EMERGING TECHNOLOGIES AND OPEN DATA

Artificial Intelligence

January 2020

GOBIERNO DE ESPAÑA
MINISTERIO DE ASUNTOS ECONÓMICOS Y TRANSFORMACIÓN DIGITAL
SECRETARÍA DE ESTADO DE DIGITALIZACIÓN E INTELIGENCIA ARTIFICIAL
red.es
Iniciativa aporta
datos.gob.es
reutiliza la información pública

Content prepared by Alejandro Alija, expert in Digital Transformation and Innovation

This study has been developed within the framework of the Aporta Initiative, developed by the Ministry of Economy and Digital Transformation, through Red.es, a Public Corporate Entity. Contents and points of view expressed in this publication are the exclusive responsibility of its author. Aporta team does not guarantee the accuracy of the data included in the study. The use of this document implies the express and full acceptance of the general reuse conditions included in the following legal notice:

http://datos.gob.es/es/aviso-legal

EXPLANATORY NOTE

Artificial intelligence is one of the technologies with the greatest evolution in recent years. The new cycle of expansion and growth of this field does not seem to have a near end. There are **two fundamental factors** that drive this growth. On the one hand, the **development of silicon-based technologies** and, on the other, **the abundance of large datasets accessible thanks to the Internet**. For this reason, this report analyses the current impact of Artificial Intelligence in our lives and its strong relationship with the availability of open datasets. This analysis is carried out using a new methodology called AIA -*Awareness Inspire Action*-. Each section of the report, according to the methodology, can be read independently. The level of depth of the content of the report is progressively increasing, from the Awareness section to the Action section.

1. INTRODUCTION

According to a report (2018) by the European Commission, **data-based innovation is a fundamental engine of growth and employment** that can significantly boost European competitiveness in the world market. If the ideal framework conditions are established, the **European data economy could double over the next few years**. In 2020, it is estimated that, in the European Union as a whole, there will be around 360,000 companies whose main business model will be based on the use of data.

A very important part of the data that have enormous value for society is generated by the public sphere. For example, data related to the planet's climatic conditions can - potentially - increase the competitiveness and yield of crops. Another example is data from public telecommunications networks such as satellites, which have the potential to improve management in the event of natural disasters or exceptional situations.

Public sector organizations produce and collect huge amounts of data, which constitute a valuable raw material for developing innovative digital services and improving public policies. In this context, the European Directive 2019/1024, regarding open data and the re-use of public sector information, defines the framework for promoting cross-border use of publicly funded data and contributing to the development of pan-European data services and products[1].

Undoubtedly, open data plays a very important role in the development of new business models built on new products and services that use Artificial Intelligence as an enabler of customer value. In this report we will see different aspects of this relationship, through examples and use cases.

[1] More information can be found in this article on the recently published European Directive.

2. METHODOLOGY

Figure 1. *Awareness, Inspire, Action collection methodology*

This report is part of a broader collection of resources on emerging technologies and open data, whose objective is to **introduce the reader to the subject through the use of practical, simple and recognizable use cases**. At the same time, it is intended to provide a practical learning guide for those readers with more advanced knowledge, who, through the development of a practical case, can self-taught experience with real tools for the analysis and exploitation of open data.

To achieve this double objective, the report is structured in three parts: Awareness, Inspire and Action (Figure 1), which can be addressed independently at any time and without previously read the other sections.

The first section, **Awareness**, is an introduction to the subject (in this report, Artificial Intelligence). This section is indicated for the reader who starts on the topic for the first time and tries to address the topic in a simple, clear way and without the use of technicalities that make reading difficult.

The second section, **Inspire**, intends to serve as inspiration to those readers who have started in the field and who wonder how the subject is affected in their daily or working life. The way to identify yourself with a technology, a field of science or any other subject is to be reflected on it. In this way, the Inspire section contains examples and cases of application of a certain technology in (more or less) everyday situations, that favours the reader to identify him/herself and begin to think of such technology as something that also affects him/her.

Finally, the **Action** section selects one of the use cases explained in the Inspire section and develops it in a practical way, using real data and technological tools. The example, developed in Action, is made available to the reader in the form of code and open data (Annex I), so that the reader can experience and develop the use case that is addressed in the Action section with their own means.

3. AWARENESS

3.1 Key Concepts

Colloquially, we can all understand Artificial Intelligence (AI) as the ability of a machine to mimic human intelligence. The ability to learn or solve **complex and diffuse problems** are characteristics that are attributed to the human mind, but, in one way, we try to transfer it to machines in the form of software programs. Talking about artificial intelligence (AI) is already part of normality today. It is common to associate artificial intelligence with the technology sector, but not only. The World Economic Forum already collects a multitude of monographic reports on the application of Artificial Intelligence in - among others - the **fight against climate change**; the **work transformation**; the **effects on the gender gap**; the **revolution in health care** and **disruption in financial markets**, without forgetting the important **ethical considerations** we face.

From a more formal or technical point of view, artificial intelligence could be defined as: the field of science that studies the possibility of automating intellectual tasks that are normally executed by humans. From a scientific perspective, AI is normally divided into two sub-fields of **computer science and mathematics** called **Machine Learning and Deep Learning**. Without going into too many technicalities, the differences between the two fields are in the type of algorithms used to predict the target variables. There is a type of algorithm used in Machine Learning called Neural Networks. Deep Learning makes deep use of neural networks, creating several layers of these networks.

