Research Article

The mediating role of big data analytics in enhancing firms' commitment to sustainability

Wail El Hilali[1*], Abdellah El Manouar[2] and Mohammed Abdou Janati Idrissi[2]
Research Scholar, Mohammed V University in Rabat, Morocco[1]
Professor, Mohammed V University in Rabat, Morocco[2]

Abstract
The era of big data has overturned the classical model of how businesses work. Companies are now dealing with data as one of their most valuable assets that fuels their digital strategies to overcome the competition. As the potential of big data analytics is endless, looking for any impact from this cutting-edge technology on sustainability could lead to new insights about this paradigm. This paper represents a quantitative study that included 41 Moroccan firms from diverse industries. By using a Partial Least Squares-Structural Equation Modeling (PLS-SEM) approach, the results showed that big data analytics may enhance companies' commitment to sustainability (p-value =0) if the firm crafted the right digital strategy fuelled by data (p-value=0), established a data-driven decision-making culture (p-value =0.008), and acquired the right talents with soft skills (p-value = 0.029). Against all expectations, the results showed that technical skills do not play a significant role and do not boost the mediating role of big data regarding sustainability (p-value=0.124).

Keywords
Sustainability, Big data, Digital transformation, Digital capabilities.

1.Introduction

Big data is one of these attractive phrases that is used everywhere nowadays. It is considered to be the driving force behind the digital revolution that companies are currently leading. According to IBM [1], the amount of data that was created in 2020 is mind-boggling as it is 300 times more than the number of bytes that were available in 2005, just 14 years ago. Overemphasising data when compared its impacts on this era to the impacts of oil on industries during the last century was being discussed [2].

Companies nowadays have all the possibilities to integrate this technology during their digital transformation journeys. As data are considered to be the fuel of the economy in the digital era, embracing this digital capability becomes necessary to survive in this volatile climate.

We are moving past the days when top management waived control of data to IT department. Data should be at the heart of any business strategy.

Discussions nowadays focus on how companies can monetize data collected from their customers, products and operations by combining it with cutting-edge technologies (advanced analytics, deep learning…) to both enhance their business processes and pivot their business models [3].

Potentials of big data are endless. If its impacts on firms' financial numbers are universally recognized, studies that link this technology to social and environmental efforts are still scant. In fact, data analytics could be "the secret ingredient" that companies are looking to own in order to differentiate brand image and to reach sustainability (a mindset that looks for an equilibrium between economy, society and environment) in this era.

This paper presents a pure quantitative study that tries to prove the impacts of big data analytics on a firm commitment to sustainability. 41 Moroccan companies from different industries were the object of an empirical investigation. The gathered data were used to validate our proposed research model which highlights the mediating role of business data analytics regarding the quest for sustainability.

*Author for correspondence

The study addresses the following research questions:
RQ1: How could companies seize the tremendous opportunities that big data offer in order to enhance their sustainability footprint?
RQ2: What are the main drivers that enhance the mediating role of big data analytics regarding sustainability?

The results using the Partial Least Squares-Structural Equation Modeling (PLS-SEM) approach show that big data analytics can boost companies' efforts to reach sustainability if they succeed in crafting a digital strategy fueled by data. Establishing a data-driven decision-making culture, and recruiting the right talents with the right soft skills. These findings will encourage firms to embrace digital capabilities and seize the opportunities that a digital transformation may offer regarding the sustainability paradigm.

This paper is organized as follows: the next section presented the literature review on both the technical and business sides of big data. It discusses some recent and related works that linked big data to sustainability. Section 3 focused on the proposed research model and formulated hypotheses, whereas section 4 explains our methodology. The results were displayed in section 5 and discussed in section 6. Finally, the paper concludes with limitations and future directions of this study.

2.Literature review
2.1Big data: definitions
Big data is not just a new buzzword, it is a concept that has emerged slowly and captured both the IT and business worlds. The appearance of data ecosystems that offer a collection of infrastructure, analytics, and applications (such as the Hadoop ecosystem) has simplified life for business to embrace this technology.

This concept is defined in many ways in the literature. Oracle [4], for example, contends that big data is the derivation of value from both structured and unstructured data. As for Microsoft [5], the giant tech company, provides a beautiful definition: big data is the term used today in describing the process of applying cutting-edge computing technologies (machine learning, deep learning, artificial intelligence etc.) to enormous and highly complex sets of data. For Intel [4], big data means data of huge and unusual size (that exceeds the capacity of a relational database) or generated at spectacular velocity. A clearer definition of this concept appeared in the MIT Technology Review magazine [6], in which big data were defined as a concept that describes the processing and analysis of huge and complex data sets using new technologies such as NoSQL, MapReduce and machine learning.

Theoretically, big data are generally concerned with five Vs [7]:
- Volume, as the generated data is in terabytes, or even petabytes, which exceed the capacity of a relational database.
- Velocity, which evokes the speed generating and processing data
- Variety, as data is collected from different sources
- Veracity that shows how truthful the collected data is.
- A value which refers to the capability to extract valuable business insights from huge sets of data

2.2The technical side of big data
Big data analytics combine solutions and means to develop algorithms that access, consume and manage data which is not supported by traditional database systems. Professionals recommend to have recourse to big data architecture in order to exploit its maximum potential. Adopted architectures could include some or all of the following elements according to reference [8]:
- Data sources: concerning real-time data sources, such as Internet of Things (IoT) devices or data stores including structured databases.
- Data storage: storage in distributed files that can hold large volumes of data. Storing data this way is often called "Data Lake" and used in case of batch processing operations.
- Batch processing: jobs that are used to clean and prepare the data for processing.
- Real-time message ingestion: methods to collect and store real-time data to be processed.
- Stream processing: the cleaning and processing of data after capturing real-time generated data.
- Analytical data store: for example, a Kimball-style relational data warehouse or it could be presented through a low-latency NoSQL technology. It could also concern an interactive Hive database that provides a metadata abstraction over data files in the distributed data store.
- Analysis and reporting: reporting is the act of transforming raw data into information while analysis transforms the obtained information into business insights. To help users analyse the data, a data modelling layer should be a component of every big data architecture. A multidimensional Online Analytical Processing (OLAP) cube or

Wail El hilali et al.

tabular data model are examples of these components.

