ARTIFICIAL INTELLIGENCE PATHWAYS *and* OPPORTUNITIES

A View from Portugal

By
JOÃO CASTRO
VASCO TELES

September 2020

FUNDAÇÃO
FRANCISCO MANUEL DOS SANTOS

Come mothers and fathers
Throughout the land
And don't criticize
What you can't understand
Your sons and your daughters
Are beyond your command
Your old road is rapidly agin'
Please get out of the new one if you can't lend your hand
For the times they are a-changin'

Bob Dylan – The times they are a-changin', 1964

Executive Summary

Interest in Artificial Intelligence research and its applications has grown considerably given the present computing capabilities. We analyse the activity dedicated to the development of this technology in the Academia, its different applications and the related industry and governance regulatory constraints in the Portuguese context. We conclude that Portugal has qualified people and is well equipped to explore this technology and that the main issues surrounding its future uses and impact are common around the world and not specific to Portugal. Most participants in this study suggest a large-scale social debate on the current and future uses of such a powerful technology.

Introduction

Artificial Intelligence technology is several decades old, but now it seems more able to fulfil some of its early promises. There are many uses for this technology and multiple examples of its applications around the world. This report analyses the Portuguese context and the role of this technology in the country.

When we first designed and planned this study, we were far from imagining the impact COVID-19 would have on the world, and society at large, in just a few weeks. Looking back at the notes from our early meetings and, now, at the end result, the opening words of Dickens' "Tale of Two Cities" inevitably come to mind: "It was the best of times, it was the worst of times…".

It was the worst of times because the original plan did not survive long enough. Nationwide, lockdowns constrained our movements, and we became unable to conduct extended interviews with the intended AI experts. As a stop-gap measure, we tried alternative teleconferencing solutions and, although they were better than nothing, we quickly realised they could not replace the richness of direct interaction.

Many of the experts on our "Wishlist" became unavailable since their companies were facing disturbances and uncertainties of potential catastrophic magnitude and, therefore, their attention was rightfully elsewhere. We kept in touch, but from a distance, which changed the early stages of data collection for this work.

Our plans to visit companies, walk the premises and register actual uses of AI were also suspended due to nationwide sanitary lockdowns. As we write this report – September 2020 – many facilities still face severe restrictions or have security measures in place, which compromise the purpose of our visits.

Nevertheless, there was still room for some "It was the best of times" optimism. The people we were able to talk to had been forced out of their daily management routines and day-to-day operations as their organisations were trying to adapt to the new circumstances. While some companies were striving to deploy VPN and conferencing solutions to enable their employees to work from home, others were inevitably forced to think in the long term. What impact will the COVID-19 pandemic have on consumers? How long will it last? Will the market rebound in a U, V, L, or W shape? What future markets will there be? How will people work in the post-COVID future? What strategies and solutions will organisations need to implement in order to thrive? With so much transformation, is this the best time to foster new ideas?

All these questions were valuable in our conversations. Our interviewees were not focused on their yearly outlook but were, instead, thinking three or more years ahead. For a study trying to collect perspectives on the future, this certainly helped.

This type of thinking was very favourable, considering the subject of this study, and the analysis of the future impacts and opportunities of AI.

This study is broken down and organised into four distinct sections:

I. Over the last few years, Artificial Intelligence (AI) has been increasingly mainstreamed and adopted by industries. This report mimics this transition, and begins with a brief overview of the current research activity in Portugal within this field, highlighting some of its main actors and research institutions.

II. In the transition from Academia to practice we review three short studies on three underlying themes in this report, namely, understanding the role of AI in society, understanding its prevalence and how to deal with some of its inherent bias, and understanding how organisations are using and implementing it.

III. Thirdly, we explore different topics concerning the future uses of AI technologies, as well as perspectives from Portugal, by interviewing practitioners in different roles, and from different organisations and industries.

IV. Finally, as a thought exercise, we try to combine all the information we came across on a brief essay.

We hope you enjoy this work and that it provokes your thoughts about the future and what society is building. We are always eager to learn more, so feel free to reach us at:

joaoc@alum.mit.edu
vasco@vascoteles.pt

Acknowledgements

The authors would like to thank and acknowledge the valuable support and contribution of the following people and institutions:

Firstly, we would like to thank the Francisco Manuel dos Santos Foundation for inviting us and giving us the opportunity to explore this topic and share our findings in this report. Without the support of Gonçalo Matias, Mónica Barreiros, and João Gaspar this work would not have been possible.

Secondly, we would like to thank all the organisations and people who shared their time, knowledge and perspectives with us, or simply put us in touch with the right people with whom we were able to engage in yet another insightful conversation. We sincerely thank Alexandre Santos, Miguel Moreira, Vanda Jesus, Ana Casaca, Pedro Viana, Joana Sá, Hugo Rosa, Hugo Macedo, Paulo Ribeiro, Luis Teixeira, João Ascensão, Pedro Faria, João Vinagre, João Moreira, and all the others who, given the nature of some conversations and our agreement with them, preferred to remain anonymous. You know who you are and we are indebted to you.

Lastly but definitely not least, we would like to thank our friends and family for their continuous support and encouragement during what has been a strange year.

We are all here because we are all together.

Part I

ACADEMIA IN PORTUGAL

As will become apparent throughout this report, Artificial Intelligence is a murky beast.

Research into Artificial Intelligence had a strong call to action when a group of pioneers gathered for a conference in New England in 1956. While historians seem to be able to pinpoint precisely "when" this research started, "what" was being researched remains somewhat elusive.

No single precise definition of Artificial Intelligence is shared between experts or practitioners, as it feeds from and entangles different concepts from disciplines such as philosophy, engineering, mathematics and computer science, and different levels of proficiency in "intelligence".

The same confusion is also apparent within the industry. The lack of a precise definition of what AI is, or is not, has also led the way for an abundant misuse and misappropriation of this technology by some companies, who claim they are using or designing products with AI when, in reality, there is no AI to be seen other than that made up by the marketing department. (MMC Ventures – State of AI 2019: Divergence, page 99).

For this reason, in this report, we do not try to: (1) clarify the definition of AI nor (2) identify and study all those who claim

to be linked with AI. However, we do list those we consider to be the most relevant academic and research laboratories in Portugal working with Artificial Intelligence.