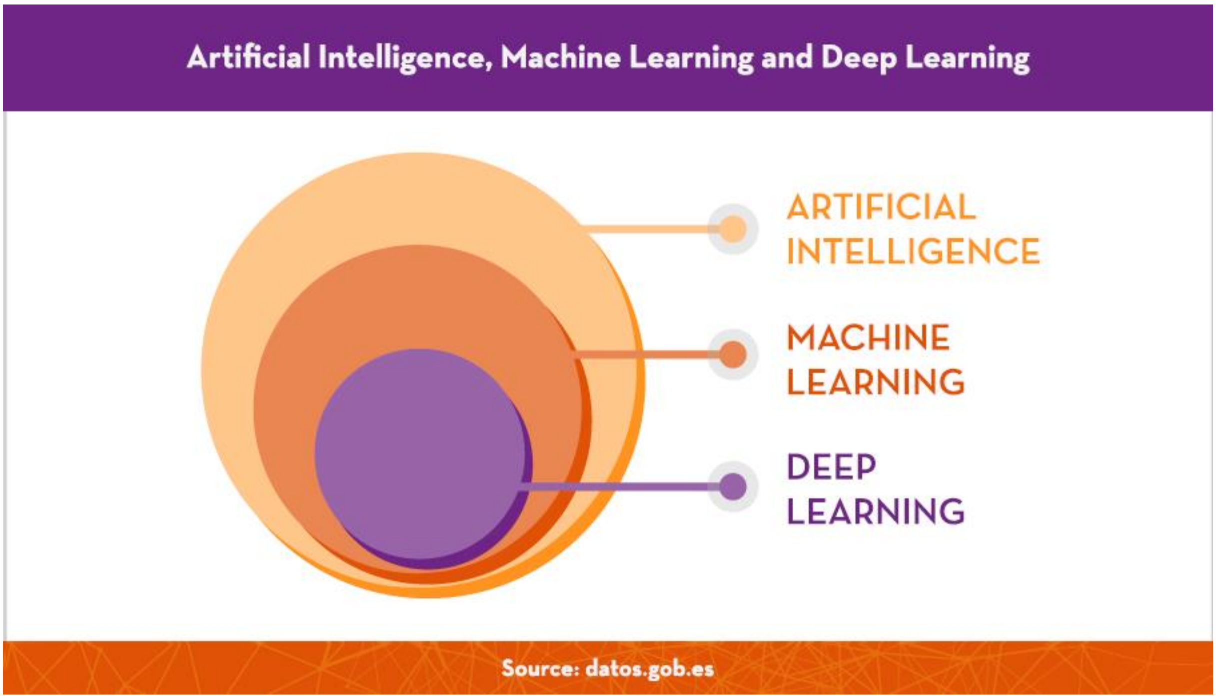

Figure 2. *Relationship between Artificial Intelligence, Machine Learning and Deep Learning*

3.2 Impact

The present and future impact of AI in our society is already analysed in many publications and reports. According to [a study](#) by the technology consultancy Accenture, AI will be the ultimate lever for economic growth in the coming years. Fundamentally, it points to three main growth paths leveraged on AI:

- An **automation of daily, routine and dangerous tasks** with intelligence to work in a diffuse, flexible and highly adaptive way. In the short-medium term, AI will take control of what are known as DDD (Dull, Dirty and Dangerous) tasks, freeing many workers for the development of *more human* activities.

- The ability **to increase and strengthen the workforce of the future** (highly specialized) so that they can exercise genuinely human tasks such as design, creation and innovation.

In a transversal way to the two ways that we have just mentioned, the **development of innovation** in a field as new as AI will trigger enormous technological and human progress that is hardly valuable today.

3.3 A bit of history

However, it is important to note that the concept of **artificial intelligence is not something new** (Figure 3. Timeline that highlights the most important milestones in ...). The first practical steps towards artificial intelligence began in the **1940s**. From the last century to the present day, the development of artificial intelligence has encountered few difficulties.

During the first years of the **1970s** it was thought that the problem of creating artificial intelligence with capabilities similar to human intelligence was almost solved. In the late **1960s**, Marvin Minsky (considered one of the parents of the AI) came to ensure that "... in the course of a generation ... the problem of creating artificial intelligence will be virtually solved ...". Only a few years later, the first major disappointment of AI would freeze this field for several years.

The overexpection generated in this field in its earliest stages resulted in a deception stage of equivalent magnitude. The same happened in the early 1990s. Huge expectations were generated between **1980 and 1985** regarding the ability of those known as expert systems[2] to create artificial intelligence. In the early **1990s**, the inefficiency of expert systems was noted due to its high maintenance cost and low scalability. These periods of decline in the history of artificial intelligence are known as

[2] Expert systems are computer programs that contain logical rules that code and parameterize the operation of simple systems. For example, a computer program that encodes the rules of the chess game belongs to the type of programs we know as expert systems.

the winters of AI. Even today we are not able to predict whether we will be able to build an AI with similar capabilities to the human mind.

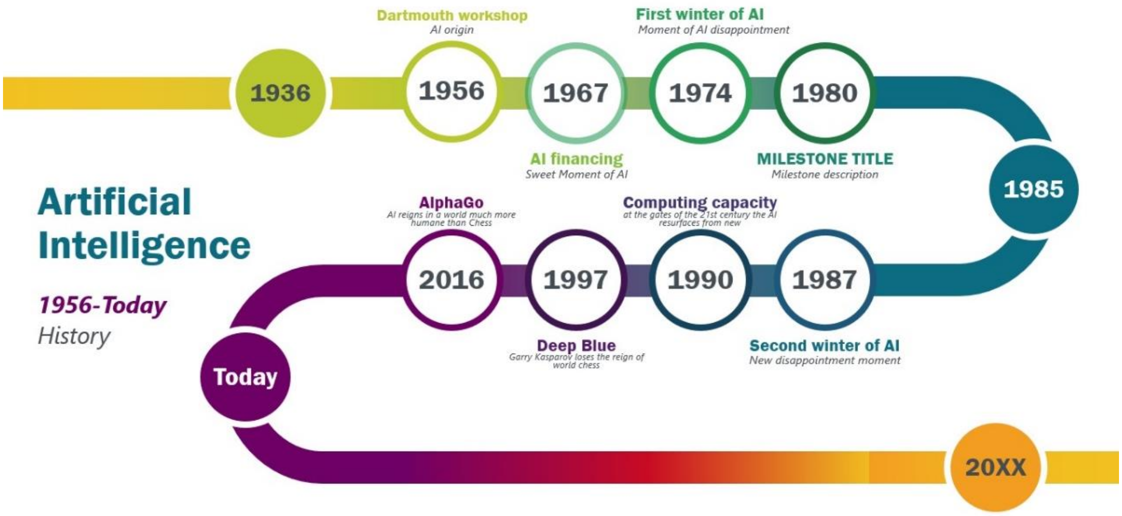

Figure 3. *Timeline that highlights the most important milestones in the development of AI from its beginnings to the present day.*

