Update of the European Data Market
SMART 2016/0063

Story 4 – How Big Data is driving AI: Selected Examples of AI Applications across European Industries

14th March 2019
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Giorgio Micheletti; Philip Carnelley; Erica Spinoni (IDC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deliverable</td>
<td>D 3.4 Quarterly Stories (Story 4)</td>
</tr>
<tr>
<td>Date of delivery</td>
<td>14&lt;sup&gt;th&lt;/sup&gt; March 2019</td>
</tr>
<tr>
<td>Version</td>
<td>2.0</td>
</tr>
<tr>
<td>Addressee officer</td>
<td>Katalin IMREI</td>
</tr>
<tr>
<td></td>
<td>Policy Officer</td>
</tr>
<tr>
<td></td>
<td>European Commission - DG CONNECT</td>
</tr>
<tr>
<td></td>
<td>Unit G1 – Data Policy &amp; Innovation</td>
</tr>
<tr>
<td></td>
<td>EUFO 1/178, L-2557 Luxembourg/Gasperich</td>
</tr>
<tr>
<td></td>
<td><a href="mailto:katalin.imrei@ec.europa.eu">katalin.imrei@ec.europa.eu</a></td>
</tr>
<tr>
<td>Contract ref.</td>
<td>N. 30-CE-0835309/00-96</td>
</tr>
</tbody>
</table>
Table of contents

EXECUTIVE SUMMARY .................................................................................................................. 4

1. INTRODUCTION ......................................................................................................................... 5
   1.1 MAIN OBJECTIVES AND SCOPE ....................................................................................... 5
   1.2 METHODOLOGY AND STRUCTURE .................................................................................... 5

2. BIG DATA SUPPORTING AI IN EUROPE – AN OVERVIEW .................................................. 7
   2.1 DEFINING ARTIFICIAL INTELLIGENCE ............................................................................. 7
   2.3 HOW IS AI GETTING REAL IN EUROPE TODAY? ............................................................... 9
   2.4 AI IN EUROPE TODAY: GREAT EXPECTATIONS AND OBSTACLES TO OVERCOME .......... 10

3. THE CASE STUDIES - WHAT WE FOUND SO FAR ................................................................. 13
   2.1 CASE STUDY 1: COYOTE SYSTEM - PREDICTIVE ROAD TRAFFIC ............................. 13
   2.2 CASE STUDY 2: VOYAGE PRIVÉ - PREDICTIVE MARKETING ........................................ 15
   2.3 CASE STUDY 3: DAZN - SCALING SMALL DATA TEAMS WITH ML ......................... 17
   2.4 CASE STUDY 4: AXA - PREDICTIVE LARGE-LOSS CASES ........................................... 20
   2.5 CASE STUDY 5: AIRBUS (D&S) - CLEANING SKIES FROM CLOUDS ....................... 22
   2.6 COMMON THEMES AND FINDINGS ACROSS THE CASE STUDIES ............................ 24

4. PRELIMINARY POLICY CONCLUSIONS ................................................................................. 29

REFERENCES .................................................................................................................................... 30
Executive summary

Artificial Intelligence (AI) is no longer a simple buzzword. From banking to retail, from the manufacturing industry to the insurance and finance sector, a myriad of AI-applications based on Big Data and Analytics (BDA) are rapidly emerging and Europe is no exception in this respect. At worldwide level, however, European organizations are still struggling when it comes to the full adoption and deployment of AI applications. According to a recent IDC survey conducted in July and August 2019 about the status of AI in Europe, the adoption of AI technologies and solutions is still relatively low in Europe — around 14% of organizations today — but it is increasing quickly with a positive trend set to continue in the coming years.

While AI can be deployed in virtually every sector and for the most disparate reasons, most organizations in Europe are using AI and BDA to enhance customer experience and satisfaction. IDC’s AI survey reveals that 66% of the respondents adopt, or are considering the adoption, of AI systems to increase quality of products and services alongside with creating improvements in customer support. Interestingly, Big Data and Analytics are viewed to serve similar purposes – 44% of the respondents maintain that the main business value driving the adoption of BDA tools and technologies is to improve business process, 43% are using BDA to improve market understanding and 42% indicated product improvement as the third main driver for adoption. Again, a customer-centric approach revolving around better business processes, enhanced products and services, increased market understanding, and improved customer services lies behind the interaction and adoption of AI and BDA.

The findings obtained through the desk research and the in-depth interviews carried out to feature the five case-studies presented in this research point to the same direction. AI solutions are adopted to advance the organization’s predictive capabilities, allowing for a better interpretation of customer needs and improved forecast accuracy. This is the case of Voyage Privé (an online travel agency) and its ability to better interpret customers’ preference signals or AXA (a global insurance company) and its advanced predictive model thanks to deeper self-learning capabilities. AI is further embraced to extend the quality of products and services (AXA claims an increase in forecast reliability of 78%, for example, and Airbus – the aerospace company – signals an error rate passing from 11% to 3% thanks to the use of AI), as well as to obtain overall business benefits in sales & marketing campaigns (+6% in one year for Voyage Privé) and customer loyalty programmes (+11% for Coyote System, a real-time road information company).

Several challenges remain to be overcome as there is still too little understanding of the elaboration mechanisms deployed by AI applications, especially when Machine Learning and Deep Learning techniques are involved and a general lack of trust in the results produced by this technology is still widespread. For these reasons, regulatory initiatives and policy-related interventions are starting to arise. A fundamental step in this direction has been taken by the European strategy recently set forth by the European Commission to promote a human-centric AI approach that places people at the centre of the development of AI. Delivering on this strategy, the Commission presented a coordinated action plan prepared with Member States foster the development and use of AI in Europe and launched the AI Watch platform to ensure quality input and inform policy-makers of the developments, uptake and impact of Artificial intelligence for Europe.

---

1. Introduction

Artificial intelligence (AI) has become an area of strategic importance and a key driver of economic development. It can bring solutions to many societal challenges from helping treating diseases to minimising the environmental impact of farming.

To build robust models at the core of AI-based systems, high quality data is a key factor to improve performances. The recently adopted EU legislation to improve data sharing and open up more data for re-use, together with the ever-increasing use of Big Data technologies, is likely to flood a wide variety of markets and industries with a new wave of data, opening up, in turn, new conditions for a broader take-up of AI technologies in Europe.

However, socio-economic, legal and ethical impacts have to be carefully addressed. For this reason, the European Commission has put forward a European approach to Artificial Intelligence based on three pillars:

1. Being ahead of technological developments and encouraging uptake by the public and private sectors;
2. Prepare for socio-economic changes brought about by AI;
3. Ensure an appropriate ethical and legal framework.

To assist the European Commission in making this approach a reality, additional research is needed. IDC has therefore prepared a qualitative piece of research investigating the interaction between Big Data and AI technologies in Europe today.

1.1 Main objectives and scope

The aim of this story is to investigate how companies and organizations can generate economic, operational and organizational benefits by leveraging Big Data and AI applications in a number of different industries and vertical markets. Through extensive desk research and a series of in-depth interviews, the story aims at:

- identifying the reasons and needs underpinning the decision to develop and deploy AI applications in Europe today;
- examining the way in which European companies and organizations across a multitude of sectors and industries apply AI technologies;
- uncovering the underlying benefits for companies and organizations that successfully deploy AI solutions - improved decision making and more impactful decisions, reduced time to market, improved services and customer satisfaction, cost optimization, additional revenue streams, etc.
- highlighting the barriers that currently hamper the adoption of AI solutions in Europe today and providing an initial set of recommendations on how to support AI use for policy makers at national and European level.

1.2 Methodology and structure

To better understand how European companies are approaching and implementing AI in a variety of different sectors, we have conducted extensive desk research across a multitude of publicly available sources and IDC existing research material.

We have further carried out in-depth interviews with representatives of Dataiku - a leading AI and Machine Learning technology provider representing a central hub for analytics and machine learning model deployment and management.
These secondary and primary research efforts led to the realization of five distinct case-studies in different sectors and across different AI use cases. In particular:

- Coyote System SAS – a France-based real-time road information provider – demonstrated the use of AI applications to improve **predictive road traffic solutions**;
- Voyage Privé SAS – an online travel agency – illustrated the use of AI to improve **predictive marketing capabilities**;
- DAZN – a live-streaming service specialized in sporting events – demonstrated how AI solutions can be used to **scale-up scarce resources**;
- AXA – a French multinational insurance firm – provided a good example of how AI can help identify and **mitigate the risks related to large-loss cases in the insurance sector**;
- Airbus – the aerospace company – illustrated how AI can prove to be a powerful means to **improve satellite imaging and mapping**.

The current document is structured along three main sections.