Artificial Intelligence has been quite an active subject of Research and Development (R&D) in Portugal for the past few years. Some of the most senior researchers in the main R&D Centres have produced relevant scientific work in numerous areas of the study of AI and its applications.

To identify these areas of R&D, we conducted a brief and non-exhaustive analysis of the relevant literature published between 2018 and 2020. This analysis comprised 13 Portuguese researchers (h-index ≥ 25) and identified 16 different and broad R&D areas of study, namely, and in alphabetic order:

1. Biotechnology
2. Consumer
3. Finance
4. Health
5. Human-robot interaction
6. Image
7. Media/News
8. Quality
9. Sensors
10. Smart Cities
11. Social Networks and Web Communication
12. Sound/Speech
13. Telecommunications
14. Text
15. Transportation and Mobility
16. Vision

From all of the above areas, health is by far the most researched one. Most researchers in this group have published, at least, one work in this area. They address major health challenges such as cancer (e.g., brain, breast, cervical, and lung cancer), by assisting on the prediction, diagnosis, surgery, reconstruction, treatment and evaluation of the progression of the disease. An example of a tool supported by AI that helps monitor diseases is medical image analysis.

Other significant health challenges addressed by AI include, among others: Parkinson's Disease, by detecting movement patterns; Diabetes diagnosis; the evaluation of cardiovascular/stroke functional outcomes; kidney failure evaluation and kidney care; forensic evaluation of sexual assault; the evaluation thyroid dysfunctions; and injury recovery.

Other transversal topics range from recognising and analysing human activity through smartphone sensors, to simulating emergency medical services, supporting policy and decision-making, developing fall detectors, or, in a broader sense, analysing medical data for a variety of purposes.

These scientific publications are more frequently developed under the scope of projects in consortia, and supported by regional, national, or European/international R&D funding. In Portugal, the INESC TEC has been broadly leading the work developed and published on AI in the health field.

Following health, many other R&D areas are receiving similar attention from researchers, with a close number of publications identified for the referred period. Namely, sound and speech

analysis, and specifically natural speech recognition and processing, have been widely researched in recent years, especially by the INESC ID. Specific applications of such work include speech analysis for clinical applications, such as Automatic Speech Therapy or the detection of Speech Affecting Diseases. In addition, research has also been focusing on natural language processing and speech applied, for example, to the translation of instant voice messages. Also worth mentioning is the research interest in music generated by AI.

Image analysis and processing are also currently being researched with the main purpose of restoring or extracting visual information contained in still or moving images. Additionally, the application of AI to human routines, such as face detection and identification, the automatic prediction of human trajectories, or the automatic detection of violent behaviour have gained increasing importance. It is also relevant to highlight the application of AI to the recognition of patterns in images, used, for example, to identify diseases in plants.

Being such a prevailing theme, social networks have also, naturally, been the subject of recent research. Examples of works being developed in this area include the identification, ranking, and tracking of community leaders in evolving social networks, sentiment classification, main factors for online marketing campaigns, stream-based recommendation systems or the automatic detection of cyberbullying.

The potential of AI applied to consumer behaviour is also widely recognised. Topics such as identifying consumption patterns

in retail, developing improved recommendation algorithms for e-commerce, understanding and improving marketing campaigns (mobile and telemarketing), as well as analysing customer relationship management systems have been receiving plenty of attention from researchers.

Another area with several published works is related to the so-called Smart Cities. Several studies aiming to improve the efficiency and quality of life in cities address topics such as the collection and analysis of human mobility data (to learn about locations and habits); the identification of points of interest and similar individuals using GPS data; the detection of urban water consumption patterns and the optimisation and automation of waste collection and sorting.

Another related area of study is transportation and mobility. Here, research has been dealing with ways to address more efficient transportation by modelling and simulating the behaviour of the network, exploring space in traffic simulators and working on vehicle routing problems.

The other identified areas, despite their importance, have, nevertheless, resulted in fewer published works in Portugal. Other AI-related scientific articles published by Portuguese researchers cover topics such as (1) quality, e.g., with industry failure prediction; (2) energy, with an artificial vision of non-conformities in power line insulators; (3) software, with predictive models for reliable and secure software; (4) telecommunications, with work on signal processing optimisation, risk management, and the detection of anomalous users in data

networks; (5) biotechnology, with big data application, e.g., drug discovery, recycling and safety; (6) fraud detection and prevention; (7) human-robot interaction, specifically educational child--robot interaction; (8) credit scoring for microfinance, with behavioural data; (9) processing and analysis of sensor data; (10) text analysis, with the classification of sentences in scientific abstracts; and (11) machine vision applied to sign language recognition and learning.

Research Centres

The most active Portuguese researchers in AI are affiliated with the main R&D Centres, namely, and in alphabetic order:

Centro Algoritmi – University of Minho

The "Centro ALGORITMI is a research unit of the School of Engineering – University of Minho, that develops R&D activity in Information and Communications Technology and Electronics (ICT&E), spreading into six major fields: (1) Computer Science and Technology (CST), (2) Information Systems and Technology (IST), (3) Computer Communications and Pervasive Media (CCPM), (4) Industrial Electronics (IE), (5) Industrial Engineering and Management (IEM), (6) Systems Engineering and Operational Research (SEOR).

The majority of PhD-level researchers of the Center are also Faculty members of four departments of the School of

Engineering, namely: Industrial Electronics (DEI), Information Systems (DSI), Industrial Engineering & Management (DPS), Informatics (DI). Furthermore, the MSc and PhD projects supervised by those Faculty members are developed at ALGORITMI.

ALGORITMI Research Center focuses its activity on projects that explore a strong link with the community, namely, the industry and the public administration. The University of Minho is located in an industrialized region with an important expression on textile and footwear industries, and an emerging field on ICT services enterprises pushed by the e-phenomena age. The automotive industrial field also gained recently an important share of the regional market. Another external factor that influences" the target field of application the Centro Algoritmi "is the growth of the cities in the region, which introduces several demands for the co-operation of researchers, namely in the field of logistics, communication, e-government, health and resource management. The results of this policy are stated by the number of applied research projects in cooperation with companies (funded by ANI and QREN) and by specific national projects, namely POSI, POCTI and POE programmes". {*link*}

CISUC – University of Coimbra

"The Centre for Informatics and Systems of the University of Coimbra (CISUC) is a large Portuguese research centre in the fields of Informatics and Communications, which was created in 1991 under the Program Science. CISUC aims at carrying out original R&D at a pre-competitive level, training highly

qualified young researchers, co-operating in national and international projects and programs, and promoting the dissemination of results by means of contracts with different companies.