- Orchestration: in big data architecture, data is moved between multiple sinks and processes. This

is why orchestration technology should be used in order to automate these workflows.

Figure 1 represents the proposed logical components discussed above which constitute to big data architecture accordingly.
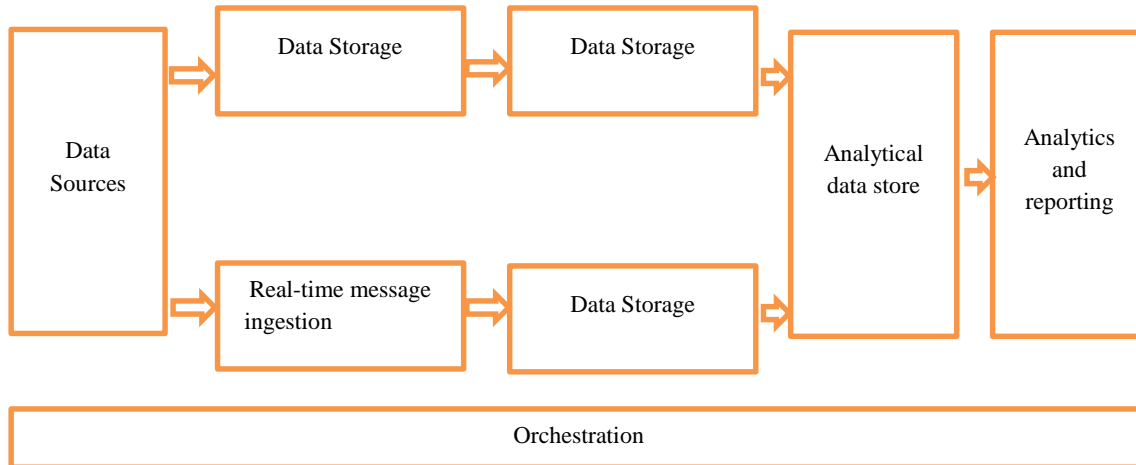

Figure 1 Big data architecture components

2.3 The business side of big data

Big data is a digital capability that enabled firms to rediscover their businesses through digital transformation. With the reporting and the analysis of structured, unstructured, real time and stored data, firms have all the abilities to boost their existing businesses, improve their customer experience, producing and monetizing new kinds of value adapted to this era through digital business models.

Each firm has their own drivers and reasons for embracing this technology. Some companies sow it in a way to reduce cost [9] while others take it as an opportunity to install a data-driven decision-making culture [10]. Big data are also considered by companies as the ultimate way to have a clearer understanding of market trends, and to preserve any competitive advantage and stay ahead of their competitors [11].

2.4 Big data and sustainability

Many papers in the literature discuss the relationship between big data and sustainability from different frames. A pure qualitative study on big-sized companies proves that embracing this technology can boost firms' quest for sustainability [12]. They highlight the positive impacts of big data analytics on sustainable business performances with a focus on the environmental pillar of the sustainability paradigm [13]. They focus on big data enhancement of firms' financial performances, whereas some scholars

consider big data analytics as an organisational capability and tested its effect on both social and environmental sustainability [14, 15]. Other researchers have proposed a reflection on the ethical implications of big data initiatives' environmental impacts [16, 17]. The privacy and the protection of personal data is also highlighted as potential risks of big data regarding the well-being of society [18]. This is why experts recommend the adoption of a big data governance framework to avoid any liabilities [19].

Although there is a sort of unanimity in the scientific community that big data may contribute to firms' efforts to reach sustainability, only few academic papers discuss the drivers that enhance the big data mediating role. In an emergent economy like Morocco where digital transformation is on everyone's lips, looking for the main drivers that enhance the impacts of big data on companies' quest for sustainability will certainly enrich the literature on this subject.

3. Research model and hypotheses

Many studies in the literature highlight the effect of big data on sustainability. As everything around us is impacted by big data today, companies try to apply data analytics to solve real word problems. Understanding operations, optimizing resource usage, assessing environmental risks and ensuring better environmental regulation are a few of many promises of big data regarding sustainability. As an example, Pirelli, the Italian tire company, has invested in the

Systems, Applications and Products for Data Processing (SAP's) big data management system, High Performance Analytical Appliance (HANA), to optimize its inventory. By collecting and assessing data from sensors in the tires, the firm succeeded in reducing waste and diverting defective ones from landfills, therefore lowering their environmental negative externalities [20]. Hence, we hypothesize that:

H1: Big data analytics generate useful insights related to sustainability issues

There is no doubt that technical skills are needed in order to take advantage of this technology. Reference[21] highlights through a deep analysis that analytical skills, data visualization skills, programming skills, skills of data mining and familiarity with technologies are the top in-demand big data skills that companies are looking for when hiring. The same skills were discussed in a report published by the United Nations [22], showing their indispensability regarding expressing and implementing public policies, projects, and management techniques and models conducive to sustainability. Confirming whether big data profiles have the required digital skills to cope is one of the more pressing issues [23]. Starting from this premise, we presume that:

H2: Technical skills are indispensable to enhance sustainability impacts through big data.

As well as the technical skills, managerial skills are of high importance for companies embracing big data technology [24]. Leadership, interpersonal skills, communication and motivations are few of many managerial skills that could enhance the role of big data analytics in improving sustainable capabilities [25]. Thus, we can assume that:

H3: Managerial skills are essential to enhance sustainability impacts through big data

The discussion on data-driven enterprises usually focusses on the technical side of big data that has made parsing and processing data faster and cheaper. While this is important, creating a data-driven culture across firms is mandatory to avoid and move beyond any successful but limited data initiatives to certain business areas.

