

Data-driven educational algorithms pedagogical framing

(El encuadre pedagógico de los algoritmos educativos basados en datos)

Daniel Domínguez Figaredo

Universidad Nacional de Educación a Distancia, UNED (España)

DOI: <http://dx.doi.org/10.5944/ried.23.2.26470>

How to reference this article:

Domínguez Figaredo, D. (2020). Data-driven educational algorithms pedagogical framing. *RIED. Revista Iberoamericana de Educación a Distancia*, 23(2), pp. 65-84. doi: <http://dx.doi.org/10.5944/ried.23.2.26470>

Abstract

Data from students and learning practices are essential for feeding the artificial intelligence systems used in education. Recurrent data trains the algorithms so that they can be adapted to new situations, either to optimize coursework or to manage repetitive tasks. As the algorithms spread in different learning contexts and the actions which they perform expand, pedagogical interpretative frameworks are required to use them properly. Based on case analyses and a literature review, the paper analyses the limits of learning practices based on the massive use of data from a pedagogical approach. The focus is on data capture, biases associated with datasets, and human intervention both in the training of artificial intelligence algorithms and in the design of machine learning pipelines. In order to facilitate the adequate use of data-driven learning practices, it is proposed to frame appropriate heuristics to determine the pedagogical suitability of artificial intelligence systems and also their evaluation both in terms of accountability and of the quality of the teaching-learning process. Thus, finally, a set of top-down proposed rules that can contribute to fill the identified gaps to improve the educational use of data-driven educational algorithms is discussed.

Keywords: teaching practice; learning conditions; sciences of education; experimental education; educational research; electronic data processing.

Resumen

Los datos procedentes de los estudiantes y de las prácticas de aprendizaje son esenciales para alimentar los sistemas de inteligencia artificial empleados en educación. Asimismo, los

datos generados recurrentemente son fundamentales para entrenar los algoritmos, de manera que puedan adaptarse a nuevas situaciones, ya sea para mejorar el ciclo de aprendizaje en su conjunto o para gestionar tareas repetitivas. A medida que los algoritmos se propagan en diferentes contextos de aprendizaje y se amplía su capacidad de acción, se requieren marcos pedagógicos que ayuden a interpretarlos y que amparen su uso adecuado. Basándose en el análisis de casos y en una revisión de la literatura científica, en este artículo se analizan los límites de las prácticas de aprendizaje fundamentadas en el uso masivo de datos desde un enfoque pedagógico. Se toman en consideración procesos clave como la captura de los datos, los sesgos en las bases de datos y el factor humano que está presente en el diseño de algoritmos de inteligencia artificial y de sistemas de Aprendizaje Automático. Con el fin de facilitar la gestión adecuada de los algoritmos educativos basados en datos, se plantea la idoneidad de introducir un marco pedagógico que permita analizar la adecuación de los sistemas de inteligencia artificial y apoyar su evaluación, considerando su impacto en el proceso de aprendizaje. En ese sentido, se propone finalmente un conjunto de reglas de enfoque heurístico con el fin de mejorar los vacíos pedagógicos identificados y que puedan apoyar el uso educativo de los algoritmos basados en datos.

Palabras clave: práctica pedagógica; condiciones de aprendizaje; ciencias de la educación; pedagogía experimental; investigación educativa; tratamiento electrónico de datos.

The ability to access directly the large amounts of data from online learning platforms is affecting the establishment of the purposes, procedures and the very consideration of educational practices based on digital data. At the same time, the growth of digital learning spaces is boosting basic research on learning processes based on the huge volume of digital data available.

In order to address the challenges of massive data analysis in the study of digital learning experiences, new disciplines –such as learning analytics (Siemens et al., 2011; Buckingham & Ferguson, 2012; Greller & Drachler, 2012)– combining computer science, mathematics and applied statistics have been introduced (Gitelman, 2013; Kitchin, 2014). Educational research is also increasingly using automatic processes that rely on available information to intervene directly in the learning cycle –i.e. predictive learning analytics, student modelling, recommendation systems, or educational process trace analysis (Breslow et al., 2013; Thille et al., 2014) are all methods that use Artificial Intelligence (AI) algorithms to adapt course design to student needs–. In addition, digital data are also used to design and train Machine Learning (ML) based applications to guide students, and monitor and evaluate learning (Hew, Qiao, & Tang, 2018; Hussain, Zhu, Zhang, Abidi, & Ali, 2019).

Along with the emergence of new methods and disciplines, there is a debate about the change involved in accessing information on student behaviour directly, without previous filters or, at least, without the type of conceptual and methodological filters used previously –i.e. statistical inference, sampling, theoretical framing, etc.–. And, in the same way, radical changes are being discussed in the epistemic conditions that

support the ethical regulation of research and intervention in students' daily lives (Crawford, 2016; Farrow, 2016; Metcalf, Keller, & boyd, 2016; Amo et al., 2019).

In that context, this document attempts to frame the main current debates on the use of AI in education by providing a pedagogical view from the educational sciences (Goksel & Bozkurt, 2019; Luckin & Cukurova, 2019; Sharma, Kawachi, & Bozkurt, 2019; Sloane & Moss, 2019; UNESCO, 2019; Zawacki-Richter, Marín, Bond, & Gouverneur, 2019). AI is the combination of a certain type of technology – an algorithm– and a large set of data; and it also includes non-human data, product design and the software used (Sinders, 2019a). AI-based systems and products can affect learning in many ways and, above all, is currently changing the face of educational research and technological interventions aimed at improving the learning cycle. Thus, applying AI in learning contexts involves addressing many of the conceptual and epistemic concerns of data-based educational research (Domínguez, Álvarez, & Gil-Jaurena, 2016).

The paper discusses the pedagogical principles associated with data-driven educational algorithms in order to provide useful rules to guide their design and application in educational spaces. According to the previous analysis by Houlden & Veletsianos (2019), a critical and relevant example-based approach together with a literature review is applied here to conduct the analysis. Firstly, the importance of the human component in the design of AI and ML systems is described. It then analyses the need to introduce a pedagogical dimension that frames the specifically educational aspects arising from data privacy, algorithmic biases and enhanced surveillance systems. Finally, based on the identified pedagogical elements, a heuristic approach is used to propose a set of rules to guide the design and evaluation of data-driven AI applications in education. It is intended to serve as a theoretical precedent to empirically validate a set of criteria for the implementation of AI-based learning systems in education.