All administrative support of the centre and the majority of the research laboratories are located in the Department of Informatics Engineering building at the new campus of the Faculty of Sciences and Technology.

CISUC is composed of about 200 researchers, including full--time university professors and graduate and post-graduate students. This young, dynamic and highly qualified team of researchers is the guarantee of success of the centre's mission.

Covering a substantial segment of Computer Science and Engineering research topics, CISUC is organized into 6 research groups (Cognitive and Media Systems, Adaptive Computation, Software and Systems Engineering, Communications and Telematics, Information Systems and Evolutionary and Complex Systems), and has registered, over the last years, an increasing activity on multidisciplinary and emergent research subjects". {*link*}

FCUP – Faculty of Science, and FEUP – Faculty of Engineering – University of Porto

The "LIACC – Artificial Intelligence and Computer Science Lab, started in 1988 at the University of Porto with three different groups coming from the Faculties of Science, Engineering and Economics".

LIACC's general objectives are three-fold:

1. "To contribute to the research on:
   1.1. Distributed and decentralized software systems and tools;
   1.2. Text mining and Information extraction;
   1.3. Human-machine intelligent cooperation;
   1.4. Declarative software programming enabling safer systems.
2. To implement proof-of-concept software and systems leading to possible prototypes and future real applications.
3. To supervise research work leading to PhD and MSc theses on the aforementioned topics.

Following these lines, LIACC aims at producing both software algorithms and prototype systems backed by relevant publications, while also assisting young researchers in developing their theses. In order to fulfill the aforementioned general objectives, LIACC research will be centered in the core activities of the unit, adjusted to the new challenges of a rapidly evolving society, of the technological development, of the proliferation of networked devices, of a world where we all simultaneously are information producers and consumers. Intelligent decision-making for individual agents but also in networks and teams, in distributed environments, including artificial (agents, robots) and humans is a major aim of the unit". The LIACC also advocates "the need for more theoretical efforts on developing methods, mostly declarative-type of programming based, for better matching software models and architectures. In order to better implement the strategic programme, LIACC is organized in three research groups:"

- The DAIAS – Distributed Artificial Intelligence and Agent-based Simulation group, which "pursues research on multi-agent systems and agent-based interoperability for networking and cooperation; agent-based simulation; Text mining; Cloud, parallel computing and applications".
- The CS – Computer Science group, whose "main general aim is software and information processing reliability with emphasis on fundamental research, ranging through computational complexity, automata and formal languages, algorithms, programming languages, semantics, and formal verification".
- The HMIC – Human-Machine Intelligent Cooperation group, which is "primarily concerned on creating methodologies that enable machines and software systems to think and interact like humans and have social capabilities similar to humans enabling them to be members of heterogeneous human/machine teams". {*link*}

Fraunhofer Center for Assistive Information and Communication Solutions – AICOS

At Fraunhofer AICOS, the Intelligent Systems research group develops, among other technologies, "cognitive systems inspired in the way the human brain works as well as its capabilities to learn from past experiences to solve highly complex tasks".

They "research different information extraction techniques such as text and natural language processing, computer vision and signal processing, to extract information from unstructured sources of information". They "combine and integrate multiple sources of

information, so that" their "cognitive systems can understand the meaning and the relation between relevant concepts in any context".

This Centre also addresses "legal compliance from [its] initial stages of development, to ensure that dataset bias and unfairness are minimised, and that the transparency of machine-decisions as well as that" of their "solutions have a positive impact in the world". {*link*}

INESC-ID

"The Instituto de Engenharia de Sistemas e Computadores, Investigação e Desenvolvimento em Lisboa (INESC-ID) is a R&D institute dedicated to advanced research and development in the fields of Information Technologies, Electronics, Communications, and Energy.

INESC-ID is a non-profit institution, privately owned by IST and INESC, officially declared of public interest. It was created in 2000, as a result of a reorganization of its parent institution. Since December 2004, the institution has the status of ["Associated Laboratory"] from FCT".

The INESC-ID currently integrates "more than one hundred PhD researchers and two hundred post-graduation students divided between nineteen research groups, organized in five main research lines:

- Computing Systems and Communication Networks
- Embedded Electronic Systems

- Information and Decision Support Systems
- Interactive Intelligent Systems
- Energy Systems

INESC-ID has participated in more than 50 research projects funded by the European Union and more than 190 funded by national entities". INESC-ID's "researchers have published more than 700 papers in international journal papers, more than 3000 papers in international conferences, and have registered 15 patents and/or brands". {*link*}

INESC TEC

"INESC TEC is a private non-profit research institution, dedicated to scientific research and technological development, technology transfer, advanced consulting and training, and pre-incubation of new technology-based companies.

As an institution operating at the interface of the academic and business worlds, bringing closer together academia, companies, public administration, and society, INESC TEC typically applies the knowledge and results generated as part of its research in technology transfer projects, seeking value creation and immediate social relevance".

"The 13 R&D Centres of INESC TEC are structured in four thematic domains – Computer Science, Industrial and Systems Engineering, Networked Intelligent Systems, and Power and Energy.

INESC TEC hosts over 700 integrated researchers (about 350 PhDs), including staff researchers, researchers from Higher Education Institutions, grant holders and affiliated researchers. INESC TEC's team also includes trainees, and technical and administrative support staff". {*link*}

Instituto de Telecomunicações

The "Instituto de Telecomunicações (IT) is a private, not-for-profit organization, of public interest, a partnership of nine institutions with research and development in the field of Telecommunications

IT mission is to create and disseminate scientific knowledge in the field of telecommunications.

IT is actively involved in fundamental and applied research both at national and international levels. Simultaneously it is committed to foster higher education and training, by hosting and tutoring graduate and postgraduate students. it also plays its role towards public society with public awareness initiatives, knowledge transfer to industry, and by providing consulting services on a non-competing basis.

IT is organized around three sites and 4 branches

- Aveiro, in the University Campus
- Coimbra, at Site II of the University of Coimbra
- Lisbon, at Instituto Superior Técnico

- Porto, at the Faculty of Engineering and Faculty of Sciences of the University of Porto
- Covilhã, at the University of Beira Interior
- Leiria, at the Leiria Polytechnic
- Lisbon, at ISCTE-Instituto Universitário de Lisboa

Some IT researchers are located also at other Portuguese higher education units namely at the universities of Algarve, Évora and Nova de Lisboa, and the polytechnics of Lisbon, Setúbal and Tomar.

