

Global Big Data Maturity Model and its Corresponding Assessment Framework Results

Soukaina Mouhib, Houda Anoun, Mohammed Ridouani, Larbi Hassouni

Abstract— Nowadays, we live in the era of Big Data. Companies have realized by now that transitioning to a data-driven business is strategic for their growth and competitiveness. To achieve that, companies should get the most value out of their data through Big Data technologies.

In the previous work, we showed how maturity models are essential to assess the ability of companies to start Big Data projects and prevent eventual failures. We also explored literature and editors' offerings in this field and proposed an exhaustive maturity model that includes Methodology, a maturity domain of high importance and impact.

In the present work, we aim to provide a detailed picture of the proposed global Maturity Model design by exploring the temporal domains and explaining how they evolve through time. In addition, we aim to introduce the assessment framework, a tool we made available for North African companies to be able to evaluate their Big Data maturity. Unlike the currently available models, which are usually detailed and complex, the global assessment tool is quick and easy. We enriched the assessment questionnaire with best practices, and more importantly, the assessment tool suggests a list of shortcomings that companies should avoid in order to succeed in their Big Data adoption journey.

Moreover, we present the technologies used to implement the global assessment tool. We also show an example of a company's assessment results via visualizations.

Furthermore, we demonstrate the importance of the methodology domain through assessment results; and we analyze which sector has mature companies and which maturity domains are more mature by industry sector.

Finally, we conclude with an opening on using Multiple Criteria Decision-Making techniques to calculate companies' maturity accurate scores.

Index Terms— Big Data framework, Maturity Models, Maturity Assessment, Multiple criteria decision-making.

I. INTRODUCTION

FOR many years now, Big Data has revolutionized the world. The last years marked by the COVID-19 pandemic reveal the importance of Big Data in the vital field of healthcare. Many papers covered this subject, for instance: how to predict COVID-19 prevalence [1], and the

implementation of Big Data to restrain the spread of COVID-19 [2], [3].

While some authors focus on a particular aspect and choose to create a new kernel for analyzing Big Data [4] or create a plug-in to predict the relationship between COVID-19 prevalence and weather factors [1], others still focus on Big Data techniques [5] or its main processes [6] of data acquisition, management, analysis, and presentation.

Big Data allows companies to harness their data and get the most valuable insight to help run their business in an optimum manner. However, according to Gartner Top Data and Analytics Predicts [7] by 2025, 80% of organizations will fail in their digitalization journey because they are not adopting a modern approach in their data and analytics strategy.

Companies should excel in many areas to make their Big Data project a success. In most cases, Authors and practitioners refer to these areas as domains or dimensions within Maturity Models. Hence the importance of Big Data Maturity Models [8].

Companies must precisely define their maturity level to decide whether they can start their Big Data journey or improve in some areas first. Indeed, this kind of assessment addresses a variety of domains: human competencies, technology, data, governance, and many others [9].

To identify these domains, we went through Big Data success factors [10]. We studied the most recent and relevant papers for our research: Saltz [11], Eybers[12], Yeoh[13], Adrian [14], Sun and al. [15], and Marijn[16]. As a result, we created a summary table of Big Data success factors categories and identified the convergence and divergence between authors. This work helped us identify the majority of our global maturity domains.

To complete the vision, we extensively analyzed the existing literature and practitioners' offerings in maturity models [8] field. This research revealed two facts: First, we do not have a complete defined list of domains employed by all maturity models. Second, project methodology was not appearing as a domain in any of these maturity models. Hence, the motivation behind proposing the Global Maturity Model.

In this paper, we are not only completing the picture of our global Maturity Model design by exploring the temporal aspect of our Maturity domains [17]; furthermore, we will reveal our global maturity assessment Tool. Moreover, we will present some useful insights related to companies' assessment of their ongoing projects.

This paper is structured as follows: following the introduction, in section I, we will recall the motivation

Manuscript received June 09, 2022; revised February 13, 2022.

Soukaina Mouhib is a PhD candidate in Hassan II University RITM Laboratory, Casablanca, Morocco (e-mail: soukaina.ine@gmail.com).

Houda Anoun is a professor at High School of Technology, Hassan II University, Casablanca, Morocco (e-mail: houda.anoun@gmail.com).

Mohammed Ridouani is a professor at the High School of Technology, Hassan II University Casablanca, Morocco (email: mohammed.ridouani@gmail.com).

Larbi Hassouni is a professor at High School of Technology, Hassan II University Casablanca, Morocco (email: lhassouni@hotmail.com).

behind Big Data maturity models and illustrate the discoveries from our previous work [8]. Then, in section III, we will describe our Global Maturity Model and focus on Maturity levels, as we give a detailed description of each one of them. Section IV will be a preview of the assessment framework and a presentation of results and insights based on the expert's feedback. Furthermore, we will demonstrate the importance of methodology and show the technologies used for the assessment tool. Finally, we will conclude and highlight some future work in section V.

II. FLASHBACK

A. Motivation behind Big Data Maturity Models

The ultimate objective behind creating a Global Maturity platform is to support organizations in their Big Data journey. It is crucial to help them properly assess their maturity level at first; to achieve that, all companies should identify the assets they own and the capabilities they need to implement Big Data successfully.

In the previous paper [8], we explored the literature and editors' offerings. We aimed to offer companies a Global Maturity Model that covers the shortcomings discovered in existing work. In addition, we aimed to propose a framework to self-assess companies' readiness to adopt Big Data. This framework will not only help companies figure out their ability to start Big Data projects, but they can also use it during the implementation phase to measure their evolution and how well they are improving in the different areas.

B. Related Work

In our paper [8], we went through maturity models proposed by practitioners and academics from 2013 to 2020. The result was a list of fifteen Big Data Maturity models (TDWI[18], MM Index [19], IDC[20], MM Veenstra[21], InfoTech [22], MM Knowledgent [23], EL Darwiche[24], Radcliffe [25], IBM[26], Zakat [27], Hortonworks[28], Comuzzi [9], Adrian [29], Value-Based MM[30], Temporal BDMM[17]). Figure1 shows all these Maturity Models with their respective years of publication.

We analyzed these maturity models and extracted nine domains (Strategy, Process, Data Analytics, Data Management, People, Culture, Governance, Information management, and IT infrastructure). However, we noted that

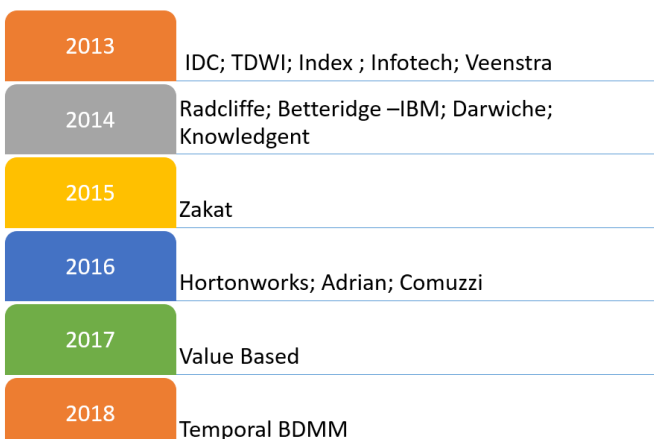


Fig. 1. Big Data Maturity Models proposed between 2013 and 2020

these maturity models are not exhaustive and only focus on a limited set of domains as we can see in figure 2. This figure also shows that none of the maturity models considers project methodology as a maturity domain, even if many publications since 2014 stressed the importance of methodologies and team management as an essential success factor of Big Data projects [23]. For instance, it has been proven that Big Data project methodologies such as STAMPEDE [31] or Process Guidance by Majchrowski [32] are critical factors in Big Data project success. Finally, we should not ignore the time dimension. As we mentioned earlier, time is a decisive factor in the Big Data journey.

Based on all these findings, we have proposed a Global Maturity Model aiming to be as exhaustive as possible. The model covers all the maturity dimensions identified in the literature and extends them with the methodology domain. However, until now, we did not explore the temporal aspect of maturity dimensions. In the next section of this paper, we will provide a detailed description of the proposed Global Maturity Model, its domains, and maturity levels.

