll first state the problem, then we'll see how to measure bias and finally we'll use some techniques to mitigate bias during pre-processing, in-processing and post-processing.

PROBLEM STATEMENT

Say you're building a scoring algorithm in the banking sector to automate the targeting of clients that will benefit or not from a premium deal. You're given a historical dataset that contains many features on your meaningful data about your customers as well as the binary target "eligible for a premium deal". Elements of PII (personal identifiable information) have been previously removed from the dataset so there won't be any privacy issue at stake (on this matter, the google cloud data loss prevention service is a great tool to perform the task of de-identification of your sensitive data).

This use case may seem somewhat fictitious but the problem is close to a real use case we dealt with in the past on a different sector.

MEASURING BIAS

The first step of the analysis is to explore the data in order to identify sensitive features, privileged value and the favorable label.

Sensitive features (or sometimes called protected attributes) are features that partition a population into groups that should have parity in terms of benefit received. These features can have a discriminatory potential towards certain subgroups. For example: sex, gender, age, family status, socio-economic classification, marital status, etc. and any proxy data derived from them (e.g. geographical location or bill amounts can act as proxies to socio-economic classification as it is observed in some situations that they can be strongly correlated) are sensitive features. A privileged value of a sensitive feature denotes a group that has had, historically, a systematic advantage.

A favorable label is a label whose value provides a positive outcome that benefits the recipient.During the data preparation phase, steps such as splitting the data, undersampling or oversampling, dealing with missing values and outliers could introduce bias if they're not carried out carefully. The proportions of missing values or outliers across subgroups on sensitive features can be a first step in identifying bias. Some imputation strategies may introduce statistical bias, for example, imputing the missing values of the client's age feature by its median.

In our scoring exemple, we drew the graph of how the training data is distributed across genders with regard to the target "eligible for a premium deal".

We can see that the distribution of the target is unbalanced in favor of the gender Male. Let us hypothesize that the privilege value is Male where gender is a sensitive feature and the favorable label is "eligible for a premium deal". Furthermore, this could correspond to a representation bias in the data. In fact, in a case where equity is respected, one could ensure that the distributions in the data are totally balanced or correspond to the distributions in the demographic data.

At this point you might be tempted to simply discard the sensitive features from your dataset but it has been shown that removing sensitive attributes is not necessarily enough to make your model fair. The model could use other features that correlate to the removed sensitive feature, reproducing historical biases. To give an example, a feature A could be strongly correlated to the age of a client so if the data is biased towards a certain age bin (historical bias could result in discrimination on the basis of age in hiring, promotion etc.) this bias will be encoded into feature A and removing the age of a client won't alleviate the problem. By keeping the sensitive feature in your data, when it's necessary, you can have greater control over bias and fairness measurements and mitigation.

BIAS METRICS

There are a variety of fairness definitions and fairness metrics. We can divide fairness into individual fairness and group fairness. Individual fairness gives similar predictions to similar individuals whereas group fairness treats different groups equally.

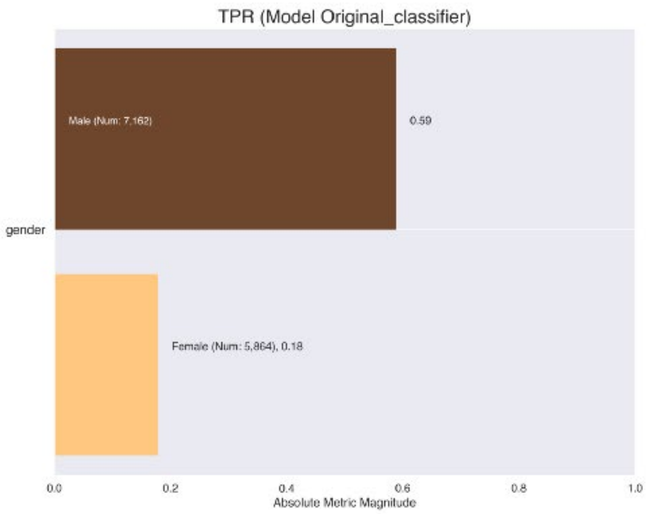
To achieve group fairness, we want the likelihood of a positive outcome to be the same regardless of whether the person is in the protected (e.g., female) group or not.