- The first section (chapter 2) introduces the concept of AI, its interaction with Big Data and Analytics technologies and present the status of AI implementation in Europe according to the latest available sources.
- The second section (chapter 3) is devoted to an overview of the real-life case studies that formed the bulk of the primary and secondary research underpinning this story.
- The final section (chapter 4 and 5) presents the common themes and main lessons learnt from the case studies and provide a few conclusive remarks with possible policy directions for the years to come.
2. Big Data supporting AI in Europe – An Overview

2.1 Defining Artificial Intelligence

There is little doubt that Artificial Intelligence (AI) has become a topic of conversation well beyond the high-tech circles today and is now firmly on the agenda of most of business leaders, regulators and policy-makers. But what is AI in the first place? Alas, the answer is not straightforward, as several definitions of AI have been put forward since the early days of computer science. Indeed, the term Artificial Intelligence dates back to 1956 when John McCarthy – an American computer and cognitive scientist – invited a group of researchers from a variety of disciplines including language simulation, neuron nets, complexity theory and more to a summer workshop to discuss what ultimately became the field of AI\(^5\). The conference endorsed a definition of AI that was essentially based on the concept of “thinking machine”, thus including cybernetics, automata theory and complex information processing. With time, the AI definition evolved from the idea of thinking machines to the notion of machines imitating human intelligence (“human-like thinking machines” rather than “thinking machines tout-court”). This shift allowed for the introduction of some nuances and a pinch of realism in the overall quest for AI. After all, the purpose of building AI systems was less the one of building machines that think exactly like humans than to use the human reasoning as a pattern to make machines perform more and more sophisticated tasks, such as visual perception, speech recognition, decision-making and language translations\(^6\). This scientific understanding of the mechanism underlying thought and intelligent behaviour, and their incarnation in machines, leads to another key feature of AI: the core ability for machines to learn how to autonomously automate processes and take self-governing action.

The definition used by the European Commission in its recent Communication on Artificial Intelligence in Europe\(^7\) and in its latest Coordinated Plan on Artificial Intelligence\(^8\) seems to encompass all the key elements above. According to the European Commission, Artificial Intelligence (AI) “refers to systems that display intelligent behaviour by analyzing their environment and taking action – with some degree of autonomy – to achieve specific goals” and can be based exclusively on software (as in the case of voice assistants, image analysis software, search engines and speech and face recognition systems) or embedded in hardware devices (e.g.: advanced robots, autonomous cars, drones or IoT applications).

Among its many merits, this definition has the advantage of highlighting the close interrelation between AI and data-driven technologies. An AI system can take action, automate processes and learn from it only if, and insofar as, it is able to ingest, combine and correlate data, thus investigating and creating logical and statistical patterns and, ultimately, interacting heavily with data-driven technologies. Throughout this document, the above-mentioned definition is therefore to be intended when referring to Artificial Intelligence (AI).

2.2 AI and Data-Driven Technologies – A fruitful Interaction

Big Data and AI technologies are driving changes in nearly every sector of the economy and every layer of society. While this is hardly a new statement, what is more interesting is that additional evidence\(^9\) shows that AI has become real and has abandoned the “Demonstration/Trial/Proof-of-Concept” phase to embrace impactful applications at operational

---


\(^7\) “Artificial Intelligence for Europe”, COM(2018) 237 final, 25.4.2018

\(^8\) “Coordinated Plan on Artificial Intelligence”, COM(2018) 795 final, 7.12.2018


and business level. A global research project recently carried out by the MIT Sloan Management Review and the Boston Consulting Group\(^\text{10}\) finds that executives around the world are increasingly looking at artificial intelligence to create new sources of business value and that this is particularly true for leading adopters of AI – the so called “Pioneers”, i.e. enterprises that have extensive understanding of AI tools and concepts associated with significant levels of AI adoption. Interestingly, Pioneers are also more mature than other group of companies in their management of data and are far more likely than their less sophisticated peers to have strong, company-wide data governance systems and effective data-driven technologies in place. Data are therefore seen and treated as a fundamental corporate asset, centrally localized, owned and managed, accessible to a wide range of stakeholders and on top of the agenda of the senior management (see Figure 1 below).

**Figure 1: Data Maturity Drives AI**

![Data Maturity Drives AI](https://example.com/data_maturity_drives_ai.png)


From banking to retail, from the manufacturing industry to the insurance and finance sector, a myriad of AI-applications based on Big Data technology is rapidly emerging. This, in turn, stresses the need of a better interaction and integration between the two fields, so to maximize the current potential of data, while obtaining the farthest-reaching results in AI. One way to do so is to exploit the opportunities offered by data platforms. Next generation data platforms, embedded with AI, Machine Learning/Deep Learning (ML/DL) technologies, tend to strengthen not only data management and analytics techniques, but have also positive impacts on wider management and operational processes, improving, for example, customer satisfaction and workforce operations. As novel forms of interaction between Big Data and AI emerge across industries and markets (from diagnosis and recommendation treatment plans to purchase recommendation for e-commerce platforms; from autonomous vehicles to more and more sophisticated home assistant devices), the practical applications (and the ensuing business and operational benefits) of the tight bond between Big Data and AI also broadens. These applications can be classified in three main fields\(^\text{11}\):

1. **Process Automation**, i.e. the automation of digital and physical tasks. Automation is not only reflected on the production shop-floor, with the deployments of robots, autonomous transporting vehicles, etc., but also across back office and administrative activities. The latter make intense use of data platforms and integrated data to obtain insightful information for a wide range of decision-making processes impacting both the operational as well as the business aspects of organizations.

2. **Cognitive insights**. AI and ML applications are used to detect patterns in big batch or flows of data. The most common fields of application are prediction of customer

---


Data platforms and data-related technologies are not enough to grasp the full potential that AI is offering to businesses and organizations alike across all sectors. A key additional element is data availability, access and sharing. Indeed, data-sharing mechanisms are of paramount importance to sustain the deployment and well-functioning of AI applications. In addition to the data produced internally (usually through IoT- and sensor-produced data, plus customer and sales information) companies need therefore to rely on insights from externally produced data to improve their decision-making processes and obtain new and more effective data-driven business models. This is possible only with an extensive data-sharing environment that can rely on technical and regulatory mechanisms that effectively encourage the exchange of data in a trusted environment.

### 2.3 How is AI getting real in Europe Today?

Artificial Intelligence (AI) is no longer a simple buzzword – it is becoming real by the day and European organizations are no exception in this respect. In August 2018, IDC conducted a survey among European AI developers to investigate their resources and technology choices. the European landscape, companies are increasingly adopting open source solutions for AI, not only to leverage cost reductions but also for the successful opportunities that open source communities give to developers to self-implement programs, APIs and platforms (Python, R, Scala, GoogleCloud, TensorFlow, etc.) The combination of cloud (platforms), Big Data, AI and open source ease the approach of organization to get a fast and scalable development. For instance, IT solutions providers are increasingly benefitting from interoperability of different data source, languages, cloud services, APIs, etc. fastening and simplifying the operation that developers should conduct to connect those structures. In addition, strong developers' communities are rising to ease and speed up the exchange and integration processes of programs and systems.

All in all, however, European organizations are slightly lagging in comparison to some of their U.S. or Chinese counterparts when it comes to the full adoption and deployment of AI applications. IDC’s 2019 Western Europe Artificial Intelligence User Survey\(^\text{12}\) highlights that the European companies still lag slightly behind, compared to other countries and regions, as only around 14% of organizations fully adopt AI technologies and solutions (figure that rises to 27% when considering also companies and organizations that are piloting or currently implementing AI solutions). When analyzing the reasons for the AI and Big Data Analytics (BDA) deployment across European companies, a common development pattern is detected: most of the organizations are using the technologies to increase customer experience and satisfaction. This is confirmed by additional ongoing research. The results currently being released by DataBench\(^\text{13}\), for example, point to an increased use of AI and Big Data Analytics across Europe’s organizations with the principal aim of improving customer experience and increasing customer value. Again, IDC’s AI survey reveals that 66% of the respondents adopt or are considering the adoption of AI systems to increase quality of products and services alongside with creating improvements in customer support. Interestingly, Big Data and Analytics are viewed to serve similar purposes – 44% of the respondents maintain that the main business value driving the adoption of BDA tools and technologies is to improve business process, 43% are using BDA to improve market understanding and 42% indicated product improvement as

---


\(^{13}\) https://www.databench.eu/public-deliverables/: www.databench.eu
the third main driver for adoption. Again, a customer-centric approach revolving around better business processes, enhanced products and services, increased market understanding, and improved customer services lies behind the interaction and adoption of AI and BDA.

