Big Data, Hadoop and Spark for Employability: Proposal Architecture

Aniss Qostal[1][a], Aniss Moumen[2][b] and Younes Lakhrissi[1][c]
[1]*Intelligent Systems, Georesources and Renewable Energies Laboratory, Sidi Mohamed Ben Abdellah University, FST Fez, Fez, Morocco*
[2]*Laboratory of Engineering Sciences, National School of Applied Sciences, Ibn Tofaïl University, Kenitra, Morocco*

Keywords: Big Data, Data Analytics, Hadoop, Spark, Employability.

Abstract: The application of big data and data analytics has reached all aspects of life, from entertainment to scientific research and commercial production. Mainly by taking advantage of the explosion of data at an unprecedented rate attain the level of exabytes per day. On the other hand, it benefits from the sophisticated analytics approaches that have been given new manners to translate the raw data into solutions and even into predictions for complicated situations. This paper aims to discover the application of big data, data analytics and technical architectures based on the Hadoop and Spark ecosystems to build employability solutions. Beginning with a literature review of previous works and proposed solutions to draw a roadmap towards new approaches and enhance the recruitment process for youth people.

1 INTRODUCTION

Today, employability is considered as a significant concern for different entities involved in the elaboration of the Labour Market Policies (LMP)(Caliendo, 2016), with the primary goal to enhance the employment opportunities and facilitate the integration of job seekers and especially the youth generation in the workplace; by improving the matching process between job offers (vacancies) and job seekers (i.e. the unemployed). Furthermore, the basic philosophy behind the LMP is to supervise and monitor the global atmosphere of the employment environment and guaranty the necessary conditions for good wellness in the workplace and ensure the development of the workforce for the future competencies required. Currently, according to the reports of the international labour organization ILO in 2020, the global rate of unemployment has shown a slight increase of 6.1% compared to 5.37% in 2019, mainly due to the COVID-19 pandemic. The same report has indicated the youth generation as the most suffering layer with a rate of unemployment reaching the highest level during a decade with 8.7% (Ryder and Director-General, 2020).

On the other hand, we found many disciplines, each in their domain of competence, try to find and give their vision on the possible solutions for the employability problems. Therefore, one of the prominent and promising fields under investigation is the use of the big data concept on employability and the marketplace. Firstly as a way to treat data coming from various sources related to employability. Secondly, to support and enhance projects related to the development of the labour market and facilitate the integration of youth people.

In this paper, we will conduct an exploratory study on the research carried out on the application of big data for employment. Firstly, to discover fundamental approaches developed on the employability context, focusing on technical architectures proposed based on Hadoop, Spark and analytics with the aim to find solutions, propositions and limitations and finally give answers to questions like:

[a] https://orcid.org/0000-0003-1016-8467
[b] https://orcid.org/0000-0001-5330-0136
[c] https://orcid.org/0000-0002-2537-8283

Qostal, A., Moumen, A. and Lakhrissi, Y.
Big Data, Hadoop and Spark for Employability: Proposal Architecture.
DOI: 10.5220/0010732200003101
In *Proceedings of the 2nd International Conference on Big Data, Modelling and Machine Learning (BML 2021)*, pages 260-266
ISBN: 978-989-758-559-3

1) Can big data provide employment Solutions?
2) Are there any approaches, initiatives or projects launched to exploit big data in the employability context?
3) Are there any initiatives in the developing countries, especially the Moroccan context?

The results obtained from the previous researches can be the base for proposal architecture and roadmap for further studies.

2 MANUSCRIPT PREPARATION

2.1 Context

The application of big data in the employability context benefited from the digitalization happened last decades and from the growth of data generated from different sources related directly or indirectly to the labour market, including all types of e-services(Qostal et al., 2020). They are producing and storing massive and heterogeneous data containing personal and professional information. Accordingly, this data can be exploited and feed any big data system built to understand variables and factors with their weights and their impacts on the workplace using different data mining techniques(Piad et al., 2016)(Mishra et al., 2017).