Scientific expertise in IT, from which follow its main research and education activities, spans the following areas:

- Wireless Communications
- Optical Communications
- Networks and Multimedia
- Basic Sciences and Enabling Technologies

IT members (about 260 PhD holders) are organized into research groups which cover most of the R&D areas in Telecommunications and supporting sciences and technologies. In 2014 IT hosted about 300 PhD students and about 200 MSc students.

Advanced laboratory facilities are available in most Scientific Areas of IT to support applied research, which is carried out in the framework of national and international projects in cooperation with similar research institutions worldwide. Each year IT is involved in more than 170 projects, of which about 30 have European funding, obtained in a competitive basis.

IT actively promotes a high level of scientific research trough a "Quality Policy", issued in 2011, which includes an annual evaluation of its members' scientific production, and a reward/sanction system. Yearly, IT scientific output nears 40 books and book chapters, 350+ papers in international scientific peer refereed journals, 450+ papers in conference proceedings and 10+ filed patents". {*link*}

Instituto Superior Técnico – University of Lisbon

"The Lisbon Unit for Learning and Intelligent Systems (LUMLIS) – the Lisbon Unit of the European Laboratory for Learning and Intelligent Systems (ELLIS) – is hosted at the Instituto Superior Técnico (IST) of the University of Lisbon (UL).

The vision of LUMLIS is to become the leading AI&ML research unit in Portugal and one of the top centers in Europe. LUMLIS aims at contributing actively to defining and executing the European AI strategy in the next decade, fostering the rapid development of AI technologies and promoting their use for addressing societal challenges and maximizing economic impact.

The research agenda of LUMLIS will bring together researchers from different AI&ML sub-fields with the common goal of designing human-interacting explainable AI systems: this involves a strong bet on human language technologies, social and cognitive robotics, and computational biology, well supported on the development of efficient and reliable ML systems with theoretical guarantees.

The unit will exploit synergies between these areas towards the unified goal of designing AI systems that interact seamlessly with humans, learning from them and for them. To this end, LUMLIS will aggregate researchers who are all faculty members at IST and affiliated with non-profit research centers associated with IST: Instituto de Engenharia de Sistemas e Computadores (INESC--ID), Instituto de Telecomunicações (IT), and Instituto de Sistemas e Robótica (ISR-Lisboa)". {*link*}

Recently, the Instituto Superior Técnico was "chosen to host one of the 30 research centres of the European Laboratory for Learning and Intelligent Systems (ELLIS). The Lisbon Unit for Learning and Intelligent Systems (LUMLIS) will be aligned with the ELLIS mission, focused on basic and applied research, playing an important role in guaranteeing Europe's sovereignty and leadership in the field of modern Artificial Intelligence (AI) research.

The ELLIS network brings together 14 European countries and includes several world-class institutions, such as the University of Cambridge, the Italian Institute of Technology and ETH Zürich. The opening of the new 30 centres took place this Tuesday morning, September 15, through a virtual ceremony. ELLIS units were selected by a committee of renowned scientists from different countries based on the scientific excellence of the institutions involved" {*link*}.

NOVA LINCS – NOVA University of Lisbon

"The scientific organization of NOVA LINCS is based on four expertise areas, each corresponding to a formal FC&T research group, and which provide the grounding knowledge basis of the laboratory. Each research group is led by a Principal Investigator (PI), responsible for steering, promoting, and coaching the general activities within the group and their alignment with the strategic program, in coordination with NOVA LINCS Board and Steering Committee.

- CS – Computer Systems
- KBS – Knowledge-Based Systems
- MMS – Multimodal Systems
- SS – Software Systems

The research roadmap of NOVA LINCS is focused on a running strategic program which defines a coordinated set of R&D activities of pluriannual duration, seeking results of high scientific and technological impact, developed within national and international academic and societal partnerships, with relevant research institutions and companies.

Within and across research groups, research activities widely span the laboratory in flexible and synergistic ways so as to better suit the strategic program objectives. To promote agility, activities are often organized in lightweight teams, involving members from several research groups to address the unit research objectives that require interdisciplinary skills in computer science and beyond". {*link*}

Researchers in Portugal have been producing very relevant and practical work on AI in the past few years. It is interesting to observe that they are affiliated with several different Portuguese R&D Centres, thus contributing to the creation of a solid AI network in Portugal, with complementary competencies and applications.

Although research on AI has global applications, Portuguese incumbents, startups, large and small companies are able to benefit from the talent and knowledge produced by their R&D neighbours to meet their business needs. Besides the more fundamental work developed by R&D Centres, they also support concrete challenges faced by companies, helping them find where to apply newly created knowledge, improve their businesses or even create new ones. It is well-known that geographical proximity between partners creates value among them, and this is evident in the Portuguese case, with the development of critical mass in an AI-based value chain of competences and business.

Part II

FROM ACADEMIA TO PRACTICE

Given the importance of the AI subject in Portuguese Academia, it is natural to find early studies that take into account the implications of this technology for businesses and the society at large. The areas of research interest go beyond the technical capabilities and methods that support AI. They also focus on the amount of understanding that is required for it to be adopted by practitioners, used and exploited in specific industries or, given the imperfections still lingering on AI solutions, on possible thresholds or trade-offs that will increase its acceptance. Below, we present one example for each of these cases.

Understanding how AI technology operates and why it makes the decisions it makes is essential for the acceptance and adoption of AI solutions. A recent study titled "The effect of knowledge about Artificial Intelligence (AI) on openness towards AI-enabled products and services" by Jana Lautenschlager, under the supervision of João Castro (2020), looked into the current understanding of AI technology, its efficiency, convenience, privacy and security by customers, and its relationship with their openness to using these solutions.

The study gathered data from around the world and presented interesting findings, suggesting a strong relationship between knowledge and understanding of how AI technology works and the willingness to adopt, interact or prescribe these types of solutions:

- Well informed customers are more open to using AI-enabled products and services, and expect higher efficiency when compared to regular non-AI offerings.
- This openness leads them to believe that AI products are more efficient and convenient than their non-AI counterparts.

