

Data analytics in education: are schools on the long and winding road?

Data analytics
in education

Mara Soncin and Marta Cannistrà
Politecnico di Milano, Milan, Italy

AQ:1
AQ:2

Received 1 April 2021
Revised 22 September 2021
Accepted 31 October 2021

Abstract

Purpose – This study aims to investigate the organisational structure to exploit data analytics in the educational sector. The paper proposes three different organisational configurations, which describe the connections among educational actors in a national system. The ultimate goal is to provide insights about alternative organisational settings for the adoption of data analytics in education.

Design/methodology/approach – The paper is based on a participant observation approach applied in the Italian educational system. The study is based on four research projects that involved teachers, school principals and governmental organisations over the period 2017–2020.

Findings – As a result, the centralised, the decentralised and the network-based configurations are presented and discussed according to three organisational dimensions of analysis (organisational layers, roles and data management). The network-based configuration suggests the presence of a network educational data scientist that may represent a concrete solution to foster more efficient and effective use of educational data analytics.

Originality/value – The value of this study relies on its systemic approach to educational data analytics from an organisational perspective, which unfolds the roles of schools and central administration. The analysis of the alternative organisational configuration allows moving a step forward towards a structured, effective and efficient system for the use of data in the educational sector.

Keywords Decision-making process, Learning analytics, Educational data scientist, School accountability

Paper type Research paper

1. Introduction

The datafication phenomenon is shaping different sectors because of the increasing number of automated systems, which store data from different sources (Jarke and Breiter, 2019). The education sector is one of the most noticeable domains affected by datafication, given the underlying potential of data for supporting effective teaching and learning and for transforming the ways in which future generations (will) construct reality with and through data (Namoun and Alshanjiti, 2021; Jarke and Breiter, 2019). Leveraging on the increasing amount of data available, schools have started to deal with analytics (Li and Zhai, 2018) as a

© Mara Soncin and Marta Cannistrà. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial & non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

The authors would like to thank Prof. Michela Arnaboldi for her expertise and comments on this study.



possible way to ensure greater quality and improve efficiency and inclusiveness, as crucial goals of any educational institution (Gaftandzhieva *et al.*, 2021).

The importance of the adoption and use of educational data analytics is highlighted by several national and supranational bodies to serve multiple purposes. For instance, UNESCO (2017) stresses how data and information may sustain a better external and internal accountability of schools. Further, data analytics plays a crucial role in detecting students at risk or left behind, thus helping identify factors that impact learning outcomes in the short and in the long term (Sepulveda, 2020). Exploiting data from educational institutions allows, in addition, to produce evidence that helps to validate/evaluate educational systems, to increase the quality and equity of education and to lay the groundwork for a more effective learning process (Romero *et al.*, 2004). Using data in education may span from studying learners' experience model to investigating the effectiveness of teaching activities, up to investigate the factors related to the quality of an educational system overall (Goldberg *et al.*, 2019). From an academic point of view, the interest in the field has grown with the availability of increasingly precise algorithms coming from several fields of computer science and statistics (Krumm *et al.*, 2018). Researchers are now focussing on how and which quantitative techniques, from the most traditional ones, coming from econometrics and statistics (Marcenaro-Gutiérrez *et al.*, 2021) to more advanced ones, such as machine learning algorithms (Masci *et al.*, 2018), may improve educational outcomes. However, the current focus is more on the accuracy of alternative algorithms (Schiltz *et al.*, 2018) and on the use of analytics to support the learning process (Walkington, 2013; Chatti and Muslim, 2019; Marienko *et al.*, 2020), rather than on their function to create value inside an organisation.

Under this perspective, the importance given by the adoption of a structured organisation of data analytics in schools affects their accountability systems and decision-making processes (Schildkamp, 2019). In this setting, the role of the policy-maker is crucial for defining the boundaries of school accountability based on data and for using that same data for designing an evidence-based educational system (UNESCO, 2017). In this regard, the present study explores the organisational dimensions that affect the exploitation of data analytics in education taking a national-level perspective. In doing so, alternative organisational configurations are proposed by analysing their challenges and opportunities. More specifically, each organisational configuration is outlined by considering the organisational layers, the organisational roles and the data management that characterise schools, central administration (CA) and external stakeholders in the exploitation of data analytics. Empirically, the study proposes three configurations based on a participant observation approach coming from research projects experienced in collaboration with schools and institutions in the period 2017/2020.

To set out our argument, the paper is structured as follows: Section 2 shows the academic research dealing with educational data analytics with a particular reference to the organisational issues within the educational domain. Then, Section 3 presents the methods adopted in the research, while Section 4 presents the main findings. Finally, Section 5 discusses the implications of the study and concludes.

2. Setting the context about data analytics in education

The organisational complexity related to data analytics in educational systems represents the mainstream literature discussed in this section. Before going at the core of the discussion on educational data analytics, however, it is necessary to define the boundaries of such a context. Within the perimeter of educational data analytics, learning analytics represents a key concept coming from the necessity of a comprehensive definition referring to the

adoption of data for educational purposes. Learning analytics has been firstly defined by [Ferguson \(2012, p. 305\)](#) as “the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs”. Hence, it represents a key concept in the domain of data used to support decision-making. Indeed, learning analytics (henceforth, educational data analytics) deals with the collection, analysis and translation of results into impactful insights, converting raw data coming into useful information that potentially has a great impact on educational research and practice ([Romero et al., 2010](#)). Along with data analytics activities, the necessity of a new professional role is clear: the educational data scientist. This key figure should own analytical competences to develop, research and apply technical methods to detect latent patterns in large collections of educational data ([Scheuer and McLaren, 2012](#)). It is also required for this role to have soft skills to effectively communicate results and insights to the educational decision-makers, who may be teachers, schools’ principals or policy-makers ([Agasisti and Bowers, 2017](#)). Also, choosing what data to collect, focussing on the questions to be answered and making sure that the data is aligned with the questions are relevant tasks of the educational data scientist ([Mining, 2012](#)).