III. GLOBAL BIG DATA MATURITY MODEL DESIGN

A. Global Big Data Maturity Model Areas

In previous work, we focused on Maturity dimensions, which we learned in depth in both: literature papers and practitioners' documentation. As a result, we proposed a Global Big Data Maturity Model (GBDMM) with six global Domains englobing ten dimensions as shown in figure 3. We can recall from the previous par paper [8] the following domains' description:

Strategy Alignment assesses principally: "Alignment of company orientations with Big Data approach; Management own project sponsorship; Business Process to integrate Big Data in decision-making; Budget allocation to Big Data project".

Data gauges: "Data sources; Data lifecycle management (collection, storage, processing and usage); Data Accessibility, Availability, quality and reporting solutions adopted by the company".

People evaluates whether or not they have: "Interdisciplinary team including experts on different technologies with different skills (Machine learning, Database, Analytics and business); Data-directed decision making company ; Organization change willingness; Culture of teamwork and data sharing (across companies' lines of business and from external entities)".

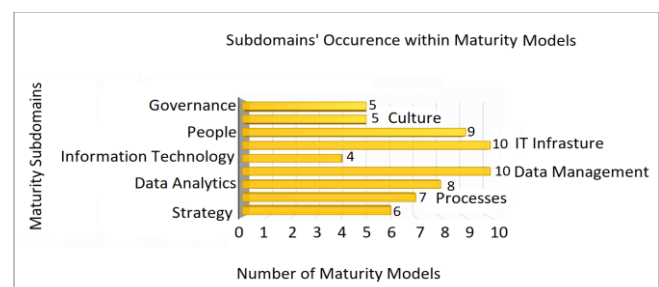


Fig. 2. The occurrence of Maturity subdomains within maturity model as mentioned in our previous paper [8]

Governance assesses “Data safety, privacy and ethical concerns”.

Technology gauges if the corporate has: “Business driven IT platform; Agile technology solutions; Flexible, scalable and secure infrastructure and platform”.

Project Methodology measures: Project outputs (if there are realistic and progressive); Project Committee with IT and Business users; Development approach (Iterative/ Linear) and clear deliverables of each iteration; Project risk (if it is identified and communicated/ not identified).

B. Global Big Data Maturity Model Levels

As mentioned in 2.2, Global Maturity Model's domains are temporal and evolve through time. This evolution represents the growth of the company through maturity levels. For our Global Maturity, we propose five stages shown in figure 4. These maturity levels are:

Ad hoc level: Company is doing only some statistic reports with spreadsheets. The company ignores Big Data capabilities.

Explore level: The company is learning about analytics and discovering Tools. Business people acknowledge analytics' value in decision-making. At this level, the IT team is initiating data integration and data consolidation to have consistent reports.

Transformation level: Data management and traditional dashboards are there. The company is extracting meaningful insights from data to facilitate the decision-making process. The company is also initiating Big Data use cases.

Adoption level: Analytics is essential for running the company's business. To have more insights, IT is using all data sources they have internally and externally.

Maturity level: An agile self-service analytics and consistent data are available to everyone. The Big Data platform is generating revenue and profits for the company.

Our Global Maturity domains evolve, as mentioned before, through those levels.

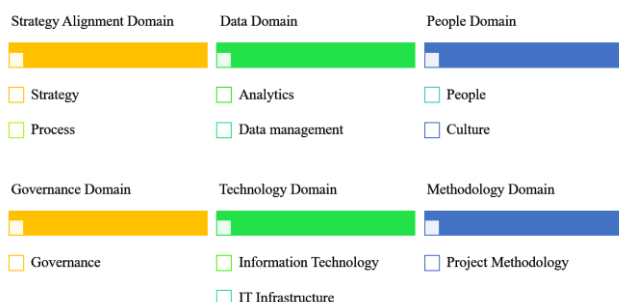


Fig. 3. : Global Maturity Model domains and subdomains

1) Ad hoc level

Strategy Alignment

The company ignores Big Data primacy to drive the business; there is no process at the company's level to incorporate data in decision-making; Business and IT analytics initiatives are not aligned with data strategy.

Data

The company uses only operational Data to create reports; IT team creates reports to monitor the business and assess activity performance; IT team provides Ad hoc analysis only when required by the business; there are Silos of Analytics across the company with inconsistent KPIs.

People

Few people in the company know analytics. At this stage, there is no collaboration.

Governance

No governance at this level

Technology

Company uses databases in silos (each department uses its data); the company has a fragmented infrastructure; the Current Platform cannot handle data growth.

Methodology

There is no methodology defined at this level; IT and the business do not work together to define use cases; Business people have no access to data without IT control.

2) Explore level

Strategy Alignment

The management is spreading the analytics culture in the company. The company is mindful of Big Data benefits and potential risks. Management considers a corporate-level process to implement Big Data and seek skills improvement.

Data

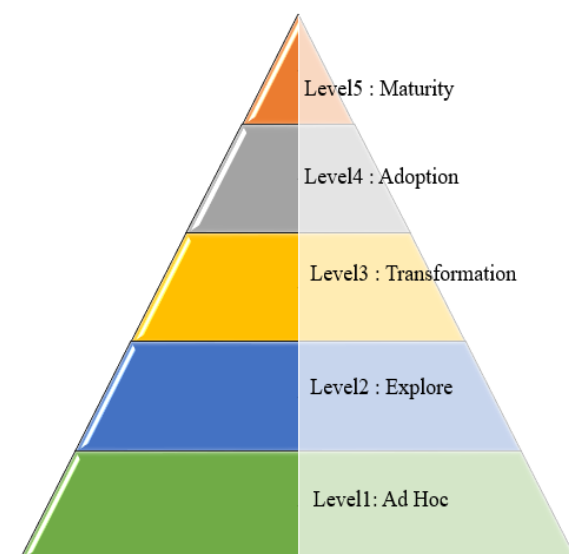


Fig. 4. Global Maturity Model levels

IT team aims to integrate and combine data sources to have consistent dashboards; the company is analyzing data quality issues and Business people get interested on visualization tools.

People

IT teams are initiated to Big Data concepts and best practices; at cultural level, management is introducing the concept of Big Data to the organization.

Governance

The company launches a discussion about data privacy and security.

Technology

Use integration solution to consolidate internal data sources and design data marts / data warehouse. The company use data visualization tools at the department level.

Methodology:

Company starts considering Big Data's use cases but no Project process is defined yet.

3) Transformation level

Strategy Alignment

Management prepares the ground to adopt Big Data. There is a financial commitment from the management; it is the start of Big Data processes' implementation at a corporate level; IT and the business are collaborating to solve some business issues and planning to cover larger scope. However, the company is not analytics-driven yet.

Data

DataMart is already in place; users have access to relevant and appropriate data for their analysis; company use advanced analytics, such as predictive modeling and does some data discovery using unstructured data.

People

The IT team is up skilled in Big Data technologies and best practices; the team also explores some business use cases. In the culture part, the company introduces the importance of knowledge sharing and collaboration.

Governance:

The company establishes departmental compliance policies and sensitive data is protected.

Technology

Analysts use data mining tools to create reports and insight with advanced analytics using ML algorithms. The company is preparing the ground for implementing a Big Data platform where all data types can coexist on the same platform (Data Lake); IT teams are studying Big Data architecture and design; the company may consider cloud services to implement the first proof of concept. The company widely uses data discovery self-service tools at this stage.

Methodology

IT and Business are collaborating to define clear objectives of the first POCs; the team sets Realistic goals for those Pilots and anticipates potential risk.

4) Adoption level

Strategy Alignment

At this level, Big Data is considered a competitive tool for the company. The management has fully defined processes to drive Big Data projects; IT successfully implemented Prototypes; Business and IT are working on new workloads.

Data

The company uses all external data types (partners' data, social media, and sensors); the company has Defined data life cycle management with a data lake and data warehouse. From an analytics perspective, we have a single truthful version of KPI across all the company

People

Dedicated team is certified in chosen Big Data Technologies to implement projects; the company is being more data-driven.

Governance

The company has implemented data governance policies; that guarantee data privacy while providing the users all the data they need for their analytics.