In contrast, those who are more knowledgeable about this technology are also those who are most sceptical about how it respects, or not, people's privacy. Nevertheless, they still consider the trade-offs advantageous. This does not mean that they are willing to let go of their privacy, only that the benefits are significant enough to outweigh the loss of such privacy.

According to Lautenschlager, educating users about AI "is of utmost importance for them to develop a more open relationship with AI and the offerings that employ this technology" but that "does not explicitly mean that society should frivolously be educated about the benefits of AI – but a well-balanced knowledge about AI technology and its capabilities should be developed" (Lautenschlager, 2020). And such capabilities should be developed with a special focus on the benefit areas of efficiency, convenience and privacy protection.

When diving deep into a sector, as was the case of a study titled "State of the Art of Artificial Intelligence in the Portuguese Food

Retail Sector" by Joana Oliveira, under the supervision of João Castro (2020), it is possible to observe the quick impact this technology is having on this specific sector.

Technology is rapidly changing the retail sector in many of its different tasks and operations, with highly competitive and low--margin solutions supporting the sector's efforts to provide the right range of products and services, at the right place and at the right price. In this study, the author found that, as is the case with other industries, AI in the retail sector enables "more accurate decision-making, improves efficiencies and reshapes customer experiences" as part of "a unique opportunity to create and capture value".

The resistance of users to move from analogue documents, with affordances that are not possible in the digital domain, has also hindered the adoption of some AI solutions by organisations. An anecdotal example we encountered was that of a woman working for a large energy company, who was responsible for collecting the clients' payments. She would refuse to let go of her pile of paper, as each paper documented the client's order and the amount to be paid. However, once the payment was processed and received, the paper would be torn up and thrown into the paper bin. Digital solutions can offer many more functionalities but, in this case, the ability to "tear up paper" was the key point, signalling that the task was complete, regardless of all of the arguments presented, including reducing the use of paper and other sustainability considerations.

Finally, there are numerous examples where the use of Artificial Intelligence solutions went astray and produced results contrary to the initially intended objective. With this in mind, another study tried to see what type of adoption threshold could there be and when would the technology still be sufficiently beneficial to be put to use, even knowing that there were imperfections and that it might fail on critical functionalities.

The study "Artificial Intelligence in Recruitment: Just because it's biased, does it mean it's bad?" by Giusy Beneduce, under the supervision of João Castro (2020), considered that AI systems failed because, in some cases, they are not able to adequately represent the values and laws of the hiring process, therefore introducing or producing biased results.

When dealing with people, our societal values require that bias is not present for issues such as, among others, gender, race, age, religion and sexual orientation. Artificial Intelligence systems, which depend mainly on input data to train their function, are doomed to repeat most of the bias present in the set of training data: "Bias in, bias out".

Two classic examples of this are:

- The 2016 ProPublica study of COMPAS, in which the system had a higher failure rate of estimating recidivist criminals if they were black.
- Face classifiers have a harder time correctly detecting gender for women, and also perform worse with darker faces, hence having their worst performance with women of darker skin tones.

Human resource management is a field that has characteristically been targeted as ripe for change with the introduction of AI. It requires the ability to process a vast set of information describing job opportunities and requirements, including workers and their accomplishments, and try to find successful matches.

While promising, the full automation of the hiring process is still met with some resistance, as professionals believe that AI is not yet able to capture the same details as face-to-face meetings. Especially, there is arguable concern that AI systems cannot exclude bias from their analysis, even when developers explicitly programme routines to avoid them.

Below, we present three cases where AI failed in recruitment or career counselling activities, even though the bias was not present:

- Click-rate optimisation algorithms from Google showed ads from a coaching agency promising higher salaries more frequently to men than women. The system used by Google tries to target ads to those they might be more relevant to. However, training data showed that men would click on those types of ads more often than women, which injected a behavioural bias into the data.
- A similar campaign run on Facebook sharing information on STEM (Science, Technology, Engineering and Math) programmes was less likely to be seen by women. However, in this case, it was determined that the bias came from the fact that the female market segment was targeted more often and, therefore, it commanded a higher price for ads, which exhausted the allocated budget quicker.

- Amazon tried to implement a tool to review resumes, which was trained under a set of previous resumes and the success of the applicants over time. The company later found that the programme had a gender bias since it had been trained on a set of mostly male resumes and, thus, classified them as more successful.

These three cases have become cautionary tales in the field, informing and educating us about some of the perils of blindly trusting the learning abilities of algorithms and their recommended decisions.

Given this bias in AI technologies, Giusy's focused his work on the extent to which "Artificial Intelligence bias is a blocker to the implementation of those technologies in recruitment". The study found that the recruitment community understands that bias is already present, inevitable and very hard to extricate from their processes. However, they also believe that AI can help diminish the presence or the strength of such biases. Although the quality of the process revealed to be paramount, the fact that recruiters are willing to accept a certain degree of bias if tools facilitate their job in terms of speed and cost is somewhat optimistic.

These three examples support the pertinence of further academic studies on AI that focus on the adoption and use of such technology by governments and organisations.

Another area where it is possible to encounter a transition from Academia into the industry is the universe of the startups. Several small companies are using AI as the foundational building block

of their value proposition, offering novel services or services at costs and levels of efficiency that were not possible before. Some notable examples are:

- Abyssal – a company that provides 3D visualisation, simulation and digitalisation capabilities for subsea operations based on AI algorithms.
- Automaise – a human-like platform for customer interaction, customised to each company.
- Box To Life – a platform based on the internet of things and on Artificial Intelligence that creates and distributes content and collects and analyses visitors' data such as behaviours, routes and interests.
- DefinedCrowd – a company that sources, structures and enriches high-quality training data in speech, text, image/video, translation and evaluation of experience to create a natural interaction between people and machines using AI.
- Feedzai, a company that specialises in transaction anomalies such as fraud detection, especially in the financial sector.
- Mindreach – a technology that facilitates the connection between the brain and external devices, by modulating brain waves and controlling external devices with thoughts.
- NILG.AI – a company that provides consultancy and develops AI solutions for several business sectors.
- Nimest – a company that is developing a travel/tourism assistant platform supported by AI.
- Overstory – a company that offers interpretations of satellite imagery and climate data based on AI to support electric utilities, forest industry, insurance companies and the management of reforestation/deforestation.