On the quantitative side, the application of algorithms and statistical modelling to educational data analytics has been investigated by a plethora of studies regarding the use of data to improve the effectiveness and personalisation of the learning process ([Walkington, 2013](#); [Chatti and Muslim, 2019](#); [Marienko et al., 2020](#)), the early detection of at-risk students ([Mac Iver et al., 2019](#)), the identification of effective teachers ([Gershenson, 2021](#)) or the assessment of different principals’ leadership styles in relationship with student achievement ([Agasisti et al., 2017](#); [Leithwood et al., 2020](#)). These studies, applying either traditional modelling or statistically advanced algorithms, are relevant to support schools in decision-making and enrich the academic literature on educational data analytics. However, they are often applied to specific quantitative domains, while less is known concerning the impact of educational data analytics on the organisational system as a whole ([Dawson et al., 2019](#)). The difficulty of setting a comprehensive organisational framework is (probably) due to the complexity of the educational system and the multiplicity of roles involved ([Siemens et al., 2018](#) with reference to the higher education sector, easily extended). Indeed, the investigation of educational data analytics from an integrated perspective requires taking into account different groups of educational agents, the technologies adopted and the specific environments within which they interact ([Ferguson, 2012](#)). The aspects that need to be taken into consideration are both technical, like data modelling and “soft”, like data culture. A further step would be represented by a shift towards holistic and integrated system-level research ([Dawson et al., 2019](#)). This latter consideration represents the main goal of this paper, which discusses possible organisational configurations for a structured approach to data analytics in the educational sector.

Despite not being applied to the educational field, [Grossman and Siegel \(2014\)](#) provide a rare example of critical reflection on the organisational design to support data analytics. In their study, the authors propose two antithetical organisational models: the centralised and the decentralised organisational design. The crucial reflection presented in their research concerns the trade-offs between achieving the critical mass of data analytics staff within an organisational unit versus having a broad view of the enterprise-wide problems and opportunities. A third organisational model, the hybrid one, arises as an intermediate solution in which a critical mass of data analysts is housed in a central unit, while the remaining staff is distributed across the organisation ([Grossman and Siegel, 2014](#)).

To summarise, past research highlighted the relevance of taking an integrated perspective towards data analytics, despite no evidence has emerged on how educational

data analytics may be integrated into the educational sector from an organisational point of view. The objective of this paper is addressed by proposing three organisational configurations for adopting data analytics, highlighting their potentialities and challenges as inputs for policy-making.

3. Methods

3.1 Methodology

The study adopts a participant observation approach (Parker, 2007; Ferreira and Merchant, 1992), based on the direct participation of the researchers in four research projects with schools and central educational bodies (like the National Evaluation Committee for Education, INVALSI), as well as two training courses that involved school principals and teachers. These occasions provided the possibility to collect a rich range of experiences and observational data that cover a four-year period, ranging between 2017 and 2020.

In all the mentioned projects and training activities, data analytics represented a central issue of investigation, while the unit of the analysis is represented by the educational system as a whole. The heterogeneity of actors involved, the diverse functions of data (define, evaluate and predict) and the data management adopted in each research project allow reflection about the alternative organisational configurations that the educational sector may adopt in dealing with data analytics.

The creation of an integrated framework and data set for the accountability of schools was the purpose of the first research project, which represented the initial and fundamental reflection on the need of schools in terms of data and competences. The project was carried out in interaction with a network of schools and ended with the definition of the areas of interest for their accountability.

The evaluation of the school value-added in the Italian educational system was the objective of the second project, which dealt with the use of standardised test scores. The importance of the value-added information relied on its recent inclusion in the set of indicators provided to schools by the ministry to support decision-making and school self-evaluation. This project was commissioned by the INVALSI, and thus represented an important opportunity for interaction with the central government, together with a special lens to observe how data are exchanged between schools and CA.

The last two research projects had the prediction as a common function of data. The first was related to the collection and the analysis of data on behalf of a network of schools to predict academic success between middle and high school. The research group had the primary responsibility of collecting administrative and student-related data by physically visiting each school, making this a relevant opportunity for investigation on the way data are collected and stored by schools. In addition, the project presented the very first evidence on the functioning of schools' networks, giving some preliminary insights about a possible successful organisational configuration of educational data analytics. Finally, the analysis of data coming from schools' electronic registers to predict students' performance represented the goal of the fourth project. The opportunity offered by this research project related the addition of an angle to the complexity of data integration when mixing "paper-based" and online data on students' characteristics and performance. In this project, researchers had the possibility to see how data are stored and used by the school, together with deepening the school principal's strategic vision about data analytics in education.

All these empirical projects contributed fundamentally to the definition of the reference dimensions of analysis to be considered when studying school accountability based on data analytics. Further, the heterogeneity in functions of data (define, evaluate and predict)

applied in the research projects allows extracting a comprehensive view of the diversified approach to data between schools and CA.

The analysis of the utility, problems and sustainability of the reference dimensions was supported by continuous interaction with teachers and school principals. This happened regularly because of the researchers' involvement in the design of the research projects, within which the topics of school accountability systems and data analytics for schools are core elements. Moreover, the contacts collected during past projects represented an important basis to get a validation of the model proposed in this study to a group of school principals, which reinforced the validity of findings.

The direct participation of the researchers in the aforementioned projects and programmes represented the main data source for a systematic collection of notes and in-depth interaction with three clusters of actors:

- (1) CA bodies, mainly represented by the INVALSI,
- (2) school principals both as single entities and as organised networks of schools,
- (3) teachers, especially those having managerial roles within their schools (as, for this reason, they were more likely to be involved in research and training projects). The list of actors involved in each project is reported in [Table 1](#).

The interaction with a rich set of actors gave us the possibility to explore the complexity of the topic under investigation by overlapping two layers of complexity. The first is that of the level of a decision supported by data analytics, which is highly interrelated with the complexity of the actors involved in school accountability and decision-making (namely, the CA, the school principals and the teachers). The second and somehow orthogonal layer is represented by the time evolution of the state and the use of data analytics for education in Italy. Indeed, the introduction of a structured approach to data analytics increased after the introduction of Law 107/2015, which prompted the use of data to support a standardised approach to self-evaluation, accountability and external evaluation of schools. The possibility to combine evidence coming from projects and interactions that covered a wide range of years (2017–2020) gave researchers the possibility to have a privileged observatory on the evolution of data analytics for schools.

3.2 Reference dimensions of analysis

The organisational configurations of data analytics in education are analysed and described under a theoretical lens that supports the interpretations of the research projects. In doing

Function of data analytics	Project	Years covered	Actors involved
Define	Accountability framework	2017–2018	Network of 10 primary and secondary schools
Evaluate	School value-added	2019–2020	National Evaluation Committee for Education (INVALSI) Dissemination to schools (national level)
Predict	Predicting student success	2019–2020	Network of 30 primary and secondary schools
Predict	Predicting student performance	2019–2020	Primary and secondary schools

Table 1.
List of projects and informants, categorised on the typology of analysis, the projects' years and the actors involved

so, related literature suggests three pillars for building an organisational framework around data analytics: its analytics culture, staff, processes and governance (Grossman and Siegel, 2014). In this study, these pillars have been combined with the dimensions used for the description of educational systems at the national level, which include the level of (de) centralisation and the empowerment of organisational roles (San Fabián Maroto, 2011). Therefore, three dimensions of analysis have been explored: the organisational layers of decision-making, the roles and the data management, respectively.