THE HUMAN FACTOR IN EDUCATIONAL ALGORITHMS

The data determine much of what educational algorithms do. The data that feed the educational algorithms are a variety of inputs that people make, such as what they choose to like online, what they comment on, how often they check something, and when they use something. They are constantly feeding into the algorithm within the myriad of existing AI-based products, such as recommendation systems, text editors, conversation robots, or activity supervisors. In this way, the data are activated: they have a particular purpose and can become as important as the code of the algorithm (Sinders, 2019b).

But the data is not the main element that determines how the algorithms behave. System design and, especially, human decisions about how to combine data sets are fundamental to understanding how an algorithm uses data.

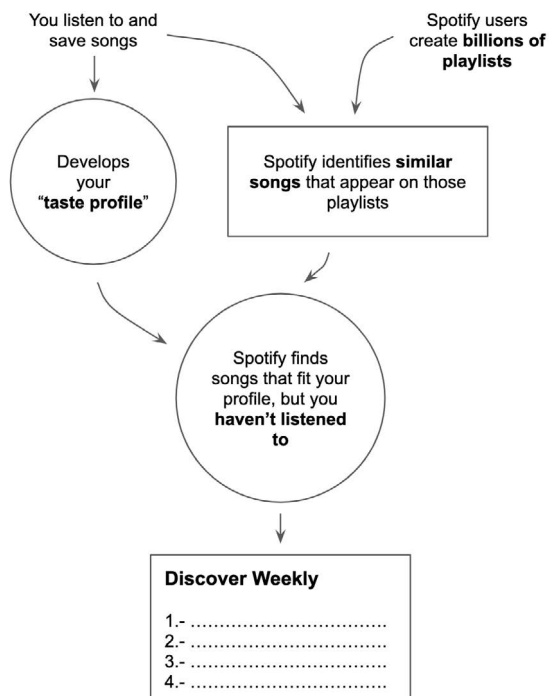
Core decisions in predictive analytics

This is the case, for example, with learning recommendation systems, which is one of the outstanding features in e-learning products and also supports institutional strategies for student recruitment and retention (Bodily & Verbert, 2017; Prabhakar, Spanakis, & Zaiane, 2017; Romero & Ventura, 2017). In general, recommendation systems are algorithms that aim to suggest relevant elements to users such as movies to watch, products to buy, text to read, learning activities to do, or courses to enrol in. In education, recommendation systems are the main product of predictive analysis, which many colleges and universities use to achieve their student recruitment objectives, focusing on enrolment strategies and adjusting scholarship policies. Demographic and performance data can help educational institutions predict whether a student will enrol in a course, whether once enrolled he/she will stay on track during his/her learning cycle, and whether he/she will require support not to fall behind before completion. Predictive analytics are also used to better tailor counselling services and to personalize learning with the goal of improving student performance (Domínguez, 2018).

To explain how these systems work, as well as the human component in algorithm modelling, the case of Spotify's recommendation app called *Discover Weekly* is described (see Figure 1). *Discover Weekly* is a playlist of songs created from a combination of user data and algorithmic inference. In order to display a suitable playlist to a target person, the system initially relies on other people's playlists. Spotify commences by looking at all the playlists created by users, which contain a reflection of their interests and sensitivities. These human-made song selections and groupings are at the heart of *Discover Weekly*'s recommendations. From there, the algorithm gives extra weight to the company's own playlists and the lists that have the most followers. It then attempts to fill in the gaps between the target person's listening habits and those with similar interests. Consequently, if Spotify detects that two of the target user's favourite tracks tend to appear in other playlists along with a third track that the target has not listened to before, it will suggest the new track. In addition, Spotify also creates a profile of each user with their particular music interests, grouped into singer sets and music genres. Finally, the algorithms are responsible for connecting the data from the millions of playlists and the personal interest profile (Pasick, 2015; Sinders, 2019c).

The approaches behind this process of configuring Spotify's algorithms include collaborative filtering and natural language processing, which are automatic selection systems, along with deep learning, which is a technique for recognising patterns in huge amounts of data using powerful computers that are *trained* by humans to improve their selections (Johnson & Newett, 2015).

Figure 1. Spotify Discover Weekly algorithm model



Note. Reprinted from “The magic that makes Spotify’s Discover Weekly playlists so damn good,” by A. Pasick, 2015 (<https://qz.com/571007/>). Copyright 2019 by Quartz & Nikhil Sonnad

In educational contexts, automatic referral systems meet the same requirements as *Discover Weekly*. To make the results fit the interests of the students, it requires previous access to the trace data generated in the interaction with the educational software, mainly with the Learning Management Systems (LMS). The decisions about which data to obtain or how to combine them do not correspond to the algorithm, but to the people in charge of modelling the information and designing the automatic processes that will later be executed by the algorithm. When it happens in learning contexts, many questions arise that have a clear pedagogical component.

On the one hand, students may wonder how a certain sequence of recommendations came to exist. Which concrete data trained the algorithm. Whether the algorithm infers only from the learning habits of a single student, or whether it takes into account the most popular patterns among the set of actions performed by all students in the LMS. If it takes into account one gender over another, or the time when the actions happen. Whether the actions made by friends –i.e. people you

have contact with within the LMS, or eventually outside on social networks– have an effect on the suggestions made.

What is more, from the perspective of the teacher who uses AI-based software in the classroom (Smith, 2019), it is necessary to know the rudiments behind the technologies employed. To improve teaching, it is equally necessary to have the ability to adapt the system to the specific learning practices that arise spontaneously. This aims to prevent the biases and issues associated with current AI-based learning systems which, as mentioned above, require human intervention –and in this case, also the application of a pedagogical vision– decisively to operate properly in a given learning context.

