2021

Big data analytics in public sector university libraries in Pakistan

Sadaf Shahid

Rafat Parveen Siddiqui

2021

Big Data Analytics in Public Sector University Libraries in Pakistan

Sadaf Shahid
Aga Khan University, sadaf.shahid@aku.edu

Rafat Parveen Siddiqui
University of Karachi, Pakistan

Big Data Analytics in Public Sector University Libraries in Pakistan

Abstract:

This study examines librarians' perceptions, capabilities, and understandings of Big Data analytics in public sector university libraries in Karachi, Pakistan. To acquire the desired results, a survey was conducted and using a quantitative approach. The study's target audience was library administrators at public sector university libraries in Karachi, all of which are recognized by Pakistan's Higher Education Commission and chartered by the Sindh Government. All respondents were sent an e-mail inviting them to participate in the survey on their own time. This study is important because it fills a large vacuum in the literature about the perspectives of Karachi's public sector university librarians on Big Data analytics. The result shows that most of the academic librarians are familiar with the concept of big data and thought that they need to develop their skills for the use of big data analytics tools and the government should also provide a sufficient budget for the professional development of the library staff.

Keywords:

Big Data, Public Sector University Libraries – Pakistan, Analytical Tools

Introduction:

In this digital time, the use of information communication technologies has resulted in a tremendous increase in data in nearly every aspect of life. (Liu et al., 2018). The difficulty of handling this vast unstructured increase in data has arisen as a result of the data eruption and widespread consumption and evolution of data. Big Data is a burning issue in corporate, trade, education, and government organizations all over the world right now, and research and development methodologies are being used to solve the tests and promise of analytic applications in Big Data. (Cuzzocrea, 2014). Big Data, it may be said, is helping to pave the way for the world's expanding digital native economy. However, there is no single definition of Big Data in the present literature. Various parties have

characterized Big Data in a variety of ways. Some of the definitions that have been chosen are listed below: “Big Data refers to such datasets whose volume is outside the capacity of traditional software’s (Manyika et al., 2011). Big Data comprises of those data sets which are so large and complex that commonly used software’s are incapable of dealing with them” (Garcia and Wang 2013). Big Data is characterized by Gartner Group Inc. (2001) as “high volume, high velocity, and high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery, and process optimization.” Exabytes (1,024 petabytes) of data including billions of records from millions of people collected from numerous sources. Laney (2001) was the first to coin the term "Big Data," claiming that it possesses certain distinct features that distinguish it from usual data. The 5Vs, or "large volume, high velocity, high variety, low veracity, and high value," are used to describe and explain these traits. (Jin et al., 2015). In comparison to typical data, the magnitude of Big Data is immense. Data volume and expansion are not constrained in any way. Big Data velocity relates to how quickly and dynamically it is created.

The features and variety of data sets make the data use and organization process tough. Data scientists and commercial organizations can organize data that has been obtained in a methodical and formatted manner. Some data, on the other hand, is discovered in an unstructured format and comes from a variation of sources, as well as electronic mails and figures gathered online. (Wang et al., 2016). The growth of Big Data generation, gaining, storing, and transfer has posed certain tests and chances for new libraries. Organizational structures are being updated to contend with environmental and societal changes as communication and information technology advances. Libraries, according to Ilesanmi (2013), are knowledge centers that organize, retrieve, and transmit information, as well as sustain information systems throughout society to meet societal demands. Though, the commencement of Big Data is prompting libraries to rethink their service arrangements, which they had previously used to carry out their operations. (Affelt, 2015). Noh (2015) stated that the current form of libraries can be turned into library 4.0 in response to the change in this digital era. Library 4.0 will be distinguished by its ability to handle large amounts of records. It also demonstrates how the developing trend of Big Data can help libraries advance and create infrastructure to deliver better services to library users and the

community. Libraries are currently confronted with the challenges of data processing and a shortage of library employees' abilities. These skills must also be enhanced in order to deal with the opportunities and difficulties that have arisen as a result of this age. (Gordon-Murnane, 2012). This movement in library infrastructure and services from traditional to contemporary presents an uncommon scenario that has to be investigated. The ramifications of big data for libraries are discussed herein paper. The current paradigm shift, as well as the review of research, help us understand librarian competencies and abilities, as well as the capability of libraries, which is also required to deal with the prospective application of big data analytics in libraries.