The first dimension concerns the level of decision-making to which data analytics are useful. Indeed, the kind of data that may support decision-making and accountability differ according to the organisational level considered, with specific reference to the contraposition between central government and administration (i.e. the ministry and its collateral bodies) and educational institutions (meaning both the school and the class level) (Reyes, 2015). At the central level, despite the growing body of literature concerning evidence-based educational policy (Slavin, 2002; Davies, 1999), the examples of the practical application of those principles are still scattered, particularly in the Italian context. Past research highlighted the relevance to collect research data to support policy-making (Slavin, 2002). However, several administrative data sources, as well as information collected through *ad hoc* surveys on the occasion of standardised test collection could complement the picture and should represent the basis for data-informed educational policy-making (Agasisti *et al.*, 2017). The central level decision-making differs substantially from the school level. The difference is not necessarily in the type of data to be collected and analysed – as information based on research and administrative data would be almost identically useful for school principals, as well as for policy-makers – but is related to the kind of decision-making and accountability supported. At school and class level, indeed, information should be used to:

- support the definition of priorities,
- monitor the advancements and highlight eventual gaps,
- report results to internal and external stakeholders, in light of an accountability process based on data (Figlio and Loeb, 2011).

In this respect, the difference between the central and school level does not refer to the type of data needed, but to the use of that data and from this standpoint to the different support needed for data interpretation, an issue that should be critical in the competences owned by the school data scientist, as detailed below.

The second dimension regard the roles involved with data analytics for education. The complexity of roles and competences involved in data analysis within public organisations has been recently highlighted with reference to the difficulty of matching policy demand and data offer (Arnaboldi and Azzone, 2020). It has been demonstrated how, in this respect, a critical role is that of the translator, namely, the role of the interpreter between the needs of policy-makers and data scientists. The educational sector is peculiar in this sense, as the final policy objective can be (over)simplified into a plain message: supporting students' success. Despite the several angles that this objective can take, its existence minimise the possibility of conflicts arising between educational agents. However, the issue of deciding which actions and policies should be prioritised to reach the aforementioned objective is an object of debate and this is exactly the point in which the support of data analytics is critical. What is missing to the educational sector to fulfil this aim are the competences to support with evidence

the decisions made at both central and peripheral (school) levels. The set of competences that are needed in data analytics for schools have been discussed by scholars, who proposed a data analytics model made of a recursive set of stages (Agasisti and Bowers, 2017; Siemens, 2013), which can be specified as follows:

- data collection, once that the source of data is identified;
- data cleaning;
- data integration, in case multiple data set, are needed;
- data analysis and finally;
- data visualisation and interpretation.

The last step, related to decision-making, should be finalised in the hands of the school manager/principal or of the policy-maker. What is critical for our investigation is that the roles in charge of the different activities are multiple and variegated. Indeed, by following the stages proposed by Agasisti and Bowers (2017), the first role is that of the data collector, who identifies the source of data and makes the database available. Secondly, the data analyst takes care of the statistical analysis and interpretation. Finally, the communicator supports the decision-maker for a correct interpretation of data. In turn, this complexity poses an issue related to whether those competences should be owned internally by schools or externally provided by the central government bodies (Cech *et al.*, 2018) to guarantee the greatest possible effectiveness of the decision-making and accountability system.

The last dimension is related to data management (Wang, 2016), which represents an important layer of complexity in educational data analytics due to the constant inter-agent interactions and flow of data (Sergis and Sampson, 2016). Despite the possibility to create a coherent framework in which sequential steps bring from data collection to decision-making, the number of interactions needed during that stages and the complexity in the flow of data may represent an important obstacle to efficient and effective use of data analytics. This dimension is strictly related to the debate on data infrastructures and on the existence of “centres of anticipation” (Williamson, 2016) that pre-elaborate the data that are then accessed and visualised by the schools to support decision-making. The organisation for economic co-operation and development is one of the main actors of this process when analysing the data collected through the Programme for International Student Assessment (PISA) and the same holds for all the national-based bodies collecting data on standardised tests (like INVALSI for Italy). The multiplicity of these agents makes the data aggregation and integration more complex to be managed by schools. In other cases, the data are stored in different software as they are collected by the electronic register, the student office or the canteen service database, with the threat of creating data silos that hardly communicate (Wang, 2016). In this respect, the need to create an integrated system for data analytics is a necessary condition for effective and efficient data management. To summarise, the drivers of complexity related to the flow of data are represented by (i) the multiple agents involved that favour the creation of loops and bounces in the flow and (ii) the unrelated sources of data, which favour the creation of data silos.

4. Results

Results from the study are presented by assuming a comprehensive perspective analysed through the reference framework, which includes the level of decision-making, the organisational roles and the data management. These dimensions are discussed within each of the alternative configurations in which educational data analytics may be structured at the national level. The following paragraphs describe the features of each configuration

together with an interpretation based on the three dimensions of interest and the main pros and cons characterising each model. A graphical representation of the three models is presented in Figure 1 while a more detailed description of each configuration is presented in Table 2. In detail, the first configuration, the centralised one, is the actual representation adopted by the Italian context. The second and the third configurations, decentralised and network-based, respectively, represent the “to-be” modelling. In this respect, a kind of evolution path emerges from the results, starting from the actual configuration to arrive at the more advanced and integrated use of data analytics with the decentralised and network-based ones. In each configuration, three actors have represented: the CA, the schools and the external stakeholders.

4.1 The centralised configuration

The centralised configuration is characterised by several bounces of data between schools and CA. In this model, data for educational analytics may derive from two sources. Firstly, they can be collected and analysed directly by the schools, on the basis, for example, of data taken from the electronic registers or from the administrative student office (as represented by the arrow entering each school). Alternatively, data are requested to the school by the CA (i.e. by the Ministry), which periodically asks schools to fill in questionnaires reporting school-level data on students, teachers and school context. After this stage of data collection, the CA elaborates data (internally) and reports them back to the school by means of an online platform developed to support school accountability at the national level (this platform, called *Scuola in Chiaro* in Italy, has existed since 2015). Data are reported in the platform as key performance indicators (KPIs) that can be accessed by the school principal and her/his staff to support self-evaluation, continuous improvement and social accountability. In turn, KPIs may be complemented with data directly collected from the school to support school accountability towards students, families and CA. At the central level, the CA makes available the KPIs aggregated at the school level on the online platform to support families and students in school choice.