- Reckon.AI, a company that offers two product lines with extensive use of AI solutions, namely: (1) automating the analysis of supermarket marketing brochures of different competing chains to assist managers in benchmarking promotions and counteracting them with other offers and; (2) using AI for computer vision in a 24/7 vending solution similar to vending machines and monitoring which products are being selected through cameras.
- Skorr – an app that delivers brands and agencies the means to explore social media influencers and establish online relationships along with advanced ML and AI qualification features to ascertain which influencers will have the best fit for each brand and campaign, along with unique analytics capabilities to monitor and track campaign and influencer performance.
- Smartex – a company that develops quality control systems for inspection in the textile industry employing AI.
- Thorly Education– a student-centric approach using AI to customise education.
- Unbabel, a company that provides translation services that combine artificial intelligence and human specialists. It offers faster and cheaper services when compared to using only human translation, and much higher quality than the traditional machine translation solutions.

Part III

PERSPECTIVES FROM PRACTICE IN PORTUGAL

The initial plan for this report was to engage organisations, large and small, from different industries to observe the extent to which their operations were using AI technology. This plan was compromised by the lockdown measures put in place in early March 2020 due to the COVID-19 pandemic.

For a while, none of the organisations were available to participate in the study as they were facing a severe and abrupt change in customer spending patterns and struggling to adjust their business practices and workplace environments.

This was a time of great agility for organisations as they endured hard times with some of them fighting for outright survival, which will be illustrated in the two following examples of adaptation through the use of technology, however not necessarily AI technology.

Before the pandemic, a single premise neighbourhood supermarket had very little motivation to have an online presence. Most of its clients lived within walking distance, and the value of a delivery service did not outweigh the advantage of personally selecting the products. This context has been previously suggested in the "The future of online grocery shopping – an explorative approach to discover current German market dynamics" by Cindy Engelmann, under the

supervision of João Castro, (2019) as one of the most important factors of why the German food retail has not made inroads into the digital world like other countries. In the context of the COVID--19 pandemic one of these local Portuguese supermarkets, with no online presence, reacted by quickly creating an excel spreadsheet with all the available products and circulating it by word-of-mouth and attaching it to its clients' emails. Orders would be placed by filling in the spreadsheet and emailing it to the store, which would then reply with a pickup time. This was a stop-gap solution only possible due to the small scale of the operations involved.

Another example of extreme agility came from DOTT, an online store aggregator, which saw physical storefronts that previously frowned upon joining the service now, amidst the pandemic, scrambling to sign up and having their products listed on the site.

Considering all the adjustments that many organisations had to undergo, it was only natural that they were not available to participate in our study. As the situation calmed down, businesses started to recover slowly, and we were able to find some time to meet with the people we had intended to meet.

Whenever possible, we avoided top management entry points and interviewed people who were directly and daily involved in AI-related projects inside each organisation. From our experience, top managers will usually restrict their conversations to what is publicly available and match their message with the company's mission and strategy. By going some levels deeper, we were able to gain richer insights into the real struggles with the adoption and implementation of AI solutions within each organisation. Due to

the candour of many of these revelations, we had to anonymise their contributions, retaining the most important findings for the benefit of this study.

We previously referred that there is no standard, correct and bounded definition of AI so, when we approached organisations, we first had to make an assessment of what kind of AI was being used. When we finished our interviews, we decided to rely on a model that divides these solutions into two categories: Simple AI and Blackbox AI. We refer to "Simple AI" when what we find is the straightforward mathematical analysis of large quantities of data that has now become possible. We use the term "Blackbox AI" to refer to systems in which it is not trivial to describe or understand how machines are working and reaching solutions.

Most of the approached organisations and their AI efforts lie within the first category. While not as sophisticated as Blackbox AI, Simple AI still adds significant value to these organisations. We also found some companies which were still in the aspirational stage, preparing to get started or looking into different AI options and figuring out how these might help them. Those who embark on "simple AI" projects are doing it not because it is "simple", as the name implies, but because it represents an extraordinary step for their growth. For the first time, these companies are able to analyse large troves of data stored in their Enterprise Resource Planning (ERP) in more than a simplistic regression form and, as low-hanging fruit projects reveal their worthiness, they start looking into other ancillary data that they already have or may actively begin to record. For these organisations, AI is a form of

extracting rich value from data which was previously considered redundant or irrelevant and left unexplored.

The level of enthusiasm of this first contact with AI led one of our interviewees to state that "almost all businesses can take advantage of the information they already hold but, in general, they are not doing it yet". Others also considered AI technology a "fantastic opportunity to free people" of some laborious tasks and educate staff for more useful positions.

The availability and access to large quantities of quality data are considered by many to be the essential building blocks to explore AI. In order to obtain this added value, people within the organisation need to determine what data is currently available, treat it for processing and detect potential gaps or additional internal or external information sources that can be added to the set. Data needs to be relevant, cleaned and organised in order to be processed in "simple AI" systems.

An example of an application of simple internal data, for instance, to plant production planning is the correlation of client purchase orders over time with weather data records, which, combined, can help producers to better predict and meet future orders.

We found that organisations playing at the "Blackbox AI" level often have teams with a highly experienced staff combining PhDs, MBAs and years of practical industry experience. This allows the teams to explore AI and what it does in more depth.

There is founded concern with both theoretical and anecdotal evidence of the dangers of blindly using any type of data and with the fact that trained systems might seem operational but actually hold biases in their analysis. This has put some restrictions on the adoption of AI systems.

In some cases, there was also evidence of systematic work and investment on what could be described as "explainable AI", where, when using "blackbox AI" there was an effort to reverse the output to better understand the behaviour of the machine and what it was producing. These efforts were mostly made to avoid blindly delegating critical decisions to a machine without understanding why it was making such decisions and to ensure compliance with laws, regulations and best practices (e.g., equal opportunity employment, sustainability acts, ethics and organisational value codes).

Given the different nature of AI approaches in the market for different sized companies and sectors of activity, we also investigated the characteristics that might suggest that an organisation is exploring AI. What seems to be the most significant driver for AI exploration in organisations is the availability of data. The currently available technology is sufficiently cheap and powerful to be tasked with making sense of massive historical data records, and new types of data which are now possible to collect, store, process and quickly add to organisations' "data lakes" or data repositories.