Machine learning pipelines in educational contexts

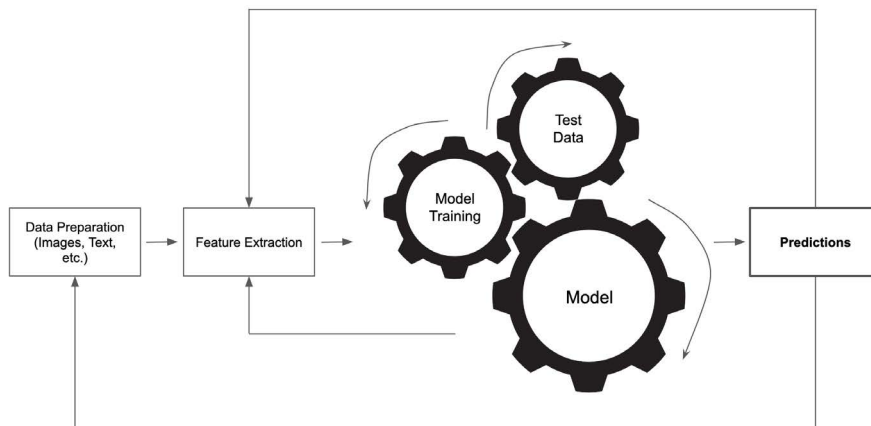
In addition to AI systems for recommendations, there are educational applications of ML –a subset of AI– especially oriented to the grading process (Alsuwaiket, Blasi, & Al-Msie'deen, 2019), predictive analytics (Uskov, Bakken, Byerly, & Shah, 2019), and identification of learning paths adapted to each student (Kurilovas, 2019). And as it happens with the data-based AI applications, also in the design of educational systems based on ML there is an outstanding human component that requires a pedagogical approach.

ML pipelines consists of the steps to train a data model. It helps to automate the workflows leading to the design of an ML algorithm. It is a cyclical and iterative process, as each step is repeated to continuously improve the accuracy of the model and to have an efficient algorithm. Many of the current ML models are *trained* neural networks, capable of executing a specific task or providing knowledge derived from *what happened* to *what is likely to happen*. They are complex models that are never completed. Rather, through repetition of mathematical or computational procedures, they are applied to the previous result and improved each time to obtain closer approximations to *problem solving*. Thus, a huge amount of data, processed iteratively, are required to provide the resources to train the ML models (see Figure 2).

One last element to consider in order to obtain good results in the processing of large volumes of data, is the value of the metadata. Metadata resides with the captured data and provides descriptive information about the digital objects –which aggregate data– and the autonomous data. Metadata extraction and correlations between them are the basis of ML models. This is due to the need to work with tags in order to associate data that considered independently would be difficult to handle with each other (Zhou, 2018).

Figure 2. A Standard Machine Learning Pipeline

A Standard Machine Learning Pipeline



Note. Reprinted from “How to Build a Better Machine Learning Pipeline,” by L. Zhou, 2018 (<https://www.datanami.com/2018/09/05/how-to-build-a-better-machine-learning-pipeline/>). Copyright 2020 by Datanami & Western Digital

Moving that process into the field of education, the main consideration relates to the types of essentially educational tasks required to work with ML systems. Designing the model, training the model, and testing and tagging the data are all human tasks. People are needed to train the models, because currently this task cannot be done without the participation of people. And it is those people who make decisions about what happens to the ML systems, where they are going to be used and for what purposes.

The main pedagogical concerns here are related to the evaluation of the whole system, so that the training of the model is properly oriented to the requirements in terms of learning improvements, without deviations, once several iteration cycles have passed. Additionally, we must also consider the adequate pedagogical approach of the whole system, in terms of fostering the adequate development of skills and competencies of students (Reich, 2014).

MISALIGNMENTS IN DESIGNING DATA-DRIVEN ALGORITHMS

Over the past few years educational sciences have developed a set of conceptual, policy and institutional resources based on how to work with data from learning practices. But AI educational systems are questioning the strict application of that framework to the case of digital data. When researching in a digital context, many open questions arise on substantive issues: whether research methods and programmes

based on digital data should be excluded from current ethical frameworks, or are required to comply with existing standards; whether these current standards should be adapted to the special circumstances of digital systems, or whether completely new standards and institutional commitments are needed.

So working on AI requires expanding the framework for educational research. Data from students' digital practices become –at least, in theory– indefinitely connectable and reusable, continuously updateable and easily removable from the context in which they were collected (boyd & Crawford, 2012; Zwitter, 2014). These features that characterize digital systems challenge the limits corresponding to analogue practices, which depend on data that are bounded in time and context, and which are highly constrained by technical infrastructure and financial cost.

A set of methodological challenges associated with the educational use of automatic data processing technologies is analysed below. The concerns involved in the socio-educational use of data-based technologies are raised (Tufekci, 2013; Pitcan, 2016; Bulger, 2016; Caplan, Donovan, Hanson, & Matthews, 2018; Perrotta & Selwyn, 2019), and from there a renewed approach is provided to improve learning based on the management of students' digital data.

Data set and platform bias

As mentioned, to suggest recommendations predictive AI systems study people's behaviour and relate it to some pattern that can explain their actions and, especially, predict their behaviour in the future. In the case of e-learning, the data analysed come from highly complex situations, with multiple meanings and whose interpretation depends largely on the context in which they have been collected. The main element that determines the context is the specific digital platform where the learning activity takes place. This is so important that the same behaviour could have different meanings depending on the platform on which it occurred.

For example, in the case of research on social behaviour on the Internet, the most analysed platform has been Twitter. However, Twitter is far from being a platform that represents the set of digital applications that allow social interaction. Each platform incorporates certain specific functionalities that may not be representative of other social platforms or of human social behaviour in general.

As for Twitter and social networks, in education the platform that has been most researched from learning analytics methodologies has been Open edX. This is mainly due to the fact that it is a free, open source tool that was originally developed for the courses of the edX project, which is the main MOOC site on the Internet.

The multiple studies and experiments on student activity in Open edX have led academics to suggest a general framework for student behaviour in online courses. The framework addresses such important issues as communication in the forums, course completion rates and teacher assignments. However, the Open edX platform does not have some of the features that are common and widely used in other

tools, such as Moodle, Canvas or Blackboard platforms, which are leaders in the LMS market. For example, Open edX differs from Moodle in aspects such as the integration of visual elements into text, the monitoring of forum discussions or the management of assessment tests. Open edX's simple interface is well suited for use on mobile devices, making it the preferred platform for studying in mobile situations or from low-bandwidth environments. The mechanism for consulting video classes also causes a particular behaviour, since it is based on a series of viewing rules that are not necessarily equivalent or correspond to the way audio-visual content is consumed on other digital platforms.