4.1.1 Organisational layers. In this configuration, the CA develops a data-informed policy on the basis of the information provided by the schools themselves. This allows the central government to make decisions owing data that can be aggregated at different levels, i.e. the single institutions, the grades, the school tracks or the regions and on the basis of the prioritisation coming from the political decision-maker. In turn, schools make decisions based on indicators computed and provided by the CA or collected and analysed autonomously. In this respect, the CA partially substitutes the school in data analysis. During the *school value-added* project, indeed, the CA designed a platform to let schools visualise their value-added (to students’ results) and whether it was in line, below or above the value expected on the basis of the school’s characteristics. School managers could

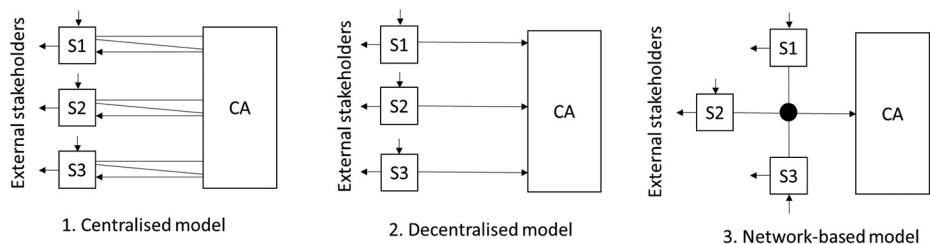


Figure 1. Configurations about data analytics supporting decision-making process in the education sector

Notes: S = School; CA = Central administration

Organisational configuration	Organisational layers	Organisational roles	Data management
Centralised model	<p>CA: the Ministry decides the analysis of national interest according to the political direction. Having an overview of the national educational system is the main goal</p> <p>The CA represents the barycentre of educational data analytics</p> <p>School: Principals decide the analysis of interest to improve the educational proposition + have access to the KPIs selected and analysed by the CA</p>	<p><i>Data collector:</i> schools, in particular secretaries and teachers</p> <p><i>Data analyst:</i> specific role in CA</p> <p><i>Communicator:</i> CA or schools' principals</p>	<p>Several bounces of data due to the following flow:</p> <ol style="list-style-type: none"> 1. Government or national agencies require schools to send them specific data based on a centrally established framework of analysis 2. Schools, in particular student offices, collect information, that is sent back to the CA 3. The CA analyses data and produces reports 4. Reporting is available for schools and families (and maybe complemented by additional school-developed analysis)
Decentralised model	<p>CA: the Ministry decides the analysis of national interest according to the political direction. Having an overview of the national educational system is the main goal</p> <p>School: Principals decide the analysis of interest to improve the educational proposition + provide the data and KPIs needed by the CA</p> <p>The school represents the barycentre of educational data analytics</p>	<p><i>Data collector:</i> Technical teachers or data experts collect, clean and store relevant data</p> <p><i>Data analyst:</i> led by schools</p>	<ol style="list-style-type: none"> 1. Government or national agencies require schools to send them specific data to implement analysis of interest. In parallel, school principals may carry out additional analysis 2. Schools, in particular teachers or technical figures, collect information that is partly shared with the CA and partly used for internal analysis
Network-based model	<p>CA: the Ministry decides the analysis of national interest according to the political direction. Having an overview of the national educational system is the main goal</p> <p>Schools' network: the principals' board define the main objectives, requiring the adoption of an integrated analytical framework</p> <p>School: Principals decide the analysis of interest to improve the educational proposition</p>	<p><i>Data collector:</i> the network's educational data scientist, based on the goals decided by principals, collect and clean data</p> <p><i>Data analyst:</i> the network's educational data scientist analyses data</p> <p><i>Communicator:</i> reports the network's educational data scientist supports the main results by supporting data interpretation</p>	<ol style="list-style-type: none"> 3. The CA and schools analyses data e produce reports 4. National and internal reporting is available for schools <ol style="list-style-type: none"> 1. Government or national agencies require the network to send them specific data to implement analysis of interest. In parallel, network's school principals may have been interested in specific analysis 2. The network's educational data scientist collects data that are partly shared with the CA and partly used for internal analysis 3. The CA and the network's educational data scientist analyse data e produce reports 4. National and internal reporting is available for schools

Note: CA stands for central administration

Table 2.
Main findings of the three configurations based on the dimensions of analysis proposed: organise rational layers, roles and competences and data flow and management

visualise their positioning but were not provided with the information necessary to explore further or replicate the analyses. Thus, at the school level, some KPIs are received as inputs already classified and visualised by the data analysts within the CA. In this case, schools can complement that data with additional sources collected internally from internal sources.

4.1.2 Organisational roles. In this configuration, the barycentre is shifted towards the CA. Indeed, the central government plays the role of the *data collector*, which periodically sends schools questionnaires requiring particular data or indicators needed to inform the national databases. In this respect, the role of schools is “only” that of providing the data, a task that may be particularly time expensive depending on the efficiency of school offices in getting access to the required information. In many school offices, the presence of many paper-based documents makes this process more complex than it would be if the information would have been stored in an informative system.

The second role is that of the *data analyst*, which is performed both by the CA and by the schools. Indeed, the CA holds the internal competences to analyse and visualise data to inform national policies, as well as to populate school-level KPIs. It is important to stress that, in this respect, the CA also organises KPIs into a coherent framework and on the basis of that framework, indicators are finally accessed by schools. Then, within schools, there may be stronger or weaker competences of data analysis and rarely a data analyst is formally recognised. As emerged during the discussion with school principals during the aforementioned research projects, the role of data analyst within the school is often informally played by a single or a group of well-meaning teachers, who make available their expertise to support the school principal in the analysis of data that may be useful for school decision-making and accountability. In this configuration, thus, the main set of competences related to the data analyst role are not held internally but are demanded by the CA. On the one hand, this simplifies the complexity of tasks carried out by the school. On the other hand, this does not sustain the development of evidence-based decision-making within the school.

The final role is the *data communicator*, who is also somehow the weakest role in this configuration, despite it would be critical in two directions. The first is the one going from the CA to the school. Given that the competences needed for data interpretation are not a basic skill owned by the school personnel, clear and straightforward communication would support schools in the more efficient and effective use of data. The second direction is the one going from the school to the external stakeholders. Indeed, data externally presented to stakeholders are barely elaborated, exchanging a principle of transparency with an absence of data elaboration and communication that sometimes makes data difficult to be interpreted by external stakeholders. This aspect clearly emerged during the *Accountability framework* project that revealed how schools presented in external documents the tables of data coming from student offices without any further elaboration, for a matter of complete transparency. This actually made data and KPIs even more difficult to be interpreted from external stakeholders, undermining the original goal.