3.4 Factors that make AI possible today

In the introduction of this report we referred to two key factors in the modern development of Artificial Intelligence. We were referring, on the one hand, to the improvement in processing technologies (silicon technologies) and on the other, to the availability of large datasets accessible on the Internet. Most of the current publications[3] agree on the importance of these two factors and extend the list to other agents that make possible the new historical moment in the development of AI. As an example, the publication 7 key factors Driving the Artificial Intelligence Revolution, or the Harnessing Artificial Intelligence for the Earth report, published by the World Economic Forum in 2018. The latter indicates that the main growth factors are:

- Big data

[3] Harnessing Artificial Intelligence for the Earth
7 Key Factors Driving the Artificial Intelligence Revolution

- Processing capacity
- Hyperconnectivity
- **Technologies and open data**
- Improved algorithms
- Greater and faster investment returns

The following table explains each of these factors:

01	**Big Data**	The democratization of computers, mobiles, sensors and other electronic devices generates vast amounts of data of all kinds that allow us to train better algorithms.
02	**Processing capacity**	The reduction of the processing technology cost together with the increase in the processing capacity of the current microprocessors allows to execute algorithms that would have been impossible in the past.
03	**Hyperconnectivity**	The development of cloud computing, the expansion of the Internet and the development of social networks demand and promote the development of AI in the same way.
04	**Technologies and open data**	An exponential development of AI fundamentally needs two ingredients: technology and large amounts of data. The communities of developers, together with organizations and companies, have created platforms to share software and data. This has accelerated the development of AI as we have never seen before.
05	**Improved algorithms**	The scientific community has made numerous advances in recent years, especially in the field of Deep Learning. The latest advances in generative adversarial networks (GANs) are giving rise to those known as deep fakes.
06	**Greater and faster investment returns**	Market pressure to extract value from consumer data is one of the factors that most contributes to the growth of AI. The need to better know the customer and increase worker productivity are two clear examples of AI use cases.

Figure 4. *Critical factors that drive the exponential development of AI technology. Adapted from the original [Harnessing Artificial Intelligence for the Earth](#).*

3.5 Examples of AI applications

After analysing the critical factors for the current growth of AI, let's now give a couple of examples of recent applications of Artificial Intelligence.

The medical imaging company Subtle Medical uses an AI-based software development for image treatment that allows substantial improvements in the quality of images from medical scanners. Shortening diagnostic times and increasing the productivity of radiologists and machines brings true value to both the healthcare system and the patient's experience.

Despite the large amount of valuable information that we can find on the Internet today, a good part of the global wisdom of humanity is found in the books. Artificial intelligence can help us read faster, in fact, much faster. In a recent academic study, published in Nature, its authors explain how using AI capabilities, they have been able to *read* 8 million books to analyse the concept of *happiness* since 1820 in different countries.

The above examples are just a small sample of how Artificial Intelligence complements and increases our human abilities to see and read. Next, in the Inspire section, we will deepen these two areas of AI with an exponential development in recent years:

- Artificial intelligence as an **extension of human vision**.
- Artificial intelligence as an **amplifier of human language**.

4. INSPIRE

In this section we will see in more detail some of the particular use cases of how artificial intelligence amplifies our capabilities as humans and gives us super powers.

In Human + Machine (Harvard Business Review Press, 2018), Paul Daugherty and James Wilson differentiate **three states of collaboration between machines and humans** (see figure 3). In the first state, AI plays no role and **genuinely human characteristics** such as **leadership**, **creativity** and **value judgments** are distinguished. The opposite state is one in which **features where machines achive better performance than humans** stand out. We talk about **repetitive, precise** and **continuous** activities. However, the most interesting state is **intermediate**. In this state, the authors identify **activities or characteristics in which humans and machines perform hybrid and complementary activities**. In this intermediate state, Daugherty and Wilson distinguish two stages of maturity.

In the first stage - the most immature - **humans complement machines**. We have numerous examples of this stage today. Humans teach machines to drive (autonomous cars) or to understand our language (natural language processing). The second stage of maturity occurs when AI enhances or amplifies our human capabilities. In the words of Daugherty and Wilson, AI **gives superpowers to humans**.