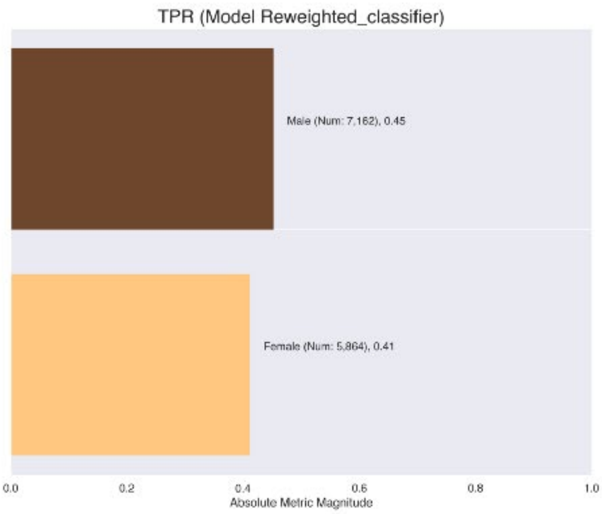
One simple group metric is to compare the percentage of favorable outcomes for the privileged and non-privileged groups (in our exemple the gender Male that are "eligible for a premium deal" compared to the gender Female that are "eligible for a premium deal"). You can compute this comparison as a difference between the two percentages which leads to the statistical parity difference metric (also called demographic parity):

```
Statistical parity difference :=
Pr(Y=1/D = unprivileged) − Pr(Y = 1/D =
privileged
```

For there to be no difference in favorable outcomes between privileged and non-privileged groups, statistical parity difference should be equal to 0. On the subject of individual fairness metric there is the consistency which measures the degree of similarity of labels for similar individuals using a nearest neighbor algorithm. You can use the handy library AIF360 that lets you compute many fairness metrics.

All you have to do is wrap your dataframe into the StandardDataset. AIF360 uses a StandardDataset that





wraps a Pandas DataFrame with many attributes and methods specific to processing and measuring ethical biases. You can then use this as an input to the BinaryLabelDatasetMetric class which will compute a set of useful metrics.

```
params_aif = {
"label_name" : "eligible_for_a_
premium_deal",
"favorable_classes" : [1],
"protected_attribute_names" :
["gender"],
"privileged_classes" : [[0]] # in our
case, 0 is Male and 1 is Female
}
# Create aif360 StandardDatasets
train_standard_dataset =
StandardDataset(df=train_dataframe,
**params_aif)
privileged_groups = [{'gender': 0}]
unprivileged_groups = [{'gender': 1}]
train_bldm =
BinaryLabelDatasetMetric(train_
standard_dataset,
unprivileged_groups=unprivileged_
groups,
privileged_groups=privileged_groups)
```

Once measured on our scoring exemple's training data, we observe a mean statistical parity difference of -0.21 which indicates that the privileged group Male had 21% more positive results in the training data set.

BIAS MITIGATION

Methods that target algorithmic biases are usually divided into three categories:

- **Pre-processing**

Pre-processing techniques act on the training data and try to transform it so that the underlying discrimination is removed.

- **In-processing**

In-processing techniques act on the learning algorithms in order to remove discrimination during the model training process either by incorporating changes into the objective function or imposing a constraint.

- **Post-processing**

Post-processing techniques that take an already-trained model and transform its predictions so that they satisfy the constraints implied by the selected fairness metric. It's particularly useful in the case where the algorithm can only treat the learned model as a black box without any ability to modify the training data or learning algorithm.

We used a pre-processing technique on the training data in order to optimize the statistical parity difference. We applied the Reweighing algorithm that is implemented in AIF360 in order to weight the examples differently in each combination (group, label) to ensure fairness before classification.

```
RW = Reweighing(unprivileged_
groups=unprivileged_groups,
        privileged_groups=privileged_
groups)
reweighted_train = RW.fit_
transform(train_standard_dataset)
```

The instance weights attribute has changed in order to re-balance the sensitive feature with respect to the target. Doing so, the Reweighing algorithm mitigated the group bias on the training data: a new measure of the statistical parity difference is completely rebalanced from -0.21 to 0.

There are other pre-processing bias mitigation algorithms implemented in AIF360, such as the DisparateImpactRemover which is a technique that edits feature values to increase group fairness while preserving rank order within groups or LFR (Learning fair representation) which is a pre-processing technique that finds a latent representation that encodes the data but obscures information about the protected attributes.