Research Questions:

- Are the librarians know about the concepts of Big Data?
- What are the perceptions of librarians towards the use of Big Data analytics?
- Are they familiar with the Big Data analytical tools?
- Are they competent enough for the carrying out of Big Data analytics?

Scope and Limitation:

This study's target audience was library administrators at public sector university libraries in Karachi, all of which are by Pakistan's Higher Education Commission and chartered by the Sindh government.

Significance of the Study:

This study is important because it fills a large vacuum in the literature about the perspectives of Karachi's public sector university librarians on Big Data analytics.

Literature review:

The volume of data produced, stored, organized, and analyzed in this digital environment has greatly expanded. According to the findings of Bedeley et al.

Academic librarians, according to Lu.et al. (2017), are well-versed in the awareness of Big Data analytics and are developing data-related accomplishments. However, in order to implement Big Data analytics, coordination between different sections of libraries is required. The internet's population was 10, 24 million in 2005, but it has already grown to 4,000 million. These figures suggest that the number of options for internet researchers is growing, posing a greater problem for the storage and handling of Big Data and privacy in businesses (Schaich, 2018). Corporations and organizations are considering redesigning their organizational processes to accommodate the creation of records for facilities and to increase quality of product, according to Taylor-Sakyi (2016). As part of this plan, rational data base management solutions are replacing old data management methods in businesses. According to a survey of the prior literature, some studies have been undertaken on the employment of Big Data analytics in libraries. However, here is a prerequisite to comprehend and investigate the familiarity of ideas and methods of big data analytics in the context of public sector libraries, particularly in the city of Karachi. The current study is an attempt to compile the literature in the relevant topic of study.

In Libraries, Big Data Analytics:

The three more Vs. of Big Data, according to DeVan (2016), are “variability,” the mixes of constantly changing data that could have a huge impact on data homogeneity. Analytics problems include data visualization through charts and graphs, as well as the value of organizational data. After retrieving data from many sources, integrating the data is tough. The types of data in a library (Goldberg et al., 2014) vary greatly, and large amounts of data must be managed and preserved in order for the library's services to be available. The desires of library patrons will continue to expand in the future as a result of the digital world (Showers, 2014). Librarians should be able to understand how data is created, managed, and preserved (Semeler et al., 2017). The importance of their role in Big Data analytics necessitates the development of librarians' skills and competence in order for Big Data analytics to be adopted. According to Xie and Fox (2017), library workers lack the necessary competence to create new value-added services using Big Data analytics.

Librarians' Big Data Competencies and Skills:

In today's digital world, librarians have more opportunity to get involved in tasks involving data. Library professionals should have the necessary expertise to provide potential data services. Executives should attempt to construct systems of strategic intelligence. The transformative core of a data manager's purpose is to change the various types of data that are generated or can be useful to academics and library customers (Ahmad, 2017). Companies are using Big Data as information assets, according to De Mauro et al. (2015), thus librarians should incorporate Big Data analytics approaches as well. Also, an other study found that librarians' expertise and abilities in new technologies are critical for providing timely and high-quality information to consumers (Ullah and Anwar, 2013). As a result, acquiring the topography of library environments in the analytics in the context of Big Data. Collaboration with others can help librarians prepare and grow their skills. Hoy (2014) made a similar point, stating that with technology improvements, librarians must be conversant with the capabilities and difficulties associated with massive data, and utilize their knowledge to assist their patrons in selecting the appropriate tools. In librarianship, there is also a paranoid culture of learning new technical abilities and emergent trends (Braganza et al., 2017). In light of this, Librarians talents and competencies must be strengthened in this practical time of Big Data.

Research Approach:

This study was conducted using a quantitative technique since the results of the quantitative approach allow for generalization to the entire population. (Daniel, 2016).

Furthermore, quantitative conclusions are presented in figures and numbers such as statistics, mathematics, and percentages. Quantitative data is gathered in the form of numbers. As a result, number provides a clear and full picture that is free of inconsistencies.

Research Method:

A survey was conducted in this study to acquire the desired results. The survey method is employed because it is efficient and cost-effective, and it allows for the collection of a great amount of data in a short sum of time.