To compensate for the shortcomings of the single-platform research models, the data sets involved should be extended to cover the emerging ecology of the contexts that are related to the phenomenon under analysis (Ruipérez-Valiente, Halawa, Slama & Reich, 2019). This does not mean that nothing valuable can be investigated from a single-platform analysis. Rather, it is to assume that these analyses are examining a closed system. And that, ultimately, the solution to this limitation of research based on specific data sets may not be solved by learning analytics methodologies alone.

Searching for tags and keywords in single case studies

Many educational studies with big data –later taken as a reference for modelling AI software– extract relevant text from a platform using tags or keywords. For example, in a course's virtual forum, messages are analysed for words such as *exam*, *query*, or *thanks*. While studies based on tags and keywords can be a powerful method to examine the flow and subject matter of conversations in a course, they are analyses built on the basis of selecting the dependent variable, which is the one that corresponds to the case under study, with all the characteristics and weaknesses that entails using such a methodological route.

In a social investigation, a sample comprising one or several cases has limited analytical power and could offer misleading results, since the variation in the dependent variable is limited (Geddes, 1990). For example, if research is conducted on the essential conditions for students to better understand a topic within a course by looking only at cases of successful courses that have occurred, the explanatory power will necessarily be limited. To improve explanatory power, it would be necessary to also include cases that might have similar characteristics, but where failures have occurred and students have not adequately understood the topics. In the same vein, in keyword-based datasets, a message is included in the dataset precisely because it has a particular outcome already associated with it. In addition, most keywords used to create large datasets are examples of successful terms, which are well known, widely distributed and generate great interest. This calls into question the capacity of this type of study and points to the need to open up the design of research by incorporating a wider variety of techniques and instruments for analysis.

Correlation does not imply causation, even for algorithms

Related to the above assumption, there is a close relationship between the selection of dependent variable features and the attribution of specific factors on which the uncorrelated sample features depend. That is, a self-selected population will not only have general characteristics different from those of the general population, but may also exhibit significantly different correlation trends. This creates –at least– two types of problems.

On the one hand, there is confusion in the variables analysed. Following the example of the tags associated with a message in a forum, these are often related to assumptions, meanings and the cultural or political structure of the context where the conversation takes place. Therefore, the use of tags, in addition to being a method of self-selection, often involves participation and commitment to the framework that the tag integrates. The biases inherent in this situation prevent the conclusions from being generalized to other contexts, which limits the research.

However, the main mistake that research designs that confuse the dependent and independent variable can make is the assumption that the correlation between the factors or traits observed simultaneously in the variables implies some kind of causality between them. This is a common fallacy in the field of statistics, which consists of inferring that there is a causal relationship between two or more events because a statistical correlation between them has been observed, and that big data studies have helped to generalise in part for the reasons given above (Muller, 2018).

Big data studies often emphasize the variations and slides that occur in large volumes of data and assign simple explanations to the complex phenomena behind those variations (Michael & Miller, 2013; Poel, Meyer, & Schroeder, 2018; Brady, 2019). One example is studies at the level of the education system, such as those that analyse the segregation of students in neighbourhoods according to socioeconomic level (Ball, Bowe, & Gewirtz, 1995; Orfield & Lee, 2005), or those that make comparisons between academic performance and other geographic variables such as the country or region of residence of the students (Coleman, 1966; Sirin, 2005). In the history of education there has been much research that has sought correlations between simple variables in order to respond to complex problems, and these have often been questioned over the years. Currently, access to large data sources has opened the door to new and increasingly creative interpretations that are closer to the theoretical approach that supports the studies than the observed evidence (Hansen & Reich, 2015; Monarrez, 2018). Limited funding and time constraints also lead researchers to find causality between factors where there is only apparent correlation that does not always explain the variance in variables analysed in the studies.

Sample limitations

When a study is based on big data, there is a risk of not sufficiently understanding the value of the underlying sample. In social research, the sample corresponds to the selection of people chosen to represent the population where the conclusions are to be applied. Since often not all of the population is available, you must choose a sample that represents it and is manageable. The study is applied to the sample with the expectation that the conclusions obtained can also be replicated in the whole population.

In the case of big data, the research is usually very extensive and the populations to which the studies are projected are often very large. For example, they may concern all Internet users (González-Bailón, Wang, Rivero, Borge-Holthoefer & Moreno, 2012; Ruths & Pfeffer, 2014; Pfeffer, Mayer & Morstatter, 2018), or in the case of education there may be studies whose findings are intended to apply to all students participating in digital courses, all university students or all schools located in a particular type of neighbourhood (Warnakulasooriya & Black, 2018). As the information available in the massive databases is very numerous, the researcher tends to think that these data are sufficient to represent the population. However, this is not always the case and, if one moves forward without an adequate sample selection, one will be assuming a certain risk. Thus, problems may arise in guaranteeing representativeness and equity when attempting to generalize results to populations that, because they are so broad, are characterized by great heterogeneity.

The lack of representativeness of the sample in the case of massive information sources can be tackled by using selection methods appropriate to the size of the population. This includes using big data also in the previous phases of the study, so that it is possible to segment the large volumes of data available. And, on the other hand, social research is called to imitate experimental sciences and incorporate scales close to 1:1 both in the process of information analysis and in the inference of results, thus expanding the commitment to the social reality which intends to study.

The network structure does not reveal everything

Most big data research uses social network analysis methods. In education, it is common for LMSs to incorporate the feature of displaying network structures created from relationships between students or from their interactions with learning resources in online courses. Social networks analysis tries to know the evolution of the information flows provided by the people who are interacting in a certain context and, for this purpose, it uses graphic representations that show the connections between the nodes that make up the network –which can be people, messages sent to a forum, interactions with a resource, etc.–, filtered by the attributes of those nodes –for example, the subject of a message, the type of interaction, etc.– and according to the weight of the links between those nodes –more or less weight depending on

the role of the person sending the message, whether the resource is autonomous or part of a learning sequence, etc.–

In many cases, researchers using social network analysis take into consideration the structural properties of the whole network to infer from them other properties of the links between the different nodes. For example, one of the most common practices is to connect the links between the alters –an individual’s network consists of an *ego* representing that individual, and his *alter*, which are the others to whom that ego is connected– to the properties of the network structure. This is true only under certain strict conditions where bridging relationships between groups of networks would be more likely to be weak links (Onnela et al., 2007). These are technical issues, but they can lead to inaccuracies as the information contained only in the network structure is limited.