4.1.3 Data management. The main feature of the data management in the centralised configuration is the several bounces required to get from data collection to data use. As described below, several data are required from the CA, which analyses, visualises and reports them back to the schools. It is possible to recognise, in these bounces, the separation of roles between schools and CA that characterises this configuration. These bounces make it difficult for schools to get full empowerment of the data management and data potentiality because the process is centred on a set of variables selected and analysed externally, thus schools hardly reflect on what data analytics should be more useful to fulfil school-level strategic goals.

4.1.4 Strengths and weaknesses. The main strengths connected to the centralised configuration relate to the simplification of the roles and competences needed for data analytics that are concentrated within the CA. Indeed, having a sound framework enriched with KPIs created and elaborated externally by the CA represents a facilitator for starting to diffuse a data-driven mentality among schools. However, this same aspect may turn into a weakness in the long run. Demanding externally the elaboration and the visualisation of data does not foster the creation of internal competences that schools may apply to a larger realm of data, to better support the use of information for decision-making and accountability. Thus, an aspect that may be a strength in the short term, can actually turn into a challenge when taking a long-run perspective.

4.2 The decentralised configuration

On the opposite strand, the decentralised configuration highlights the case in which every task related to data analytics is demanded of the schools. This configuration may be seen as an evolution of the current centralised setting and as an alternative to the network-based configuration. In this setting, educational institutions share the data needed for external accountability with the CA or with external stakeholders, depending on the interest or request. Compared to the previous configuration, where the schools are settled and organised to store and analyse data internally, and therefore competences for data analytics must be internally owned. Hence, in this context, the focus for data analytics is on schools, which are empowered with the possibility to explore and the responsibility to share their data and indicators. Along with this new aspect, the previously described process of data exchange and reporting with CA is still in place.

4.2.1 Organisational layers. Within this configuration, the two organisational layers involved are still the central (CA) and the school levels. However, in this perspective, the barycentre is very displaced towards schools, which become the fulcrum of all the activities related to data analytics – from data collection to communication. Indeed, schools are in charge of collecting data for both their internal analysis and for national assessments of the CA. Two different types of processes correspond to the organisational layers. From the schools' perspective data are relevant to support principals' decision-making process, while, on a central perspective, data support national policy-making. Compared to the previous configuration, the centrality of schools requires their organisational empowerment, with relevant implications for organisational roles and data management.

4.2.2 Organisational roles. As schools represent the fulcrum of this configuration, the main organisational roles reside within them. Therefore, the data collector is inside schools and can be represented by a teacher or a specific person hired for this purpose. The main task of the data collector is the storage and cleaning of the school's data into an organised and structured infrastructure. In fact, today most of the data owned by schools are kept in different repositories, which rarely are linked to one another and this makes it difficult to take full advantage of the possible data sources. In this configuration, both schools and CA are required to have analytical competences to support the corresponding decision-making processes. Thus, despite referring to different organisational layers, both the CA and the schools should have a data analyst role within them. On a similar strand, the communicator role is present in both the layers despite being relevant especially within schools. Today, this latter role is usually played by school principals as the main communicators about schools' results to external stakeholders. In this configuration, therefore, data analytics roles are highly decentralised towards schools, which should internally own the necessary competences for data collection, analytics and finally, communication.

4.2.3 Data management. In the decentralised setting, data management is biased towards schools, which are in charge of defining data needs, collecting them, creating the data infrastructure and querying databases according to the purpose of investigation. When needed for the process of policy-making, CA asks for data and information from schools, which collect and send them back to the Ministry. While, for the process of decision-making, schools, principals, together with teachers with organisational roles, define the analysis of interest themselves. For instance, schools may be interested in deepening students' retention rates. Thus, the person in charge of data collection stores the necessary data into a database and then proceeds with the analysis. Schools' principals use the insights from the analysis for accountability decisions, which can range from implementing new projects to attracting prospective students. Each step is carried out internally to the school and this represents the main feature of this configuration.

4.2.4 Strengths and weaknesses. The main strength of the decentralised setting relies on the centrality that data gains to make informed decisions even at the local (school) level. In this context, the school can reach full self-awareness on how crucial data can be for improving internal processes and for supporting students' learning paths. During the *predicting students' performance* project, school managers and teachers involved in data collection increased their perception of usefulness around data analytics applied to electronic registers, gradually overcoming the feelings of worthlessness perceived during the long process of data cleaning and preparation. Further, in this setting, the school is totally autonomous in deciding how to use data and for which purpose, having complete autonomy about it. However, the main weakness is related to the fact that schools rarely have an internal technical figure able to organise and analyse data with sophisticated techniques. Hence, for the majority of cases, the solution would be to hire an external figure or to train an internal one, which, in turn, requires extra resources for schools. The last challenge regard the still-in-place data bounce between school and CA: in fact, this configuration allows schools to be more agile in sending required data and information, but this does not fully streamline the process.

4.3 The network-based configuration

The last configuration, the network-based, represents an interesting evolution for the educational sector to efficiently deal with data analytics. The creation of a network of schools to jointly carry out data analytics projects generates a positive commitment among school principals and facilitates the sharing of knowledge. This organisational modelling is placed midway between the centralised and decentralised configurations, representing a sort of hybrid model, which adds an intermediate organisational layer, embodied by the network. Its hybridity aims at overcoming the trade-off between having a broad view of the system, typical of the centralised configuration and having a point of observation closer to the object of investigation (i.e. the student), which characterises the decentralised configuration. The network specifically deals with the aggregation, analysis and communication from raw data to actionable insights. The network is composed of a group of schools, aggregated by similarity, by proximity or by the similar didactical offer, within which a data analytics facilitator, i.e. a network's educational data scientist, is appointed. Periodically, the principals' board defines the strategic objectives, together with the analyses of interest. Then, following the directions defined, the network's educational data scientist starts collecting and analysing data, then reporting the main insights to the board. Together with network-based analysis, the educational data scientist is in charge of exchanging data and information with CA, for annual national reporting.

4.3.1 Organisational layers. As anticipated, compared to the previous configuration, this one includes a new and fundamental layer of analysis: the network. Even if it represents a conceptual layer more than a physical one, its relevance is critical. In fact, the bundling of schools into a network, with a specific figure that collects their data and transforms them into relevant strategic information, represents an innovative approach to educational data analytics. Starting from the highest layer, the CA still receives data to annually report them into the public platform (i.e. *Scuola in Chiaro*) and to provide inputs for data-based policymaking. Then, the network aims at deeply understanding specific and underlying patterns of its schools and students. It is worth noting that the analyses of interest may be of different nature, from the adoption of advanced analytics, such as Machine Learning techniques, to more simplified yet effective statistics for communication, such as the creation of informative dashboards or infographics. The final organisational layer is represented by the school, which may propose strategically relevant matters to be investigated within the network and has to support student-level data collection.