Data should be considered as a strategic asset. According to reference [26], a data-driven culture will promote this mindset. Taking advantage of big data pushes companies to craft and adopt a data strategy. Aligning it with the firms' business strategies will avoid companies from possessing the big but useless "white elephant" [19]. As a reference [26] affirmed, "data should be for a purpose, not the purpose".

As developing metrics that measure progress towards sustainability is essential for companies in order to embed sustainability in the organizational culture [27], we could say that:

H4: Building a data-driven culture is essential for companies that bet on big data to enhance their sustainability commitment.

A data strategy helps to unlock the power of data. It is considered to be one of the first pillars of the digital strategy, resulting from the fusion of IT and business strategies in the digital era. In fact, by crafting digital strategies, companies succeeded in moving IT from supporting to enabling business strategies.

Digital strategies have strengthened the position of IT as a value creator [28]. Big data can affect perspectives on sustainability only if companies are considering big data as a digital capability that fuel and power a digital strategy. In this case, we can assume that:

H5: Big data can affect perspectives on sustainability only through a digital strategy

Figure 2 shows the proposed research model which highlights the mediating role of business data analytics regarding sustainability.
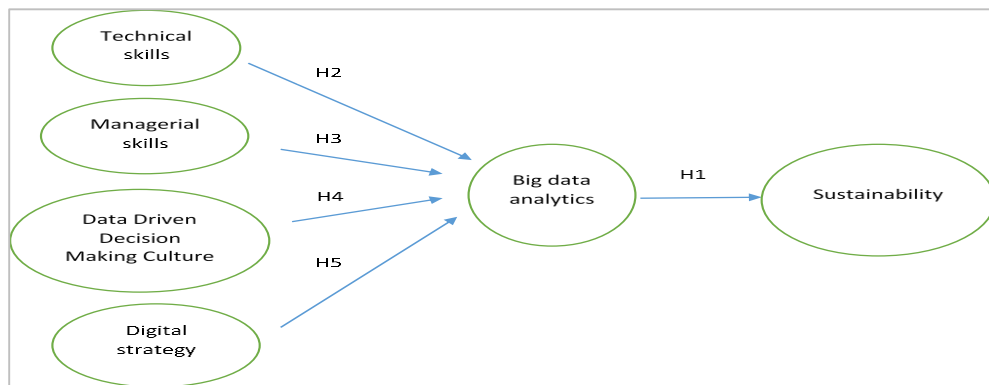

Figure 2 The proposed research model

4.Methods
4.1Research design and Sample
To understand and assess the mediating role of big data regarding sustainability, a survey was carried out, targeting 80 big Moroccan companies from different industries. 41 answers were received from firms operating in telecom, distribution, wholesale, banking, consulting, pension funds, and logistics. In order to honour ethical values, the participants gave their permission to share and publish the collected data anonymously for scientific purposes.

4.2Measures
The measurements were established after a deep analysis of previous works published on both big data and sustainability. The survey research was divided into two sections: the first part was dedicated to the demographic profile data of the 41 respondents, while the second section was devoted to measurements on the theoretical constructs of this quantitative study. *Table 1* shows the constructs and items of our hypothetical model.

The survey was composed with items that covered the six axes discussed above, namely technical skills, managerial skills, data-driven decision-making culture, digital strategy, big data analytics and sustainability. The participants were asked to quantify their agreement regarding each statement through a five-point Likert scale. 1 illustrates a strong disagreement and 5 shows strong agreement. Below are the details of the six constructs and their items which form the questionnaire:

Table 1 Constructs and items of our hypothetical model

Construct	Item
Technical skills	TS1: Content knowledge: Knowledge of how to use technology within a specific content area is important as a technical skill
	TS2: Analytical skills are crucial in order to embrace big data technologies
	TS3: Familiarity with technologies is more that needed to exploit big data capabilities
	TS4: Capacity to complete tasks with precision, respecting the timeframe are important while dealing with data
Managerial skills	MS1: Great leadership skills are mandatory to succeed in a data challenge
	MS2: Communication skills are more than needed to work in a context that involves dealing with data
	MS3: Interpersonal and negotiation skills are important
	MS4: Ability to analyse business problems is useful in a project that deals with data.
Data-Driven Decision-Making Culture	DDDMC1: Data is seen as a strategic and organizational asset
	DDDMC2: Metrics are clearly defined and shared with the teams regularly
	DDDMC3: Decisions are made using data rather than intuition. Hypotheses are tested and validated thanks to data.
Digital Strategy	DS1: IT is not considered as a support activity
	DS2: IT and business strategies were merged inside the company
	DS3: Data is considered as a digital capability and fuels the digital strategy of the firm
Big data analytics	BDA1: Big data analytics are used to generate predictive insights
	BDA2: Big data analytics are used to generate descriptive and prescriptive insights
Sustainability	S1: Economical performance is important for the sustainability mindset
	S2: Social performance is important for the sustainability mindset
	S3: Environmental performance is important for the sustainability mindset

4.3Methods
In order to assess our theoretical model, a PLS-SEM analysis was accomplished. This method of structural equation modeling allows estimating complex cause-effect relationship models through latent variables.

Structural equation modeling was originally developed to examine the multiple causal relationships. Then, their use gradually extended to analyze the validity of latent constructs and to conduct multi-group analysis [29].

We used SMART PLS 3.3.2 as a tool to analyze our data to check the validity of our hypotheses and our research model. We decided to perform a PLS-SEM analysis considering its ability to estimate causation among all latent variables simultaneously, dealing at the same time with measurement errors at the level of the structural model [30].