Therefore, the big data concept(De Mauro et al., 2016), as defined by De Mauro, Greco, & Grimaldi "information assets characterized by such a high volume, velocity and variety that they require specific technology and analytical methods for its transformation into value". The fundamental definition of big data is based on the 3Vs characteristic:

- Volume: refers to the size of data generated. Currently, a study conducted by IDC showed that in 2020 the data created, captured, copied, and consumed has reached 59 (ZB). Furthermore, there are 69444 users creating information in the professional world and applying for jobs on the LinkedIn platform for each minute. Based on information from Data Never Sleeps 8.0.
- Variety: Generally, the data forming the big data systems come from different sources, which means different types and varieties of data, generally can be grouped into three types: structured, semi-structured, and unstructured. Today, 90% of this data is semi-structured and unstructured, varying from text, video, audio metadata (data about data).
- Velocity: The speed at which data is generated on the web varies from platform to platform, requiring appropriate strategies to manage data as it enters the big data system. Generally, the applied approaches vary from batch processing to stream processing.

Accordingly, for the labour market context, the 3Vs dimensions benefitted from the emergence of SNS, e-learning, e-recruitment. They were giving the base for the application of different types of analytics in the employment context (Mishra et al., n.d.) (Saouabi and Ezzati, 2019) (descriptive, diagnostic, predictive and prescriptive).

2.2 Big Data for Labour Market Intelligence System (LMIS)

Previously, collecting and treating data related to the labour market was done through the traditional Labour Market Information (LMI), where the main sources of data for LMI were based essentially on surveys and administrative records from public employment agencies to construct the information about the labour market and help decision-makers to analyze and take actions on the LMP process (Sorenson and Mas, n.d.). Although, the traditional LMI suffered from low accuracy, outdated classifications techniques and the lack of the ability to turn labour market information into intelligent decisions in a short time (Johnson, 2016). To overcome those difficulties especially with the era of real-time data generation with high heterogeneity and velocity, an alternative represented with the real-time LMI approach is developed to support data coming from web platforms and using solutions provided with the big data ecosystem.

Furthermore, the analytic part of the data collected from traditional and real-time LMI is handled by projects grouped under a big field of research known as the Big Data for Labour Market Intelligence System (LMIS) (Mezzanzanica and Mercorio, 2018) (table1). Mainly, with the initiative of projects launched by the European Centre for the Development of Vocational Training (Cedefop) agency since 2016 with the aim to support the use of big data into the employment context and fellow the undergoing process where the combination of the big data ecosystem and the computation techniques including machine learning algorithms is applied to analyze labour market.

Table 1: Some projects supported for LMIS context.

Study/Project	Description	Methodology	Results
WoLMIS :(Boselli et al., 2018)	A system developed for the Cedefop Agency; with the aim to create a KDD and LMIS on the European labour market by using a data coming from European e-recruitment platforms	Using scarping techniques on e-recruitment platforms of the countries involved in the project (at first stage: 6 countries) using the batch mode for the ingestion phase (once per week over 4 months) to loads the outcomes of this step into a data warehouse. Classify each job vacancy with respect to the International Standard Classification of Occupations (ISCO) taxonomy. Using supervised learning approach as machine learning techniques with three algorithms to classify the job vacancies based on SVMs, ANN Artificial Neural Networks and Random Forests algorithms	Evaluation on n-grams extracted both from titles and full description windows best results was with SVM Linear: Precision 0.910, Recall 0.909, F1-Score 0.908, Accuracy 0.909.
DataOps for Societal Intelligence(Tamburri et al., 2020)	Based on the DataOps pipeline in the context of project Werkinzicht, author propose as LMIS system using big data as base to create KDD for job seekers of the Dutch-Flemish labour market	The dataset was provided by UWV agency containing millions of vacancies posts from year 2014, the training and validation steps was on 10K vacancies and the validation was under supervision of experts from UWV. DataOps pipeline combining Big Data and Machine Learning models to process vacancies and resumes, Six domain experts from UWV annotated a total of 3K sentences – around 500 sentences per person The baseline pre-trained model used both for BERT Cloud and BERT Local for text classification based on ISCO and ESCO taxonomy.	The best performing model in this case is BERT Cloud, with an accuracy of 0.816 after 210 training steps (epochs), followed by BERT Local (0.8) and Simple-NN (0.78)