We then trained two classifier models one on the original training data and the other on the reweighed data. We observe that Reweighing had only a weak impact on the performance, losing 1% of F1-score.

We also tried an in-processing algorithm on our example use case: adversarial debiasing that significantly improved the group bias metrics (statistical parity difference was divided by 2) with little deterioration in model performance (about 1% on the F1 score).

There can therefore be a trade-off between performance and bias metrics. Here the deterioration is quite small but in some situations the compromise could be more acute. This information must be brought to light to the team and appropriate stakeholders who can make decisions on how to deal with this issue.

Now that we have trained models we can explore their predictions and investigate for imbalance toward the favorable outcome across genders. There are many tools such as What-if tool or Aequitas that let you probe the behavior of trained machine learning models and investigate model performance and fairness across subgroups.

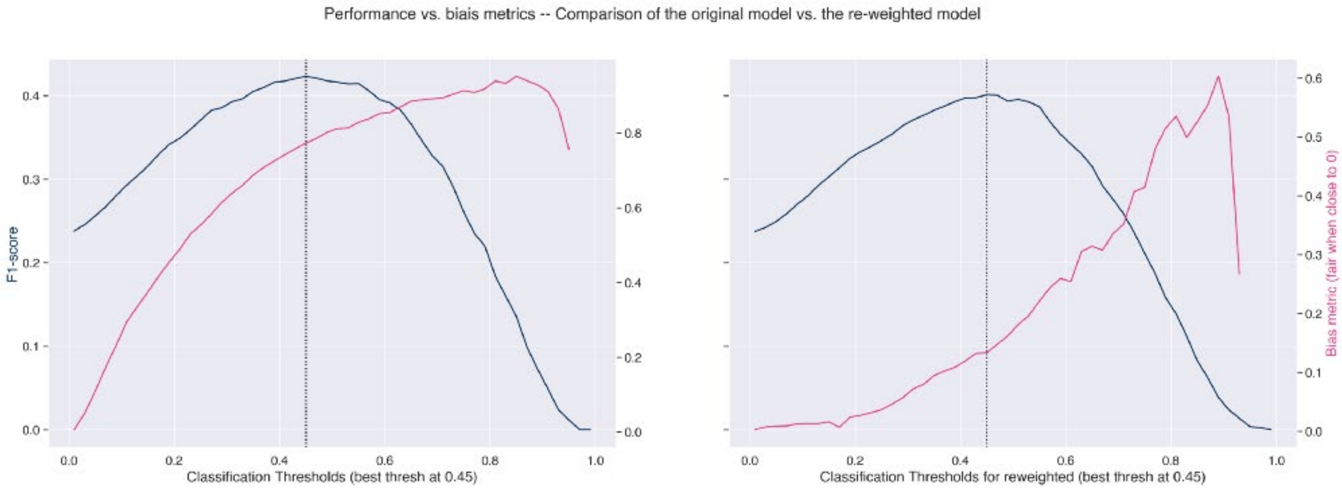
As an illustration, you can use Aequitas to generate crosstabs and visualisations that present various bias and performance metrics distributed across the subgroups. For example we can quickly compare the true positive rates of the classifiers that were trained on the original data and on the re-weighted data. We see that this rate has been balanced and therefore allows for greater gender equity towards the model's favorable outcome of being eligible for a premium deal.