The interplay between AI and BDA and, more importantly, its substantial market potential over the next few years in Europe, is confirmed by IDC’s latest research in the field of cognitive and Artificial Intelligence software platforms and Big Data and Analytics (BDA) Defined as a complex aggregation of Business Intelligence (BI), analytics tools and platforms, data management applications and performance management applications, BDAIs expected to expand at a compound annual growth rate (CAGR) of 9.2% over the next five years in Western Europe to reach $21.9 billion (approximately €19.4 billion) by 2022 according to IDC. In parallel, IDC predicts cognitive and AI software platforms (the AI component of the ICT market) to exhibit a buoyant 49.9% CAGR through 2022 in the same geography reaching $1.6 billion (approximately €1.4 billion) by 2022. The Cognitive/AI software market is still relatively small compared to the general European software market, and its potential development is broad and still largely unexploited.

2.4 AI in Europe Today: Great Expectations and Obstacles to overcome

The AI reality is loaded with fascinating expectations and is bound to herald substantial benefits for organizations and businesses alike. AI capabilities, coupled with enhanced automation processes, have already made the way to new methods of addressing the challenges of data management and analysis.

The joint deployment of AI and ML applications, for example, has led to the creation of “self-managing” platforms, where the burdensome tasks of data cleansing and data preparation are performed automatically by the machine, leaving data scientists the freedom to fully focus on making sense of the results. What is more, conventional machine learning techniques (such as random forest regression\(^{14}\), nearest neighbour interpolation\(^{15}\), naïve Bayes classifiers\(^{16}\), etc.), together with the most recent developments in deep learning technologies based on neural networks, allow for the collection and analysis of data at any point of the data value chain. For instance, AI enables machines to gather information and analyse it from the core data storage/core infrastructure (cloud on-premise, public or hybrid) or at the edge, right exactly where the data is produced and sourced (phones, wearables, self-driving vehicles, etc.). This is going to significantly speed up not only the process of data insight but also the need of supplementary and more powerful computational mechanisms. In other words, the increased adoption of innovative machine components (such as: GPUs, FPGAs, ASICs, etc.), coupled with the development of new communication technologies (i.e. 5G), is going to accelerate the race to AI and make its impacts even more momentous.

Big Data and Analytics have already introduced substantial changes across several value chains – in manufacturing, for example, predictive maintenance\(^{17}\) and self-healing infrastructures\(^{18}\) are features already at play since the beginning of the 2010s. AI deployment though (and even more so the use of Machine Learning technologies) is bound to further strengthen the predictive power and accuracy of BDA. Machine Learning, in particular, will make the prediction of a crash system, or of a component or part failure, or of the potential

---

\(^{14}\) A random forest regression is a technique, which is able to perform a regression model alongside with a classification procedure. It is commonly used in multiple decision environment, where the decisions are shaped as the crown of a tree and each decision is a branch of the tree.

\(^{15}\) This is the simplest interpolation method, based on clustering/approximating value of a function to the closest value possible to the one shown by the function.

\(^{16}\) This statistical technique creates models to assign different labels or descriptions to a specific problem-instance to assess the probability that the instance is exactly what researchers expect to be.

\(^{17}\) Predictive maintenance is a machine learning-native implementation to maintenance issues. Sensors positioned on machines or devices capture and register usage information (temperature, speed, air condition, etc.) which are sent to a core data centre that analyse the data gathered. Lately, organizations are adopting a more fashionable solution for which data are analyzed on the edge – exactly at the source where data comes from. The machine learning analyses the data and creates plans for maintenance to avoid unforeseeable devices and machines break down.

\(^{18}\) Self-healing infrastructures are IT infrastructures built as Software-as-a-Service (SaaS) platforms able to diagnose and fix problems within the infrastructure itself, without human intervention, thanks to AI and machine learning.
churn of a customer\textsuperscript{19}, or the probability that a specific driver will cause a relevant loss to an insurance company even more precise. This is possible because data-nurtured and data-trained machines are able to notice variations and deviations from standards that humans are not able to detect. This reduces the risk of human enhancing customer satisfaction, improving cost reduction, reducing production downturn and the workload for monitoring and tuning infrastructures. All of which, can be turned into a measurable business and operational benefits for companies and organizations.

Enhanced computational devices and better ways to manage data and information streams are not the only advantages of AI applications. AI is in fact more and more deployed in the so-called ‘hands-free’ processes. In this respect, AI is shifting the technology paradigm from screens and visual interfaces towards voice, audio and self-understanding of events in the wider surrounding. A hybrid form of AI between computational devices and self-understanding is represented by Augmented Reality (AR) and Virtual Reality (VR) – two technologies allowing users (to have displayed information on spectacles instead that on screens, thus considerably enhancing the human-computer interaction.

A further AI deployment, that will generate benefits and on which the society is increasingly betting, is the ‘Affective Computing’ or ‘Emotional AI’. Through natural language processing (NLP) and natural language understanding (NLU), AI-enabled machines and interfaces are able to detect and react to human emotions measuring feelings (skin tone, heart rate, etc.), behaviours (voice alteration detection), and cognitive states. As an example, AI-empowered vehicles are already in a position to perceive drivers’ tiredness or drowsiness and put in place corrective measures.

While steadily growing and increasingly accepted, AI still faces a number of obstacles before being fully embraced by businesses and organizations. Indeed, an overall lack of trust\textsuperscript{20} in the technology and, more specifically, in the results produced by the technology, is observed. The first hindrance is directly tied to the scarce or little understanding of the data source and elaboration mechanisms deployed by AI applications – especially when Machine Learning and Deep Learning techniques are involved. The second issue is related to the fear of the potential bias in the data selection and elaboration that AI could exert on inputs originally set by humans\textsuperscript{21} and is further linked to issue of data privacy when gathering, storing, elaborating and transmitting data.

This is not to mention the ethical problems tied to AI. In this respect, two opposing theories have been emerging over the past few years. On the one hand, AI technology is not capable of producing outcomes in the way a “human conscience” does, whereas, on the other hand, AI is extremely influenced by human prejudices and stereotypes. Considering the first ethical objection, as AI does not have a human conscience, its choices and undertaken actions may be questioned and the trust in the overall AI technology may suffer. The second ethical issue focuses on the emergence of the Artificial General Intelligence (AGI) – an extremely advanced form of AI that capable to perform all the cognitive activities that are related to humans. As the machine is fed by humans and the first general rules are mankind-provided, what if the AI selection of decision-making process is biased by race, gender, or religious prejudices? Both these ethical issues can be mitigated and overcome by looking in more detail at the so-called “AI decision-making process black box”. According to a recent study published by PWC\textsuperscript{22}, a deeper and look into the “AI black box” is made of several steps:

- Transparency - first of all, there is a need of understanding; the AI decision making model needs to be fully apprehended and understood in its functioning;
- Explainability - secondly, the context and background reasoning of each decision taken by the AI system needs to be overtly explained;

\textsuperscript{19} In the business of service providers, the churn (or “churn rate”) is the number of customers or subscribers who discontinue the service during a given time period http://www.churn-rate.com/
\textsuperscript{20} The lack of trust partially comes from a still limited understanding of what AI is and how it works. People do not fully grasp how a machine can process information and have human-like reactions. Do another aspect that undermines the trust in AI is ethics, or, better, the advocated lack thereof. The alleged inability of AI and AI embedded devices to follow rules, policies and standards follows under this second aspect. https://blog.idcuk.com/augmenting-or-substituting-how-the-workforce-should-tackle-the-fear-of-ai/
\textsuperscript{21} Augmenting or Substituting? How the Workforce should tackle the Fear of AI (http://blog.idcuk.com/augmenting-or-substituting-how-the-workforce-should-tackle-the-fear-of-ai/)
• Provability – As the last steps, the understanding of the mathematical certainty of the decisions and the processes undertaken is needed.

While clear in theory, the practical applicability of these principles to the everyday business and to operational processes in the real world may prove to be much more complex and additional research in the field is needed.

The AI and BDA world is not straightforwardly black or white, where benefits/expectations and obstacle/issues are clearly defined. For instance, the IT implementation projected towards the future, with AI playing an increasing relevant role in leveraging automation and reducing the impact on workers of lower-level, manual-intensive tasks, is seen as a positive aspect. On the other hand, researchers are expecting an increasing need for more ICT and digital skills, an increasing competency and technology/technological skills of the workers. This will be translated into the need to look for a more/different talented pool of workers. The coupling of AI-based automation, employees and increasingly intelligent machines will leverage the need to re-evaluate and enhance or change business processes in new ways, alongside with the upskill and/or reskill of the current pool of workers.
3. The Case Studies - What We Found So Far

This story was developed through ad-hoc desk research and in-depth interviews with the aim to unveil concrete, real-life examples of AI and Big Data applications in Europe today. Five significant case studies were selected and are presented below. For each case study, we provide an introduction of the company/organization and the type of AI and BDA solution adopted, and explanation of the challenges that had to be tackled by the adoption of AI and BDA, and, more importantly, a short overview of the benefits achieved through the applications of AI and BDA.