Population of the Study:

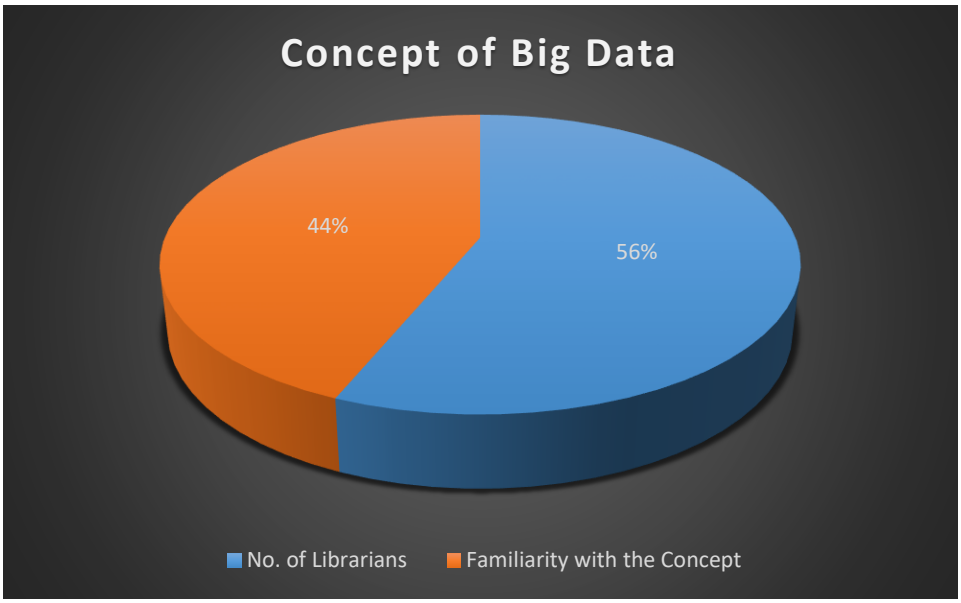
The study's target demographic was the main library managers of public sector university libraries in Karachi. Because they are the most educated as regards the state of expertise and facilities in their libraries, administrative librarians were chosen to collect data. At top management positions in university libraries, these professional librarians play a part in the decision-making and execution of new technologies. Administrative librarians take part in the development of policy for all types of libraries at university. All of these institutions have been approved by the Pakistani Higher Education Commission and chartered by the Sindh Government.

Data Collection and Analysis:

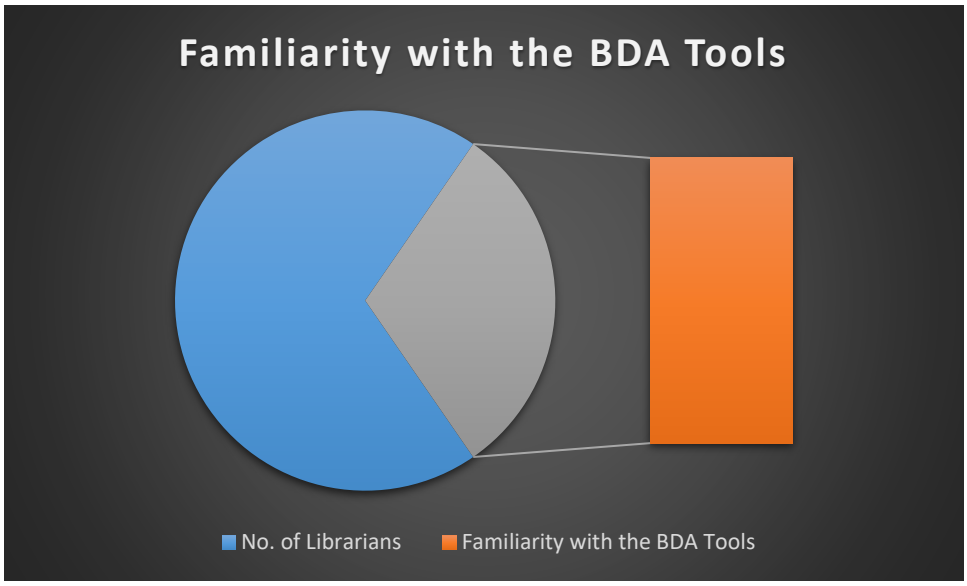
A structured questionnaire with several closed-ended items was created using a five-point Likert scale. All respondents (academic librarians) were freely invited to participate in the survey by e-mail. MS Excel was used to analyze the survey data.