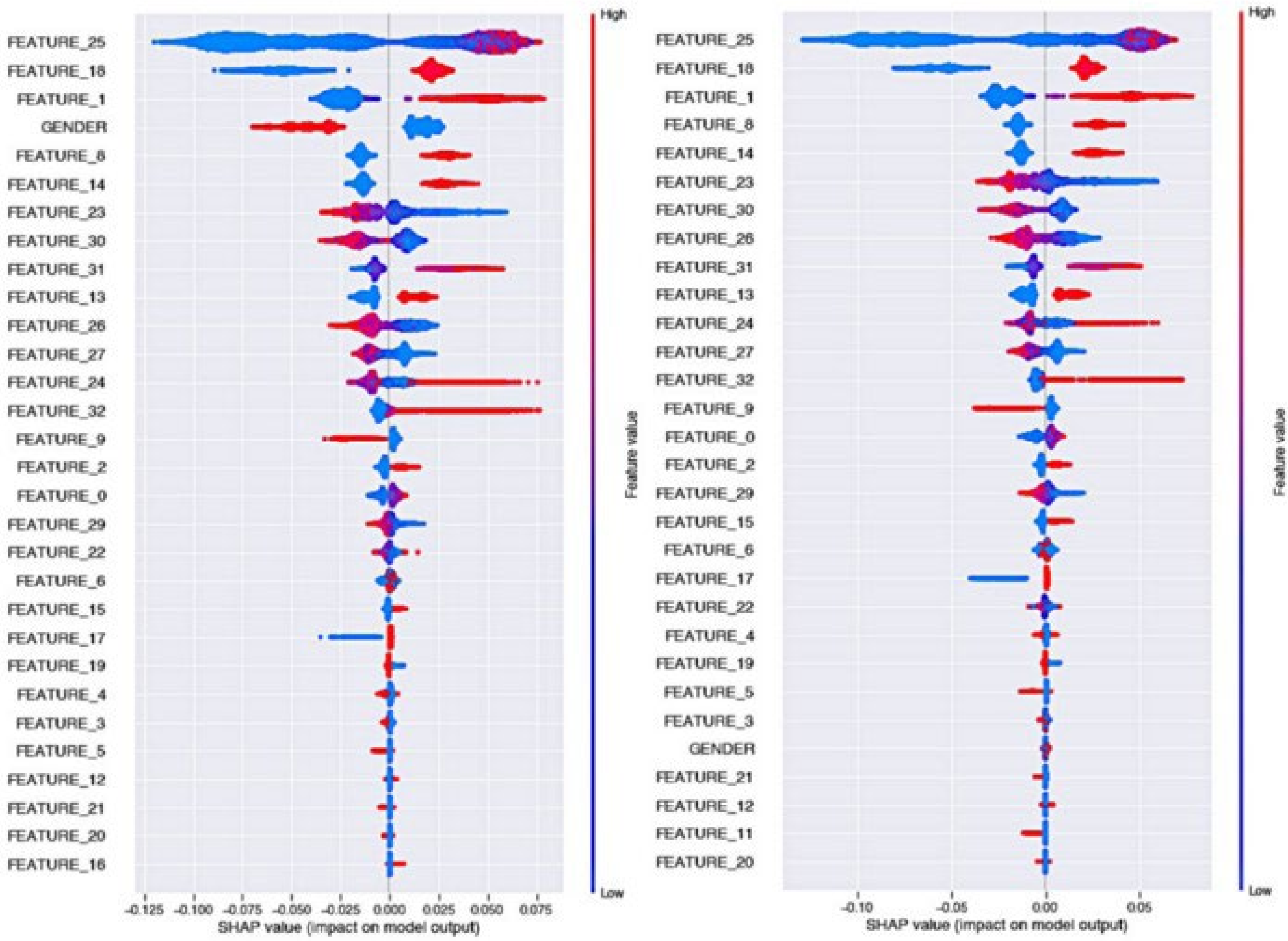
As a post-processing technique we interacted on the classification threshold. A classification model usually provides us with the probabilities associated with the realization of each class as a prediction. This probability can be used as is or converted into a binary value.

In order to identify the class corresponding to the obtained probabilities, a classification threshold (also called decision threshold) must be defined. Any value above this threshold will correspond to the positive category "is eligible to a premium deal" and vice versa for values below this threshold.

By plotting the performance metric and the bias metric (here 1 – disparate impact) across all classification thresholds, we can define the optimal threshold. This helps us choose the appropriate threshold in order to maximize performance and minimize bias.

On the left figure we see that if we push the threshold to the left, thereby lowering the performance a bit, we can improve on the bias metric.





Also, as expected, we observe a clear improvement of the group bias metrics on the re-weighted model (right figure) which could be further improved by choosing another classification threshold but at the expense of performance.

EXPLAINABILITY

Another core pillar to build trustworthy machine learning models is the explainability. Explainability is the ability to explain both the technical processes of the AI system and the reasoning behind the decisions or predictions that the AI system makes, therefore being able to quantify the influence of each feature/attribute on the predictions. Using easily interpretable models instead of black-box models as much as possible is a good practice.