The commonalities and horizontal themes pertaining to all case studies are further analysed in the subsequent chapter 4 below.

3.1 Case Study 1: Coyote System - Predictive Road Traffic

Background Information

Founded in 2005 by Fabien Pierlot and Jean-Marc Laethem, Coyote System SAS develops solutions and services for real-time road information. In 2014 the company employed 240 people generating a turnover of €105 million\(^23\). In 2017 Coyote Systems acquired Traqueur - a French organization specialised in tracking and managing stolen vehicles\(^24\). Coyote Systems describes itself as a real-time provider of information on speed limits, danger locations, traffic hazards and traffic conditions to an ever-growing community. The company operates a multiple platform serving connected devices, applications, and vehicle embedded systems.

Initially positioned on the market to help drivers detect speed radars as its core product, in 2009, the company shifted the commercial scope of its device from a radar detector to a detector of dangerous road and traffic conditions. For this purpose, Coyote used IoT devices and mobile applications to enable drivers to receive constant notifications of traffic hazards and conditions. The technological solutions developed by Coyote System SAS have been awarded several prizes, including Frost & Sullivan/ 2008 European Telematics Entrepreneurial Company Award (2008)\(^25\), Trophée de la Décennie (2015)\(^26\) and Blue Ocean Awards (2016)\(^27\) – mentor category. The organization spends 5-10% of the turnover on R&D, and in 2017 it opened the Coyote Lab in Québec, Canada. The research platform has the aim to accelerate the innovation of the road-information devices and market.

Business Challenge(s) and Solution(s) Adopted

Since the beginning of operations, Coyote relied on accuracy and speed on incoming real-time data to provide its services. Of particular interest were all the traffic management data such as traffic obstruction, accidents and related queuing, fixed and mobile speed cameras and, normal or temporary changes in speed limits. In 2015, Coyote received more than 18 million signals per month, with its service being used by 4,8 million users for more than 25 million of hours, covering more than 800 million of km and registering roughly 1,5 billion of GPS positions per month\(^28\). To manage all this information, Coyote needed an automated algorithm-based solution, which was able to leverage the increasing amount of incoming data from all the IoT connected devices. From data leveraging, Coyote could create effective and efficient previsions and insights of the real-time situation on the European roads.

\(^23\) Datiku – Case Study/Transportation: Anomaly detection

\(^24\) Communiqué de dépôt d’un projet d’offre publique d’achat visant les actions de la société Traqueur:

\(^25\) Awards Banquet Commends Stellar Performers in Multiple Industries, November 28th, 2008

\(^26\) Communiqué de presse: Trophée de la Décennie: Coyote monte sur le podium:

\(^27\) Blue Ocean Awards 2016: 8 PME récompensées, November 30th, 2016
https://www.entreprises.gouv.fr/services/blue-ocean-awards-2016-8-pme-recompenses

\(^28\) https://corporate.moncoyote.com/coyote-aujourd'hui/
In the light of these difficulties, Coyote started a partnership with Dataiku in early 2015. The solution offered to Coyote was Dataiku Data Science Studio (DSS). The Machine Learning capabilities of Dataiku DSS, based on Random forest modulization, enabled Coyote to leverage the data acquired through the IoT network, displaying anomalies in the speed limit referential. The embedded predictive model allowed Coyote to estimate the speed limit of each road section, previously identified, and the ML process simplified the detection of anomalies in the speed limits, to have a real-time and accurate estimation of the displayed speed limit on the connected devices. The Dataiku DSS, adopted by Coyote, is a logical all-in-one solution to: (1) connect, integrate & prepare data, (2) analyze, visualize data and forecast, and (3) automate the data packaging and deployment processes. For instance, the software can analyze millions of real-time data to predict behaviors, anomalies and insights. Alongside the adoption of Dataiku DSS, the company has leveraged different technologies as Python, HP Vertica, PostgreSQL implementing the random forests model. Coyote, within four months from the first roll out of the projects, was able to implement and start the whole process of Dataiku DSS adoption.

Benefits Obtained with AI and Machine Learning Deployment

The solution provided by Dataiku translated into four main improvements for Coyote:

- a single platform of easy usage by data scientists, analysts and business users alike;
- real-time collection, ingestion and analysis of IoT data to better follow the real-time traffic evolution;
- fast data iteration to optimize the random forests algorithm, and;
- constant (daily) re-training to improve forecasts accuracy and real-time data insights.

With the deployment of Dataiku platform, Coyote experienced an overall improvement in efficiency thanks to enhanced cooperation among its employees, (even between those with different skill-sets) as well as an improved ability to exchange data and information with partners and customers. More concretely, Coyote claimed of having been able to enhance its speed limit detection of 9% (on the dataset considered and tested against the previous method of data management and forecast) by fully automating the speed limit correction process. These results have had a positive effect on the company customer loyalty: according to Coyote, the company achieved an 11% increase of the high-performance outbound campaign, improving also the adoption rate for its products, thanks to the higher reliability of its products and the increased ease-of-use of its services. From a predictive analytics perspective, Coyote was able to improve its customer knowledge and extend the application of the prediction and forecasts to other business areas, such as online insurances and parking lot availability.

---

29 Dataiku – Case Study/Transportation: Anomaly detection

30 See footnote 9.

31 Dataiku – Case Study/Transportation: Anomaly detection

32 Valoriser les données – Symposium Présentation Coyote-Dataiku, June 4th, 2015
https://fr.slideshare.net/Dataiku/symposium-coyote-dataikufinal?next_slideshow=1
3.2 Case Study 2: Voyage Privé - Predictive Marketing

Background Information
Voyage Privé SAS is a French travel agency and e-tourism company headquartered in Aix-en-Provence. Founded in 2004, Voyage Privé provides offerings and discounts to their members in 4- and 5-star hotels and has today more than 31 million members worldwide. The final aim of the company is to offer to its members the same customer experience as a physical travel agency could offer. In addition, the organization is active in the innovation environment. For instance, in 2013 it announced the launch of an incubator for a summer university to support the emergence of new and innovative projects in the e-tourism market. This led to the creation of the first start-up company accelerator called "Les Tremplins de l'innovation E-tourisme"34 (the springboard of the E-tourism innovation)35. In 2016, the French organization acquired Splendia, a Spanish company operating in the reservation of luxury hotels and villas. This acquisition allowed Voyage Privé to increase the turnover and the presence on the market, but also to acquire a technology for CMS (Content Management System) and to ease and empower the relationship with the hosts and hotels36.

"We focus our energy on a very unique business activity that we are experts in - travel. This is a decisive element in our success and goes hand-in-hand with the passion that drives our teams and the power of teamwork” (Sébastien Tricaud – Co-founder of Voyage Privé Groupe)37.

Business Challenge(s) and Solution(s) Adopted
Voyage Privé has been using AI-enabled smart user segmentation to offer and display personalized offers to its members since the beginning of its operations. The organization's needs, though, were to fully understand members' desires and build customer-tailored offers to improve customer experience and satisfaction. To this aim, Voyage Privé needed to expand the ability of the Big Data-ingested recommendations system through the increased ability in capturing most of the preference signals from customers. To satisfy its needs, the e-travel agency needed a software solution, not only to extend the range and type of captured data but also to make sense of the increasing number of customers' data received that was contributing to daily enlarge the Voyage Privé's pool of data (over 100 GB of data per day). With the data extracted, Voyage Privé created an AI platform called ‘collaborative filtering’ through the

---

33 https://www.voyage-prive.co.uk/help/about
35 http://les-tremplins.com/
deployment of Dataiku Data Science Studio (DSS).\textsuperscript{38,39,40} The platform enabled the travel company to improve the ability to understand customers’ needs according to a mechanism of click path and bookmarking (a record of all bookings looked at and then the chosen ones), thanks to the deployment of AI and statistical techniques to elaborate and make sense of historical data. Alongside the platform, the software adopted allowed Voyage Privé to improve customer segmentation and obtain a non-rule-based approach to analyze the increasing amount of data collected, as well as the historical database retained. With this, Voyage Privé significantly improved client satisfaction through customer-tailored offers while augmenting each transaction’s profitability.\textsuperscript{41} A further boost to the company’s performance was the creation of a machine learning-derived score for each customer based on different criteria such as, for example, the time spent on the website, the number and type of bookings made, and more importantly, which offers were looked by customers but were not chosen in the end. Considering the ease of use of the AI-embedded platform (based on Drag & Drop interface), the processes of data diagnostic and iterations were significantly simplified. The high-level understanding of the online customer behavior, derived from the development of an AI and machine-learning approach to customer data, made it easy for Voyage Privé to offer travel opportunities with the highest likelihood of customer acceptance. Alongside the adoption of Dataiku DSS, the company has leveraged different technologies such as Python, HP Vertica, and Impala.\textsuperscript{42}