4.3.2 Organisational roles. In this configuration, the main roles defined so far to describe the approach to data analytics in schools are all embodied by the network's educational data scientist. In fact, he/she is the data collector, data analyst and communicator. Among his main tasks, there is the deep comprehension of strategic guidelines and interests set by principals, which are needed to be translated into effective data collection and modelling. Further, the educational data scientist is required to communicate results with simple and easy-to-understand reporting. This does not mean that the analyses should be simple, but it implies that data communication needs to be understandable by non-technicians. Also, managerial and organisational skills are required for the educational data scientist to effectively handle the interaction with the network of schools.

4.3.3 Data management. The schools' network, in this configuration, acts as a sort of data hub, where information is stored and used when needed. In this setting, the bounce of information with the CA is simplified for two main reasons: on the one hand, schools' data are stored at the network level, from which they can be shared with the CA for national assessment; on the other hand, this configuration allows to streamline the process of communication between schools and CA. In fact, compared to the centralised and decentralised configurations, the advantage for the CA is clear: the number of its interlocutors is smaller when dealing only with the network, instead of dealing with the single school. It is worth underlining, from a data management perspective, the importance of building integrated information systems managed by the network. This is one of the most relevant operative actions required to have an effective data network of schools. This critical point emerged particularly from the *Predicting student success* project, in which the collaboration within the network of schools was initially hindered by the difficulties in creating a shared information system.

4.3.4 Strengths and weaknesses. The network-based configuration presents several strengths. Firstly, the new and highly qualified professional figure of the educational data scientist is an investment that can be sustained by a cluster of schools, while it could have been hardly faced by a single institution. Then, having all information accessible in the same "place" allows shortening the phase of collection and cleaning, which are the longest phases when dealing with data. Thirdly, the data bounce among schools and CA is simplified and streamlined because of the mediation of the network, which has the competences and the information to easily share the data required. Further, this configuration pushes schools' principals to reflect and discuss the importance of data and an information-driven approach, which can be adopted along with more traditional decision-making processes. It is worth adding, if the previous configuration is more plausible for big

schools or schools with important resources, the network-based setting is more “democratic”, giving to schools of all typologies, dimensions or locations to be part of the network.

On the other hand, the weaknesses related to the network-based approach are related to the ability of schools to collaborate and centralise their information: only through a real understanding of the value of this activity this configuration may be a success. This implies that principals need to collaborate and set common guidelines to improve their schools’ management. This is the most critical part, as schools are sometimes in competition with each other to attract students. On the technical side, the main challenge is the creation of a centralised information system, requiring an important investment in terms of resources. This represents the main switching (and fixed) cost to move from a decentralised setting to a network-based one. On this issue, the creation of school networks should be endorsed at the central level, by supporting schools in the initial stage of network creation.

5. Discussion and concluding remarks

The study analyses the organisational implications of data analytics in education, by proposing three alternative organisational configurations that can be adopted at the national level. The configurations differ in their balance between organisational layers, in the distribution of roles and competences between the schools and the CA and in their approach to data management, as needed to support evidence-based decision-making. Despite the importance that data analytics has increasingly gained in education (Agasisti and Bowers, 2017; Cech *et al.*, 2018), to the best of our knowledge, this is the first study that proposes an integrated organisational framework for the adoption of data analytics in education. Similar reflections on the possibility to create either a centralised or a decentralised structure to support data analytics have been proposed by Grossman and Siegel (2014), while other studies considered the organisational implications of data analytics in the different contexts (Wang *et al.*, 2018; Yu *et al.*, 2021). However, their analyses moved from the perspective of a single organisation, while the purpose of this study is to highlight the organisational complexity that arises when analysing the school in connection with the CA and the external stakeholders, within the context of data analytics. The current study contributes specifically to this gap, by means of a participant observation study contextualised within the Italian education system. Despite being empirically applied to the educational sector, the organisational configurations developed in the study may be adapted and generalised to alternative fields of application, given the relevance of data analytics in several contexts (Vidgen *et al.*, 2017).

The three organisational configurations can be interpreted as alternative frameworks characterised by different levels of complexity. The “as-is” framework is the centralised configuration, which may be indeed more suitable for an early stage of application, in which the need to gain the technical competences, retrieve data and build the integrated infrastructure can be managed more efficiently and effectively from the central level. However, the distance from the specific data needs of schools can represent a limit to the full exploitation of data analytics. The decentralised and the network-based configurations belong to a more mature stage of data analytics, in which the organisational configuration gets closer to the specific needs of schools, yet responding accountability needs at the central level. Both the alternative configurations still present some criticalities. In the decentralised model, they are related to the critical mass of resources (both human and technological) that would be needed to implement the configuration. On the other hand, the centralised configuration presents the cons of not being close to the students, the final subjects of interest. The network-based model represents a middle way between the other two configurations. In this model, the resource issue is much less critical, while the element of

attention is related to the creation of the integrated data system. Still, the network-based approach gives schools the possibility to achieve a critical mass by leveraging on the strength of the network and in particular on the role of the educational data scientist to support the use of data.

From this perspective, the network-based model overcomes the usual critics addressed to the accountability of school networks, namely, the unclarity and the difficulty to assess network outcomes and functioning (Ehren and Perryman, 2018). In the proposed model, instead, each school remains accountable for its activities, while the network supports only their ability to use data analytics for decision-making. Thus, they would share competences for data analysis, while remaining accountable for their individual accountability duties towards the CA. In this process, the educational data scientist would create synergies in the database infrastructure, by simplifying the integration of the multiple data sources that each school can access to King Smith *et al.* (2020). The removal of a data silos approach is not only relevant to increase the efficient use of data but also is critical to foster its effectiveness and support student learning (Supovitz and Klein, 2003).

Within this setting, the definition of the organisational configuration should be in chief on the CA on the basis of three main considerations, which represent the policy implications of this study. Firstly, the choice of the configuration comes after the definition of the intended level of maturity to be reached by the data analytics system. As anticipated, the decentralised and the network-based models require greater maturity at the system level as the result, first of all, of larger school autonomy. To make the most out of a decentralised system, the policy-maker should invest in human and technological resources for a structured adoption of data analytics at the school level. This would include training for the educational data scientist, as well as for school managers and principals, in addition to technological devices and servers. In this respect, a structured decentralised approach to data analytics and data-informed decision-making requires financial and human resources, as well as autonomy in setting priority based on data. While the extent of school autonomy may vary depending on the national context, the availability of resources from central government to school to support school investments in data analytics is generally poor (Agasisti and Bowers, 2017) and this represents a challenge of primary importance.