There are many methods to obtain explainability of models. These methods can be grouped into 2 categories:

Intrinsic explainability where the model itself gives the feature importance or feature weights.

Post-hoc explainability where small input changes are leveraged to infer feature importance.

In-processing techniques act on the learning algorithms in order to remove discrimination during the model training process either by incorporating changes into the objective function or imposing a constraint.

Here we'll apply a famous post-hoc method, namely SHAP (SHapley Additive exPlanations), for more info we recommend exploring this very comprehensive resource on the subject. Shap is a library that implements a game theoretic approach to explain the output of any machine learning model.

Quick reminder on how to read Shap's beeswarm plots:

- The features are sorted from top to bottom from the most important to the least important.

- The color corresponds to the amplitude of the values of the feature. The more red the color, the lower the value and vice versa for blue.

- The horizontal axis corresponds to the direction of influence of the feature on the prediction of the target. For instance, in our scoring exemple, negative values will have the impact of influencing the prediction towards the class "not eligible to a premium deal" and vice versa for positive values.

On the left, we have the original model's explainability where we observe that in this case the gender variable has a very strong predictive power and that the gender Female has an impact that influences the decision towards the "not eligible to a premium deal" target with a big gap with respect to the gender Male.

We can see on the right graph, in this case where the model was trained on the re-weighted data, that the importance of the gender feature has strongly decreased. It is now part of the least important features. Moreover, the influence of the female vs. male class on the prediction of the target is much more balanced (the colors are close to 0 in Shapley value).

TRACEABILITY

Another essential aspect in the process of creating trustworthy machine learning algorithms is the traceability of results and good reproducibility of experiments. This makes it easy to identify which version of a model has been put into production so that it can be audited if its behavior causes harm and no longer conforms to the company's ethical values.

To do this, one must be able to track and record each model version and its associated training data, hyperparameters and results. Several tools can accomplish this task: Mlflow is one great option that allows you to quickly generate a web interface that centralises all runs, while saving your artifacts into the storage of your choice. Each version of the experiment can be tracked with the hash of the associated commit. Each of these versions will contain all the elements recorded by MLflow.

Here is a tool that we've open sourced at Artefact that lets you deploy a secure MLflow on a GCP project with a single command.

It is also a good practice to create a FactSheet for each model, which corresponds to a model identity card that summarizes various elements tracing the pre-processing steps, performance metrics, bias metrics etc.

These identity cards are delivered by the data scientists to the model operating teams, allowing them to determine whether the model is suitable for their situation. The FactSheet can also be stored, in tabular form for example, in MLFlow alongside the associated model.

"Another essential aspect in the process of creating trustworthy machine learning algorithms is the traceability of results and good reproducibility of experiments. This makes it easy to identify which version of a model has been put into production"

How to follow up on ethics once deployed

Once your model is deployed, you have to make sure that it is used for the purpose for which it was thought, designed and built. Deployment bias occurs when there is a mismatch between the problem a model is intended to solve and the way it is actually used. This frequently happens when a system is developed and evaluated as if it were totally self-contained, whereas in reality it is part of a complex socio-technical system governed by a large number of decision-makers.

Production data can drift over time which can result in algorithm performance degradation that could inject bias. Tracking production data quality and data drift by monitoring the distributions of new data compared to the data used to train the models, should be a step in the production pipeline to raise the proper alerts when necessary and define when retraining is mandatory.

The production pipeline should be designed so that there is a way to turn off the current model or roll back to a previous version.

Conclusion

This article has barely scratched the surface of the vast subject that is ethical AI and only merely touched on of the interesting tools being developed that are now available.

As we've seen, the most logical way to explicitly address fairness problems is to declare a collection of selected features as potentially discriminatory and then investigate through this prism ethical bias. This straightforward technique, however, has a fault in that discrimination can be the outcome of a combination of features that aren't discriminatory on their own. Moreover, in many cases you won't have access to any sensitive feature.

Fairness assessment is a complex task that is dependent on the nature of the problem. Approaching a scoring problem based on tabular data won't be the same as mitigating bias in natural language processing.

CASE STUDY

PIERRE FABRE GROUP

Accelerating online growth with a global e-Retail upskilling programme

CHALLENGES

Building an e-Commerce culture – fast

Despite its overall success and a flourishing online dermo-cosmetic market, Pierre Fabre is facing strong competition from more digitally-advanced competitors such as L'Oréal Cosmetic Active. As the Group's online business target is 20,5% for 2021, new growth policies needed to be implemented.