A PEDAGOGICAL FRAMEWORK FOR EDUCATIONAL ALGORITHMS

A set of heuristic top-to-bottom dimensions aimed at filling the gap detected in the design of algorithms in the educational context is proposed below (see Table 1). In social sciences, heuristic-based analytical frameworks are associated with dynamic and open assessment methodologies. Their main utility lies in the formulation of simple evidence-based rules that provide a wide margin for the analysis of cases that depend on a large number of variables, helping to limit the high degree of complexity in those cases. Based on these rules, key performative indicators can be proposed that function under the logic of criteria satisfaction. The criteria are considered satisfied if a minimum percentage of achievement associated with the indicator is covered, which makes the analysis process more open and flexible than control methods based on dichotomous criteria such as A/B type (Gigerenzer & Selten, 2002; Sundar & Singh, 2013; Mousavi & Gigerenzer, 2014).

The proposed scheme is based on the principles already presented and also benchmarked the existing frameworks on the appropriate use of AI systems in other non-educational settings (Saurwein, Just, & Latzer, 2015; Caplan, Donovan, Hanson & Matthews, 2018; Bunker & Thabtah, 2019;; Floridi & Cows, 2019; Jobin, Ienca, & Vayena, 2019;). The aim is to introduce a pedagogical layer in the general rules that guide the design of data-based algorithms (Reif, n.d.), for which dimensions and questions are posed to guide the action, here following the model of Diakopoulos et al. (2017) and US-ACM (2017).

Table 1. Key dimensions for educational algorithms

Dimensions	Guiding questions
<p>Accountability: In education, Algorithms are used to make decisions and allocate learning resources based on large datasets. And algorithmic accountability is the process of assigning responsibility for harm when algorithmic decision-making results in discriminatory and inequitable outcomes.</p>	<ul style="list-style-type: none"> • Are interested audiences informed about the algorithmic decision-making? • Is there a system of internal rules on transparent behaviour? • Are users warned about taking responsibility when interacting with the system? • Are there public measurement criteria for the system's performance?
<p>Biases: When algorithms produce unfair results, we refer to them as biased. Algorithmic biases can occur in many ways: by the social context in which an algorithm is created, as a result of technical constraints, or by the way the algorithm is used in practice.</p>	<ul style="list-style-type: none"> • Is the system design focused on trust? • Is there a decision review mechanism? • Is there a system for social/automatic monitoring of bias? • Is there a system for modification in case of bias?
<p>Data provenance: The data within the algorithms are symbiotic with the algorithm itself. So in product design, the data entered into the algorithms determine the characteristics of a product. When data sets are opaque, there is no way to accurately evaluate the results of digital products.</p>	<ul style="list-style-type: none"> • Is the data properly tagged? • Is the algorithm trained to discriminate cultural variants in the data? • What data is used to feed the suggestions? • Does the data of others affect the suggestions in particular cases?
<p>Explainability: Ensure that both the algorithmic decisions and the data that drive them can be explained to end users and learning management stakeholders in non-technical terms.</p>	<ul style="list-style-type: none"> • Who are the end-users and who are the stakeholders? • What part of the system can be explained to users and stakeholders? • How much of the data sources can be disclosed? • How many of the decisions assumed by the algorithm can be explained?
<p>Fairness: Ensure that algorithmic decisions do not generate discrimination or unfair impacts when different social profile variables are considered.</p>	<ul style="list-style-type: none"> • Is there control of users who may be favoured over the disadvantaged? • Is there control of potentially harmful effects generated by the mistakes of other users? • Is there control over the context in which the system operates? • Are cultural rules taken into account?

Dimensions	Guiding questions
Harmful content: The design of an AI-based product has to consider the type of content that users can add to a repository. It is detrimental not to check and verify whether that content is harmful or not. Policies are required that define the possible damages caused to third parties in terms of containment and actions.	<ul style="list-style-type: none"> • Is there control of false identities? • Is there verification of suspicious content? • Is there a social/automatic damage control system? • Is there a protocol against possible damages?
Pedagogical approach: Not only, but mainly, the design and use of educational data-drive algorithms requires a pedagogical approach. This means addressing, at least, essential issues such as the learning theories behind the AI model, attention to the context of the data, and the usefulness of the output to improve learning.	<ul style="list-style-type: none"> • Have the features of the people involved in the proposal that are of pedagogical interest been properly framed? • What is the educational theory behind the algorithmic decision-making scheme? • Have evidence-based alternatives in the field of learning been considered? • Have the attributes that are of pedagogical interest been adequately contextualized in the data used?
Privacy: The data used in educational algorithms come from individuals. They are intimate data, because conversations and social interactions are various forms of intimacy. So the lack of privacy gradients in the design of the algorithms can facilitate harassment and violations of student privacy.	<ul style="list-style-type: none"> • Have privacy gradients been defined? • Have intimate, personal, social and public spaces been delimited? • Are there mechanisms for user consent? • Is the user allowed to move between the variations of public and private?

The dimensions and guiding questions of the framework are intended to provide operational shortcuts to educational professionals on how to incorporate a pedagogical approach as well as student sovereignty into the practice of algorithm design. It also aims to focus on the orientation of algorithms to the achievement of student competencies and skills, on the basis that decisions about recommendations and nudges should be guided by pedagogical evidence. All this seeks to foster safer and more inclusive learning spaces and interactions with IA.

DISCUSSION AND FUTURE WORK

The arguments provided in this paper are intended to complement existing evidence in the scientific literature about providing educators with resources to face the introduction of AI in learning spaces. Proposing key methodologies and guidelines grouped in heuristic rules is considered an appropriate way, since this allows for the management of resources in particularly complex situations.

As a non-technical theoretical proposal, the ability to implement the presented framework in practice will depend on further empirical validations referred to in future studies. Thus, the discussion on the construction of the heuristic scheme points to a set of research references on the design of theoretical frameworks and the subsequent empirical validation of rules and constructs.

Another issue with heuristics that can be discussed concerns the so-called consistency of the context. Heuristics are a great contribution when the assumptions on which they are based are sufficiently consistent in the contexts where they are applied. Therefore, the proposed scheme should also be validated in the variety of phases/territories where it is intended to be applied: either in the design of an algorithm, or in the implementation in practice situations, or if it is a technical development context, or one of educational instructional design, etc.