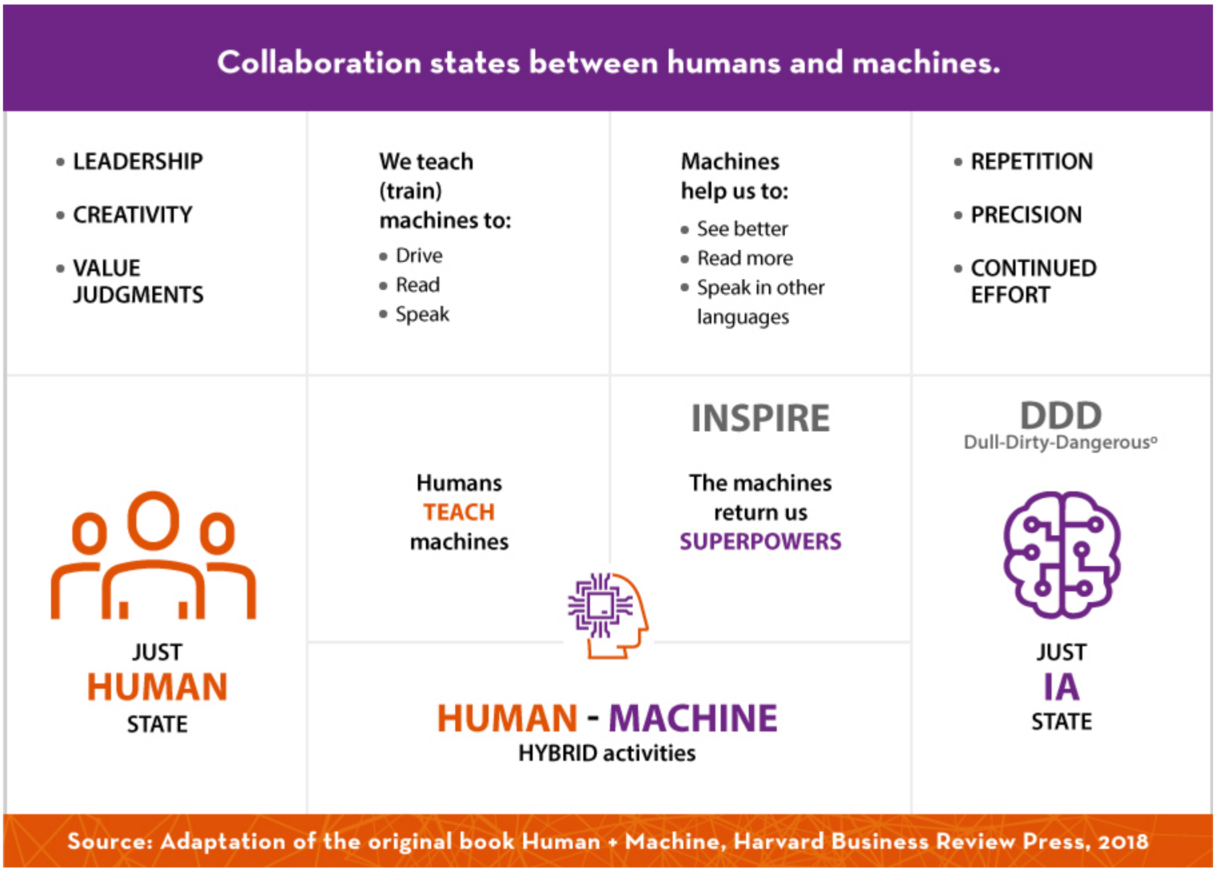

Figure 5. *The three states of collaboration between humans and machines.*

Next, we will see two examples of how AI gives us super powers by amplifying our brain's ability to process what our eyes see or what our ability to communicate is capable of.

4.1 AI as an amplifier of human language

In the field of the extension of human language, Artificial Intelligence helps to promote different disciplines, among others:

- Language translation
- The conversion of written to spoken language.
- The conversion of spoken language to writing.
- Detection of writing errors in texts.
- Related information search engines.
- Spam detectors.

Let's detail some of the previous cases:

4.1.1 Languaje translation

The field of **language translation** is one of the most developed in recent years thanks to the field of artificial intelligence. No one can miss that the translation of texts has evolved incredibly in just a few years. In the era of pre-internet computing, we could only aim to translate word for word in a very small number of languages. At present, we can translate complete texts in which the subjective characteristics such as the **tone** of the message, the **intention** and the **style** of the writing are considered by the translation algorithms to achieve the best result. If we talk about voice translation instead of written texts, we are getting closer to reaching the translation of spoken language in real time to many languages and without delays.

4.1.2 Search for related information

Our ability to read and understand texts as human is also amplified thanks to artificial intelligence. Now, a team of researchers from the Lawrence Berkeley National Laboratory has shown that, thanks to a combination of unsupervised machine learning (AI) and 'text mining', you can process millions of scientific articles and find relationships that, until today, remained hidden. In particular, this research team has applied this technique to 3.3 million summaries of articles on materials science published between 1922 and 2018. When analysing the text bank[4],scientists discovered that **the algorithm was able to identify possible materials related to different physical and electrical properties**.

When we talk about our ability to write as humans, AI also has much to say and help in this task. We may be still far from getting an AI to write a good part of our daily chores, but today, we already find a good set of examples in which artificial intelligences are beginning to write for us. For more information on the latest achievements in AI applied to the automatic creation of texts we recommend reading these two recent articles: The first book created by an artificial intelligence and JPMorgan will replace its advertising editors with an artificial intelligence.

As we saw in Figure 2, factors such as **big data** (we have hundreds of thousands of digitized texts); the **processing capacity** (a current mobile phone is hundreds of times more powerful than a personal computer 20 years ago); **hyperconnectivity** (the development of 4G and 5G low latency wireless communication technologies) and the opportunity for **greater and faster investment returns** (the promise of smart headphones with real-time audio translation capability can create a business of billions of dollars) enhance the development of AI applied to language translation.

[4] Most of the summaries of scientific publications are publicly accessible data, although they cannot be strictly considered as open data.

4.2 AI as an extension of human vision

Image recognition and classification

Probably the most developed field in recent years, enhanced by the most recent advances in AI, is the recognition and classification of images. This discipline brings together dozens of use cases in the most diverse sectors. Let's see some examples.

The agribusiness sector has been greatly benefited by the advances in automatic image classification. Images taken by satellites and, more recently by drones, allow precision agriculture applications. The classification of images (not only in the visible spectrum) makes it possible to determine the concentration of certain substances in the soil or the level of irrigation required.

The planning of the territory improves considerably when we introduce the automatic cataloguing of plots and land.

Nature conservation and the fight against climate change also benefit from the latest advances in artificial intelligence. There are few apps that allow you to take pictures of flora and fauna in order to perform an automatic classification of the photographed object and determine the type of animal or plant species. In the same way, the analysis of the historical (temporary) series of climatological data (using artificial intelligence techniques) allows to predict the speed of desertification of the ecosystems or the future rises of the sea level due to climate change. Of course, these types of AI applications would never be possible without open datasets in the form of image databases that allow training algorithms that identify objects.

Natural environment aside, automatic image recognition is transforming many other sectors.

The security and defense sector is extracting valuable information through automatic image classification. A clear example is the controversial application of citizen security of the Chinese government that applies AI to easily recognize a person in less than two seconds.

The health sector is not left out. The enormous potential of AI in medicine drives multimillion-dollar investments in the development of use cases. From the identification of pathologies by medical imaging to the early detection of cardiac abnormalities by wearable devices.

Again, the development of these applications of enormous social impact will not be viable without the contribution of open data communities. Although the medical field has not been historically characterized by promoting open use of data (mainly for privacy reasons) there are already numerous open data repositories of invaluable value in the training of AI models for medical science.

5. ACTION

In the Methodology section, we introduced the reader on the way in which this report is structured. The AIA (Awareness, Inspire, Action) tour allows us to gradually enter into the subject of artificial intelligence, from the most basic concepts to the development of a practical case indicated for those readers who want to move on to the Action. In this section we have decided to use the **superpower** that artificial intelligence gives us to improve our ability to interpret images. AI will see through our eyes and analyse through our brain the content of some photographs. Perhaps AI is still not much superior to humans by identifying everyday objects in images but it is much more efficient if we try to classify hundreds of thousands of images in a few seconds.