"In today's market, we knew that prioritising e-Commerce was crucial for this client. But to succeed in this new ecosystem, they needed to adapt and improve their capabilities in order to compete with the best digital players",

says Thomas Faure — *E-Retail Lead at Artefact.*

We had to find a way to support their commercial and marketing teams in their digital acceleration. And so we proposed a company-wide upskilling programme.

39

SOLUTIONS

Teaching teams the secrets of e-Retail

To meet the objectives of Pierre Fabre, we structured the e-Retail upskilling solution around five key principles:

- **Link the e-Retail programme** to a concrete business achievement / target support by top management
- **Adapt the programme to the regional** needs and maturity of Pierre Fabre Group trainees
- **Identify and empower e-Commerce champions**
- **Provide multiple learning formats**, from digital self-service to one-to-one coaching
- **Build actionable tools**, not just theoretical content

To set things in motion, we performed an e-Retail audit to identify knowledge gaps across the company. We analysed Pierre Fabre's relationship with online business, including their organisation, processes, tools, and people. The audit revealed a need for training in digital basics and specific skills within various teams at different levels throughout the company.

The information we drew from the audit revealed five topics to be tackled:

Module #1 – Omnichannel strategy

Module #2 – e-Retail acculturation

Module #3 – Brand animation

Module #4 – Performance measurement

Module #5 – E-Retail negotiation

The upskilling programme is composed of eLearning modules, bootcamp sessions, and one-to-one coaching. Engagement is ensured by attractive and varied learning formats (from motion design and interactive videos, infographics and quizzes, to physical training and individual coaching).

We launched the programme to approximately 600 employees in 27 markets in June 2020.

Nicolas Mouton
Learning & Development Manager

"Learner feedback was extremely positive. It was good to see that there was general awareness of our need to be present in e-Retail. Pierre Fabre doesn't historically have an e-Commerce culture, so upskilling is viewed as necessary and desirable."

RESULTS

Making eLearning engaging – and profitable

Multiple stakeholders were addressed by each module in all markets, with on-demand programmes offering generic courses for all, and specific topics for marketing and/or commercial teams. Three regional adaptations were produced for Europe, Asia, and the Americas.

Bootcamp sessions were held in 26 markets, and individual coaching sessions targeted some 50 trainees. About 750 hours of one-to-one coaching sessions will be eventually delivered in total (ongoing).

The satisfaction survey sent to all trainees – from operational to top management – received high notes on all indices, including assessment of content and format quality, programme usefulness in daily activity etc.

Data Governance, a prerequisite for AI project success

Justine Nerce
Managing Partner

Data and its applications are being increasingly integrated into business activities. They're at the heart of the search to improve productivity and overall efficiency. Through specific processes and an adapted organizational structure, data governance enables companies to organize data, enhance its quality and meet the ethical and regulatory challenges of data processing. An interview by the Hub Institute with Justine Nerce, Managing Partner at Artefact.

What are the challenges of data governance today?

The amount of data and the number of use cases around data is constantly increasing. First, companies have to deal with the challenge of getting the most possible value out of their data and democratizing it. Good quality, well documented data should allow it to be accessible to the end user.

All of this applies within the framework of ethics and data protection. Data governance is becoming essential to ensure compliance with certain privacy laws. In Europe, the General Data Protection Regulation (GDPR) is in effect and is tending towards becoming the global standard.

This means that the organization must first be able to demonstrate that it knows what data is flowing through its infrastructure. It must be fully transparent about what data it is collecting from its users and be able to delete all data linked to any individual immediately.

Second, migration to the cloud is essential. Three main uses are emerging: Business Intelligence, Artificial Intelligence, and data exploration. By structuring data based on data governance, businesses will be able to offer these three types of uses as a service.

These data products constitute a common, cross-disciplinary good, which requires a dedicated team. These products must be of high quality, but also visible and usable by all. The challenge for companies is data democratization. These data products must also be secure and protected to comply with various regulatory and ethical issues.

How does Artefact support companies in implementing data governance?