\textbf{Benefits Obtained with AI and Machine Learning Deployment}

With the adoption of the new AI-embedded platform, Voyage Privé was able to optimize the marketing and sales campaigns tailored for each customer. The optimization was translated into an increase in the likelihood of purchases, in the improvement of customer satisfaction and in a higher value per transaction. According to Voyage Privé, a 6% increase in total transaction value per member was achieved since the introduction of the new platform. With Dataiku solutions, Voyage Privé could re-run the collaborative filtering model and re-create over 100 million recommendations daily. As a result, Voyage Privé could significantly increase the revenue per member and open and set the scene for further operational and business improvements such as savings optimization, enhanced customer segmentation to further personalize offers and improved interactions with the members across different sites and devices.\textsuperscript{43}

According to Laurent Dupé, Head of International Marketing, the next step for Voyage Privé is to implement and extend the external dataset to refine customer and potential customers' scorings.\textsuperscript{44, 45}

---

\textsuperscript{38} Dataiku DSS – Success story: Advanced Analytics for Targeted Recommendations

\textsuperscript{39} As described by Laurent Dupé, “[at Voyage Privé], we tried through the recommendation system to offer our customers seamless experience on our website. We select only 300 key products each week. The idea is to pop up the 10 key products they were expecting even before they know they were looking for these products

\textsuperscript{40} Voyage Privé and Dataiku: Smart User Segmentation for Targeted Recommendations (published September 8th, 2016)

https://www.youtube.com/watch?v=hXiXXZzdIw

\textsuperscript{41} The AI platform is not only able to derive insights from historical data. For instance, upon the information driven by historical analysis, the algorithm within the platform is able to create recommendations and suggestions for each single customer, improving the satisfaction of each single person.

\textsuperscript{42} How a Travel Club Gets Personal Boost Sales, July 6th, 2016

\textsuperscript{43} Dataiku DSS – Success story: Advanced Analytics for Targeted Recommendations

\textsuperscript{44} Dataiku DSS – Success story: Advanced Analytics for Targeted Recommendations

\textsuperscript{45} Voyage Privé and Dataiku: Smart User Segmentation for Targeted Recommendations (published September 8th, 2016)

“With Dataiku we have plenty of projects on course. We are working on Fraud system, to optimize our savings (and costs). We are also working on Segmentation, to be able to personalize our offer and interaction with our members across different sites and devices”

https://www.youtube.com/watch?v=hXiXXZzdIw
3.3 Case Study 3: DAZN - Scaling Small Data Teams with ML

Background information

DAZN is a subscription video streaming service owned by Perform Group and founded in August 2016. The company has branches in Tokyo and New York and operates direct subsidiaries in Berlin and London. In terms of service delivery, DAZN first launched services in Switzerland, Germany, Japan and Austria (2016) and it later expanded into Canada (2017), the U.S. (2018) and Italy (2018). For 2019, the plans are to reach and penetrate two new markets - Spain and Brazil.

DAZN’s model of content provision is structured as an Over-the-Top (OTT) service - a video-on-demand service which provides the viewers with streaming media services bypassing most traditional content distributors (multichannel televisions, broadcast television platforms, etc.). The organization streams sports globally, offering more than 8,000 sporting events every year across a wide range of devices\(^{46}\). The words of DAZN’s Director of Data Science and Advanced Analytics best summarize the company’s mission and vision: "We work in the sporting arena and we deliver a live stream service of sports globally. We have decided to adopt a B2C model, where we’ve gone for live streaming. I guess the best way to think about our company is simple as this: we are the Netflix of live sports."\(^{47}\) (Shaun Moate – Director of Data Science and Advanced Analytics at DAZN).

Business Challenge(s) and Solution(s) Adopted

When DAZN started up, it was small in terms of human resources (a tiny team of 2-3 analysts) but very rich in database. The company's challenge was then to figure out what AI based technology could have been used to enable a small analysts’ team to derive insights from huge amounts of data and information. To tackle the challenge, the company adopted Dataiku DSS platform and so that, for instance, a small team of half a dozen analysts was able to get a full end-to-end process and had it running as quickly as possible. Once DAZN had set everything

---


\(^{47}\) How a Start-up is Scaling Analytics Globally – Shaun Moate @DAZN (published July 30th, 2018) https://www.youtube.com/watch?v=IFlgXt8ngb8
up, it got a full feedback loop ready to go. The team started to really think about the key questions to be asked and engaged with the first analytics stage, the exploratory modelling. The questions relevant for a subscription service like DAZN are:

- Why do people “churn”, i.e. why do they discontinue their subscription to the service?
- What types of behaviors do they exhibit?
- Why do they join the service in the first place? For example: Did they join for the El Clásico, which is the name given to any football match between Real Madrid and FC Barcelona or did they join for the World Cup?

The following step was to engage in survival analytics (why do people leave, how long it takes them to leave, and which might be the solutions to reduce or avoid this churn?).

With all those questions in mind, DAZN wanted to explore data and built an AI-native process around its customer-base in 5 main steps, creating a proper Customer journey48 along the following steps:

1) **Pre-acquisition**: every customer starts by not knowing the organization and the services offered, this entails empowering branding and marketing activities, the ROI from the campaigns, etc.
2) **Early customer experience**: then an ad-hoc model investigated the reasons and modalities of customers joining the service. Key questions in this step are: why did the customer joined, what devices are they using, when they join where do they come from? What channel did they come through, was it just through a paid search or through other types of channels?
3) **Forming habits**: once customers have subscribed to the service and had a positive early user experience, they start using the service more and more.
4) **Segmentation**: what sports they tend to watch? Are they American football fans, soccer fans, or darts fans? For instance, DAZN reported 57 customers in Canada that watch nothing but golf, and about 6000 users in Germany who watch nothing than darts.
5) **Survival**: once understood why and how customers are consuming the service (i.e. what they're watching, how they're watching, etc.), DAZN started to tag them up and tried to understand for how long they survive, how much value the are generating (Life Time Value – LTV), etc. As matter of fact, the English organization discovered that people streaming only on living room devices are much sticker and therefore generate more and more revenues on the service versus people on their mobile devices.

At the very start, DAZN had deployed Dataiku DSS on the Amazon Web Services (AWS) cloud, because it wanted to build an ever-increasing scalability and a full feedback loop. The AI-embedded technological solution adopted eased every process within DAZN. For example, every time the company faces the need to create a new segmentation model (what types of sports customers want, etc.), it takes to the team only two days of work to fully create an end-to-end (E2E) response on a system like Dataiku DSS, as it is a system where multiple technologies are embedded in a single, unique and extendible interplay. On a later stage of the transformation process, the types of analytics put in place were simplified for people without a strong and solid analytics background delivering ever-increasing benefits for company's business. At DAZN, many of the analysts employed had a poor technical background (i.e. lack of SQL coding knowledge), but for the organization this is not strictly relevant as Dataiku DSS enabled them to understand how to manage data, to transform it, and actually start asking relevant business questions. With AWS and Dataiku DSS, DAZN was able to (1) create the desired end-to-end feedback loop, (2) design the full data insights process, and (3) start to scale up and see some preliminary benefits.

---

48 This journey is firstly based on studying and understanding historical data on customers, deriving trends and common patterns of behavior, consumption model, preferences, geographical peculiarities, etc. On a second stage, the AI system enables DAZN to offer personalized offerings to increase customer satisfaction and the user experience. Lastly, the platform is able to study and predict future customers behavior according to the ‘clusterization’ process performed in the first stage of understanding past behavioral trends (i.e. churn prediction, etc.).
Benefits Obtained with AI and Machine Learning Deployment

Through the deployment of Dataiku DSS, DAZN has been able to deliver a huge amount of value to the business with small teams. According to Shaun Moate, "the organization has been able to: (1) have small, agile and scalable teams, (2) have the right tools in the right place given to the right people, enabling teams to deliver full end-to-end products, (3) have each team both asking and understanding the core and right questions to gain insights and transform them into business value, and (4) have a fully connected collaboration environment across countries and time zones". These four conditions enabled the creation of a rich selection of analytics, model and products that are the key strengths of DAZN. As Shaun Moate highlights⁴⁹, the deployment of Dataiku and AWS by DAZN opened up several opportunities:

a) **Basic modelling** allowing DAZN to better analyze:
   - Propensity to churn;
   - Customer segmentation;
   - Survival analytics;
   - Attribution modelling;
   - Early experimentation of Natural Processing Language (NPL);
   - Customer Life Time Value (LTV);

b) **Data products** allowing DAZN to better perform:
   - Survival Cave Plug-in: creation of a single, worldwide applicable and standardized plug-in for non-scientists to deploy mobile analytics into their workflows;
   - Customer targeting (retention activity);
   - DMP segmentation/Look-alike Models (marketing);
   - Content recommendations (engagement);
   - Automatic control and test groups;

By applying AI technologies, DAZN brought about an interesting paradigm change on how employees work together. For instance, at DAZN each data scientist is not required to produce and deliver every single model and data product, nor he or she must be proficient in complex coding techniques or fully knowledgeable of APIs functioning. As a result, the complementarity of different sets of skills among employees is significantly brought up and the company creates a working environment where the learning curve is shorter and where it is easier to find talents within a wider candidates' pool.