5.1 The dataset

In this use case we will use a set of images available in the data catalogue of datos.gob.es. In particular, we will use the Photographic Archive of the Basque Government: images about Euskadi and the Government activity. This photographic archive contains thousands of images in jpg format about nature, everyday life, common objects, etc. With the development of this use case, we exemplify how AI helps humans classify images and write down their description. Without the help of an AI, image classification can be a long and boring task for a human. With the help of AI, humans can free up time to do human activities such as developing creative or artistic tasks, activities still outside the real scope of AI. Already done the introduction to our exercise. Let's get started!

The photographic archive is available for download in the following link. The contents of the Irekia portal are subject to a Creative Commons Attribution license unless otherwise indicated. An example of the images that we can find is illustrated in Figure 6.

***Figure 6**. Example set of photographs available on the website*
http://argazki.irekia.euskadi.eus/es/photos

As we mentioned, AI can help us in the classification of images. As we observe when we click on a particular image (as in figure 7), **the website asks us for help to improve the classification**. In this particular case, we see how the image is labelled with the description *birds*, in addition to the additional tags *Animals*, *Wildlife*, *Birds*, etc. We can imagine how the work of opening image by image and adding description and labels manually is a task that can leave any human exhausted. However, let's now see how AI performs this task effortlessly as many times as we want.

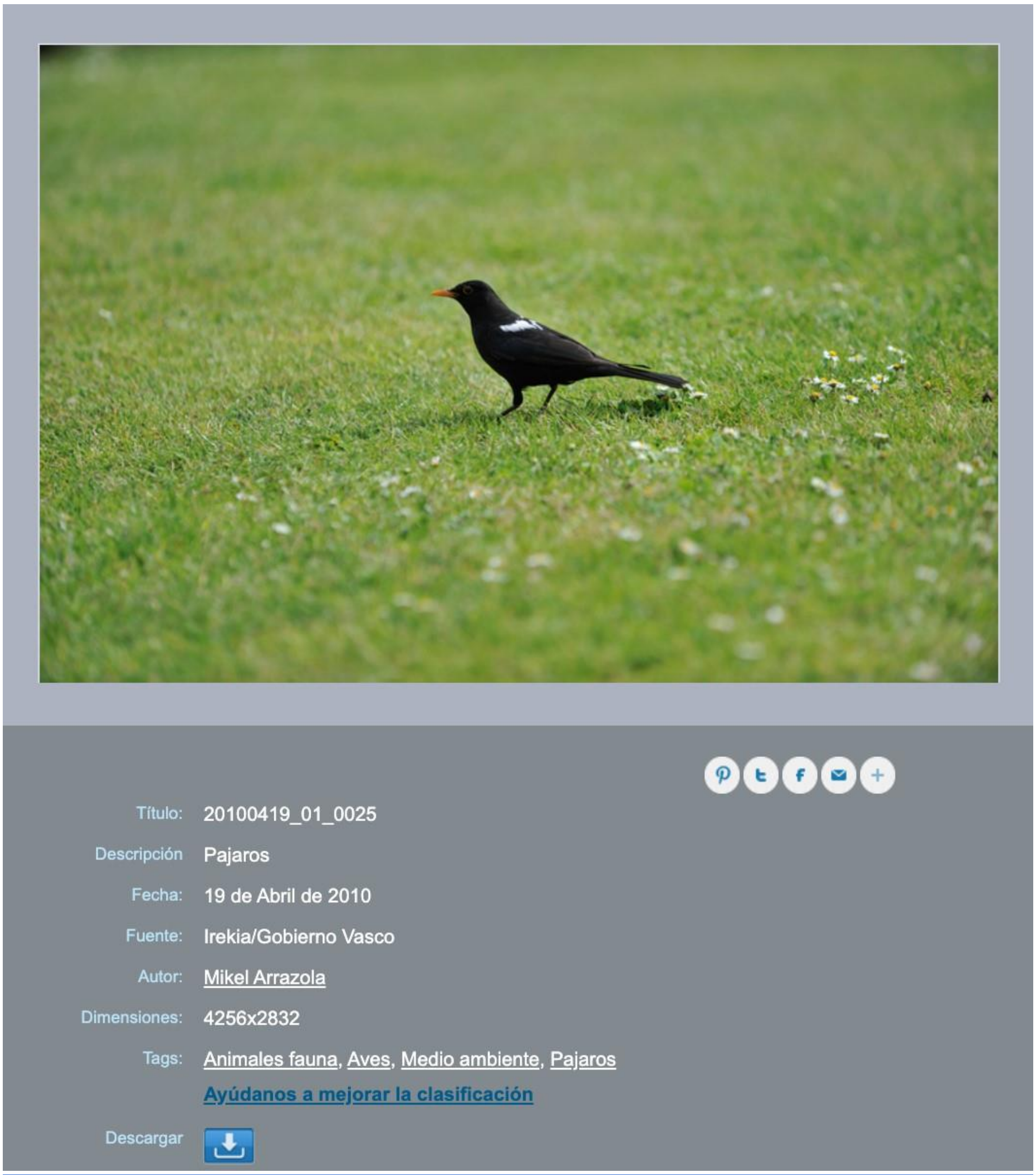

***Figure 7**. Particular image of the image repository http://argazki.irekia.euskadi.eus/es/photos*

5.2 The technology

We have talked about artificial intelligence more as a concept than as technology. As we have previously introduced, the technical discipline that makes AI possible in the world of algorithms is known as Deep Learning. However, below, we describe very briefly the technological tools that will allow us to load and classify images automatically.

To develop our example of automatic image classification, we will use some tools that will considerably facilitate the necessary effort. The Deep Learning algorithm for classifying images is very complex and we could not dedicate this section to creating a classification algorithm from scratch. However, we can use pre-trained models, programming frameworks and Deep Learning development that make our work easier. In particular, in this use case, we will use R as a programming language, R Studio as an environment or IDE (*Integrated Development Environment*) and Keras as a high-level neural network API, written in Python and capable of running on TensorFlow. The Keras package itself incorporates several **pre-trained models** that can be used for this and many other use cases. The selection of the most appropriate model (scoring models) for each type of problem is one of the most important tasks of the data scientist.