The first results of the implementation of Cedefop projects over the Union European were published in 2019 with the participation of 16 Member States (Austria, Belgium, Czech Republic, Denmark, Germany, Hungary, Spain, Finland, France, Italy, Ireland, Luxembourg, the Netherlands, Poland, Portugal, Sweden, and Slovakia). Otherwise, since 2019 the European Training Foundation ETF launched imitative to enhance the use of LMIS within NON-EU countries such as Morocco and Tunisia to analysis the web labour markets and job vacancy (OJV) portals(Vaccarino et al., n.d.). Therefore, an initiative launched as sa result of the cooperation between the Moroccan government and Millennium Challenge Corporation (MCC) since 2015, via the MCA-Morocco in October 2020, a competition to set up a digital labour market information platform based on artificial intelligence and big data.

2.3 Big Data Architectures for the Employment

The solutions proposed to deal with the employability must take into consideration the identifications process of variables with their complexity and their impact in the labour market, the stakeholder analysis process(job providers, job seekers, etc.) with their actions and reactions towards the workplace, and the complexity of data generated related to employability with the determination of different data sources. Secondly, develop an advanced analytics method to process the information collected from LMIS to discover hidden patterns, extract relevant features, and propose new insights to make appropriate decisions. Accordingly, ETF proposed architectures based on big data ecosystem, including Hadoop and Spark solutions (Figure 1), to guarantee scalability, distribute tasks, and store data over Multi-node architectures with the possibility of using advanced analytics approaches.

The solutions proposed for LMIS have a primary objective to overcome the previous challenges and create the Knowledge Discovery in Data KDD related to employability (Fayyad et al., 1996).

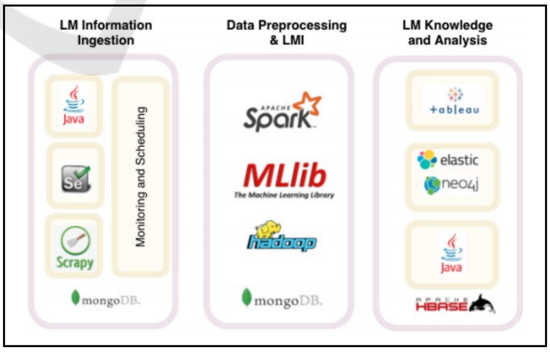

Figure 1: an overview of LMIS based on Hadoop & spark solutions proposed by ETF (Vaccarino et al., n.d.).

The proposition adopted by ETF (figure 1), divided the implementation of big data for the LMIS into three significant steps using the Hadoop ecosystem and apache spark.

2.3.1 LM Information Ingestion

The ingestion process is considered as a vital step with the aim to integrate valuable information into the proposed big data system (Meehan et al., n.d.). The process is about transferring data from various data sources (NoSQL stores, web pages using scraping and crawling techniques, databases, SaaS data, etc.) into the distributed storage unit of the system (data warehouse, database, data mart or filesystem), usually in the Hadoop ecosystem is represented with the Hadoop Distributed File System HDFS, or throw other propositions of NoSQL databases (e.g. MongoDB, HBASE). Mainly, we can group the data sources related to the labour market into four types:

- Employment websites: regroup all e-recruitment platforms and professional social networks like LinkedIn, where information can be gathered for job providers and seeker's profiles, resumes, experiences and responsibilities.
- Education and training platforms: represented by educational platforms, including e-learning and MOOCs like Udemy, Coursera, GitHub, Slack and Sourceforge. Those platforms provide valuable information of training process, skills acquired, performance and contributions on the share of technical knowledge. The information collected can be a way to measure the quality of skills within a community.
- Social media platforms: Such as Facebook and Twitter, many studies have shown the vital information can be extracted, including indicators on personality traits using the big five model using NLP and sentimental analysis to decide on the recruitment process (Satyaki Sanyal et al., 2017).
- Government e-services: with the era of open data adopted by the governments, the official data can be gathered on the employability context and the development of the labour market to identify the weak points and skills gaps to build a global strategy and LMP.