Technology

The company has a Big Data ecosystem fully implemented; backup and recovery procedures are in place; the choice may be a Big Data Appliance versus DIY Cluster or on premise versus cloud; analytics tools are generalized across the company.

Methodology

The company is mature enough to choose Big Data methodologies or Combine more than one methodology to drive Big Data project (CRISP-DM [33], IBM-Stampede [31], Architecture-centric Agile Big Data Analytics (AABA)[3], KDDS [34] or Process Guidance by C.Ponsard [36]).

5) Maturity level

Strategy Alignment

Big Data is fundamental for the company. It is a valuable asset not only to drive decisions; but also to create new opportunities and generate revenue; the company establishes best practices and Big Data processes at the corporate level; the company successfully implemented Big Data projects. Everyone in the company uses Big Data insights.

Data

Company at this level can integrate any new data sources (unstructured, real-time data, Geo-special...) and get insight on the spot; data and insight are accessible from any device, any time for everyone in the company.

People

At this level, the company has a multi-disciplinary collaborative team to run Big Data projects; Business people are proactive and work actively with data analytics on a daily basis.

Governance

At this stage, well defined Governance policies and committees are in place.

Technology

The company has an agile infrastructure adapted to any new workload; or any new business use case; the new infrastructure is scalable and can handle any amount of data; hybrid architecture is often used (Cloud and on premise Big Data solutions) in this kind of solution; containers and Kubernetes can be good options for an agile scalable and automated ecosystem.

Methodology

Well-defined Project methodology is in place. The company has already tested this methodology in previous pilots.

C. Global Big Data Maturity Model Design

Figure 5 shows the GBDMM conception. We converted this model to an assessment tool. At the end of the assessment, the company has a score that determines whether they are able to start their Big Data project or they should address the shortcomings listed by the Tool before retaking the assessment a second time. We will talk more in detail about the assessment framework in the next section.

IV. GLOBAL BIG DATA MATURITY ASSESSMENT FRAMEWORK

A. Global Big Data Maturity Model Tool

The first goal of our Global Maturity Model is to provide companies with a snapshot of their current state and list all the gaps to address.

We created an assessment framework that follows the same categories as the Global Maturity Model (Strategy, Data, People, Governance, Technology and Methodology). For each domain, we have a list of questions shown in Figure 6.

Each company can answer this questionnaire with the

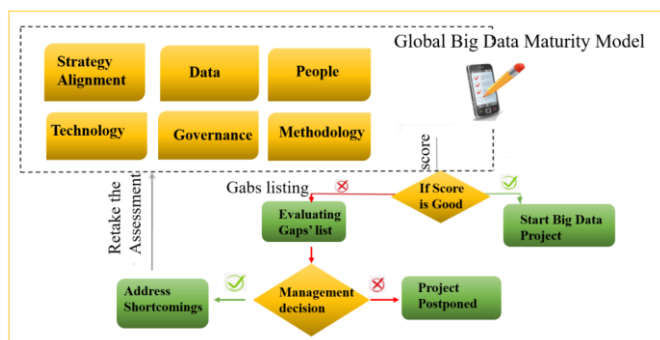


Fig. 5. Global Big Data Maturity Model Design[8] proposed by M. Soukaina .

following choices (Disagree, Somewhat disagree, Neutral, Agree, and fully agree).

The application will return a score from 1 to 5 for each answer, then sum up the answers' scores to provide the domains' scores and final result.

We added recommendations for assessment questionnaire to give the companies an idea about the best practices for each domain, and to help them assign the correct rating in the assessment.

Companies take the assessment by responding to all the questions and then submitting the poll. If the submission is successful, the company can view their responses and the results proposed by the framework.

After finishing the assessment, the framework provides three outputs:

- 1) Final score, which is the linear sum of maturity domains' scores;
- 2) Radar chart representing scores by maturity domains.
- 3) Areas of improvement for the specific company :

When the domain's score is low, the tool lists the shortcomings of this particular domain.

When scores are high, the tool lists the company's strengths.

When the global score is low, the company has to improve its competencies in all areas of improvement and then retake the assessment if needed. Once the company is well scored, they can start their implementation.

For better score interpretation, we present in Table I the maturity level corresponding to each global score and the breakdown of scores per dimension.

We can see that companies can analyze their maturity in two ways: globally or per domain.

Per domain:

If the score is ≤ 4 , the maturity level is Ad-hoc

If the score is $> 4 \ \& \ \leq 8$, the corresponding domain is at explore level, and so on. As explained in "Table I".

If the company wants to assess its maturity globally, we will have:

The global score is ≤ 24 means that the company is at an Ad-Hoc maturity level.

If the score is $> 24 \ \& \ \leq 48$, the company is at Explore level.

If the score is $> 48 \ \& \ \leq 72$, the company is at Transformation level and so on until global score > 96 , this final score is mapped with the most mature level which is Maturity level.

TABLE I. MATURITY LEVELS WITH CORRESPONDING GLOBAL AND DOMAIN SCORES

Maturity Levels	Domain's Score	Global Score
Ad-hoc	≤ 4	≤ 24
Explore	$> 4 \ \& \ \leq 8$	$> 24 \ \& \ \leq 48$
Transformation	$> 8 \ \& \ \leq 12$	$> 48 \ \& \ \leq 72$
Adoption	$> 12 \ \& \ \leq 16$	$> 72 \ \& \ \leq 96$
Maturity	$> 16 \ \& \ \leq 20$	$> 96 \ \& \ \leq 120$

Strategy Alignment <ul style="list-style-type: none"> Is there alignment of the business strategy with the big data Strategy ? Is there management support and sponsorship? As an enterprise are you data-driven ? Is the project cost known and approved 	People <ul style="list-style-type: none"> Are you , as company, ready for change? Do you have Multi-disciplinary collaborative team ? Do you have experienced technical team to work on the project? Is there culture of coordination and share understanding?
Data <ul style="list-style-type: none"> Do you have access to all data needed for use cases ? Are you experienced to use diversity of data sources Relational Semi –structured and Non- Relational data ? Did you have well defined Data lifecycle management (how Data will be collected, stored, processed and analyzed)? Are intern entities and partners ready for sharing useful data? 	Technology <ul style="list-style-type: none"> Do you have Agile technological tools with analytical and data visualization capabilities? Do you have data platform that enables you to access and process all data types coming from all sources (sensors, Events, social media, Enterprise applications ..etc.)? Are you considering hybrid architecture for you analytics platform in your Roadmap? Is you data management current solution capable to handle data volumes needed for your big data uses cases ?
Governance <ul style="list-style-type: none"> Are you planning to have centralized data lake and data warehouse to have consistent insights ? Do you have strict known rules to manage your sensitive and private data? You Data platform includes data encryption, key management and other practices to protect and secure all your data? Do you have processes and tools to ensure data Quality? 	Methodology <ul style="list-style-type: none"> Are you aware of big data methodologies or are you adopting an agile and business driven process to implement big data project? Is your process based on iterative development approach? Do you have well defined deliverables for each phase of the project. How you evaluate collaboration between IT and Business on defining the project use cases ?

Fig. 6. Assessment questions by global domain

B. Global Big Data Maturity Assessment Results

As explained in the previous section, each company responds to 24 questions, submits results, and gets the platform outputs. In figure 7, we have X4 Company results. As outputs for X4 Company, we have:

- 91 as Global Score;
- Radar Char illustrating X4 Company scores by Global domains with: Data score =12, Governance score =17, Methodology score=20, People score = 16, Strategy Alignment score = 14 and Technology score is 12;
- X4 Company areas of improvement versus areas of strengths which are :
 - Big Data is aligned with your strategy.
 - Top management is sponsoring your project.
 - The project budget is allocated.
 - You have tools and processes to centralize, clean, and protect your data.
 - You have a skilled team to handle the project/
 - Readiness to change and to be a data-driven organization.
 - You are aware of Big Data methodologies. You are adopting an agile iterative approach to implement your Big Data use cases.
 - However, not all data needed for Big Data use cases are identified and accessible.
 - You still need to implement your Big Data platform and integrate it with your existing IT infrastructure.