At Artefact, we act as a consulting firm. We support all our clients throughout data governance implementation, from strategy to deployment. First, we perform an audit to see where they stand, then define a roadmap to identify areas to work on. Finally, we build a data asset structure into data products and help them choose the technical tools they need.

In our consulting approach, we insist on the importance of data as a vector of value for the company, then we work on deployment, quality tool selection and documentation of governance to give substance to the strategy and make it feasible.

We've also set up our own Artefact School of Data, which lets us train data stewards and data owners, essential roles in the implementation of data governance for businesses. Along with this professional training, we also intervene directly in companies to acculturate them to the need for advanced and supported data governance in order to succeed in their AI projects.

What is unique about Artefact's global vision?

Our strength is that we propose a global data governance model, focusing on end-use cases first. We position data governance as an "asset" of this transformation. We're able to transcribe use cases into tangible value and be part of a global transformation program.

We also have multidisciplinary experts. There are about 20 of us in France who specialize in data governance, with profiles from different backgrounds: data product owners who model products, data stewards who document and improve quality, but also data engineers and data analysts.

We also have an ecosystem of technology partners with whom we collaborate in an agnostic way. We're proficient in all the new tools that appear on the market. We have both technical and strategic DNA, and are able to link all of these subjects together to treat them in a holistic and comprehensive way and deploy them to many clients.

Have you got a concrete example of support that you've provided?

We assisted one of our major clients with very extensive data assets in their data transformation. The project concerned a redesign of their data governance. When we arrived in mid-2017, we saw that their governance had been approached from a too-technical and not sufficiently "business" perspective. This resulted in a lack of adoption of the necessary tools. To correct this, we linked their governance to their strategic use cases. To do so, we documented the use cases, democratized their access, and improved data quality to ensure good results. The first pilots were a success! We then faced the challenge of scaling up.

In 2020, we assisted this same company in launching a program to accelerate Artificial Intelligence programs and migration to the Google Cloud Platform (GCP). Governance had been positioned as a strategic asset of their transformation and this launch was performed in two stages:

- structuring their "Data Governance Office" and setting up an operating model with data stewards and data custodians, etc.
- structuring their data assets into a large "business domain", with the choice of tools to operate, etc.

We're now entering a third phase of industrialization and extension of this AI program. As part of the migration to the cloud, we're analyzing how we can structure, rationalize and pool our data assets. At the moment, we've moved on to the second stage, which consists of structuring our data assets according to these major business families. Next, we're going to start thinking about the development of tomorrow's data products, which will serve different categories of use cases.

What can we expect in the future, once everyone has implemented their data governance?

The availability of data will allow the implementation of even more use cases, particularly in the area of Artificial Intelligence. This will accelerate value creation within organizations. It will also allow us to support all the issues surrounding data democratization and decentralization, especially in terms of bringing data closer to the business. Artefact's mission is to create this bridge between data and business, and we carry it out on a daily basis with our clients. If the data is well structured and clean, if the products are available, and if we have the push-button tools to manipulate them, theoretically in five years, everyone will be able to use data in their daily work!

ARTEFACT | EVERY COMPANY TALKS ABOUT DATA.
AT ARTEFACT, WE DON'T TALK, WE ACT.

DATA ACCELERATION PROGRAMS

- Data Strategy
- Data Governance
- Data Platform
- Data Lab
- Change Management

DATA INDUSTRY SOLUTIONS

- Data-driven Sales
- eCommerce Services
- Demand Forecasting
- Data for Supply Chain
- AI for Customer Care

DATA & DIGITAL MARKETING

- 1st Party Data Strategy
- Customer Data Platform
- Consumer & Market Insights
- Advanced Analytics
- Digital & Retail Media

DATA FOR IMPACT

- AI Ethics
- Data for Sustainability
- School of Data (B2C & B2B training)

WE OFFER END-TO-END DATA & DIGITAL SERVICES

Brazil • China • Dubai • France • Germany
Lebanon • Singapore • South Korea • Spain

Malaysia • Morocco • Netherlands • Saudi Arabia
Switzerland • United Kingdom • United States

+1000 EMPLOYEES | 16 COUNTRIES | +300 MAJOR BRANDS

ARTEFACT

VALUE BY DATA

CONTACT

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

ARTEFACT HEADQUARTERS

19, rue Richer
75009 – Paris
France

artefact.com