This step raises big challenges with the high velocity and large volume of data generated, requiring real-time approaches, high fault tolerance, and simultaneous multi-connection from variant sources. The ingestion process from those sources can be made in two possible ways:

- Batch ingestion: considered as the most used method. The data collection operation from the sources is applied periodically within a predefined time interval. The popularity of this method comes from the optimization that can be made on the speed and quality of data transferred from the source in a controlled and efficient manner.
- Real-Time Streaming: The data ingestion is made in real-time from sources as soon as data is available and recognized by the data ingestion layer without respecting a predefined schedule. The loading in real-time provides valuable information to monitor variables and indicators on critical situations where the need for fast decisions.

In the Hadoop ecosystem, many solutions are offered as an interface and broker between sources and destinations, such as KAFKA, FLUME, and SQOOP (Mccaffrey, 2020).

2.3.2 Data Preprocessing & LM

The data collected and retrieved from different sources related to LM are usually in raw and noisy forms, where the step of data preprocessing is about the preparation and transformation of this data into a suitable form with high-quality values to increase the predictions and accuracy of LMIS and avoid the unclean information which will lead to wrong results "garbage in garbage out" (García et al., 2016) (He et al., 2010). Fundamentally, the major task of this phase is to deal with duplicated data such as duplicated profiles on different platforms, missing values like competence, skill and diplomas information, treat outliers, encoding categorical value, features selection and features extraction before deciding to deploy a machine learning solution (García et al., 2016):

- Noise reduction: the data retrieved from platforms such as social networks contain information that is usually unrelated to professional activities such as emojis, hashtags, URLs. That information can be considered irrelevant for the analytics step; by removing this data, accuracy and precision can be improved.
- Merging process: this step aims to merge duplicated information from different sources such as vacancy announcements on different platforms, profiles of seekers, and statistics from various sources published on OJVs. This step will give the actual situation of the needs and offers in the workplace in terms of profiles and skills.
- Normalization and standardization: the objective of this step is to unify values and information within a norm of measure, especially with the data related to

employability become very large and the fields of studies become verb y diversified, pushing to have unified norms of skills, job description, occupation and competence. In LMIS, there are many norms, Occupational Information Network (O*NET), ESCO and ISCO.

2.3.3 LM Knowledge and Analysis

The LM knowledge and analysis aims to use Machine learning techniques on the data processing to generate the KDD about labour market performance, help in the talent management process, give recommendations to stakeholders with answers about the quality of skills, professions required, sectors needed to be improved. Various analytics approaches target employability in different stages using big data. Many application beside LMIS has objective to enhance employability such as Vocational Education and Training Systems (TVET) (S. M. S. Azman Shah et al., n.d.), HRIS systems(Kaygusuz et al., 2016) (Shahbaz et al., 2019)where the objective use the people analytics approach(Isson and Harriott, n.d.; Shrivastava et al., 2018) to answers questions on employee behavior, workforce trends, and performance at the job as a way of hiring and hunting talents and predict turnover probabilities (Tamburri et al., 2020). Mainly, text classification, natural language processing and texts Job Posting Analytics JPA are widely used as a way to builds model capable to help us predict on match capabilities to market needs (table 2), identify valuable talent like never before, match capabilities to market needs. In the applicative context, many solutions proposed the application of the mapreduce where the mapper and reducer function on hadoop context or pyspark on the spark context can be used as a way on the analyze of resume, and as results is the helps in the process of LMP throw different models of recommendations (Collaborative, user or hybrid recommendation)(Li et al., 2017).

Table 2: Some studies and research for JPAs supported for LMIS context.