Simple rule frameworks provide shortcuts that assist both the algorithm design process and the use of digital tools in teaching. However, they cannot be directly applied. It is necessary to previously analyse the effective practices of the subjects in the digital spaces, trying to understand their behaviour in a global way. It is assumed that large data sets –either inherently or as a result of their size– do not have direct answers to the most interesting questions. That is why heuristic rule-based approaches advocate simplifying decision-making in complex learning situations, while optimizing the effect by placing the greatest emphasis on analysing the set of actions that produce a given learning.

The next steps in the field of data-driven educational algorithms aim at deepening from a pedagogical perspective the implementation of derived technologies in real educational practice situations, so that the implications of AI in decision making and in the enrichment of learning processes are fully understood. Also, to advance in the analysis of the challenges that AI implies for educational research. And equally, to be open to the validation –both theoretical and empirical– of schemes such as the one proposed here, which serve as a guide for professionals and academics to manage data-driven digital technologies in learning processes.

REFERENCES

- Alsawaiket, M., Blasi, A. H., & Al-Msie'deen, R. A. (2019). Formulating module assessment for improved academic performance predictability in higher education. *Engineering, Technology & Applied Science Research*, 9(3), 4287-4291. Retrieved from <https://www.etasr.com/index.php/ETASR/article/view/2794/pdf>
- Amo, D., Fonseca, D., Alier, M., García-Peñalvo, F. J., Casañ, M. J., & Alsina, M. (2019). Personal data broker: A solution to assure data privacy in EdTech. In P. Zaphiris & A. Ioannou (Eds.), *Learning and collaboration technologies. Design, experiences. 6th International Conference, LCT 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26-31, 2019. Proceedings, Part I* (pp. 3-14). Cham, Switzerland: Springer Nature.

- Ball, S. J., Bowe, R., & Gewirtz, S. (1995). Circuits of schooling: a sociological exploration of parental choice of school in social class contexts. *The Sociological Review*, 43(1), 52-78.
- Bodily, R., & Verbert, K. (2017). Review of research on student-facing learning analytics dashboards and educational recommender systems. *IEEE Transactions on Learning Technologies*, 10(4), 405-418. Retrieved from <https://dblp.org/rec/journals/tlt/BodilyV17>
- boyd, D., & Crawford, K. (2012). Critical questions for big data. Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5), 662-679. <https://doi.org/10.1080/1369118X.2012.678878>
- Brady, H. E. (2019). The challenge of big data and data science. *Annual Review of Political Science*, 22, 297-323. <https://doi.org/10.1146/annurev-polisci-090216-023229>
- Breslow, L., Pritchard, D. E., DeBoer, J., Stump, G. S., Ho, A. D., & Seaton, D. T. (2013). Studying learning in the worldwide classroom: Research into edX's first MOOC. *Research & Practice in Assessment*, 8(1), 13-25. Retrieved from <http://www.rpajournal.com/dev/wp-content/uploads/2013/05/SF2.pdf>
- Buckingham Shum, S., & Ferguson, R. (2012). Social Learning Analytics. *Journal of Educational Technology & Society*, 15(3), 3-26. Retrieved from <https://drive.google.com/file/d/1fu8JL6t8pwfGskAnk tZ4AEWChPjRnbdI/view>
- Bulger, M. (2016). Personalized learning: The conversations are not having. Retrieved from Data & Society Research Institute website: https://datasociety.net/pubs/ecl/PersonalizedLearning_primer_2016.pdf
- Bunker, R. P., & Thabtah, F. (2019). A machine learning framework for sport result prediction. *Applied Computing and Informatics*, 15(1), 27-33. <https://doi.org/10.1016/j.aci.2017.09.005>
- Caplan, R., Donovan, J., Hanson, L., & Matthews, J. (2018). Algorithmic accountability: A primer. Retrieved from Data & Society Research Institute website: https://datasociety.net/pubs/alg_accountability.pdf
- Coleman, J. S. (1966). *Equality of educational opportunity*. U.S. Dept. of Health, Education, and Welfare, Office of Education.
- Crawford, K. (2016). Can an algorithm be agonistic? Ten scenes from life in calculated publics. *Science, Technology & Human Values*, 41(1), 77-92. <https://doi.org/10.1177/0162243915589635>
- Diakopoulos, N., Friedler, S., Arenas, M., Barocas, S., Hay, M., Howe, B., Jagadish, H. V., Unsworth, K., Sahuguet, A., Tech, C., Venkatasubramanian, S., Wilson, C., Yu, C., & Zevenbergen, B. (2017). Principles for accountable algorithms and a social impact statement for algorithms. FAT/ML. Retrieved from <https://www.fatml.org/resources/principles-for-accountable-algorithms>
- Domínguez, D. (2018). Big Data, educación basada en datos y analítica del aprendizaje. In A. Sacristán (Coord.), *Sociedad digital, tecnología y educación* (299-329). Madrid, Spain: UNED.
- Domínguez, D., Álvarez, J. F., & Gil-Jaurena, I. (2016). Learning Analytics and Big Data: Heuristics as Interpretive Frameworks. *DILEMATA, International Journal of Applied Ethics*, 22, 87-103. Retrieved from <https://www.dilemata.net/revista/index.php/dilemata/article/view/412000042>
- Farrow, R. (2016). A Framework for the ethics of open education. *Open Praxis*, 8(2), pp. 93-109. <http://dx.doi.org/10.5944/openpraxis.8.2.291>
- Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data*