Result of the Study:

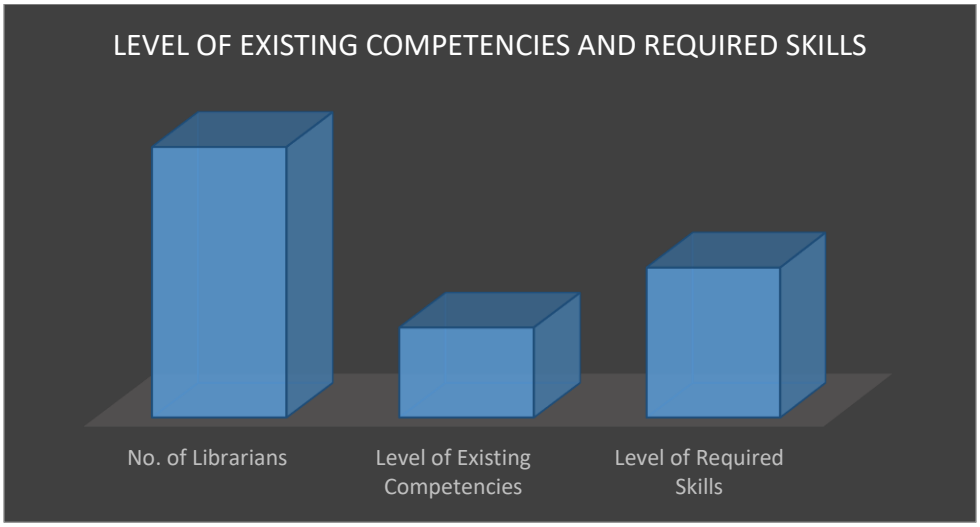
The result shows that most of academic librarians are familiar with the concept of big data, they thought that big data analytics could be very useful tools for libraries in the future for the deep insight decision-making process, but now at the present scenario, they are not using this type of tools at their libraries.



Some of the librarians are familiar with big data analytics tools but they don't know how to use them. The result also shows that most of the librarians are using MS Excel for data-related works. They believed that big data analytics tools improve work productivity and enable to control over their work and also improves the library service quality towards the users and it brings transparency and stops unfair means in library systems and provide support to library professionals to enhance the library services.



They are also facing some challenges for the execution of big data analytics like; lack of data management skills and data cleaning skills and other technical skills. They thought that librarians need to develop their skills for use of big data and the government should also provide a sufficient budget for the professional development of the library staff.



Discussion:

Big Data analytics is getting more popular by the day, both as a concept and as a way of application. Internal competitiveness has a plethora of options in the Big Data era. In this environment, librarians' jobs are shifting, and they are now serving as data scientists, data overseers, and digital services representatives, from this viewpoint, librarians play a role in data collecting, data controlling, metadata expertise, data protection, examination, and imagining. Librarians used to be in charge of information acquisition, organization, retrieval, collection, dissemination, and hoarding in their libraries. Traditionally, their methods for organizing statistics have taken the shape of scholarly dealings of concepts, such as books and serials. In this digital setting, they have now changed their processes to include all types of information, independent of format. This complete paradigm shift has broadened an area that was previously solely centered on textual scientific papers to now incorporate data analysis by academics and library users.

Conclusion:

In the current information world, academic organizations, governmental and private sector consortiums, and businesses generate enormous amounts of records. The abilities of librarians based on big data analytics enable academic libraries to focus on certain aspects of analytics. For businesses and educational institutions, managing enormous and complicated data is a huge burden. The responsiveness and perception of academics librarians with Big Data were investigated in this study. According to the findings, librarians are acquainted with the concept of Big Data and big data analytics tools but they need skills to cope up with this change to survive in this information age.

References:

Affelt, A. (2015), The Accidental Data Scientist: *Big Data Applications and Opportunities for Librarians and Information Professionals*, Information Today, Medford, NJ.