Study	Methodology/Sample	Results
an hybrid recommendation system to improve the job search process, based on the theory of graphs(nodes/edges) by exploiting several relevant values provided form job seekers profiles(Shalaby et al., 2017).	The introduction of homogeneous graph-based recommendation architecture (GBR) based on graph theory. Using a real data set (jobs and users) from CareerBuilder.com on a sample of 500,000 jobs. Using the Apache Spark GraphX framework. Job data is stored in Hive databases, and SparkSQL.	Results show an improvement of jobs obtained via different combination of these parameters reaching 90% of accuracy.
An optimal algorithm was developed to predict and give recommendation concerning employment in the big data scenarios (S. Dong et al., 2017).	A Contextual Online Learning Approach was developed (such as the preference or other characteristics in registration information and browse records). dissimilarity function $\mathscr{L}$ which can indicate the similarity of different items contexts are similar to handle the diversification of user personality. Using as a sample database from Work4 data(100 million instances) contains the candidates profile information based on Facebook and Linkedin.	The proposed algorithm ITFACP compared to algorithm(ACR, UCB, HCT and OLCR) perform better when the number of instances exceed 500,000 the regret reach 52% and 78% of the OLCR and ACR algorithm.

3 PROPOSED ARCHITECTURE

As explained in the previous sections, the objective is to adopt architecture capable of exploiting and processing the fast-growing data generated over electronic platforms such as SNS, e-government services, e-learning systems of universities, employment platforms, and any potential platform that could generate information related to youth concerns. Accordingly, many initiatives and projects are trying to create a KDD oriented to the employability context using Big Data. Mainly, our approach tries to combine the process adopted and showed in Figure 2 to create KDD and the implementation using the Hadoop ecosystem and spark with their components:

- Hadoop Ecosystem: Includes packages and subcomponents to manage Big Data throw all stages of processing of data and support building models that address youth concerns. The storage part of data can be managed using HDFS where can be dispatched over clusters as a way of fault tolerance such as HBase and MongoDB. For the analytics, parts can be handled directly throw MapReduce as a way of data processing on large clusters or throws other alternatives where the functionality of Mapper and Reducer can be done throw query language, for example, Hive and Pig.
- Spark Ecosystem: Apache Spark is an open-source alternative to Hadoop and MapReduce, where the computation is 100 times faster. The use of Spark within the step of preprocessing data repose on the RDD (Resilient Distributed Datasets) to adapt data for the processing and machine learning to develop recommendations, predictive insights and personalized, where Spark

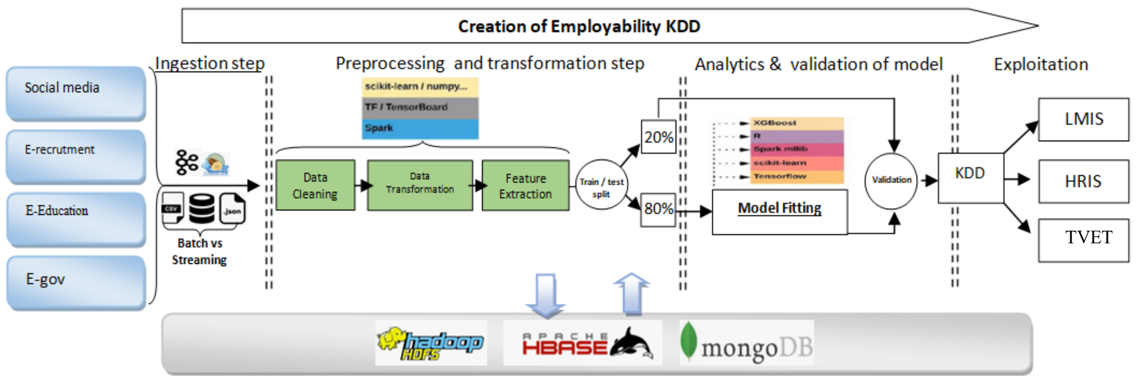

Figure 2: the proposed architecture based on Hadoop & spark and the KDD model (Fayyad et al., 1996).

is equipped MLlib library includes a package aimed to address these concerns. Combining the Hadoop ecosystem and Spark on the management and processing layer will improve and enhance data processing and provide real-time analysis of youth data. The model (Figure 2) will help to identify hidden patterns and potential solutions in many areas, such as education and employment(He et al., 2010), offering new opportunities to improve policy design for the youth generation and LMP.