Moreover, according to [31, 32], the PLS-SEM method is recommended when the number of subjects included in the sample size is small, the data are not normally distributed, and the objective of the study is to identify key drive variables of a model. A model of structural equations traditionally consists of two parts: the measurement model (made up of all the relationships between the indicators and variables or latent constructs that they contribute to their measurement) and the structural model, which deals with the set of relationships between constructs and represents the network of causal relationships that the researcher wishes to establish [29]. *Figure 3* displays the way to conduct a PLS-SEM analysis according to [33].
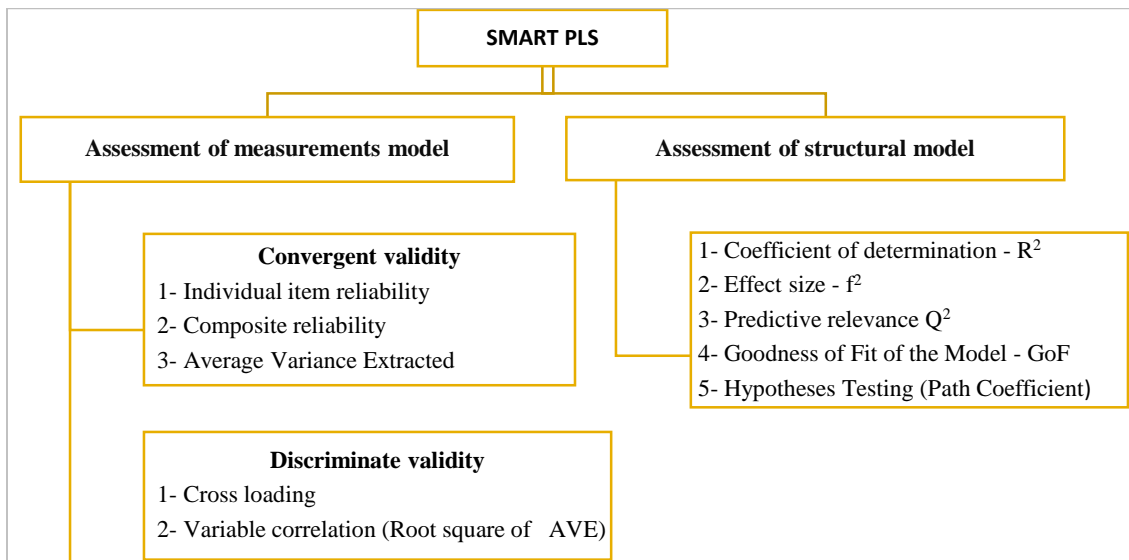

Figure 3 PLS-SEM steps

The two sections below explain each element of both the measurement and structural model.

4.3.1 Assessment of the measurement model

• **Convergent validity**

The convergent validity reflects how strongly each item correlates with its construct [34]. It is measured by three elements:
- The individual item reliability that denotes the percent of the indicator variance that is explained by the latent variable
- The composite reliability to evaluate internal consistency reliability
- The average variance extracted that reflects the correlation of alternative measures of the same construct[33]

• **Discriminate validity**

The goal of proceeding a discriminant validity assessment is to confirm that the relationship between any reflective construct and its items is the strongest [33]. This is reflected by two indicators as shown by reference [35]: the Fornell-Larcker and the cross-loading criterion. Cross-loading demonstrates how weakly each item correlates with all other variables except for the one with which it is affected [36]. Moreover, the Fornell-Larcker criterion detects those specific variations correlate more with another construct than with their own measures.

4.3.2 Assessment of the structural model

• **Coefficient of determination R^2**

R^2 is a metric that is used to examine the prediction power of the model. It is computed as the squared correlation between actual and predicted values of a specific endogenous construct [30]. R^2 could also be seen as the exogenous latent variables' combined effects of the endogenous latent variable.

• **f2 : The effect size**

The effect size f2 is a metric that reflects the change in the R2 value when one of the exogenous variables is withdrawn from the model [30]. This measure is used

to evaluate whether the real impact of the withdrawn variable on the endogenous constructs.

- **Predictive Relevance Q2**

Q^2 is a metric that reflects how predictive our model is. Q^2 greater than 0 means that the PLS-SEM model is predictive [33].

- **Goodness of Fit of the model**

The 5.2.5Goodness of fit of the model (GoF) is used to reflect for how much we can account on the studied model [37]. This metric is calculated as follows:

$$GoF= \sqrt{\overline{R^2} \times \overline{AVE}}$$

Baseline values were published to consider in order to explain GoF of any model. These values lie between 0 and 1. A cutoff value of 0.36 was declared by the same reference to consider the model with large effect sizes of R^2.

- **Hypothesis testing (path coefficient)**

In order to verify the hypotheses' validity, we estimated the paths between our independent constructs (technical skills, managerial skills, data-driven decision-making culture, digital strategy) and the purpose (big data analytics, sustainability).