For more insights, we created a data visualization report for each company. Figure 8 shows company X4 scoring results. We have four appealing visuals for the X4 company:

- The Global score is the sum of maturity domains' scores.
- The table displays scores by domains and their corresponding maturity level.
- The radar chart shows scores by domain
- In the Sankey chart, we have domains with their corresponding maturity levels.

X4 company global score is 91. Based on the description in table I, this score corresponds to the adoption maturity level. When we break it down to maturity levels, X4

company has two domains (governance and methodology) at the Maturity level, two domains (data and technology) at the transformation level and two domains (people and strategy alignment) at the Adoption level.

We showed until now only results for X4 company. We have thirteen Big Data experts who provide us with a correct assessment of their ongoing projects. The following dashboards give us more details about companies' maturity levels and scores. Furthermore, it provides a deep analysis of domains' maturity levels by industry sectors.

We will analyze the companies' maturity levels, especially by industry sector, and define which sector has mature companies

In figure 9, we have an analysis of companies' levels. We are analyzing 13 companies:

- Bar chart showing companies' number by industry sector. Companies are distributed as follows: seven companies in utilities sector, two in technology sector, two in financial sector and two others in transportation sector
- Parallel coordination chart showing companies' global maturity level. X1, X10, X12, X3, X4, X8 and X9 are on the Adoption level. X11 and X6 are on the Transformation level. X2, X5 and X7 are on maturity level and X13 is on Explore level.
- Bar chart showing companies' number having the following global levels: Seven companies are on Adoption level. That is more than half of companies' total number. Then we have three companies on the Maturity level. Two companies on the Transformation level and one company on the Explore level.

In figure 10, we focus on companies' maturity by industry. The tables shows that for financial services and utilities we mostly have a high maturity level; companies are either in Adoption or Maturity levels. After comes the technology sector with companies on Transformation Maturity levels. Then it is the transport sector with companies on Explore and Adoption levels.

In figure 11, we focus this time on companies' scores and create clusters of companies based on their scores.

- Parallel coordination chart showing global scores for 13 companies: X7, X10, X5, X2 have the biggest scores,

GLOBAL BIG DATA MATURITY FRAMEWORK

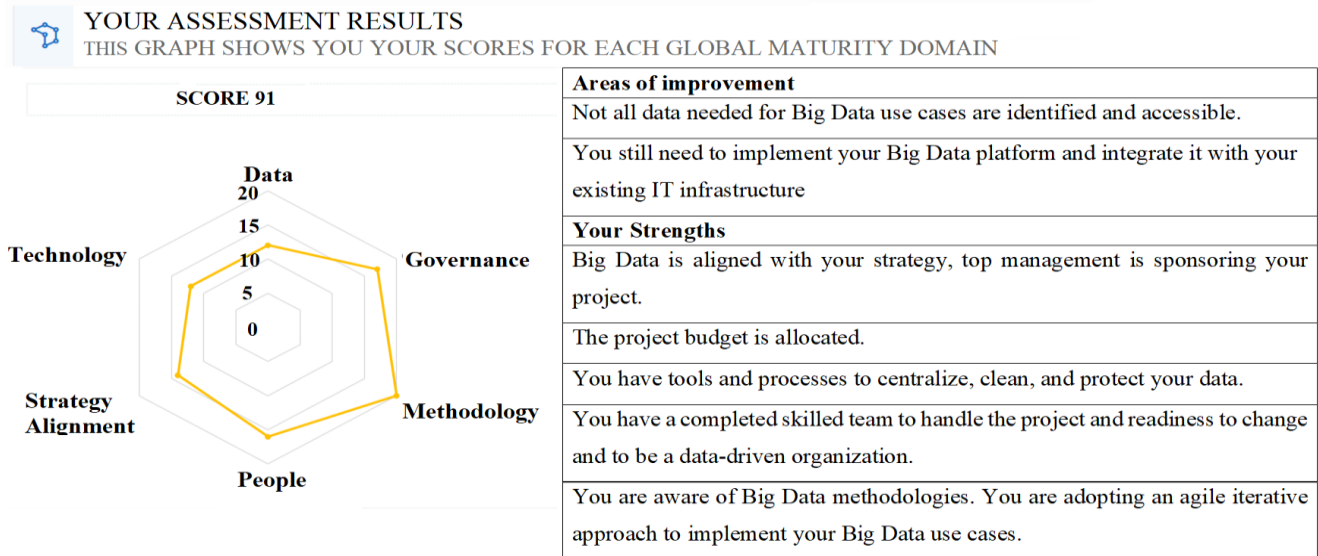


Fig. 7 Assessment framework outputs for X4 Company

with 109.25 on average; X3, X8, X4, X9, X1, X12 are the next six in terms of global score, with 88 on average; X6 and X11 are the next two, with 69 on average. The last one is X13 with a score of 28.

- 2) Scatter chart showing companies' scores clusters. We applied the clustering algorithm, and the results were:
- Cluster1: contains only X13 with scores less than 40.
- Cluster2: X1, X3, X4, X6, X8, X9, X11, X12 with scores between 60 and 100.
- Cluster 3: X2, X2, X5, X10, X7 with scores higher than 100. Now we will explore the most mature maturity domains among North African companies.

In Dashboard figure 12, we drill down to industry sectors. We have sectors where domains are almost equally mature and other sectors with significant differences: For instance, in the transportation sector, all domains are almost equally mature. In the Utilities sector, methodology is the most mature domain, followed by Governance, people, strategy alignment, and then others. In the financial sector, people followed by methodology then data are the most mature domains. In Technology sector, people and methodology are the most mature domains followed by other maturity domains. This shows that from sector to another, companies can be mature on specific domains compared to others.

In this last dashboard figure 13, we have a trellis bar chart, we can tell that Governance, Methodology, and People are the most mature domains within North African companies followed by Data and Strategy Alignment and Technology.

This dashboard confirms the value of the methodology domain in Big Data maturity models. As we can see, only one company is at the Ad-hoc level in this domain, which means that all other companies are already aware of team methodologies or already adopting one. This proves that methodology is a pillar domain within maturity models.

C. Technology choice for implementation

While choosing technologies to implement the Global Maturity Assessment, we also considered the requirements to implement the final global framework. However, we had to make an important choice before listing the platform specifications. Which deployment type to choose? These three deployment types:

On-premise:

When the company has its own data center with all servers and data stored locally.

A public cloud:

When the company uses cloud provider services and data centers to host its data, applications, or even servers.

Hybrid cloud:

This choice is a mix of both on premise and public cloud. The company can keep some workloads in-house and host others in Cloud Provider data centers.

For instance, the Hybrid cloud is very suitable for Disaster Recovery scenarios where companies can duplicate their critical data and applications in cloud storage and failover to the cloud in case of a disaster.

Analytics Domain can also be a great candidate for this deployment. A company can maintain its critical data sources on-premises (for instance, banking data) and use analytics as a service (Paas: Platform as service).

- 1) According to Gartner [37], "by 2022, public cloud services will be essential for 90% of data and analytics innovation".
- 2) The quick provisioning of cloud services makes it very easy to begin any development without spending time on software installation and configuration, and without worrying about the infrastructure behind it.
- 3) Agility is yet another advantage. For example, data storage in a cloud platform can be increased accordingly

with growing data [38]. Memory and CPU metrics also can be scaled up or down.

If we come back to the technical requirements, the target solution should provide:

- 1) Low code and secure development solution to implement the assessment component.
- 2) Capabilities to create forms and charts and integrate dynamic actions natively.
- 3) Flexibility to personalize the look and feel and apply themes and templates.
- 4) Native integration with a database where companies' inputs are stored.
- 5) The database itself should integrate prediction algorithms.
- 6) The database should also integrate ML notebooks to take advantage of ML capabilities.

To meet all these technical requirements, we choose to work with Oracle Apex and Oracle Autonomous Database. Oracle Apex is the most popular low-code development platform for enterprise applications. It enables developers to quickly develop and deploy compelling applications.

Oracle Autonomous database is a cloud database integrated with apex. Users can provision ADW (Autonomous Data Warehouse) in minutes. We can focus only on the framework development because the entire configuration, indexing, patching, and security are there automatically. Moreover, ADW provides machine-learning notebooks, which is user interface when data scientist can collaborate, create and train ML algorithms. More

importantly algorithms are executed directly in ADW data.