4 CONCLUSIONS

The application of big data in different contexts has proven to be efficient and effective; for that, our research tries to find different approaches to apply big data and machine learning within the employability context for the youth generation. Therefore, the ongoing process and research to create LMIS based on big data based on different solutions and visions can be a road map for further research about creating an intelligent system based on interaction with different stakeholders and data coming from different sources.

REFERENCES

Boselli, R., Cesarini, M., Marrara, S., Mercorio, F., Mezzanzanica, M., Pasi, G., Viviani, M., 2018. WoLMIS: a labor market intelligence system for classifying web job vacancies. J Intell Inf Syst 51, 477–502. https://doi.org/10.1007/s10844-017-0488-x

Caliendo, M., 2016. Youth Unemployment and Active Labor Market Policies in Europe 30.

De Mauro, A., Greco, M., Grimaldi, M., 2016. A formal definition of Big Data based on its essential features. Library Review 65, 122–135. https://doi.org/10.1108/LR-06-2015-0061

Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., 1996. The KDD process for extracting useful knowledge from volumes of data. Commun. ACM 39, 27–34. https://doi.org/10.1145/240455.240464

García, S., Ramírez-Gallego, S., Luengo, J., Benítez, J.M., Herrera, F., 2016. Big data preprocessing: methods and prospects. Big Data Anal 1, 9. https://doi.org/10.1186/s41044-016-0014-0

He, Q., Tan, Q., Ma, X., Shi, Z., 2010. The High-Activity Parallel Implementation of Data Preprocessing Based on MapReduce, in: Yu, J., Greco, S., Lingras, P., Wang, G., Skowron, A. (Eds.), Rough Set and Knowledge Technology, Lecture Notes in Computer Science. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 646–654. https://doi.org/10.1007/978-3-642-16248-0_88

Isson, J.P., Harriott, J.S., n.d. People Analytics in the Era of Big Data 15.

Johnson, E., 2016. Can Big Data Save Labor Market Information Systems? RTI Press. https://doi.org/10.3768/rtipress.2016.pb.0010.1608

Kaygusuz, İ., Akgemci, T., Yilmaz, A., 2016. The Impact of HRIS Usage on Organizational Efficiency and Employee Performance: A Research in Industrial and Banking Sector in Ankara and Istanbul Cities. IJoBM IV. https://doi.org/10.20472/BM.2016.4.4.002

Li, Z., Lin, Y., Zhang, X., 2017. Hybrid employment recommendation algorithm based on Spark. J. Phys.: Conf. Ser. 887, 012045. https://doi.org/10.1088/1742-6596/887/1/012045

Mccaffrey, P., 2020. Overview of big data tools: Hadoop, Spark, and Kafka, in: An Introduction to Healthcare Informatics. Elsevier, pp. 291–305. https://doi.org/10.1016/B978-0-12-814915-7.00020-X

Meehan, J., Tatbul, N., Aslantas, C., Zdonik, S., n.d. Data Ingestion for the Connected World 11.

Mezzanzanica, M., Mercorio, F., 2018. Big Data Enables Labor Market Intelligence, in: Sakr, S., Zomaya, A. (Eds.), Encyclopedia of Big Data Technologies. Springer International Publishing, Cham, pp. 1–11.

Mishra, S.N., Lama, D.R., Pal, Y., n.d. Human Resource Predictive Analytics (HRPA) for HR Management in Organizations 5, 3.

Mishra, T., MewarUniversity, Chittorgarh – 312901, Rajasthan, India, Kumar, D., Department of Computer Science, G. J. University, Hisar – 125001, Haryana, India, Gupta, S., Guru Nanak Institute of Management, West Punjabi Bagh – 110026, Delhi, India, 2017. Students' Performance and Employability Prediction through Data Mining: A Survey. Indian Journal of Science and Technology 10, 1–6. https://doi.org/10.17485/ijst/2017/v10i24/110791