According to DAZN⁵⁰, the introduction of AI solutions made its data team become 2.5X more efficient and productive in putting models together. This, in turn, translated in an overall improvement of the company’s Return on Investments with significant huskiness and operational advantages. In particular, DAZN has used AI to:

- cluster, understand and make sense of data relating sports events streamed by customers to obtain more detailed insights on users’ behaviors and to build better processes for customer segmentation. These insights on customers’ behaviors led to a better recommendation system. Armed with this additional in-depth knowledge, DAZN can create customized messages and recommendations just like other media and entertainment services do (i.e. Netflix – an American media service provider – sends weekly emails with the object "What do you want to watch this evening?");
- generate detailed and more accurate forecasts of the basic modelling activities, such as churn, Customer Life Time Value, etc. AI prediction and trend spotter abilities allow for both a deeper analysis of past data and information (historical analysis) and a better understanding of most likely future developments.

⁴⁹ How a Start-up is Scaling Analytics Globally – Shaun Moate @DAZN (published July 30th, 2018) https://www.youtube.com/watch?v=IFlgXt8ngb8
⁵⁰ How a Start-up is Scaling Analytics Globally – Shaun Moate @DAZN (published July 30th, 2018) https://www.youtube.com/watch?v=IFlgXt8ngb8
3.4 Case Study 4: AXA - Predictive Large-Loss Cases

Background Information

AXA SA is a French multinational insurance headquartered in Paris. The company deals with insurance, investment management and other related financial services. AXA operates in Life, Health, Protection and property and causality (P&C) segments and is also active in the funding of scientific R&D projects and innovation laboratories - AXA Lab for technology, for example, monitors a series of innovation processes in the Silicon Valley51 with the aim to anticipate digital disruptions, identify the most promising start-ups, test new possible usage of mobile technologies and solutions and verify the employment of connected objects and devices (including the use of social networks on the insurance market and the first-class deployment of big data).

Considering its extensive business, AXA needed a simple and effective solution to manage the insurance procedures. One of the most important segments for the company was the car insurance where approximately 7-10% of clients cause car accidents, but only 1% of them are the so called large-loss cases, which require a payout of more than $10,000 per accident. AXA wanted therefore to better understand in the behaviour and risks related to these large-loss cases and turned to AI technologies52.

Business Challenge(s) and Solution(s) Adopted

The main challenge for AXA was to understand whether a policy subscriber might risk causing a large car accident so to better optimize the pricing and conditions of the policies offered and/or renewed. The company R&D team based in Japan started with an AI and machine-learning technique called Random Forest53. The technique was built on an algorithm using multiple decision trees as predictive model. Even if the model was functioning and was well

---

51 Behind the scenes at AXA Lab in Silicon Valley, June 18th, 2015
52 Using machine learning for insurance pricing optimization, March 29th, 2017
53 See footnote 19.
adopted/implemented, the prediction accuracy was inadequate. For instance, only less than 40% of the total large-loss cases were detected\textsuperscript{54}. To fill the gap and improve its predictive power, AXA decided to build a POC (Proof of Concept) using TensorFlow\textsuperscript{55} (an open source library released in 2015, which is able to democratize machine learning for any user), which is part of Google Cloud’s Machine Learning Engine. With the deployment of this solution, AXA gained significant advantages in policies pricing and cost optimization, in addition to the possibility to manage real-time pricing at point of sales. The Google Cloud based tool takes information, from the extensive pool of data on AXA's customers, for 70 parameters, already analyzed by the previous random forest model such as: age of the driver, age of the car, annual insurance premium range, number of actual driving, region, information of previous insurance policies, etc., and, based on the value of the parameters, the AI-embedded computing engine models the values of the parameters to obtain more reliable answers. The model, designed as a neural network with three hidden layers, is backed by a ReLU (Rectifier Linear Unit) as activation function\textsuperscript{56}. The deployment of the ReLU as activation function allowed the model to better train the deeper networks, thus increasing the non-linearity of the model. The deployment of Google Compute Engine was dedicated to the training of the TensorFlow model while the Cloud ML Engine’s HyperTune add-on was dedicated to tune the hyperparameters – used to run the model\textsuperscript{57}.

**Benefits Obtained with AI and Machine Learning Deployment**

The deployment of Google Cloud solutions allowed AXA to increase the reliability of information up to 78%, figure much higher than the previous 40% forecast provided by the random forest model\textsuperscript{58}. With the increased accuracy, AXA was put in a position to significantly optimize its pricing processes, an interesting example being the real-time insurance pricing at any point of sale. Along with the increase in accuracy, the AI and ML solutions deployed by the company were used to predict potential litigious cases, possible exaggeration of damages (billing frauds) and to better identify compliance policies that match a particular damage. In addition, the deployment of TensorFlow and Google Cloud Machine Learning Engine relieved AXA from the need to build and maintain an on-premise ML infrastructure, allowing engineers to focus on machine learning applications and use cases\textsuperscript{59}. Again, as shown in other case studies put forward in this story, Big Data and Analytics technologies were instrumental in AXA for the effective deployment and implementation of AI technologies: the seamless and real-time provision of fresh data allowed the company to feed the AI/ML-embedded model and more generate more accurate predictions of large loss causes.

---


\textsuperscript{55} In order to validate the model and the ensuing designed predictive solution, AXA preferred a first trial on TensorFlow prior to a full-out investment on a complex and expensive AI and ML architecture.

\textsuperscript{56} In the generalized context of artificial neural networks, the linear rectifier is an activation function, that aligns and allows the process of knowledge incrementation to start and this process is useful to better train Deep Neural Networks. A rectified linear unit (ReLU) is the unit that is employing the rectifier.

\textsuperscript{57} Using machine learning for insurance pricing optimization, March 29\textdegree, 2017 https://cloud.google.com/blog/products/gcp/using-machine-learning-for-insurance-pricing-optimization

\textsuperscript{58} Using machine learning for insurance pricing optimization, March 29\textdegree, 2017 https://cloud.google.com/blog/products/gcp/using-machine-learning-for-insurance-pricing-optimization

\textsuperscript{59} ACAMS, Financial Services: Machine Learning Use Cases and Solutions (Google Cloud), May 15, 2018 http://files.acams.org/pdfs/2018/ai-presentation-by-google-5.15.18.pdf
3.5 Case Study 5: Airbus (D&S) - Cleaning Skies from Clouds

Background Information

Airbus Defence and Space GmbH (Airbus D&S) is a division of Airbus that deals with defense and aerospace products and services. The company was previously known as EADS Deutschland GmbH but changed name in 2014, when it was branched out to better manage the defense and space business units. The organization offers several products and services, including military aircrafts, border security, secure land communications, geo-intelligence, radars and identification friend-or-foe systems, electronic warfare systems, high-security national solutions, consulting services, etc. It is the second largest space company at world level and is ranked among the first ten defense companies in the world. Together with the IT services corporation Atos, Airbus D&S’ CyberSecurity unit was selected in October 2018 by the Council of the European Union to grant cyber security expertise, products and solutions to help 17 European institutions, services and agencies.

Business Challenge(s) and Solution(s) Adopted

Among the projects rolled out by Airbus, One Atlas is one of the most recent and important. The project was created to deliver demonstrable value to Airbus customers planning defense or security missions, but also to agriculture, maritime, and other markets. Airbus does not use One Atlas for internal purposes only, but it also serves B2B markets, providing data and data-related services also to companies whose business is related to cartography, mapping, and other geo-localization services. Airbus’ One Atlas is a repository of high-quality and high-definition satellite imagery of the hearth – currently managed through Google Cloud. Although images have always been provided in high-quality and high-definition standards, the issue of imagery cleaning and understanding was a constant requirement for Airbus. For a prolonged period, Airbus D&S’ engineers did sort manually all the piece of information contained into the

---

61 https://www.airbus.com/company/we-are-airbus.html
62 https://www.airbus.com/defence.html
images received from the satellites. The main issue was to manually detect and define clouds (and land features like snow cover) in more than the 10,000 satellite images sent daily and simultaneously. The task had to be performed in a near real-time to scale up as fast as possible the size of imagery queue. In addition, the irregularity of the workload, defined as seasonal – in case of the normal map update – or ‘on-spot’ – when a customer asks for a particular task – had a strong impact on the queuing issue. The process required high skilled engineers and was extremely time consuming. The number of pictures funneled into the pipeline required more than 50 hours of human work, that considering a standard working day of 8 hour, meant at least 6 engineers full-time employed to interpret images. In addition, Airbus had to take into account the human error involved in this task.