In close connection with this first element, there is a second factor that can be considered transversal to all configurations and relates to the need of reinforcing the organisational culture towards data analytics. Indeed, the success behind the adoption of a certain configuration is highly dependent on the techno-enthusiasm or skepticism among school managers and principals, *in primis* (Guenduez *et al.*, 2020). This factor, thus, plays the role of a relevant enabler that goes beyond the mere training on technology and modelling, as it requires a positive environment in which the policy-maker can play a leading role (Frisk and Bannister, 2017).

A final factor relates to privacy issues, which may be an element of concern for all the configurations and for the network-based configuration in particular given the need to build an integrated data infrastructure. It has been acknowledged for many years (Kobsa, 1990) that personalised interaction and user modelling have significant privacy implications because personal data about users needs to be collected for proper modelling. Big data and analytics present a number of ethical considerations, particularly around privacy, informed consent, data protection and raise relevant questions about what kinds of data to combine and analyse and the purposes behind it (Eynon, 2013). These concerns should be carefully considered by the policy-maker when designing national configurations and specifically refer to the data management dimension investigated in this study.

In conclusion, the paper proposes relevant suggestions for policy-makers, whose commitment and decisional power is fundamental to structure the educational system in the direction of a more structured approach to data analytics. Indeed, by highlighting the main features affected by each configuration, as well as the strengths and weaknesses of each model, the study aims to support the policy-makers in figuring out how to define a systemic approach to data analytics in education.

The study has a main limitation related to the type of information supporting the analysis, which is only based on the direct involvement of the researchers in the aforementioned research projects. As an avenue for future research, it would be useful to complement the current analysis with additional quantitative data on how schools are using data to support decision-making. In detail, further research could deal with the collection of data based on a survey to catch the perception and the use of data analytics by school principals. Quantitative evidence on the status quo about data analytics could be then complemented by additional qualitative insights based on focus groups involving policy-makers, school principals and educational data scientists for validating the organisational configurations proposed by the present study.

References

- Agasisti, T. and Bowers, A.J. (2017), "Data analytics and decision making in education: towards the educational data scientist as a key actor in schools and higher education institutions", *Handbook of Contemporary Education Economics*, Edward Elgar Publishing, Cheltenham.
- Agasisti, T., Ieva, F., Masci, C., Paganoni, A.M. and Soncin, M. (2017), "Using statistical analytics to study school performance through administrative datasets", *Data Analytics Applications in Education*, Vol. 181, pp. 183-209.
- Arnaboldi, M. and Azzone, G. (2020), "Data science in the design of public policies: dispelling the obscurity in matching policy demand and data offer", *Heliyon*, Vol. 6 No. 6, p. e04300.
- Cech, T.G., Spaulding, T.J. and Cazier, J.A. (2018), "Data competence maturity: developing data-driven decision making", *Journal of Research in Innovative Teaching and Learning*, Vol. 11 No. 2, pp. 139-158.
- Chatti, M.A. and Muslim, A. (2019), "The PERLA framework: blending personalization and learning analytics", *International Review of Research in Open and Distributed Learning*, Vol. 20 No. 1, pp. 243-261.
- Davies, P. (1999), "What is evidence-based education?", *British Journal of Educational Studies*, Vol. 47 No. 2, pp. 108-121.
- Dawson, S., Joksimovic, S., Poquet, O. and Siemens, G. (2019), "Increasing the impact of learning analytics", *Proceedings of the 9th international conference on learning analytics and knowledge*, pp. 446-455.
- Ehren, M. and Perryman, J. (2018), "Accountability of school networks: who is accountable to whom and for what?", *Educational Management Administration and Leadership*, Vol. 46 No. 6, pp. 942-959.
- Eynon, R. (2013), "The rise of big data: what does it mean for education, technology and media research?", *Learning, Media and Technology*, Vol. 38 No. 3, pp. 237-240, doi: [10.1080/17439884.2013.771783](https://doi.org/10.1080/17439884.2013.771783).
- Ferguson, R. (2012), "Learning analytics: drivers, developments and challenges", *International Journal of Technology Enhanced Learning*, Vol. 4 No. 5/6, pp. 304-317.
- Ferreira, L.D. and Merchant, K.A. (1992), "Field research in management accounting and control: a review and evaluation", *Accounting, Auditing and Accountability Journal*, Vol. 5 No. 4, pp. 3-34.

-
- Figlio, D. and Loeb, S. (2011), "School accountability", *Handbook of the Economics of Education*, Vol. 3, pp. 383-421.
- Frisk, J.E. and Bannister, F. (2017), "Improving the use of analytics and big data by changing the decision-making culture: a design approach", *Management Decision*, Vol. 55 No. 10, pp. 2074-2088.
- Gaftandzhieva, S., Docheva, M. and Doneva, R. (2021), "A comprehensive approach to learning analytics in Bulgarian school education", *Education and Information Technologies*, Vol. 26 No. 1, pp. 145-163.
- Gershenson, S. (2021), "Identifying and producing effective teachers", IZA Discussion Paper No. 14096.
- Goldberg, P., Sümer, Ö., Stürmer, K., Wagner, W., Göllner, R., Gerjets, P., Kasneci, E. and Trautwein, U. (2019), "Attentive or not? Toward a machine learning approach to assessing students' visible engagement in classroom instruction", *Educational Psychology Review*, Vol. 33, pp. 27-49.
- Grossman, R. and Siegel, K. (2014), "Organizational models for big data and analytics", *Journal of Organization Design*, Vol. 3 No. 1, pp. 20-25.
- Guenduez, A.A., Mettler, T. and Schedler, K. (2020), "Technological frames in public administration: what do public managers think of big data?", *Government Information Quarterly*, Vol. 37 No. 1, p. 101406.
- Jarke, J. and Breiter, A. (2019), "The datafication of education", *Learning, Media and Technology*, Vol. 44 No. 1, pp. 1-6.
- King Smith, A., Watkins, K.E. and Han, S.H. (2020), "From silos to solutions: how one district is building a culture of collaboration and learning between school principals and central office leaders", *European Journal of Education*, Vol. 55 No. 1, pp. 58-75.
- Kobsa, A. (1990), "User modeling in dialog systems: potentials and hazards", *AI and Society*, Vol. 4 No. 3, pp. 214-231.
- Krumm, A., Means, B. and Bienkowski, M. (2018), *Learning Analytics Goes to School: A Collaborative Approach to Improving Education*, Routledge, New York, NY.
- Leithwood, K., Sun, J. and Schumacker, R. (2020), "How school leadership influences student learning: a test of 'the four paths model'", *Educational Administration Quarterly*, Vol. 56 No. 4, pp. 570-599.
- Li, Y. and Zhai, X. (2018), "Review and prospect of modern education using big data", *Procedia Computer Science*, Vol. 129, pp. 341-347.
- Mac Iver, M.A., Stein, M.L., Davis, M.H., Balfanz, R.W. and Fox, J.H. (2019), "An efficacy study of a Ninth-Grade early warning indicator intervention", *Journal of Research on Educational Effectiveness*, Vol. 12 No. 3, pp. 363-390, doi: [10.1080/19345747.2019.1615156](https://doi.org/10.1080/19345747.2019.1615156).
- Marcenaro-Gutiérrez, O.D., González-Gallardo, S. and Luque, M. (2021), "Evaluating the potential trade-off between students' satisfaction and school performance using evolutionary multi objective optimization", *RAIRO – Operations Research*, Vol. 55, pp. S1051-S1067.
- Marienko, M., Nosenko, Y., Sukhikh, A., Tataurov, V. and Shyshkina, M. (2020), "Personalization of learning through adaptive technologies in the context of sustainable development of teachers education", arXiv preprint arXiv:2006.05810.
- Masci, C., Johnes, G. and Agasisti, T. (2018), "Student and school performance across countries: a machine learning approach", *European Journal of Operational Research*, Vol. 269 No. 3, pp. 1072-1085.
- Mining, T.E.D. (2012), "Enhancing teaching and learning through educational data mining and learning analytics: an issue brief", *Proceedings of conference on advanced technology for education*, pp. 1-64.
- Namoun, A. and Alshamqiti, A. (2021), "Predicting student performance using data mining and learning analytics techniques: a systematic literature review", *Applied Sciences*, Vol. 11 No. 1, p. 237.
- Parker, L.D. (2007), "Internal governance in the nonprofit boardroom: a participant observer study", *Corporate Governance: An International Review*, Vol. 15 No. 5, pp. 923-934.