Some businesses are in better condition to implement new technologies than others. Almost all tech-based companies or those who already process large amounts of data (e.g., banks, insurance companies, utility companies) have either made this transition or

have projects in a very advanced stage of implementation. Another factor that seems to determine the exploration and adoption of AI solutions is the stability of a specific market. For example, food retail in Portugal has an extremely competitive landscape with two very large players and several smaller but also fast competitors. AI has quickly become a part of the arms race in this industry, e.g., in customer behaviour, customisation of recommendations and marketing solutions, as well as in daily operations such as automatic replenishment supported by algorithms with sophisticated sales forecasts.

However, some industries, such as (traditional) banking, despite their vast amounts of available data on customers, transactions and markets, still seem conservative regarding the exploration of AI's potential. They observe the trends and opportunities, but have a somewhat *follower* behaviour, waiting until a specific technology is mature enough and has almost no implementation risks, or until a competitor implements such technology. However, other business models in this sector, such as purely online banking, are taking more chances in this field and applying it, for example, to management and risk analysis processes.

SME's are lagging in the potential use of AI solutions and, given the context in which they are operating in 2020, with many of them struggling to keep afloat, this is not one of their top priorities. Nevertheless, some evidence suggests that recently founded companies, with little or no legacy processes, are part of those who are most aware of AI´s potential and more willing to exploit it, especially companies in the technological sector and dealing with vast amounts of data.

We can find several examples of AI applications in this sector, such as fraud detection and prevention, with several layers of AI algorithms working to ensure correct, legal and swift transactions. Business-to-business tech providers are either actively working in this area or have plans to do so in the short term. Other examples of applications include the prevention of financial crime, money laundry and account appropriation. Some technological companies are also developing AI-based projects to improve infrastructure efficiency, especially in critical components, by predicting if and when a failure is likely to occur in order to act beforehand, thus minimising downtimes and the correspondent costs and revenue losses.

Online business to consumer companies (B2C e-commerce) are also applying AI to fraud detection and prevention. They have developed AI solutions, such as recommendation algorithms and the selection of specific customer types, for example, to detect and prevent account hacking and increase the value of their orders and purchases. E-commerce companies also heavily invest in tools that improve the efficiency of their supply chain and help predict and fulfil customer's orders, with solutions ranging from supplier negotiation and selection to customer delivery (location, satisfaction, etc.).

Another interesting example is the manufacturing industry, where several AI applications are being deployed, among others, in, the quality area to achieve the "zero-defect" goal, and in systems that make recommendations to the commercial area or manage inventory and the supply chain. These applications are only functional when based on good quality data. Even if manufacturing

companies are able to collect large amounts of data from their central systems (e.g., Enterprise Resource Planning – ERP, Manufacturing Execution System – MES), these applications would only be complete if they also collect data at the shop-floor level, which is not an easy task. With the emergence of distributed and smart systems and sensors at the machine and process levels (the Internet of Things – IoT within Industry 4.0), such data collection is becoming more widespread and potentiating more sophisticated AI applications, thus extracting its full value.

Portugal has a decent know-how of technology at this level of understanding and continues to benefit from several initiatives promoted by private and public entities. Specialised training centres are focusing on programming and data science skills, and the Portuguese Government also has in place, or is about to start, different upskilling/reskilling programmes for different educational levels. Several organisations stated that many of the top talents in Portugal are world-class quality, but very hard to keep in the country due to the limitations of the salaries or the work package deals in the Portuguese market.

This reality is quickly leading to a scarcity of qualified professionals in the field and, in the short term, it may become an obstacle for further adoption and exploration of AI technology in the Portuguese industry.

In addition to the scarcity of experts (an issue that is being addressed but that is still looming), another often-cited problem for organisations is the regulatory ecosystem. Early pioneers in the field complain they see too little being done and that what is being

produced is lacking, late and hindering innovation in Portuguese companies. The precautionary principle is in full force, and little space is given for early-stage exploration of novel uses and opportunities. Those who not only use but also develop AI technology see this as a major gap between them and companies located in other countries/markets where the conditions of use and access to data are quite different.

An issue where there is still debate and a need for further clarification is that of data ownership. While synthetic or non-personal data can be freely used, personal data is a very sensitive and dangerous issue. Some of our interviewees stated that no matter the nature of the project, there is a heavy bias towards considering that any use of personal data is immoral until proven otherwise. This gives innovators in the EU space very little incentive to invest in and develop solutions when compared with other geographical regions.

The perspective of some of our interviewees is that there are mainly three paths for AI, namely:

- The EU path mentioned above, built on the premises of GDPR and on the protection of personal data.
- The USA approach, with private companies amassing large quantities of consumer data ("data is the new oil").
- The Chinese approach, with government-sponsored initiatives at scale, which are deemed unstoppable as an authority rule.

Our interviewees also mentioned additional, more general concerns, such as the bias induced from past learning data sets, which might perpetuate and propagate old prejudices or create a shell of

respectability for those that do not represent the current societal values. A classic example of this is the use of zip codes to determine insurance rates, condemning people in different areas to be offered conditions that are not necessarily those that should apply to them individually.

Another layer of fear for organisations and governments is that, if these systems cannot be inspected and audited, then, there is no way of knowing for sure if they are doing what they are intended to do. The case of the Amazon recruitment system is an often-cited example. Two theoretical solutions were presented to tackle this problem, with some suggesting that systems could be checked on-the-fly and in real-time using sample scenarios as inputs to test if the desired outcomes are being achieved, and others suggesting that all AI systems' results could be recorded for later analysis, if necessary, just like aeroplanes' black boxes.

Privacy and trust are additional concerns that bring us back to the premises of GDPR and data ownership. Users and organisations state that they feel uncomfortable with the unbalanced playing field when AI systems are using data about them or their organisations to explore and induce behaviours and decisions. Since those who are harvesting data do not share it with others, this imbalance is aggravated. Given the current value of data, those who already have their significant share are in a better condition to operate than those with fewer data and, therefore, continue to widen the gap between players, culminating with only a few players concentrating most of the data on the market.

Finally, healthy scepticism and trust in AI systems is also growing. The plausibility of deepfakes – often images and videos generated by computers which appear to be actual footage – has raised concern for the fact that AI may not only be a tool used to extract certain "truths" from data, which were not apparent for previous analysis tools, but also a tool used for manipulation. Facebook's manipulation of individual users' timelines to maximise their session time, and the study in which the same company managed to induce different moods or was used to distort elections were the most common examples used to point to the "tip of the iceberg" that started this discussion.