Ahmad, K. (2017), The perspective of library and information science (LIS) professionals toward knowledge management in university libraries*, Journal of Information & Knowledge Management*, Vol. 16 No. 2, pp 1-11, available at: https://doi.org/10.1142/S0219649217500150

Atkinson, J. (2018*), Reflections on Collaboration and Academic Libraries. Collaboration and the Academic Library: Internaland External, Localand Regional, National and International*, Elsevier, available at: https://doi.org/10.1016/B978-0-08-102084-5.00020-1

Atkinson, J. (2018), Reflections on Collaboration and Academic Libraries. *Collaboration and the Academic Library: Internal and External, Local and Regional, National and International*, Elsevier, available at: https://doi.org/10.1016/B978-0-08-102084-5.00020-1

Bedeley, R.T., Ghoshal, T., Iyer, L.S. and Bhadury, J. (2018), Business analytics and organizational value chains: a relational mapping, *Journal of Computer Information Systems*, Vol. 58 No. 2, pp. 151-161.

Braganza, A., Brooks, L., Nepelski, D., Ali, M. and Moro, R. (2017), Resource management in Big Data initiatives: processes and dynamic capabilities, *Journal of Business Research*, Vol. 70, pp. 328-337, available at: https://doi.org/10.1016/j.jbusres.2016.08.006

Burton, M., Lyon, L., Erdmann, C. and Tijerina, B. (2018), Shifting to data savvy: the future of data science in libraries, available at: http://d-scholarship.pitt.edu/id/eprint/33891 (accessed March 11, 2019).

Cuzzocrea, A. (2014), Privacy and security of Big Data: current challenges and future research perspectives, *Proceedings of the First International Workshop on Privacy and Security of Big Data*, pp. 45-47.

De Mauro, A., Greco, M. and Grimaldi, M. (2015), What is Big Data? A consensual definition and a review of key research topics, AIP Conference Proceedings, Vol. 1644, pp. 97-104.

DeVan, A. (2016), The 7 V's of Big Data, available at: www.impactradius.com/blog/7-vs-big-data/

Djafri, L., Ammar Bensabeur, D. and Adjoudj, R. (2018), Big Data analytics for prediction: parallel processing of the big learning base with the possibility of improving the final result of the prediction, *Information Discovery and Delivery*, available at: https://doi.org/10.1108/ IDD-02-2018-0002

Fan, J., Han, F. and Liu, H. (2014), Challenges of Big Data analysis, *National Science Review*, Vol. 1 No. 2, pp. 293-314

Federer, L. (2018), Defining data librarianship: a survey of competencies, skills, and training, *Journal of the Medical Library Association*, Vol. 106 No. 3, pp. 294-303.

Fink, L., Yogev, N. and Even, A. (2017), Business intelligence and organizational learning: an empirical investigation of value creation processes, *Information and Managemen*t, Vol. 54 No. 1, pp. 38-56.

Garcia, T. and Wang, T. (2013), Analysis of Big Data technologies and method – query large web public RDF datasets on amazon cloud using hadoop and open source parsers, 2013 IEEE Seventh International Conference on Semantic Computing, pp. 244-251.

Goldberg, D., Olivares, M., Li, Z. and Klein, A.G. (2014), Maps & GIS data libraries in the era of Big Data and cloud computing, *Journal of Map & Geography Libraries*, Vol. 10 No. 1, pp. 100-122

Gordon-Murnane, L. (2012), Big Data: a big opportunity for librarians, available at: https://doi.org/ 1039559884

Gorman, M. (2015), Our enduring values revisited: librarianship in an ever-changing world, available at: http://books.google.com/books

Gunasekaran, A., Yusuf, Y.Y. and Adeleye, E.O. (2018), Agile manufacturing practices: the role of Big Data and business analytics with multiple case studies, *International Journal of Production Research*, Vol. 7543, pp. 1-13.

Hoy, M.B. (2014), Big Data: an introduction for librarians, *Medical Reference Services Quarterly*, Vol. 33 No. 3, pp. 320-326.

Ilesanmi, T.C. (2013), Roles of the librarian in a research library in the digital era: challenges and the way forward, *New Review of Academic Librarianship*, Vol. 19 No. 1, pp. 5-14, available at: https://doi.org/10.1080/13614533.2012.740437

Islam, A. Y. M. A., Ahmad, K., Rafi, M., & JianMing, Z. (2021). Performance-based evaluation of academic libraries in the big data era. *Journal of Information Science*, *47*(4), 458–471, available at: https://doi.org/10.1177/0165551520918516

Jin, X., Wah, B.W., Cheng, X. and Wang, Y. (2015), Significance and challenges of Big Data research, *Big Data Research*, Vol. 2 No. 2, pp. 59-64, available at: https://doi.org/10.1016/ j.bdr.2015.01.006

Johnson, V. (2017), Leveraging technical library expertise for Big Data management, *Journal of the Australian Library and Information Association*, 0158, pp. 1-16, available at: https://doi.org/10.1080/24750158.2017.1356982

Katal, A., Wazid, M. and Goudar, R.H. (2013), Big Data: issues, challenges, tools and Good practices, 2013 6th International Conference on Contemporary Computing, pp. 404-409.