-
- Reyes, J.A. (2015), "The skinny on big data in education: learning analytics simplified", *TechTrends*, Vol. 59 No. 2, pp. 75-80.
- Romero, C., Ventura, S. and De Bra, P. (2004), "Knowledge discovery with genetic programming for providing feedback to courseware author", *User Modeling and User-Adapted Interaction*, Vol. 14 No. 5, pp. 425-464.
- Romero, C., Ventura, S., Pechenizkiy, M. and Baker, R. (2010), *Handbook of Educational Data Mining*, Taylor and Francis, New York, NY.
- San Fabián Maroto, J.L. (2011), "The role of school organization in educational change: the inertia of the status quo", *Revista de Educación*.
- Scheuer, O. and McLaren, B.M. (2012), "Educational data mining", *Encyclopedia of the Sciences of Learning*, Springer, pp. 1075-1079.
- Schildkamp, K. (2019), "Data-based decision-making for school improvement: research insights and gaps", *Educational Research*, Vol. 61 No. 3, pp. 257-273, doi: [10.1080/00131881.2019.1625716](https://doi.org/10.1080/00131881.2019.1625716).
- Schiltz, F., Sestito, P., Agasisti, T. and De Witte, K. (2018), "The added value of more accurate predictions for school rankings", *Economics of Education Review*, Vol. 67, pp. 207-215.
- Sepulveda, A. (2020), *The Digital Transformation of Education: Connecting Schools, Empowering Learners*, UNESCO report.
- Sergis, S. and Sampson, D.G. (2016), "School analytics: a framework for supporting school complexity leadership", *Competencies in Teaching, Learning and Educational Leadership in the Digital Age*, Springer, Cham, pp. 79-122.
- Siemens, G. (2013), "Learning analytics: the emergence of a discipline", *American Behavioral Scientist*, Vol. 57 No. 10, pp. 1380-1400.
- Siemens, G., Dawson, S. and Eshleman, K. (2018), "Complexity: a leader's framework for understanding and managing change in higher education", *Educause Review*, Vol. 53 No. 6, pp. 27-42.
- Slavin, R.E. (2002), "Evidence-based education policies: transforming educational practice and research", *Educational Researcher*, Vol. 31 No. 7, pp. 15-21.
- Supovitz, J.A. and Klein, V. (2003), *Mapping a course for improved student learning: how innovative schools systematically use student performance data to guide improvement*, Report, Consortium for Policy Research in Education, Philadelphia.
- UNESCO (2017), *Accountability in Education: Meeting our Commitments*, Global Education Monitoring Report 2017/8, UNESCO Publishing, Paris.
- Vidgen, R., Shaw, S. and Grant, D.B. (2017), "Management challenges in creating value from business analytics", *European Journal of Operational Research*, Vol. 261 No. 2, pp. 626-639.
- Walkington, C.A. (2013), "Using adaptive learning technologies to personalize instruction to student interests: the impact of relevant contexts on performance and learning outcomes", *Journal of Educational Psychology*, Vol. 105 No. 4, p. 932.
- Wang, Y. (2016), "Big opportunities and big concerns of big data in education", *TechTrends*, Vol. 60 No. 4, pp. 381-384.
- Wang, Y., Kung, L. and Byrd, T.A. (2018), "Big data analytics: understanding its capabilities and potential benefits for healthcare organizations", *Technological Forecasting and Social Change*, Vol. 126, pp. 3-13.
- Williamson, B. (2016), "Digital education governance: data visualization, predictive analytics and 'real-time' policy instruments", *Journal of Education Policy*, Vol. 31 No. 2, pp. 123-141.
- Yu, W., Zhao, G., Liu, Q. and Song, Y. (2021), "Role of big data analytics capability in developing integrated hospital supply chains and operational flexibility: an organizational information processing theory perspective", *Technological Forecasting and Social Change*, Vol. 163, p. 120417.

Further reading

- Agasisti, T., Bowers, A.J. and Soncin, M. (2019), "School principals' leadership types and student achievement in the Italian context: empirical results from a three-step latent class analysis", *Educational Management Administration and Leadership*, Vol. 47 No. 6, pp. 860-886.
- Jeladze, E. and Pata, K. (2017), "Change management: digital transition in the public schools of GA", International Workshop on Learning Technology for Education in Cloud, Springer, Cham, pp. 171-184.
- Sampson, D. (2017), "Teaching and learning analytics to support teacher inquiry", *2017 IEEE Global Engineering Education Conference (EDUCON)*, IEEE, pp. 1881-1882.

Corresponding author

Mara Soncin can be contacted at: mara.soncin@polimi.it