Airbus A&D adopted Google Cloud Machine Learning Engine to tackle these issues. The first step was to connect the satellites with Google Cloud where the images were stored and then analyzed within the internally built Intelligence Digital Platform. The Airbus team combined multiple Convolutional Neural Networks and fully connected Neural Network to perform the main task of clouds and interfering objects detection in satellite images. Throughout the platform, images were automatically scanned and cleaned from clouds, planes and other interfering objects thanks to the deployment of AI TensorFlow and the Machine Learning tools. Subsequently, the organization deployed the HyperTune in the ML engine to automate all hyper-parameter tuning (task previously done manually). The team needed only one month to start getting the first promising results. In addition, Airbus recently started to deploy the GPU instance to reduce the training time for the predictive model by two orders of magnitude.

**Benefits Obtained with AI and Machine Learning Deployment**

To tackle the issues, the first move undertaken by Airbus D&S was to move from the on-premise imagery storage to the cloud, adopting Google Cloud.

With the application of the Google Cloud ML Engine and the HyperTune feature, Airbus experimented a 40x time reduction in the imagery analysis, from 50 hours to 30 minutes, with an evident cost reduction in engineers’ expenses and time consumption. The combination of multiple convolutional neural network with a fully connected neural network also allowed a reduction of the detection error rate from 11% (mainly based on the human error) to 3%. The adoption of Google Cloud ML solutions scaled up Airbus performances and reduced costs.

The next step for Airbus is to ‘update’ the online catalogue with the 30 million images shot in the past and to increment the uploading process from the current 500 terabytes up to 2 petabytes of images per year – in practice, quadrupling their yearly storage capacity. Moreover, the One Atlas service allowed Airbus D&S to enlarge its customer base, opening up the service also for smaller companies (mainly in the agricultural sector) and to NGOs. These organizations can now take advantage of the new and cheaper solutions, such as pay-as-you-go/consume data streaming rather than the more common and expensive subscription models. With the flexibility and scalability provided by Google Cloud Platform, One Atlas has been enriched by incremental features as Data-as-a-Service (DaaS) compared to the old and simple BaseMap service. Finally, One Atlas will offer solutions such as, Software-as-a-Service (SaaS) – image-processing tools and customizable analytics on satellite and drone imagery –, and Platform-as-a-Service (PaaS) – offering to less AI and ML literate customers own tools and third-party analytics.

---

65 How Airbus scales pools of worker for satellite images processing – Google Cloud Next’17 (published March 8th, 2017) [https://www.youtube.com/watch?v=Xe0eVdx-tvE](https://www.youtube.com/watch?v=Xe0eVdx-tvE)


68 How Airbus scales pools of worker for satellite images processing – Google Cloud Next’17 (published March 8th, 2017) [https://www.youtube.com/watch?v=Xe0eVdx-tvE](https://www.youtube.com/watch?v=Xe0eVdx-tvE)
3.6 **Common Themes and Findings across the Case Studies**

The table below summarizes the common themes and findings across the 5 case studies featured in this piece of research. For the clarity of the analysis we propose a horizontal overview taking into account the specific challenges triggering the use of AI, ML and BDA technologies, the immediate technical or process-oriented benefits obtained through the application of AI, as well as the wider impacts that were possible to observe in terms of operational or business benefits at the level of the overall organization.

<table>
<thead>
<tr>
<th>Specific Challenge</th>
<th>Specific Benefits Obtained through AI/ML/BDA</th>
<th>Wider Impact Obtained through AI/ML/BDA</th>
<th>Area of AI/ML/BDA Deployment and (Case Study)</th>
<th>BDA-related Usage and Benefits</th>
<th>AI-related Usage and Benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Huge amount of diversified data, across a vast geographic area and from multiple sources, to be analysed in real time for a large number of users</td>
<td>• Ease of use internally (for data scientist) and externally (for customers and users). • Faster data iteration and optimized algorithm. • Improved forecast accuracy through constant machine data feeding and retraining</td>
<td>• Improved teamwork and employee cooperation; • Improved customer loyalty and marketing efforts (e.g.: +11% improvement in marketing outbound campaign) • Improved quality of products and services (e.g.: +9% improvement in speed limit detection)</td>
<td>Predictive Road Traffic (Coyote System)</td>
<td>• Real-time feed of notification of hazards and road conditions/traffic management information • Ever accurate prediction of road traffic based on previous experiences/past data (analyses of historical data)</td>
<td>• Data preparation and data cleaning to: create real-time suggestions and notifications of road traffic/conditions/car accidents, etc. improve recommendation system of speed limits/parking lot availability</td>
</tr>
<tr>
<td>Need to further customize offering through better interpretation of (new) streams of customer data and improved segmentation</td>
<td>Expanded type and range of data captured and interpreted; Better interpretation of customer’s preference signals (ML-based score for each customer); Simplified data diagnostic processes and better ease of use data analysis techniques</td>
<td>Optimized sales &amp; marketing campaigns; Increase in the likelihood of purchase; +6% of total transaction value per customer; New opportunities in adjacent business areas</td>
<td>Predictive Marketing (Voyage Privé)</td>
<td>More granular data from customers (who purchased holidays, saw but did not purchase, used keywords, etc.) to feed customer segmentation and recommendation system</td>
<td>Further improve customer segmentation accuracy and the recommendation model to offer the best holiday package to each customer.</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Enable a small team of data scientist to make the most and extract the highest value from huge amounts of data</td>
<td>Simplified analytics exploitable by analysts with limited technical background; Improved and shortened learning curve for analysts and talent-oriented work environment; Free analysts time and efforts from process complexity to focus on key questions of data analysis</td>
<td>Data scientists and teams becoming way more efficient (2.5x); Small agile and scalable data scientists’ teams; More than 30 models easily scaled and managed;</td>
<td>Scalability of scarce resources (DAZN)</td>
<td>Better understanding of customer behaviours (churn, consumption options, sport/types of sports preferred, team(s)/player(s) supported, Improved evaluation of Life Time Value per customer/cluster of customers</td>
<td>Improved modelling capabilities of customer behaviours to predict trends of consumption models (what customers are going to watch and how)</td>
</tr>
<tr>
<td>Identify, intercept and predict the emergence of “large-loss cases” to reduce company’s risks</td>
<td>Constant and simultaneous tune-in of 70+ parameters to increase predictive power; More advanced predictive model with deeper self-learning capabilities</td>
<td>Increased forecast reliability (+78%); Optimized policy pricing; Real-time pricing at PoS; Improved fraud detection</td>
<td>Predicting large-loss cases (AXA)</td>
<td>Ability to cluster data into the 70+ parameters and feed the predictive model for large loss causes</td>
<td>Effective and more accurate prediction of large loss causes resulting in significant cost reduction for the insurer</td>
</tr>
<tr>
<td>Process huge amount of satellite images in near-real time, as fast as possible and according to an irregular workload</td>
<td>Identified and removed clouds and interfering objects in satellite images; Reduced training time for teams operating the</td>
<td>Reduction of detection error rate from 11% to 3%; 40x reduction of time to analyse imagery (from 50 hours to 30 minutes)</td>
<td>Cleaning skies from clouds (Airbus)</td>
<td>Usage of additional amount of data, with a higher level of granularity to create cartography, mapping and geo-localization services through the high-definition satellite imagery</td>
<td>AI/ML remove objects as clouds and airplanes/other objects from the high-definition satellite pictures to provide a good/clear picture or set of pictures</td>
</tr>
</tbody>
</table>
In terms of the triggering challenges, the common theme emerging from the case studies is twofold:

- From the one hand, we observe the need of processing, storing, accessing and treating an ever increasing and vast amount of complex data, from multiple sources and different types of devices, from different geographic areas, and according to irregular incoming batches or timelines. This is primarily the case of Coyote Systems and its predictive road traffic solutions, DAZN and its need to face huge quantity of data with a relatively small team of data scientists, and Airbus and the complexity related to the processing of large numbers of satellite images according to irregular workloads.
- From the other hand, the identified challenges focus on the need to better interpret and extract predictive value from the available data. While not necessarily discounting the fact that data have to be acquired and processed in a timely and efficient manner, this challenge stresses the need that, once processed, the data have to be interpreted, understood and “acted upon” so to provide better services or products to customers and other stakeholders. This is the case, for example, of Voyage Privé and its optimized customer segmentation and AXA and the need to reduce large-loss cases.