V. CONCLUSION AND FUTURE WORK

The goal of this paper was to put at the disposal of companies a Global Maturity Model with a framework to track their overall progress toward Big Data. This global maturity model fulfills the gaps discovered in the current models by adding the methodology dimension and having temporal domains.

When it comes to the framework, we insisted that the assessment should be as simple and exhaustive as possible and the technology to implement the platform should be intuitive and flexible, hence the cloud choice. The cloud allow companies move faster with any implementation, be more agile, and use the most innovative technologies and services. Finally, we talked briefly about the assessment outputs and analysis deeply assessment results. We showed some interesting results, for instance, we explored maturity by industry; when we discovered that among the companies of our study, financial services and utilities have higher maturity level; companies are actually either in Adoption or in Maturity levels. In future work, we will provide a special focus on the score generated. We will use Specific formula to generate the score, based on the maturity domain's weight. The final score finds its roots in multiple criteria decision-making models [39]. We will provide more details on this part in future work since it needs a special assessment and study of weighting and ranking techniques. We will elaborate more about introducing machine learning algorithms in the framework.

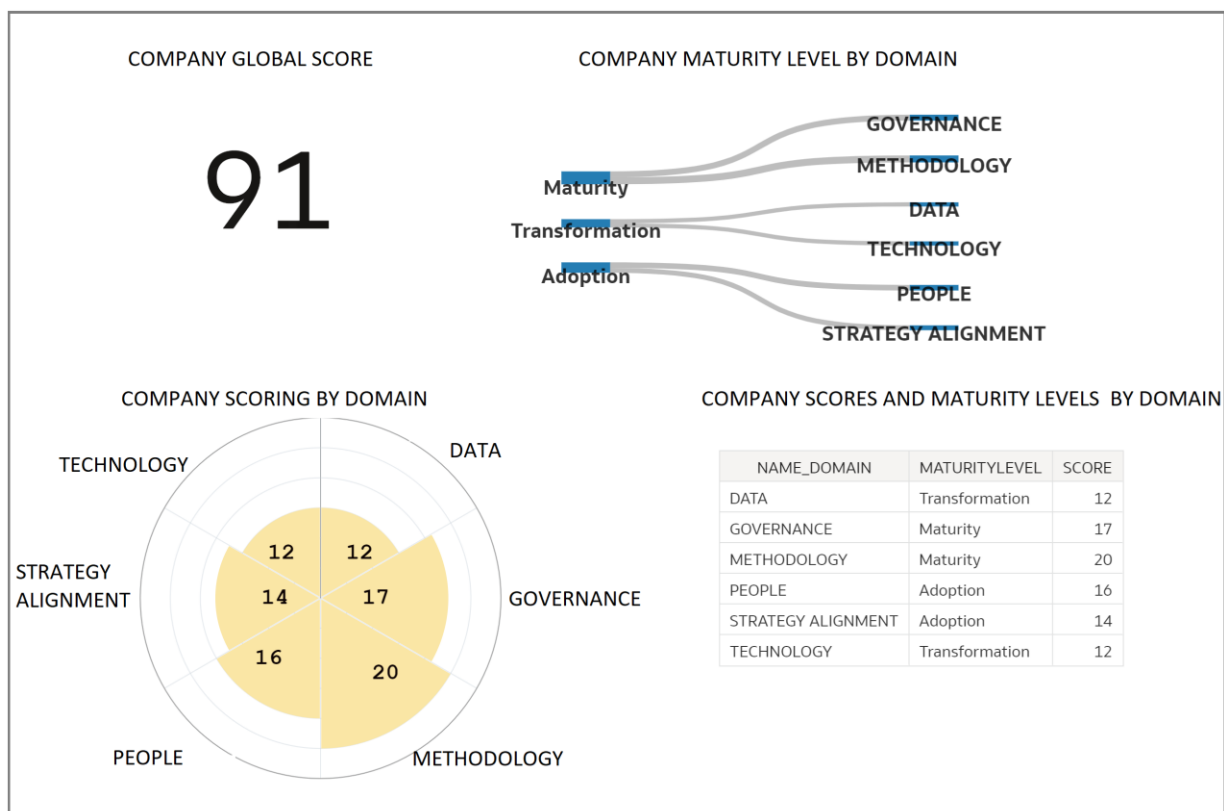


Fig. 8. Assessment framework outputs for X4 Company

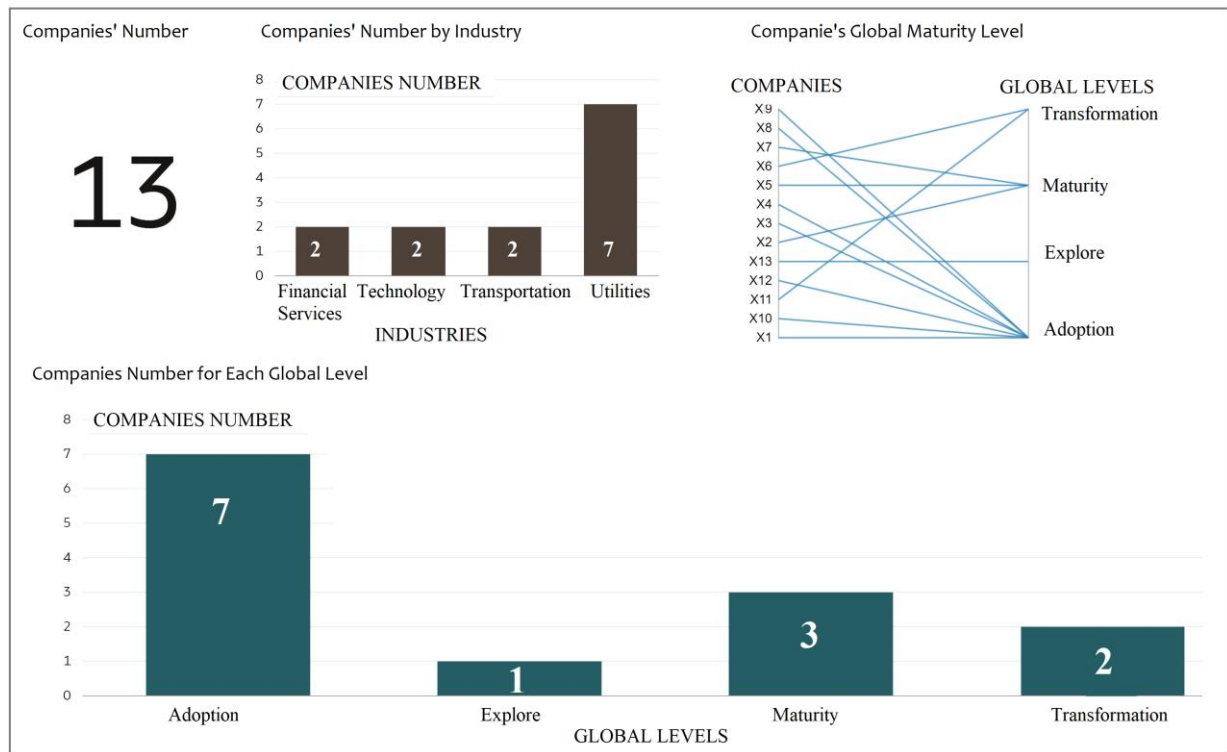


Fig. 9. Analysis of companies' maturity levels

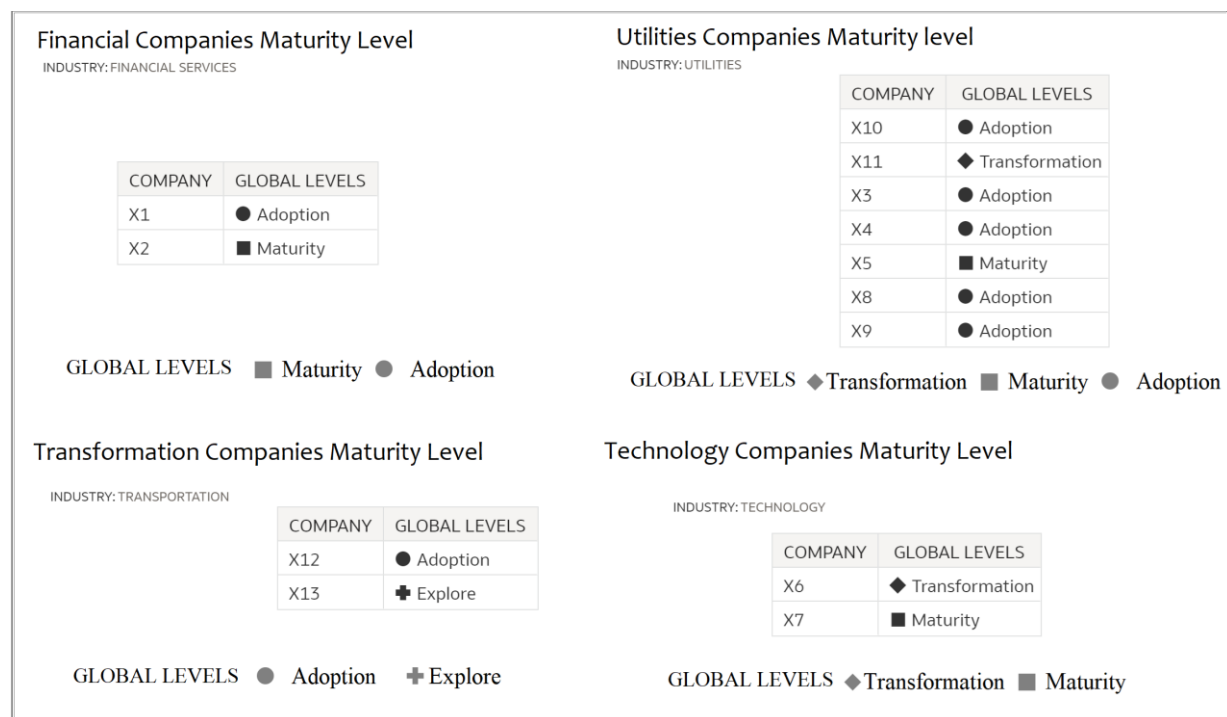


Fig. 10. Analysis of companies' maturity by industry

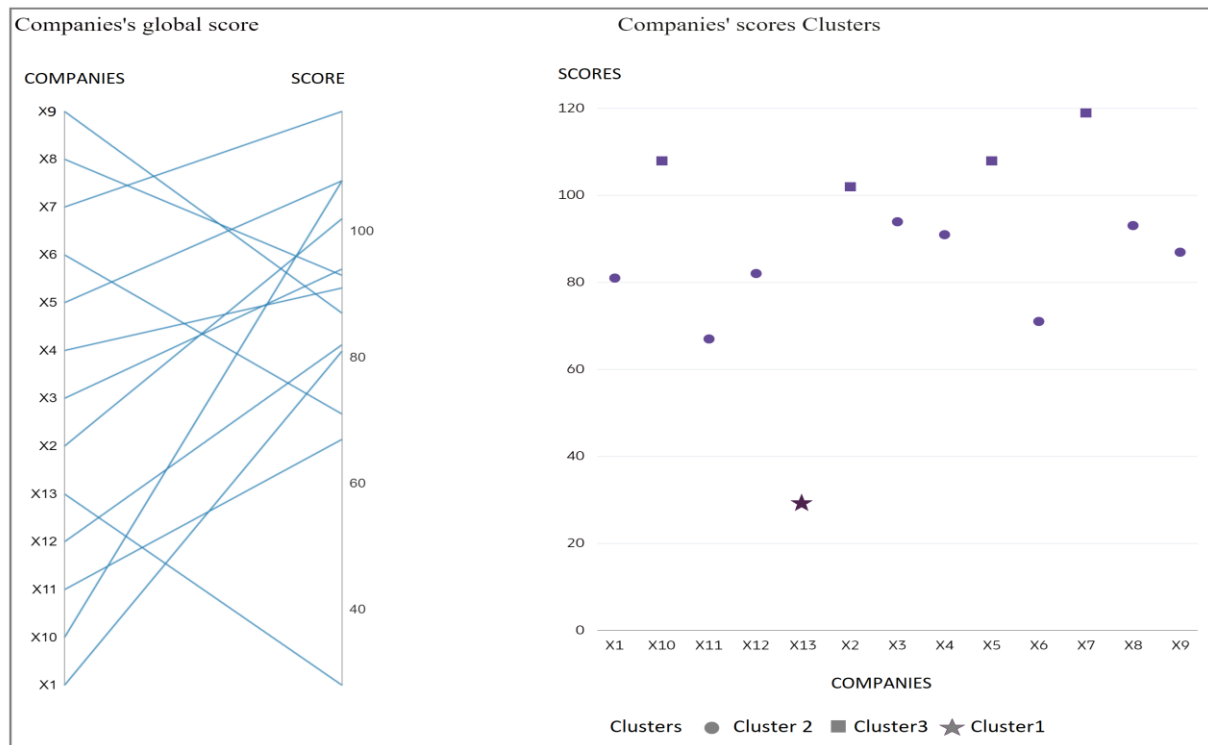


Fig.11. Analysis of companies 'maturity scores

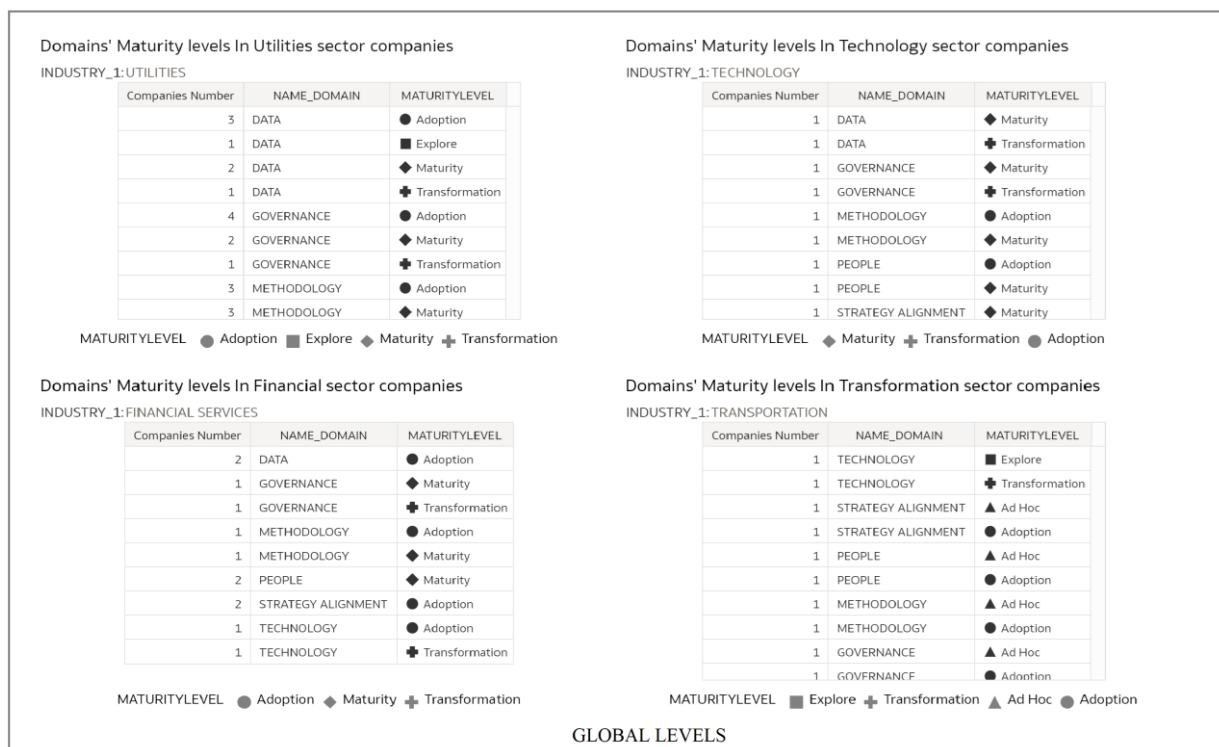


Fig.12. Domains maturity by companies' scores

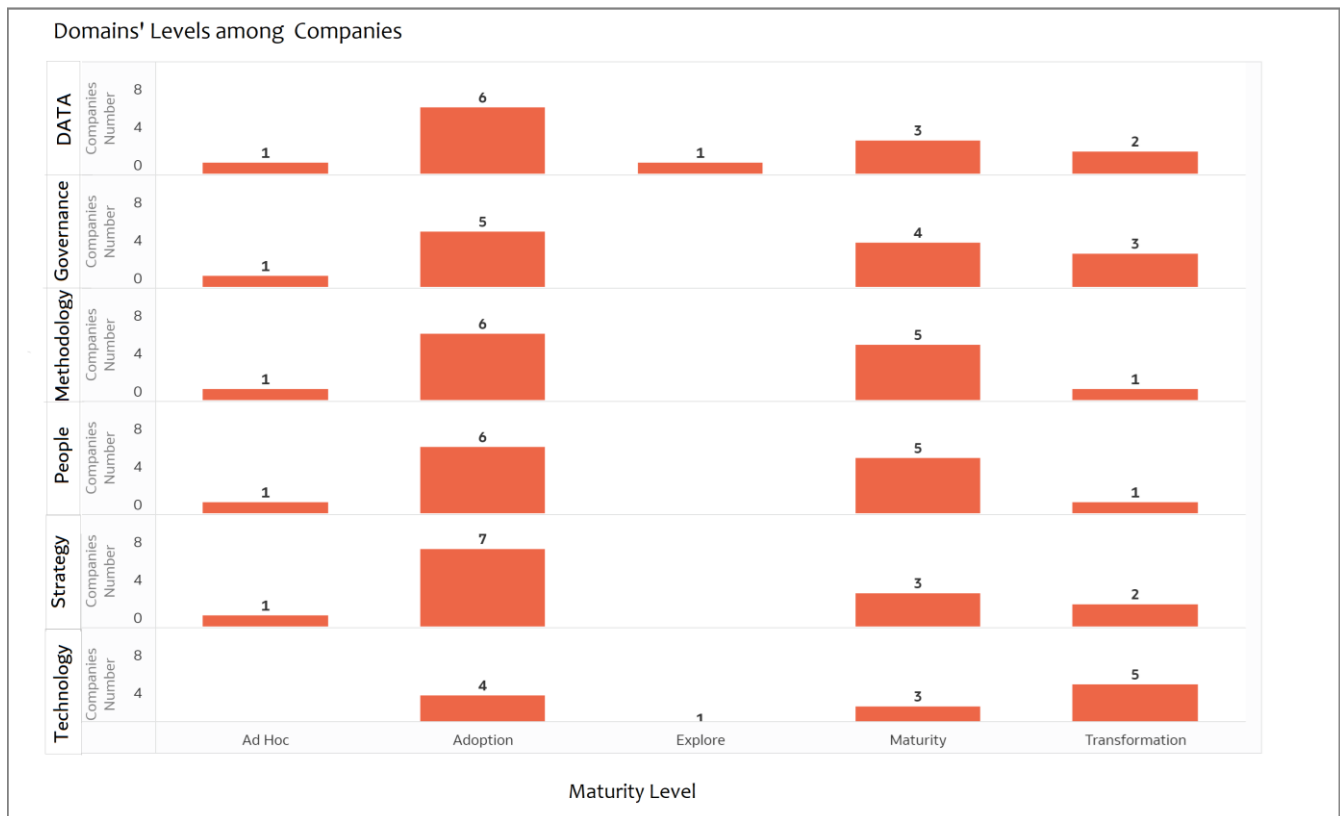


Fig.13. Analysis maturity domains among companies

REFERENCES

- [1] K. ElDahshan, E. K. Elsayed, and H. Mancy, "Enhancement Semantic Prediction Big Data Method for COVID-19: Onto-NoSQL", *IAENG International Journal of Computer Science*, vol. 47, no.4, pp613-622, 2020