In any data science process that involves whether machine learning or deep learning, the initial dataset is separated into several parts. One of the parts is used to tell the training algorithm to look for hidden relationships in the data that link some variables (features) with that variable we want to predict (target). In this case, the variable we want to predict is the label that classifies the content of the image. The variables (features) involved in the prediction are related to the information contained in the pixels of the image. This process, known as **model training**, is the most demanding (computationally speaking) and consumes extensive computing resources. Usually, these computing resources are machines with the ability to perform very complex mathematical calculations (tensor algebra) in graphic processors or GPUs (Graphic Processor Unit).

Depending on the complexity of the training algorithm and the volume of the dataset used in this training, **calculations can take hours or even days**. For this reason, in this example, we are going to use an already trained algorithm, which only has to examine the new data (new images) and use the trained algorithm to extract the classification of the image. By using this approach, we can classify images from the beginning without devoting time and effort to training. In return, the accuracy of image classification will never be as high as if we used part of the set of images from the open data repository to train the algorithm.

5.3 The solution to the problem

The steps we have to take to build our simple automatic image classification system are:

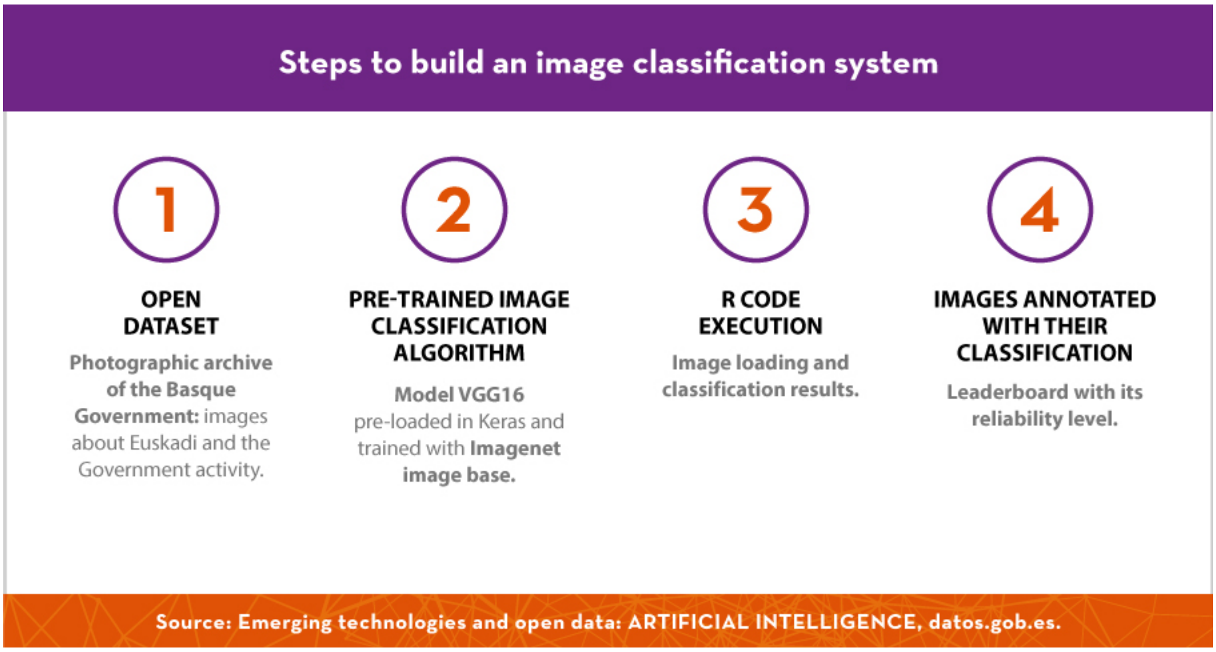

Figure 8. *Diagram of the steps to build our simple classification system.*

Once we have the programming environments and the necessary libraries installed [5], we start loading the keras library.

[5] Annex I details the installation instructions for the required environment.

```
library(keras)
```

We clean the previous environment to eliminate previous executions and start in a clean workspace without previously stored variables and objects. This ensures a correct execution of the following lines of code.

```
# Clear out the session
k_clear_session()
```

We load the image classification model that has been previously trained with the [Imagenet](#) image base. Imagenet is an image bank organized around a hierarchy of names. Imagenet images include English tags. Hence the results of our classification model are descriptions of the images in English.

```
model <- application_vgg16(weights = "imagenet")
```

Once the trained model that we are going to use to classify our images is loaded, it is time to load the first images and perform the first tests.

```
# The local path to our target image
img_path <-
"http://argazki.irekia.euskadi.eus/photos/p740/20100522_01_0203.jp
g"
```

We have loaded the following image (figure 9)[6] from the open data repository:

Figure 9. *Image titled 20100522_01_0203. CC BY-3.0-ES 2012/EJ-GV/Irekia-Gobierno Vasco/Mikel Arrazola'*

The labels entered manually for this image are:

[6] If you receive an error message when loading the image is due to the fact that the web http://argazki.irekia.euskadi.eus/ forces to mark a checkbox before downloading the image and this causes that there is no Direct link to download the image. In some cases, it may be necessary to download the image to a local directory and upload from there.

Bee-eater, Wildlife animals, Birds, Environment, Birds

The following code fragment adapts and prepares the image to be understood by the algorithm that will be in charge of classifying. It is recommended that the model work with smaller images to ensure reasonable execution times. In addition, the model needs the image in the array form with the 3 basic colours separated by channels.

```
# Start witih image of size 224 × 224
img <- image_load(img_path, target_size = c(224,
224)) %>%
# Array of shape (224, 224, 3)
image_to_array() %>%
# Adds a dimension to transform the array into a
batch of size (1, 224, 224, 3)
array_reshape(dim = c(1, 224, 224, 3)) %>%
# Preprocesses the batch (this does channel-wise
color normalization)
imagenet_preprocess_input()
```

After these previous steps, we have reached the crucial moment in which the algorithm will try to classify the image using the vgg16 classification model trained with the Imagenet image bank.

```
preds <- model %>% predict(img)
imagenet_decode_predictions(preds, top = 3)[[1]]
```

The two previous lines of code call the model prediction and filter the results of the classification, restricting the output of the algorithm to the 3 results (labels) with greater relevance. Image classification models assign a probability to the classification results. That is, the model returns the label that it think corresponds to the image along with the probability that this label is correct.