Below is a block diagram that structures our complete quantitative study (*Figure 4*).
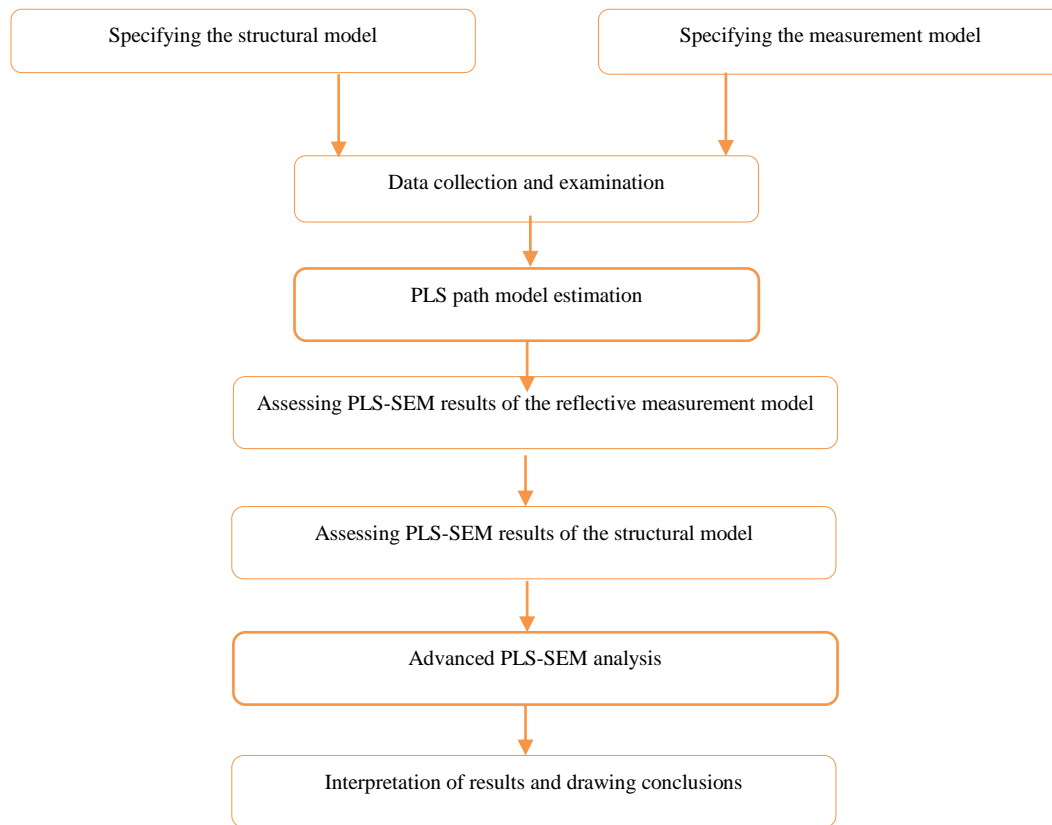
Figure 4 A block diagram that structures the study

5.Results
5.1The measurement model assessment
5.1.1The convergent validity
The results of the convergent validity indicators of our model are displayed in *Table 2:* The assessment of the convergent validity of the constructs is conclusive. The individual reliability of the items is all above 0.70, which validates their use to measure our constructs.

The internal consistency of our measurement model is guaranteed thanks to the values of Composite Reliability (CR) which are all between 0.7 and 0.95. The convergence of the validity of each construct, measure was evaluated by the Average Variance Extracted (AVE) metric which is higher than 0.5.

Table 2 The assessment of the convergent validity

Variables	Items	Individual item reliability	CR	AVE
Technical skills	TS1	0.879	0.915	0.729
	TS2	0.896		
	TS3	0.772		
	TS4	0.863		
Managerial skills	MS1	0.803	0.858	0.605
	MS2	0.882		
	MS3	0.669		
	MS4	0.742		
Data Driven Decision Making Culture	DDDMC1	0.726	0.846	0.649
	DDDMC2	0.889		
	DDDMC3	0.793		
Digital Strategy	DS1	0.896	0.885	0.721
	DS2	0.872		
	DS3	0.776		
Big data analytics	BDA1	0.774	0.833	0.715
	BDA2	0.912		
Sustainability	S1	0.778	0.879	0.709
	S2	0.918		
	S3	0.825		

5.1.2 Discriminate validity

The results of the cross-loading and the Fornell-Larcker criterion were also satisfactory. As highlighted in bold in *Table 3,* each item correlates weakly with variables other than the one with which it is affected. As for *Table 4*, it demonstrates that no variables correlate more with another construct than with their own measures.

These results confirm that our measurement model is both reliable and valid and could be used to assess our structural model.

Table 3 The cross-loading approach

	Big data analytics	Data driven decision making culture	Digital Strategy	Managerial skills	Sustainability	Technical skills
BDA1	**0.774**	0.633	0.545	0.506	0.39	0.242
BDA2	**0.912**	0.648	0.679	0.777	0.796	0.564
DDDMC1	0.535	**0.726**	0.484	0.56	0.574	0.326
DDDMC2	0.685	**0.889**	0.5	0.605	0.511	0.312
DDDMC3	0.577	**0.793**	0.238	0.584	0.372	0.136
DS1	0.593	0.376	**0.896**	0.408	0.536	0.343
DS2	0.621	0.404	**0.872**	0.403	0.526	0.334
DS3	0.637	0.505	**0.776**	0.616	0.763	0.312
MS1	0.612	0.489	0.657	**0.803**	0.708	0.324
MS2	0.562	0.486	0.422	**0.882**	0.55	0.388
MS3	0.291	0.494	0.11	**0.669**	0.308	0.095
MS4	0.773	0.714	0.412	**0.742**	0.648	0.621
S1	0.54	0.476	0.561	0.645	**0.778**	0.387
S2	0.705	0.604	0.614	0.726	**0.918**	0.432
S3	0.619	0.424	0.65	0.547	**0.825**	0.505
TS1	0.54	0.378	0.503	0.553	0.574	**0.879**
TS2	0.463	0.329	0.343	0.512	0.474	**0.896**
TS3	0.276	0.124	0.171	0.227	0.295	**0.772**
TS4	0.39	0.183	0.213	0.409	0.366	**0.863**

Table 4 The Fornell-Larcker criterion

	Big data analytics	Data driven decision making culture	Digital Strategy	Managerial skills	Sustainability	Technical skills
Big data analytics	**0.846**					
Data driven decision making culture	0.748	**0.806**				
Digital Strategy	0.73	0.508	**0.849**			
Managerial skills	0.782	0.722	0.564	**0.778**		
Sustainability	0.742	0.599	0.722	0.759	**0.842**	
Technical skills	0.51	0.321	0.389	0.526	0.523	**0.854**

5.2 The structural model assessment
5.2.1 Hypothesis testing (path coefficient)
The bootstrapping method through the software SMART-PLS was employed with 5000 resamples. The results are displayed in *Table 5*.