- [2] S. J. Alsunaidi, al. M. Almuhaideb, N. M. Ibrahim, F. S. Shaikh, K. S. Alqudaihi, F. A. Alhaidari, I. Ullah Khan, N. Aslam, M. S. Alshahrani "Applications of Big Data Analytics to Control COVID-19 Pandemic," *Sensors*, vol. 21, no.7, pp2282-2306, 2021
- [3] N. L. Bragazzi, H. Dai, G. Damiani, M. Behzadifar, M. Martini, and J. Wu, "How Big Data and Artificial Intelligence Can Help Better Manage the COVID-19 Pandemic," *International Journal of Environmental Research and Public Health*, vol. 17, no. 9, pp 3176-3184, 2020
- [4] C.-H. Huang, K.-C. Yang, and H.-Y. Kao, "A Cloud Business Intelligence System for Visual Analytics with Big Data," *Proceedings of the International MultiConference of Engineers and Computer Scientists 2019*, 13-15 March , 2019, Hong Kong , pp262-266
- [5] C. Costa, M. Y. Santos, "Big Data: State-of-the-art Concepts, Techniques, Technologies, Modeling Approaches and Research Challenges," *IAENG International Journal of Computer Science*, vol. 44, no.3, pp285-301, 2017
- [6] J. Cao, M. Lin, X. Ma, "A Survey of Big Data for IoT in Cloud Computing," *IAENG International Journal of Computer Science*, vol. 47, no.3, pp585-592, 2020
- [7] B. A. says, "Our Top Data and Analytics Predicts for 2021," Andrew White, Jan. 12, 2021. https://blogs.gartner.com/andrew_white/2021/01/12/our-top-data-and-analytics-predicts-for-2021 (accessed Sep. 15, 2022).
- [8] S. Mouhib, H. Anoun, M. Ridouani, and L. Hassouni, "Towards a Global Big Data Maturity Model," *Fourth International Conference On Intelligent Computing in Data Sciences 2020*, 21-23 October 2020, Fez, Morocco, pp1-5
- [9] M. Comuzzi and A. Patel, "How organisations leverage Big Data: a maturity model," *Industrial Management & Data Systems*, vol. 116, no. 8, pp1468-1492, 2016
- [10] M. Soukaina, H. Anoun, M. Ridouani, and L. Hassouni, "A study of the factors and methodologies to drive successfully a big data project," *Third International Conference on Intelligent Computing in Data Sciences 2019*, 28-30 Oct 2019, Marrakech, Morocco pp1-6
- [11] J. S. Saltz and I. Shamshurin, "Big data team process methodologies: A literature review and the identification of key factors for a project's success," *IEEE International Conference on Big Data (Big Data)*, 2016, 05-08 December 2016 Washington DC, USA, pp2872-2879.