In our specific case, the model returns the following table with the results of the classification:

class_name	class_description	score
<chr>	<chr>	<dbl>
n01828970	bee_eater	0.9995731711
n01530575	brambling	0.0001940502
n02011460	bittern	0.0001244307

The analysis of the output table after the execution of the model (in less than 5 seconds) indicates that, this model predicts that the image is a *bee eater* (literal transcription) with a 99.95% probability. The model has also calculated that the image could be a chaffinch with a 0.019% probability and an *avetoro with a* 0.012%. That is, the model classifies this image as a *bee eater* with almost total security. Unfortunately, we are not experts in birds, so to check if the model has been successful, we perform an image search on Google for the word bee_eater. This is what we found (figure 10).

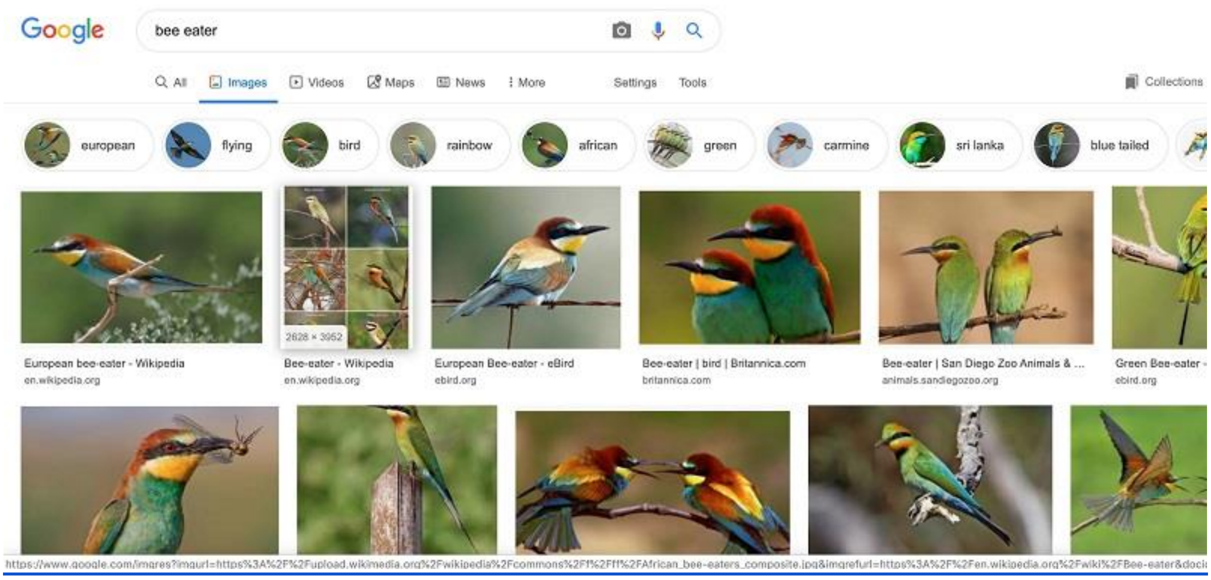

Figure 10. *Search results for images of Bee Eater on Google.*

We can see how, effectively, Google search returns hundreds of images similar to the one we used as an unknown model entry.

In this case, the model predicts the label of the image and is rigth. Let's look at another example in which the result is not so accurate.

Let's take the following image (figure 11) of the same repository whose manual tags are:

Animals fauna, Hunting, Sport, Digital, Daytime, Epagneul breton, Exterior, Format, Mammals, Environment, Dogs, Medium shot, Photo shoots.

Figure 11. Image titled 20030930_01_0026. CC BY-3.0-ES 2012/EJ-GV/Irekia-Gobierno Vasco/Mikel Arrazola'

Let's run the model again to get the results.

```
# The local path to our target image
img_path <-
"http://argazki.irekia.euskadi.eus/photos/p740/20030930_01_0026.jpg"
```

```
# Start witih image of size 224 × 224
img <- image_load(img_path, target_size = c(224, 224)) %>%
  # Array of shape (224, 224, 3)
  image_to_array() %>%
  # Adds a dimension to transform the array into a batch of size (1, 224, 224, 3)
  array_reshape(dim = c(1, 224, 224, 3)) %>%
  # Preprocesses the batch (this does channel-wise color normalization)
  imagenet_preprocess_input()

preds <- model %>% predict(img)
imagenet_decode_predictions(preds, top = 3)[[1]]
```

class_name <chr>	class_description <chr>	score <dbl>
n02100735	English_setter	0.2331140
n02091244	Ibizan_hound	0.1794144
n02101388	Brittany_spaniel	0.1744492

In this case we see how the first three results of the classification model only add up to 58.7% of the total probability. The best result of the model (with a probability of 23.3%) corresponds to the label of the English Setter image. If we perform the same Google search (figure 12) the results are more disparate, although they still seem quite correct.