Accepted hypothesis are the ones with P-value lower than 0.05

Table 5 The model hypotheses and their path coefficients

Hypothesis	Path	Standardized Beta	Standard Error	T- value	P- value	Decision
H1	Big data analytics -> Sustainability	0.742	0.076	9.783	0	Supported
H2	Data-driven decision-making culture -> Big data analytics	0.322	0.121	2.658	0.008	Supported
H3	Digital Strategy -> Big data analytics	0.36	0.08	4.512	0	Supported
H4	Managerial skills -> Big data analytics	0.284	0.13	2.185	0.029	Supported
H5	Technical skills -> Big data analytics	0.117	0.076	1.538	0.124	Not Supported

5.2.2 R^2: The coefficient of determination R^2
Using the SMART-PLS software, the results of R2 are regarding our model are displayed in *Table 6*. Although the acceptance rate of R2 depends on the context [30], the threshold of 0.10 as a minimum

acceptable level was proposed by the reference [39]. R2 value of the complete model is 0.551 (see *Table 6*), making our model a good one.

Table 6 R-Square of the endogenous latent constructs

Endogenous latent variables	R^2	Results
Big data analytics	0.786	Strong
Sustainability	0.551	Strong

5.2.3 $f2$: The effect size
The interpretation of f2 was explained in details by reference [40]: if f2 is above 0.35, the variable is considered to have an important effect size. If f2 is between 0.15 and 0.35, it is interpreted with medium

effect size. If f2 is between 0.02 and 0.15, the effect size is small while f2 values lower than 0.02 are considered with no effect size. Results are displayed in the *Table 7:*

Table 7 Effect size f^2

	Big data analytics	DDDMC	Digital strategy	Managerial skills	Sustainability	Technical skills	Results
Big data analytics					1.227		**Large effect**
DDDMC	0.222						**Medium**
Digital Strategy	0.392						**Large**
Managerial skills	0.136						**Medium**
Sustainability							
Technical skills	0.045						**Small effect**

5.2.4 Predictive relevance Q2

In our case, calculations made by SMART-PLS (*Table 8*) show that $Q^2 > 0$, which means that our model is predictive.

Table 8 Predictive Relevance Q^2

Constructs	Q^2
Big data analytics	0.477
Data driven decision making culture	
Digital Strategy	
Managerial skills	
Sustainability	0.365
Technical skills	

5.2.5 Goodness of fit of the model (GoF)

For our study, the computed GoF was 0.50, exceeding the cutoff value of 0.36 announced by [38] for large effect sizes of R^2. In this case, we can conclude that the proposed model performs well in comparison to the cutoff value. *Figure 5* displays the PLS test of our structural model using SMART PLS software. Complete list of abbreviations is shown in *Appendix I*.
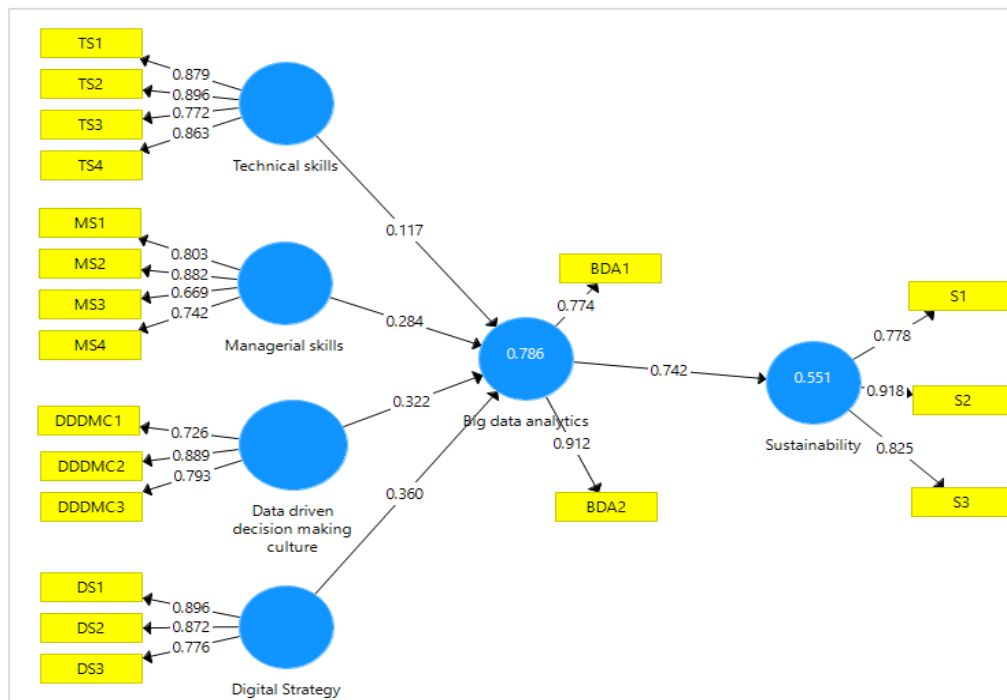

Figure 5 Results of the PLS test of the proposed structural mode using SMART PLS

6. Discussion

As we are living in a digital and connected world, the amount of data we produce every day is truly astounding. The corporate world should embrace the technology of big data in order to continue producing value in the digital era, answering their customers' social needs, understanding the environmental impact of their operations, and optimizing the use of resources.

This paper displays the results of an empirical study conducted within Moroccan context. The goal was to understand the mediating role of big data analytics in enhancing firms' commitment to sustainability. Managers of 41 companies were invited to answer questionnaires in order to prove our hypothetical model.

Four out of five hypotheses were validated through our empirical study. For hypothesis 1, the proposed model has brought evidence that big data could enhance the

sustainability footprints of companies. Big data's usefulness lies in its ability to help firms to rediscover the customer experience and to go beyond productivity frontiers by crafting blue ocean strategies and investing in adjusting businesses.