None of our interviewees had a straightforward solution for these challenges. In order to work through these issues, organisations are suggesting an open model for data, where ownership and access are levelled and the playing field rewards contributions instead of the ability to scavenge data. The interviewees do not believe that self-regulation is a possibility yet, given the current advantage held by large companies and, instead, propose the development of additional regulatory frameworks like the GDPR and the penalisation of abusive uses of data and AI technology.

The Portuguese government, through the Ministry of Science and its INCoDe.2030 initiative, published in 2019 the report "AI Portugal 2030". This work attempts to present a Portuguese position on AI as well as a public policy driver the country should follow in the next few years in this matter. It addresses the need to foster increased application of AI in industry and public administration towards added economic growth, to improve advanced education towards scientific excellence, and to invest in labor

force technological qualification and specialization. Within these 3 pillars, it provides specific objectives and actions for the upcoming years. This work is periodically revised by a committee of experts and practitioners.

This report started with an overview of the academic research on AI, transitioning into applied studies in IA, and then into its industrial applications. We end this section by coming full circle. Several of our interviewees suggested that the Academia helps lead the broad discussion on the topic, pointing to the example of the USA, where institutions like the MIT promote the sharing and the investigation of not just the technology but the impacts it can have on the social fabric.

If AI discussions in Portugal remain inside a small circle of technology experts, then only their ideas and opinions will be considered. Those whose perspective is mostly focused on the benefits will open the way for the adoption of AI technologies, while those who think the major impact of AI and other automation technologies is to create unemployment will try to stifle its possible progress. Most people consider that AI is more than another fancy consultancy buzzword, and that it has the potential to make a significant impact on society. Just like with cloning technology or other previous significant technologies and discoveries, AI also demands a large-scale ethical and social debate. Industry wants to have a voice; governments want to have a voice, and Academia should also have a voice.

Part IV

FINAL REFLECTION

It is September 2020 and, at the time of this writing, we are in the midst of a very odd year. We are facing growing tensions, at all levels, as for which paths to take in society. It is a presidential election year in the United States, with the American society in a seemingly insurmountable rift. The pandemic has now claimed the lives of 1 million people around the world. Countries are debating between closing businesses, coping with middle ground restrictive measures or ignoring the warnings altogether and trying to evade the problem, as has been the case in other crisis.

Over the past six years, the transformative power of digital technology has been gaining increasing attention and is now at the centre stage. Different technologies such as autonomous driving, blockchain, the Internet of Things and Artificial Intelligence seem to be coming of age and ready to impact the world.

The COVID-19 pandemic has partly undermined the impact and all the benefits promised by technology. For those who appreciate a stable world, this is not the best of times. For those who believe in the potential to achieve something greater, 2020 is the time when the windows, the door and the roof have opened up to allow for change to happen.

Of all the digital technologies that are currently being made available, Artificial Intelligence is most likely the one with the potential

to cause the greatest impact, either by its direct use in the analysis and understanding of systems that humans were unable to decode so far, or by empowering machines and automation to levels that have not been possible before.

If change is happening, we must ask ourselves what type of change we want to see. As shown in this report, either through examples or the testimony of different academic and industry experts, there are different paths we can follow. While some paths clearly challenge our current societal values, others work to protect those values. And then, there are possibilities and scenarios that society has not even been able to envision.

In a time of change, it seems most urgent that society does not merely let the future happen without a say in it. It is time for an urgent, pondered debate on rules and procedures on how we want to "predict the future", as Alan Kay once did.

Given the magnitude of AI's probable impact, we need to find space for an annexe, an amendment, an addition to the Universal Declaration of Human Rights. And if we are given such space, then, we better know what we want to put in it.

In this time of change, we have the ability and the opportunity to make choices. Future generations will look back at the beginning of the 21st century and judge us for our ability, or lack thereof, to foster a bright future and, hopefully, see a new moment of enlightenment.

The goal of this report was to portray the Portuguese perspective on Artificial Intelligence and, while we remained true to the viewpoints shared with us, it became clear that this is not solely a Portuguese issue, but a worldwide topic that concerns all of us. And Portugal is only one of the many contributors to the future of AI. Let us, then, get out there and create a better future.

Portugal, September 2020

João Castro
Vasco Teles

About the Authors

João Castro

João Castro holds an Undergraduate Degree in Electrical Engineering and Computer Science from the University of Porto, a Masters' Degree (ABD) from the Instituto Superior Técnico and a PhD in Engineering Systems from the Massachusetts Institute of Technology.

He currently works as an Associate Professor at the Nova School of Business and Economics.

He has an extensive career in the industry as a project manager and director for technology and innovation in different projects for different sectors of activity.

He is a non-executive board member of large and small companies, assisting them in topics such as product development, innovation and technological evolution.

Liliana Ferreira

Liliana Ferreira is the director of the Fraunhofer Research Center for Assistive Information and Communication Solutions (AICOS) in Portugal and Full Professor Invited at Department of Informatics of the Faculty of Engineering of the University of Porto.

Her research focuses mainly on Natural Language Processing, Medical Knowledge Representation, Information Extraction, and Health Informatics. Liliana holds a degree in Applied Mathematics and Technology by the Faculty of Sciences of the University of Porto (2002), an MSc in Electronics and Telecommunications

Engineering, University of Aveiro (2005), and a PhD in Informatics (2011) also from the University of Aveiro, Portugal.

Liliana has developed her research in industry and research organisations as Philips Research Eindhoven (NL), IBM Research and Development Böblingen (DE), the Institute of Electronics and Telematics Engineering of Aveiro (PT), the Ubiquitous Knowledge Lab of the Technical University of Darmstadt (DE) and the University of Tübingen (DE).

Vasco Teles

Vasco Teles holds an Undergraduate Degree and an M.Sc. in Electrical Engineering and Computer Science from the University of Porto, and a PhD in Engineering Design and Advanced Manufacturing by MIT Portugal from the University of Porto, the University of Lisbon and the University of Minho.

He is a Senior Researcher at the INESC TEC, where he works as a specialist in innovation management, digitalisation and technology strategy since 2010.

He has also worked with leading organisations in ICT, retail, manufacturing and the public sector.

He is a European, national and regional expert who participates in the definition of public policies for the technology sector.