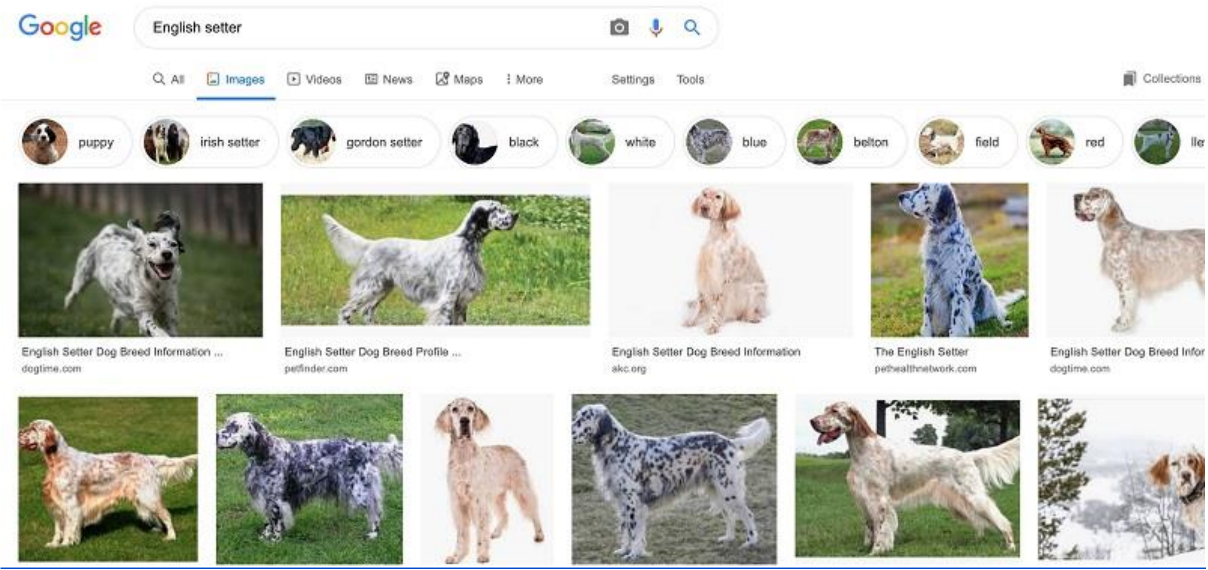

Figure 12. *Search results for English Setter images on Google.*

In this example, we have seen how to easily implement a small code in R language that allows us to classify common images available in an open data repository. The results seem very reasonable even when the model used has been trained with a set of images that has no relationship with the one we use here. We encourage readers to try to reproduce the example with images available in the open data repository or with any other image they consider interesting.

6. NEXT STOP ...

If you have not had enough with Artificial Intelligence through the different chapters of this report, we have done well. We had to stop here, but we hope you don't do it. Therefore, here is a collection of highly recommended readings for you to become an expert in Artificial Intelligence.

6.1 Complete collections on AI

If you want to explore in an interactive way all the possibilities of AI at present you have to visit this website enabled by Forbes. The format is simply spectacular and the amount of information you can access is simply impressive.

- https://www.forbes.com/insights-intelai/ai-issue-1

McKinsey & Company makes available a collection of resources on AI with resources

- https://www.rev.com/blog/artificial-intelligence-machine-learning-speech-recognition

In the same way, Price Waterhouse Coopers, Accenture and the World Economic Forum have their own AI collections.

- https://www.pwc.com/gx/en/issues/data-and-analytics/artificial-intelligence.html
- https://www.accenture.com/us-en/insights/artificial-intelligence-index
- https://www.weforum.org/platforms/shaping-the-future-of-technology-governance-artificial-intelligence-and-machine-learning

Do you need to know more about Machine Learning and Deep Learning? Surely you love this post in Medium:

https://medium.com/activewizards-machine-learning-company/artificial-intelligence-vs-machine-learning-vs-deep-learning-what-is-the-difference-a5e2bc8b835f

6.2 Artificial Intelligence for applications on language

- https://heartbeat.fritz.ai/a-2019-guide-for-automatic-speech-recognition-f1e1129a141c
- https://blog.insightdatascience.com/how-to-solve-90-of-nlp-problems-a-step-by-step-guide-fda605278e4e
- https://medium.com/@mattkiser/an-introduction-to-natural-language-processing-e0e4d7fa2c1d

6.3 Artificial Intelligence for applications on vision and image recognition

- https://towardsdatascience.com/train-image-recognition-ai-with-5-lines-of-code-8ed0bdd8d9ba
- https://medium.com/deeplearningsandbox/how-to-use-transfer-learning-and-fine-tuning-in-keras-and-tensorflow-to-build-an-image-recognition-94b0b02444f2
- https://medium.com/density-inc/ai-is-not-magic-its-manual-labor-math-how-we-built-an-accurate-people-counter-e00408ea30de
- https://towardsdatascience.com/how-to-do-everything-in-computer-vision-2b442c469928

7. ANEX I. DETAILED INSTRUCTIONS FOR REPLICATING THE EXAMPLE OF THE ACTION SECTION

In order to test the example that we have illustrated in the Action section, it is necessary to install the necessary software packages.

As in most cases, the software for data analysis and Deep Learning in particular, depends on the base platform we use (usually we understand by platform the operating system we use such as Windows, Linux, macOS, etc.)

Without a doubt, the most convenient platform for working in data science is a Linux system. Most of the tools we have used in the Action section are native to Linux or have abundant documentation for installation in Linux environments. This does not mean that it is not possible to reproduce the same example on Windows platform, macOS or other system, but it can be relatively more expensive if we encounter problems during installation or configuration.

That said, in our specific case we will use a different approach to having to choose a specific platform for this example. We have decided to use **a container approach** to download an image of the required software in a macOS Catalina 10.15.1 system. For those users who do not have previous experience with container systems, we recommend the following web resources. It is necessary to have some basic notions of containers to be able to perform the steps that follow.

What are the containers?

Basic Container Terminology

Why Docker for data science?

Once we understand the basics of container technology and the concept of an isolated environment to run software, we proceed with the process to reproduce Action.

Step 1. Download and install Docker (in our case for MacOS)

https://docs.docker.com/docker-for-mac/install/

Step 2. Download and Install Kitematic.

Kitematic is a visual user interface that allows us to locate and download Docker images available in the Docker Hub.

Step 3. Download an available RStudio image.

Open Kitematic and locate the image **rocker / tidyverse** and press **create**. Once the download process of the image is complete, we execute the following command in our terminal:

```
sudo docker run -d -p 8787:8787 -e PASSWORD=<your_password> --name deeplearning rocker/tidyverse
```

Step 4. We execute our container.

Once here, we have a container running an isolated version of RStudio on our computer.

Step 5. We access the RStudio environment from our browser.

We open a new tab in our browser and write in the address bar http://localhost:8787. The login screen appears where we enter the user (rstudio) and the password that we have chosen in the previous step.

Step 6. Associate a storage volume with the container.

As we have commented in the introduction of this annex, containers are isolated environments that, a priori, do not have an interface with our system. Practically it means that a container cannot see our hard drive if we do not allow it. To be able to save files on our computer so that Rstudio could be able to find them, we have to configure the container so that it can see a part of our hard drive. This will allow us to download the images from the open image repository and then execute the classification algorithm.

To associate a storage volume with our container (this is called the way our computer sees our disk) we simply go to the Kitematic Volumes section and configure the directory path we want to see from our container.

Step 7. Installation of Keras

To install Keras in our isolated Rstudio environment, we execute the following commands in the Rstudio console:

```
install.packages("keras")
require(keras)
install_keras()
```

Step 8. We execute the example of the Action section, loading the pre-trained model