In regard to the third hypothesis, our model confirmed that leadership, communication, interpersonal and negotiation skills, and the ability to analyze business problems are essential soft skills that companies should look for while hiring big data profiles. These skills are mandatory and foundational for the rigorous application of data science to business problems. Team members should seek to master these skills in order to help their firms benefit from the tremendous opportunities that this cutting-edge technology could offer regarding sustainability.

The fourth hypothesis were also validated. It is reasonable and logical that a data-driven decision-making culture is more than needed to exploit all the potentials of big data analytics. This mindset treats data as a strategic and a crucial asset that will differentiate a business from its rivals regarding sustainability. Thanks to this culture, firms will avoid possessing the big but useless "white elephant" [19].

With regards to the fifth hypothesis, we succeeded in proving that crafting a digital strategy is mandatory for companies to fully benefit from digital capabilities. Through a data strategy, and as a part of their global digital strategy that deals with IT as a business enabler, companies could enhance their competitiveness, increase their impact on society and lower their environmental negative externalities.

Unfortunately, and unlike what we expected, we did not succeed in proving that technical skills, enhance the mediating role of big data analytics regarding sustainability inside these companies. Perhaps the ability to delegate and outsource the technical management of IT solutions to a third party, reduces the mandatory aspect of having in-house and available technical skills.

Study limitations

This paper has a few limitations. First, a convenient sampling method was used to select participants, which may skew the findings and influence the results sought. Second, the number of participants remains relatively small. More companies could be questioned to enforce the results. Third, there is more focus in the academic world on the technical side of big data than on its other aspects and impacts. This lack of prior research studies on the topic hinders us to compare our

results with other studies. Finally, all the questioned companies are from Morocco which could impede any attempt to generalize the results.

7.Conclusion and future work

Big data analytics are the future of business. Firms from all industries will have to learn in the near future how to manage their data to avoid being disrupted. This paper presented empirical proofs that big data analytics can help companies to go beyond the traditional business thinking and invest in sustainability. We succeeded in confirming that hiring leaders, implementing a data-driven decision-making culture and crafting a digital strategy can leverage big data analytics' mediating role regarding sustainability. However, the same study showed that technical skills do not enhance firms' commitment to sustainability.

Future works will consider broadening the scope of our study by inviting companies from other industries and countries to participate in our survey. It will also be interesting to verify the impacts of other digital capabilities (such as cloud computing, IoT, blockchain…) on a firm commitment towards sustainability.

Conflicts of interest
The authors have no conflicts of interest to declare.

References
[1] https://www.ibm.com/cloud/blog/structured-vs-unstructured-data. Accessed 11 April 2021.
[2] Economist T. Data is giving rise to a new economy. The Economist. 2017.
[3] Schmarzo B. Big data MBA: driving business strategies with data science. John Wiley & Sons; 2015.
[4] https://www.forbes.com/sites/steveolenski/2015/03/19/big-data-solving-big-problems/?sh=2584f2cb5b8e. Accessed 11 April 2021.
[5] Howie T. The big bang: how the big data explosion is changing the world. The Microsoft Enterprise Insight Blog. [Internet]. 2013.
[6] https://www.technologyreview.com/2013/10/03/82990/the-big-data-conundrum-how-to-define-it/. Accessed 11 April 2021.
[7] Gudivada VN, Baeza-yates R, Raghavan VV. Big data: promises and problems. Computer. 2015; 48(3):20-3.
[8] https://docs.microsoft.com/en-us/azure/architecture/data-guide/big-data/#components-of-a-big-data-architecture. Accessed 11 April 2021.
[9] https://www.wiley.com/en-us/Big+Data%3A+Concepts%2C+Technology%2C+a

nd+Architecture-p-9781119701828. Accessed 11 April 2021.

[10] Gorelik A. The enterprise big data lake: delivering the promise of big data and data science. O'Reilly Media; 2019.

[11] Walker R. From big data to big profits: success with data and analytics. Oxford University Press; 2015.

[12] Idrissi MA. Big data for sustainability: a qualitative analysis. In 5th international conference on cloud computing and artificial intelligence: technologies and applications 2020 (pp. 1-4). IEEE.

[13] Raut RD, Mangla SK, Narwane VS, Gardas BB, Priyadarshinee P, Narkhede BE. Linking big data analytics and operational sustainability practices for sustainable business management. Journal of Cleaner Production. 2019; 224:10-24.

[14] Ali Q, Salman A, Yaacob H, Zaini Z, Abdullah R. Does big data analytics enhance sustainability and financial performance? the case of ASEAN banks. The Journal of Asian Finance, Economics, and Business. 2020; 7(7):1-13.

[15] Dubey R, Gunasekaran A, Childe SJ, Papadopoulos T, Luo Z, Wamba SF, et al. Can big data and predictive analytics improve social and environmental sustainability? Technological Forecasting and Social Change. 2019; 144:534-45.

[16] Lucivero F. Big data, big waste? a reflection on the environmental sustainability of big data initiatives. Science and Engineering Ethics. 2020; 26(2):1009-30.

[17] Stahl BC, Wright D. Ethics and privacy in AI and big data: implementing responsible research and innovation. IEEE Security & Privacy. 2018; 16(3):26-33.

[18] Cuquet M, Vega-gorgojo G, Lammerant H, Finn R. Societal impacts of big data: challenges and opportunities in Europe. arXiv preprint arXiv:1704.03361. 2017.

[19] Marr B. Data strategy: how to profit from a world of big data, analytics and the internet of things. Kogan Page Publishers; 2017.

[20] Missbach M, Staerk T, Gardiner C, Mccloud J, Madl R, Tempes M, et al. SAP and the internet of things. In SAP on the cloud 2016 (pp. 139-51). Springer, Berlin, Heidelberg.