- Science Review*, 1(1). <https://doi.org/10.1162/99608f92.8cd550d1>
- Geddes, B. (1990). How the cases you choose affect the answers you get: Selection bias in comparative politics. *Political Analysis*, 2(1), 131-150. <https://doi.org/10.1093/pan/2.1.131>
- Gigerenzer, G., & Selten, R. (Eds.) (2002). *Bounded rationality: The adaptive toolbox*. MIT press.
- Gitelman, L. (Ed.) (2013). *Raw data is an oxymoron*. MIT Press.
- Goksel, N., & Bozkurt, A. (2019). Artificial intelligence in education: Current insights and future perspectives. In S. Sisman-Ugur, & G. Kurubacak (Eds.), *Handbook of Research on Learning in the Age of Transhumanism* (224-236). Hershey, IGI Global.
- González-Bailón, S., Wang, N., Rivero, A., Borge-Holthoefer, J., & Moreno, Y. (2012). Assessing the bias in communication networks sampled from twitter. Retrieved from <https://arxiv.org/abs/1212.1684>
- Greller, W., & Drachsler, H. (2012). Translating Learning into Numbers: A Generic Framework for Learning Analytics. *Educational Technology & Society*, 15(3), 42-57. Retrieved from <https://drive.google.com/file/d/1R84FXoT3W3X6C2JV1BBXhagtCoOQiQ7l/view>
- Hansen, J. D., & Reich, J. (2015). Democratizing education? Examining access and usage patterns in massive open online courses. *Science*, 350(6265), 1245-1248. <https://doi.org/10.1126/science.aab3782>
- Hew, K. F., Qiao, C., & Tang, Y. (2018). Understanding student engagement in large-scale open online courses: A machine learning facilitated analysis of student's reflections in 18 highly rated MOOCs. *International Review of Research in Open and Distributed Learning*, 19(3). <https://doi.org/10.19173/irrodl.v19i3.3596>
- Hussain, M., Zhu, W., Zhang, W., Abidi, S. M. R., & Ali, S. (2019). Using machine learning to predict student difficulties from learning session data. *Artificial Intelligence Review*, 52(1), 381-407. <https://doi.org/10.1007/s10462-018-9620-8>
- Houlden, S., & Veletsianos, G. (2019). A posthumanist critique of flexible online learning and its "anytime anyplace" claims. *British Journal of Educational Technology*, 50(3), 1005-1018. <https://doi.org/10.1111/bjet.12779>
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389-399. <https://doi.org/10.1038/s42256-019-0088-2>
- Johnson, C., & Newett, E. (2015). From idea to execution: Spotify's discover weekly. Retrieved from <https://www.slideshare.net/MrChrisJohnson/from-idea-to-execution-spotifys-discover-weekly/>
- Kitchin, R. (2014). Big data, new epistemologies and paradigm shifts. *Big Data & Society*, 1(1), 1-12. <https://doi.org/10.1177/2053951714528481>
- Kurilovas, E. (2019). Advanced machine learning approaches to personalise learning: learning analytics and decision making. *Behaviour & Information Technology*, 38(4), 410-421. <https://doi.org/10.1080/0144929X.2018.1539517>
- Luckin, R., & Cukurova, M. (2019). Designing educational technologies in the age of AI: A learning sciences-driven approach. *British Journal of Educational Technology*, 50(6), 2824-2838. <https://doi.org/10.1111/bjet.12861>
- Metcalf, J., Keller, E. F., & boyd, d. (2016). Perspectives on big data, ethics, and society. Retrieved from The Council for Big Data, Ethics, and Society website: <https://bdes.datasociety.net/council-output/perspectives-on-big-data-ethics-and-society/>
- Michael, K., & Miller, K. W. (2013). Big data: New opportunities and new challenges. *Computer*, 46(6), 22-24. <https://doi.org/10.1109/2.1161161>

- ieeecomputersociety.org/10.1109/MC.2013.196
- Monarrez, T. (2018). *Segregated neighborhoods, segregated schools? Methodology*. Washington, DC: Urban Institute. Retrieved from https://www.urban.org/sites/default/files/segregated_neighborhoods_methodology.pdf
- Mousavi, S., & Gigerenzer, G. (2014). Risk, uncertainty, and heuristics. *Journal of Business Research*, 67(8), 1671-1678. <https://doi.org/10.1016/j.jbusres.2014.02.013>
- Muller, J. Z. (2018). *The tyranny of metrics*. Princeton University Press.
- Onnela, J. P., Saramäki, J., Hyvönen, J., Szabó, G., Lazer, D., Kaski, K., Kertész, J., et al. (2007). Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences*, 104(18), 7332-7336. <https://doi.org/10.1073/pnas.0610245104>
- Orfield, G., & Lee, C. (2005). Why segregation matters: Poverty and educational inequality. Retrieved from <https://civilrightsproject.ucla.edu/research/k-12-education/integration-and-diversity/why-segregation-matters-poverty-and-educational-inequality/orfield-why-segregation-matters-2005.pdf>
- Pasick, A. (2015, December 21). The magic that makes Spotify's Discover Weekly playlists so damn good. Quartz. Retrieved from <https://qz.com/571007/>
- Perrotta, C., & Selwyn, N. (2019). Deep learning goes to school: toward a relational understanding of AI in education. *Learning, Media and Technology*. <https://doi.org/10.1080/17439884.2020.1686017>
- Pfeffer, J., Mayer, K., & Morstatter, F. (2018). Tampering with twitter's sample API. *EPJ Data Science*, 7(50). <https://doi.org/10.1140/epjds/s13688-018-0178-0>
- Pitcan, M. (2016, July 13). Student Data Privacy: An Overview [Blog post]. Retrieved from <https://medium.com/enabling-connected-learning/student-data-privacy-an-overview-ea41ebd99095#.8jv3n83w2>
- Poel, M., Meyer, E. T., & Schroeder, R. (2018). Big data for policymaking: Great expectations, but with limited progress? *Policy & Internet*, 10(3), 347-367. <https://doi.org/10.1002/poi3.176>
- Prabhakar, S., Spanakis, G., & Zaiane, O. (2017). Reciprocal recommender system for learners in massive open online courses (MOOCs). In H. Xie, E. Popescu, G. Hancke & B. . Manjón (Eds.), *Advances in Web-Based Learning-ICWL 2017* (157-167). Cham: Springer. https://doi.org/10.1007/978-3-319-66733-1_17
- Reich, J. (2014, March 30). Big data MOOC research breakthrough: Learning activities lead to achievement [Blog post]. Retrieved from http://blogs.edweek.org/edweek/edtechresearcher/2014/03/big_data_moooc_research_breakthrough_learning_activities_lead_to_achievement.html
- Reif, J. H. (n.d.). Rules for algorithm design [Lecture notes]. Retrieved from <https://users.cs.duke.edu/~reif/courses/alglectures/skienna.lectures/lecture6.2.pdf>
- Romero, C., & Ventura, S. (2017). Educational data science in massive open online courses. *Data Mining and Knowledge Discovery*, 7(1). <https://doi.org/10.1002/widm.1187>
- Ruipérez-Valiente, J. A., Halawa, S., Slama, R., & Reich, J. (2019). Using multi-platform learning analytics to compare regional and global MOOC learning in the Arab world. *Computers & Education*, 146. <https://doi.org/10.1016/j.compedu.2019.103776>
- Ruths, D., & Pfeffer, J. (2014). Social media for large studies of behavior. *Science*, 346(6213), 1063-1064. <https://doi.org/10.1126/science.1246213>
- Saurwein, F., Just, N., & Latzer, M. (2015). Governance of algorithms: options and