- [12] S. Eybers and M. J. Hattingh, "Critical success factor categories for big data: A preliminary analysis of the current academic landscape," *IST-Africa Week Conference (IST-Africa) 2017*, 30 May- 02June 2017, Windhoek, Namibia, pp1-11
- [13] W. Yeoh and A. Popović, "Extending the understanding of critical success factors for implementing business intelligence systems," *Journal of the Association for Information Science and Technology*, vol. 67, no. 1, pp134-147, 2016
- [14] C. Adrian, R. Abdullah, R. Atan, and Y. Y. Jusoh, "Factors influencing to the implementation success of big data analytics: A systematic literature review," *International Conference on Research and Innovation in Information Systems (ICRIIS) 2017*, 16-17 July 2017, Langkawi, Malaysia pp1-6
- [15] S. Sun, C. G. Cegielski, L. Jia, and D. J. Hall, "Understanding the Factors Affecting the Organizational Adoption of Big Data," *Journal of Computer Information Systems*, vol. 58, no. 3, pp193-203 , 2018
- [16] M. Janssen, H. V. D Voort, A. Wahyudi "Factors influencing big data decision-making quality - ScienceDirect," *Journal of Business Research*, vol. 70, pp338-345, 2017
- [17] M. Mach-Król, "Conceptual Foundations for the Temporal Big Data Analytics (TBDA) Implementation Methodology in Organizations," *Towards Industry 4.0 — Current Challenges in Information Systems*, vol. 887, M. Hernes, A. Rot, and D. Jelonek, Eds. Cham: Springer International Publishing, pp235-247, 2020
- [18] F. Halper, D. Stodder "TDWI Analytics Maturity Model Guide: Interpreting Your Assessment Score," *TDWI*, pp1-22, 2020
- [19] "Big Data Business Model Maturity Index | EMC." <https://www.kdnuggets.com/2016/06/big-data-business-model-maturity-index-iot.html> (accessed Dec 29, 2022).
- [20] "IDC Future IT Maturity Assessment" <https://www.idc.com/itexecutive/planning-guides/maturity-assessment> (accessed Dec 29, 2022).
- [21] Veenstra, A.F.E. van, Bakker, T.P., and Esmeijer, J., "Big data in small steps" 2013, <https://repository.tno.nl/islandora/object/uuid:6e0e4f90-13ce-4069-a365-5c787518270a> (accessed Dec 29, 2022).
- [22] "Assess Big Data Readiness by Infotech." <https://www.infotech.com/research/leverage-big-data-by-starting->