The case studies presented in this research highlight a number of immediate benefits obtained by the organizations adopting AI/ML and BDA technologies. In essence, they can be summarized as follows:

- **Ease of use and a simplified technology**, freeing data scientists from performing burdensome tasks is a recurring theme in companies such as DAZN (where a very small group of analysts with limited technical background have been able to turn-around the amount of data at their disposal), Coyote (with an increased ease of use of its devices thanks to the adopted technology) and Airbus (with its reduced training time for teams operating the new predictive models).
- **More advanced predictive capabilities**, allowing for a better interpretation of customer needs and improved forecast accuracy is another immediate benefit obtained through the direct application of AI/ML in different case studies, for example Voyage Privé in its ability to better interpret customers’ preference signals or AXA and its more advanced predictive model with deeper self-learning capabilities.

As for the wider impacts achieved through the deployment of AI, ML and BDA technologies, they seem to follow a three-way path:

- The first path is the **efficiency obtained by data scientists and the internal users of data** - DAZN for example reports an increase in data scientist productivity of 2.5x with the ability to manage and scale more than 30 data model by a very small and agile team of data scientist. In a similar fashion, Coyote maintains that the application of AI and ML solutions has considerably improved the teamwork and cooperation across its employees.
- The second path is the **impact exerted on the quality of products and services** – AXA claims an increase in forecast reliability of 78%, for example, Airbus a reduction of error rate from 11% to 3% and Coyote an improvement in speed limit detection of 9%.
- The third path revolves around the **overall business benefits** achieved through AI, ML and BDA. Voyage Privé registered a marked improvement in sales & marketing campaigns and a 6% increase in total transaction value per customer; Coyote, again, has been able to strengthen its customer loyalty and marketing efforts claiming an increase of marketing outbound campaign of 11%. 


26
The impacts outlined above are in line with the results obtained by ongoing market research in the field of Artificial Intelligence, Machine Learning and Cognitive Solutions. IDC's European Cognitive and AI Practice conducted a survey in September 2018 to understand the attitudes, concerns, and plans of decision makers and influencers involved in cognitive/AI initiatives within European companies across Respondents were drawn from C-level, IT, and line-of-business (LOB) functions.

The goal of the survey was to identify companies' short- to medium-term plans for cognitive/AI technology and solutions and revealed that AI is already impacting a considerable number of businesses. More than a third (37%) of firms had already quantified benefits, while 37% expected benefits to come in two years. What is more, among those respondents that did expect some form of benefits out of the AI deployment, another 36% reported benefits either exceeding or meeting their original expectations (see Figure 7 below).

**Figure 7 Expected and quantifiable benefits from AI deployment in Europe**

![Figure 7](image)

Source: IDC, September 2018

AI-Based Automation is shaping across industries and range from task to activity to process level automation balancing human-machine collaboration between producing insights, making decisions and acting on decisions. In another study investigating IT spending forecast in AI and Cognitive Solutions for the period 2018-2022, IDC found that technologies such as mobile solutions, Big Data, cloud, and social media are now 'the new normal' and that the new spotlight is on next-generation technologies, Thanks to their ability to streamline back-office processes and improve customer-facing tasks, artificial intelligence (AI), together with Internet of Things (IoT), blockchain, augmented and virtual reality (ARVR), and robotics is now a key focus for all industries. Some sectors, such as retail and financial services, are already more impacted than others. As an example, Sephora – a France-based multinational chain of personal care and beauty stores – implemented AI solutions by operating Facebook messenger chatbots with the aim to increase shop visits and sales. In the U.S. this translated into an increase of 11% in booking rates in 2018 and with an average of US$ 50 (approximately EUR 44) spent by every chatbot driven customer in the same period. In a similar fashion, NatWest - a major retail and commercial bank in the United Kingdom – was able to detect almost than 7 million GBP (more than 8 million EUR) in corporate fraud and avoid substantial related fines thanks to the adoption of AI-powered anti-fraud solutions (see Figure 8 below).
Figure 8: Examples of impactful benefits of AI adoption in retail and banking

**Sephora's Facebook Messenger Chatbot**
- Goal: **Increase store visits and sales**
- +11% in booking rates
- $50 average spent from chatbot-driven customers

**NatWest AI-Powered Anti Fraud Solution**
- Goal: **Detect suspicious activities**
- £7 million corporate fraud detected
- Avoided millionaire fines
4. Preliminary Policy Conclusions

While still at an emerging stage, AI is a reality in rapid development that is expected to revolutionise and heavily disrupt nearly every industry over the next few years. From banking to retail, from the manufacturing industry to the insurance and finance sector, countless AI-applications based on Big Data technology are gaining momentum, impacting both the business environment, as well as the everyday life of citizens worldwide.

For this to happen, though, several challenges remain to be overcome. There is still too little understanding of the elaboration mechanisms deployed by AI applications, especially when Machine Learning and Deep Learning techniques are involved and a general lack of trust in the results produced by this technology is still widespread. This, in turn, leads to different forms of cultural and ethical resistance to AI approaches as well as to numerous security concerns and unclear consequences on the labour market and working environment resulting from the AI adoption. There no doubt, however, that the stakes around AI are high and that Governments and industries around the world need to compete to secure the necessary investments and attract, acquire and develop the right AI talents.

Regulatory initiatives and policy-related interventions with the aim to offer guidance towards a successful implementation of AI approaches are therefore starting to arise. In April 2018, the European Commission published a European strategy proposing a human-centric AU approach that places people at the centre of the development of AI. The strategy was followed by coordinated plan detailing the actions to be started in 2019-2020 and prepare the ground for common activities to foster AI investments at EU and national levels over the next few years.

Among the major objectives and efforts outlined by the plan, the European Commission advocates:

- A significant increase in public and private investments in AI to reach the target of 20 billion Euro per year over the next decade with the aim of giving Europe the means to fully benefit of the opportunity offered by AI and continue develop AI solutions in the continent;
- A renewed collaboration between academia and industry fostering new research and innovation partnerships on AI;
- A strong support for the up-skilling of workers through ad-hoc learning and training programmes to better prepare the society to AI;
- A reinforcement of the Digital Single Market’s objectives to remove the obstacles due to the current market fragmentation in Europe and make it easier for businesses to harness and foster digital trade across borders;
- The creation of common European data spaces to make sure that an efficient data ecosystem is at play in a number of European industries such as manufacturing or energy;
- The development of ethics guidelines to overcome the current lack of trust in AI developments and foster a technology that is predictable, responsible, verifiable and respectful of fundamental rights and basic ethical rules.

The select case studies outlined in this research feature all, to a different extent, a successful combination of BDA and AI technologies. Whether in the form of more internal efficiency for analysts and data scientist teams, improved quality of products and services, or wider business benefits in terms of sales and marketing, BDA appear to be instrumental in providing seamlessly large and accurate amount of data to AI technologies, which, in turn, are of paramount importance in bringing about concrete and quantifiable benefits in different industry sectors. The present AI expansion, as well as its significant growth potential which is expected over the next few years, is directly dependent on the availability of huge data sets combined with increases in computing power, connectivity, and data-related technologies. Further action to facilitate the sharing of data held by public and private sectors in Europe, as well as a coordinated strategy in Research & Innovation investments to ease access to connectivity, interoperability, aggregation of public data, or to support the development and implementation of a data infrastructure to enable the management and sharing of data in real-time should be fostered at EU level, as well as at national level in coordination with the European Commission.

---

Augmenting or Substituting? How the Workforce should tackle the Fear of AI (http://blog-idcuk.com/augmenting-or-substituting-how-the-workforce-should-tackle-the-fear-of-ai/)


Manufacturing in Need of AI Experts: Where do they struggle finding the right talent? (Doc #EMEA44230518, Aug 2018 – Spotlight https://www.idc.com/getdoc.jsp?containerId=EMEA44230518)


Future Challenges for European Leadership in the Global Data Economy and Data-Driven Society: Input to Framework Program 9 (BDVA) – March 20118


https://aitopics.org/search

AI Momentum, Maturity & Models for Success – Based in findings from a global executive survey

The Technology Stacks of High Performance Computing and Big Data Computing – What they can learn from each other (European Technology Platform for High Performance Computing & Big Data Value Association)

How a Start-up is Scaling Analytics Globally – Shaun Moate @DAZN (published July 30th, 2018) https://www.youtube.com/watch?v=IFlgXt8ngb8


http://www.performgroup.com/brands/opta/