- limitations. *Info*, 17(6), 35-49. <https://doi.org/10.1108/info-05-2015-0025>
- Sharma, R.C., Kawachi, P., & Bozkurt, A. (2019). The Landscape of Artificial Intelligence in Open, Online and Distance Education: Promises and Concerns [Editorial]. *Asian Journal of Distance Education*, 14(2). Retrieved from <http://asianjde.org/ojs/index.php/AsianJDE/article/view/432>
- Siemens, G., Gasevic, D., Haythornthwaite, C., Dawson, S., Buckingham, S., Ferguson, R., Duval, E., Verbert, K., & Baker, R.S.J.d. (2011). Open Learning Analytics: an integrated & modularized platform. Retrieved from Society for Learning Analytics Research website: <https://solaresearch.org/wp-content/uploads/2011/12/OpenLearningAnalytics.pdf>
- Sinders, C. (2019a, November 12). Reimagining privacy online through a spectrum of intimacy [Blog post]. Retrieved from <https://www.are.na/blog/reimagining-privacy-online-through-gradients-of-intimacy>
- Sinders, C. (2019b). Making critical ethical software. In L. Bogers, & L. Chiappini (Eds.), *The Critical Makers Reader:(Un) learning Technology* (86-94). Amsterdam: Institute of Network Cultures.
- Sinders, C. (2019c). Data ingredients: A provocation towards making algorithms human readable. Retrieved from <https://privacy.shorensteincenter.org/data-ingredients>
- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, 75(3), 417-453. <https://doi.org/10.3102/00346543075003417>
- Sloane, M., & Moss, E. (2019). AI's social sciences deficit. *Nature Machine Intelligence*, 1(8), 330-331. <https://doi.org/10.1038/s42256-019-0084-6>
- Smith, C. S. (2019, December 18). The machines are learning, and so are the students. *The New York Times*. Retrieved from <https://www.nytimes.com/2019/12/18/education/artificial-intelligence-tutors-teachers.html>
- Sundar, S., & Singh, A. (2013). New heuristic approaches for the dominating tree problem. *Applied Soft Computing*, 13(12), 4695-4703. <https://doi.org/10.1016/j.asoc.2013.07.014>
- Thille, C., Schneider, E., Kizilcec, R. F., Piech, C., Halawa, S. A., & Greene, D. K. (2014). The future of data-enriched assessment. *Research & Practice in Assessment*, 9(2), 5-16. Retrieved from <http://www.rpajournal.com/dev/wp-content/uploads/2014/10/A1.pdf>
- Tufekci, Z. (2013). Big data: Pitfalls, methods and concepts for an emergent field. <http://dx.doi.org/10.2139/ssrn.2229952>
- UNESCO. (2019, February 12). How can artificial intelligence enhance education? Retrieved from <https://en.unesco.org/news/how-can-artificial-intelligence-enhance-education>
- US-ACM. (2017). Statement on algorithmic transparency and accountability. Retrieved from https://www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf
- Uskov, V. L., Bakken, J. P., Byerly, A., & Shah, A. (2019). Machine learning-based predictive analytics of student academic performance in STEM education. In A. K. Ashmawy & S. Schreiter, *Proceedings of 2019 IEEE Global Engineering Education Conference* (pp. 1370-1376). Piscataway, NJ: IEEE.
- Warnakulasooriya, R., & Black, A. (Eds.) (2018). *Beyond the Hype of Big Data in Education. Practical lessons and illustrative examples of how to derive reliable insights in learning analytics*. MacMillan Learning. Retrieved from <http://prod-cat-files.macmillan.cloud/>

- [MediaResources/instructorcatalog/legacy/BFWCatalog/uploadedFiles/Beyond-the-Hype-of-Big-Data-in-Education.pdf](#)
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 16. <https://doi.org/10.1186/s41239-019-0171-0>
- Zhou, L. (2018, September 5). How to Build a Better Machine Learning Pipeline. Datanami. Retrieved from <https://www.datanami.com/2018/09/05/how-to-build-a-better-machine-learning-pipeline/>
- Zwitter, A. (2014). Big data ethics. *Big Data & Society*, 1-6. <https://doi.org/10.1177/2053951714559253>

ACADEMIC AND PROFESSIONAL PROFILE OF THE AUTHOR

Daniel Domínguez Figaredo. [EN] Professor and researcher at National Distance Education University (UNED, Spain). His research focuses on digital mediation and theories that support open and connected lifelong learning. In his recent work he has delved into the analysis of data-based open education and knowledge management in digital and mixed environments. He is founding member of the teaching innovation group CO-Lab: Open & Collaborative Laboratory for Teaching Innovation, and Board of Directors of CyberPractices Foundation. <http://orcid.org/0000-0002-7772-1856>
E-mail: ddominguez@edu.uned.es

[ES] Profesor e investigador de la Universidad Nacional de Educación a Distancia (UNED, España). Su investigación se centra en la mediación digital y en las teorías que apoyan el aprendizaje abierto y conectado a lo largo de la vida. En su trabajo reciente ha profundizado en el análisis de la educación abierta basada en datos y la gestión del conocimiento en entornos digitales y mixtos. Es miembro del grupo de innovación docente CO-Lab: Laboratorio abierto y colaborativo para la innovación docente, y patrono de la Fundación Prácticas en la CiberSociedad.

Address:
Dpto Teoría de la Educación y Pedagogía Social
Facultad de Educación–UNED
C/ Juan del Rosal, 14
28040 Madrid (España)

Date of receipt: 18/01/2020
Date of acceptance: 07/03/2020
Date of layout: 26/03/2020