- [small-phase-2-assess-big-data-readiness?initiate_downloads=true](#) (accessed Dec 29, 2022).
- [23] "Levels of Big Data Maturity Knowledgent." <https://www.smartdatacollective.com/five-levels-big-data-maturity-organisation> (accessed Dec 29, 2022).
- [24] B. El-Darwiche, V. Koch, D. Meer, R. T. Shehadi, and W. Tohme, "Big Data Maturity: An Action Plan for Policymakers and Executives," The Global Information Technology Report 2014 pp 43-50, 2014.
- [25] H. T. Braun, "Evaluation of Big Data Maturity Models – A Benchmarking Study to Support Big Data Maturity Assessment in Organizations," 2015, Tampere University of Technology, pp1-119 <https://trepo.tuni.fi/handle/123456789/23016> (accessed Dec 29, 2022).
- [26] "Big Data & Analytics Maturity Model | IBM Big Data & Analytics Hub." <https://www.lightsondata.com/data-governance-maturity-models-ibm/> (accessed Dec 29, 2022).
- [27] H. Sulaiman, Z. C. Cob, and N. Ali, "Big data maturity model for Malaysian zakat institutions to embark on big data initiatives," 4th International Conference on Software Engineering and Computer Systems (ICSECS) 2015, 19-21 August 2015, Kuantan, Malaysia, pp 61–66
- [28] "HORTONWORKS BIG DATA MATURITY MODEL." <https://infotech.report/whitepapers/hortonworks-big-data-maturity-model> (accessed Dec 29, 2022).
- [29] C. Adrian, R. Abdullah, R. Atan, and Y. Y. Jusoh, "Towards Developing Strategic Assessment Model for Big Data Implementation: A Systematic Literature Review," Int. J. Advance Soft Compu. Appl, vol. 8, no.3, pp 174-190, 2016
- [30] B. Farah, "A Value Based Big Data Maturity Model," JOURNAL OF MANAGEMENT POLICY AND PRACTICE, vol. 18, pp. 11–18, 2017
- [31] "Maturity Model from Nott and Betteridge (IBM)" ResearchGate. https://www.researchgate.net/figure/The-IBM-Stampede-method-source-28-Table-1-Maturity-Model-from-Nott-and-Betteridge_fig3_321944704 (accessed Dec 29, 2022).
- [32] C. Ponsard, A. Majchrowski, S. Mouton, M. Touzani "Process Guidance for the Successful Deployment of a Big Data Project: Lessons Learned from Industrial Cases" Proceedings of the 2nd International Conference on Internet of Things, Big Data and Security 2022, pp350-355
- [33] J. S. Saltz, "CRISP-DM for Data Science: Strengths, Weaknesses and Potential Next Steps," IEEE International Conference on Big Data (Big Data) 2021, 15-18 December 2021, Orlando, FL, USA, pp 2337–2344
- [34] H. M. Chen, R. Kazman, S. Haziyeve, "Agile Big Data Analytics Development: An Architecture-Centric Approach" 49th Hawaii International Conference on System Sciences (HICSS) 2016, 05-08 January 2016, Koloa, HI, USA, pp5378-5387
- [35] N. W. Grady, "KDD meets Big Data" IEEE International Conference on Big Data (Big Data) 2016, 05-08 December 2016, Washington, DC, USA, pp1603-1608
- [36] C. Ponsard, A. Majchrowski, S. Mouton, and M. Touzani, "Process Guidance for the Successful Deployment of a Big Data Project: Lessons Learned from Industrial Cases" Proceedings of the 2nd International Conference on Internet of Things, Big Data and Security, 2017, Porto, Portugal, pp350–355
- [37] "Gartner Identifies Top 10 Data and Analytics Technology Trends for 2020," Gartner. <https://www.gartner.com/en/newsroom/press-releases/2020-06-22-gartner-identifies-top-10-data-and-analytics-technolo> (accessed December 29, 2022).
- [38] S. Velayutham, Ed., "Challenges and Opportunities for the Convergence of IoT, Big Data, and Cloud Computing" IGI Global, pp1-381, 2021
- [39] J. S. Dyer, P. C. Fishburn, R. E. Steuer, J. Wallenius, and S. Zionts, "Multiple Criteria Decision Making, Multiattribute Utility Theory: The Next Ten Years," Management Science, vol. 38, no. 5, pp645–654, 1992

Soukaina MOUHIB is a doctoral student at Hassan II University, Morocco. She received an engineering degree in software engineering from Institut National des Postes et Telecommunication (INPT). She worked in Analytics and Big Data domain for more than ten years with hundreds of customers from Europe, the Middle East, and Africa. She is now Cloud Adoption Manager. Her research interest is Big Data frameworks and methodologies.

Pr. Houda ANOUN is a professor in the Department of Computer Science at Ecole Supérieure de Technologie (Hassan II University), where she has been since 2009. She received her engineering degree in Software Engineering from ENSEIRB Bordeaux in 2003 and her Ph.D. in computational linguistics from Bordeaux I University in 2007. Her current research interest lies in the area of artificial intelligence especially machine learning, deep learning, and Big Data.

Mohammed RIDOUANI received an M.Sc in Telecommunications from the Faculty of Sciences and Techniques (Fes, Morocco) in 2003. He received his degree of engineer "diplôme d'ingénieur d'état" in telecommunications from the National Institute of Post and Telecommunications (Rabat, Morocco) in 2005. He obtained his Ph.D. degree in Applied Mathematics, Computers sciences, and Telecommunication at the Hassan I University (Settat, Morocco) in 2016. He has been an Engineer-Teacher at the Computer Sciences department in the High School of Technology at Hassan II University (Casablanca, Morocco) during 2006-2018, and he was appointed as a Professor in 2018. He was certified on CISCO in 2009 and Linux LPIC-1, LPIC-2, and LPIC-3 (senior/expert level) in 2008, 2012 and 2017 respectively. He is an LPI and Cisco instructor and he has served as an international instructor at the Francophone University Agency. He serves as a reviewer for prestigious international journals, and TPC member roles in many international conferences (IEEE GLOBECOM, IEEE PIMRC, IEEE VTC2018-Spring, IEEE ISWCS, IEEE IWCM, IEEE EuCNC & 6G Summit - 6ET, IEEE 5G-WF, ...etc.). He was a recipient of the 2021 Best Paper Award of the International Conference on Smart Systems and Data Science 2021 (ICSSD'21). His research interests are in the physical and the MAC layers of design-based wireless communications systems and mathematical modeling, including cognitive radio, resource allocation strategies, compressed sensing, and cooperative. In addition, his current research interests are data sciences and smart city, especially machine learning applications for healthcare and Telecommunication, and Big Data.

Pr. Larbi Hassouni got his engineer degree in 1983 from the "École Centrale de Marseille". He prepared for his Ph.D. degree in 1987 at the University of Aix Marseille III in France. Among his research work, there is the development of a list unification software with the LeLisp language of INRIA to contribute to the realization of the inference engine of an expert system for the digital circuits' diagnosis. He also developed a behavioral symbolic simulator of digital circuits using the C, FRL, and LeLisp languages to contribute to the development of a "formal proof" tool for correcting a design of material produced by a Hardware Design Language (HDL). Currently, his research work mainly concerns Data Sciences.