[21] Gardiner A, Aasheim C, Rutner P, Williams S. Skill requirements in big data: a content analysis of job advertisements. Journal of Computer Information Systems. 2018; 58(4):374-84.

[22] Cepal N. Big data and open data as sustainability tools: a working paper prepared by the economic commission for Latin America and the Caribbean. Economic Commission for Latin America and the Caribbean. 2014.

[23] Mckay E, Mohamad MB. Big data management skills: accurate measurement. Research and Practice in Technology Enhanced Learning. 2018; 13(1):1-24.

[24] Etzion D, Aragon-correa JA. Big data, management, and sustainability: strategic opportunities ahead. Organization & Environment. 2016; 29(2): 147-55.

[25] Singh SK, El-kassar AN. Role of big data analytics in developing sustainable capabilities. Journal of Cleaner Production. 2019; 213:1264-73.

[26] https://www.cio.com/article/3571792/how-to-create-a-data-driven-culture.html. Accessed 11 April 2021.

[27] Bertels S, Papania L, Papania D. Embedding sustainability in organizational culture. a systematic review of the body of knowledge. London, Canada: Network for Business Sustainability. 2010.

[28] Hilali WE, Manouar AE. Unlocking digitalization's possibilities: reaching sustainability by adopting the right digital strategy. International conference on wireless technologies, embedded and intelligent systems 2019 (pp. 1-74)

[29] Lacroux A. The advantages and limitations of the "partial least square" (PLS) method: an empirical illustration in the field of HRM. Human Resources Management Review. 2011; 2(80): 45-64.

[30] Hair JJF, Sarstedt M, Hopkins L, Kuppelwieser VG. Partial least squares structural equation modeling (PLS-SEM): an emerging tool in business research. European Business Review. 2014; 26(2):106-21.

[31] Richter NF, Sinkovics RR, Ringle CM, Schlägel C. A critical look at the use of SEM in international business research. International Marketing Review. 2016; 33(3):376-404.

[32] Hair JF, Risher JJ, Sarstedt M, Ringle CM. When to use and how to report the results of PLS-SEM. European Business Review. 2019; 31(1):2-24.

[33] Garson GD. Partial least squares. Regression and Structural Equation Models. 2016.

[34] Gefen D, Straub D. A practical guide to factorial validity using PLS-graph: tutorial and annotated example. Communications of the Association for Information Systems. 2005; 16:91-109.

[35] Hair JF, Ringle CM, Sarstedt M. Partial least squares structural equation modeling: rigorous applications, better results and higher acceptance. Long Range Planning. 2013; 46(1-2):1-12.

[36] Henseler J, Ringle CM, Sarstedt M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. Journal of the Academy of Marketing Science. 2015; 43(1):115-35.

[37] Henseler J, Sarstedt M. Goodness-of-fit indices for partial least squares path modeling. Computational Statistics. 2013; 28(2):565-80.

[38] Wetzels M, Odekerken-schröder G, Van OC. Using PLS path modeling for assessing hierarchical construct models: guidelines and empirical illustration. MIS Quarterly. 2009; 33(1):177-95.

[39] Falk RF, Miller NB. A primer for soft modelling. University of Akron Press. Akron, Ohio. 1992.

[40] Cohen J. Statistical power analysis for the behavioral sciences. Academic Press; 2013.

Wail El Hilali is currently a Ph.D. candidate at ENSIAS (National Higher School for Computer Science and System Analysis), Mohammed V University in Rabat, Morocco. His research thesis focuses on the use of information systems to reach sustainability. Mr El hilali obtained a National Engineering Diploma and a Master of Computer Science Engineering from ENSIAS Engineering School in 2009, and an MBA in general management from the Ponts business school in Paris.

Email: wailelhilali@gmail.com

Abdellah El Manouar holds a PhD in economics from Montreal University in 1991, Canada. He is Professor of Higher Education and Head of Business Intelligence Option and IT department and Decision Support. He is a member of the TIME research team at ENSIAS Engineering School (National Higher School of Computer Science and System), Mohammed V University in Rabat, Morocco. His research topics cover Finance and Investment, Management, New Economy (Information Economy, Net-Economy), Financial Management, Financial Engineering, Green IT, CSR

Email: manouar@ensias.ma

Mohammed Abdou Janati Idrissi holds a PhD in mathematics and computer science from the Rabat Sciences University in 2002, Morocco. He received a PhD of 3rd cycle in computer science from National Polytechnic Institute of Grenoble (INPG) in 1985. He is Professor of Higher Education and Deputy Director of Academic Affairs. He is the leader of the TIME research team at ENSIAS Engineering School, Mohammed V University in Rabat, Morocco. His research focuses on e-management, Business Analytics, Social Networks Analysis, Decision Aid, Multicriteria Analysis, Project Management, Operations Management, Supply chain Management, Green IT, CSR, Simulation, Optimization, Intelligent transport systems, Networking, Security, Cloud Computing, IOT, Wireless Networks, Performance evaluation, Vehicular Networks, Wireless sensor networks, DTN, SDN, Cognitive radio, Mobile smart Networks, IP Networks Security

Email: a.janati@um5s.net.ma

Appendix I

S.No.	Abbreviations	Description
1	AVE	Average Variance Extracted
2	CR	Composite Reliability
3	DDDMC	Data-driven Decision-Making Culture
4	GoF	Goodness of Fit
5	HANA	High Performance Analytical Appliance
6	IBM	International Business Machines Corporation
7	IoT	Internet of Things
8	IT	Information technology
9	MIT	Massachusetts Institute of Technology
10	OLAP	Online Analytical Processing
11	PLS	Partial Least Squares
12	PLS-SEM	Partial least squares structural equation modelling
13	RQ	Research Question
14	SAP	Systems, Applications and Products for Data Processing