Khurshid, A., JianMing, Z., & Rafi, M. (2019). An analysis of academic librarians competencies and skills for implementation of big data analytics in libraries. *Data Technologies and Applications, 53*(2), 201-216.

Laney, D. (2001), Meta delta, Application Delivery Strategies, Vol. 949, February, p. 4, available at: https://doi.org/10.1016/j.infsof.2008.09.005

Lee, C. and Kim, H. (2018), The evolution trajectory of an ICT ecosystem: a network analysis based on media users' data, *Information and Management*, Vol. 55 No. 6, pp. 795-805

Liu, Y., Yang, L., Sun, J., Jiang, Y. and Wang, J. (2018), Collaborative matrix factorization mechanism for group recommendation in Big Data-based library systems, *Library Hi Tech*, Vol. 36 No. 3, pp. 458-481

Mahesh, G. T., & Nandeesha, B. (2021). *Big Data Applications for Improving Library Services* (1st ed.). IGI Global.

Moorthy, J., Lahiri, R., Biswas, N., Sanyal, D., Ranjan, J., Nanath, K., & Ghosh, P. (2015). Big Data: Prospects and Challenges. *Vikalpa*, *40*(1), 74–96, available at: https://doi.org/10.1177/0256090915575450

Noh, Y. (2015), Imagining library 4.0: creating a model for future libraries, *Journal of Academic Librarianship*, Vol. 41 No. 6, pp. 786-797.

Oakleaf, M. (2016), Getting ready & getting started: academic librarian involvement in institutional learning analytics initiatives, *Journal of Academic Librarianship*, Vol. 42 No. 4, pp. 472-475.

Schaich, M. (2018), Information professionals, Huguenot Networks, 1560–1780, pp. 75-91, available at: https://doi.org/10.4324/9781315188959-6

Showers, B. (2014), Developing a shared analytics service for academic libraries, Insights: *The UKSG Journal*, Vol. 27 No. 2, pp. 139-146.

Taylor-Sakyi, K. (2016), Big Data: understanding Big Data, available at: http://pdf/abs/1601.04602% 5Cnhttp://pdf/pdf/1601.04602.

Thomas, C. and Urban, R. (2018), What do data librarians think of the MLIS? Professionals' perceptions of knowledge transfer, trends, and challenges, *College & Research Libraries*,Vol.79 No. 3, p. 401, available at: https://doi.org/10.5860/crl.79.3.401

Ullah, M. and Anwar, M.A. (2013), Developing competencies for medical librarians in Pakistan, *Health Information and Libraries Journal*, Vol. 30 No. 1, pp. 59-71

Wang, C., Xu, S., Chen, L. and Chen, X. (2016), Exposing library data with Big Data technology: a review, 2016 IEEE/ACIS 15th International Conference on Computer and Information Science, pp. 1-6.

Xie, Z. and Fox, E.A. (2017), Advancing library cyber infrastructure for Big Data sharing and reuse, *Information Services and Use*, Vol. 37 No. 3, pp. 319-323

Yaqoob, I., Hashem, I.A.T., Gani, A., Mokhtar, S., Ahmed, E., Anuar, N.B. and Vasilakos, A.V. (2016), Big Data: from beginning to future, *International Journal of Information Management*, Vol. 36 No. 6, pp. 1231-1247.

Zhan, M. and Widén, G.(2017), Understanding Big Data in librarianship, *Journal of Librarianship and Information Science*, 096100061774245, available at: https://doi.org/10.1177/0961000617742451.

Zhuge, H. (2015), Mapping Big Data into knowledge space with cognitive cyber-infrastructure, pp. 1-59, available at: http://pdf/ftp/arxiv/papers/1507/1507.06500.pdf.