Piad, K.C., Dumlao, M., Ballera, M.A., Ambat, S.C., 2016. Predicting IT employability using data mining techniques, in: 2016 Third International Conference on Digital Information Processing, Data Mining, and Wireless Communications (DIPDMWC). Presented at the 2016 Third International Conference on Digital Information Processing, Data Mining, and Wireless Communications (DIPDMWC), pp. 26–30. https://doi.org/10.1109/DIPDMWC.2016.7529358

Qostal, A., Moumen, A., Lakhrissi, Y., 2020. Systematic Literature Review on Big Data and Data Analytics for Employment of Youth People: Challenges and Opportunities:, in: Proceedings of the 2nd International Conference on Advanced Technologies for Humanity. Presented at the International Conference on Advanced Technologies for Humanity, SCITEPRESS - Science and Technology Publications, Rabat, Morocco, pp. 179–185. https://doi.org/10.5220/0010432201790185

Ryder, G., Director-General, I., 2020. World Employment and Social Outlook - Trends 2020. World Employment and Social Outlook 108.

S. Dong, Z. Lei, P. Zhou, K. Bian, G. Liu, 2017. Job and Candidate Recommendation with Big Data Support: A Contextual Online Learning Approach, in: GLOBECOM 2017 - 2017 IEEE Global Communications Conference. Presented at the GLOBECOM 2017 - 2017 IEEE Global Communications Conference, pp. 1–7. https://doi.org/10.1109/GLOCOM.2017.8255006

S. M. S. Azman Shah, , H. N. Haron2, , M. N. Mahrin, n.d. The Trend of Big Data in Workforce Frameworks and Occupational Standards towards an Educational Intelligent Economy. JTET. https://doi.org/10.30880/jtet. 2021.13.01.019

Saouabi, M., Ezzati, A., 2019. Proposition of an employability prediction system using data mining techniques in a big data environment. International Journal of Mathematics and Computer Science 14, 411–424.

Satyaki Sanyal, Souvik Hazra, Neelanjan Ghosh, Soumyashree Adhikary, 2017. Resume Parser with Natural Language Processing. https://doi.org/10.13140/RG.2.2.11709.05607

Shahbaz, U., Beheshti, A., Nobari, S., Qu, Q., Paik, H.-Y., Mahdavi, M., 2019. iRecruit: Towards Automating the Recruitment Process, in: Lam, H.-P., Mistry, S. (Eds.), Service Research and Innovation, Lecture Notes in Business Information Processing. Springer International Publishing, Cham, pp. 139–152. https://doi.org/10.1007/978-3-030-32242-7_11

Shalaby, W., Alaila, B., Korayem, M., Pournajaf, L., Aljadda, K., Quinn, S., Zadrozny, W., 2017. Help me find a job: A graph-based approach for job recommendation at scale, in: Nie J.-Y., Obradovic Z., Suzumura T., Ghosh R., Nambiar R., Wang C., Zang H., Baeza-Yates R., Baeza-Yates R., Hu X., Kepner J., Cuzzocrea A., Tang J., Toyoda M. (Eds.), Proc. - IEEE Int. Conf. Big Data, Big Data. Institute of Electrical and Electronics Engineers Inc., pp. 1544–1553. https://doi.org/10.1109/BigData.2017.8258088

Shrivastava, S., Nagdev, K., Rajesh, A., 2018. Redefining HR using people analytics: the case of Google. HRMID 26, 3–6. https://doi.org/10.1108/HRMID-06-2017-0112

Sorenson, K., Mas, J.-M., n.d. A Roadmap for the Development of Labor Market Information Systems 2.2.3 LMIS in Select Countries 74.

Tamburri, D.A., Heuvel, W.-J.V.D., Garriga, M., 2020. DataOps for Societal Intelligence: a Data Pipeline for Labor Market Skills Extraction and Matching, in: 2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI). Presented at the 2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI), IEEE, Las Vegas, NV, USA, pp. 391–394. https://doi.org/10.1109/IRI49571.2020.00063

Vaccarino, A., Mezzanzanica, M., Mercorio, F., Castel-Branco, E., n.d. Methodological overview and Analytics insights on Tunisian Web Labour Market 45.