Nigerian Journal of Applied Physics (NJAP)
ISSN Print: 3115-5871
ISSN Online: 3115-588X
DOI: https://doi.org/10.62292/njap-v2i2-2026-91
Volume 2(2), June 2026

Implication of using Data Science in Nuclear Physics for the Advancement of Nigerian Society

***[1]Dahiru Dahuwa, [2]Jude O. Anaegbu, [3]Abdu Ibrahim, [4]Lawal Abba and [5]Sadik A. Lawal**

[1]Department of physics, Federal university of Health sciences Azare, Bauchi state, Nigeria.
[2]Department of Physical Science, Greenfield University, Kaduna State, Nigeria.
[3]Department of Physics, Alqalam University Katsina, Katsina State, Nigeria.
[3]Department of Physics, Aminu Saleh College of Education, Azare , Bauch State, Nigeria.
[5]Department of Physics, Nigerian Defence Academy, Kaduna, Kaduna State, Nigeria.

[*]Corresponding author's email: dahiru.dahuwa@fuhsa.edu.ng Phone: +2348039766138

ABSTRACT

The Increasing complexity of nuclear physics experiments has generated unprecedented volumes of data, creating opportunities for data-driven approaches that complement conventional theoretical and experimental methods. This review examines recent advances in the application of data science—including machine learning, artificial intelligence, statistical modelling, and big-data analytics—to nuclear physics, with particular emphasis on their relevance to Nigeria's emerging nuclear science ecosystem. The review synthesizes published literature on the use of data science for reactor monitoring, radiation detection, nuclear medicine, particle physics, and predictive modelling, and discusses how these approaches can improve research efficiency, safety, and decision-making. It further evaluates the opportunities and challenges associated with adopting data-driven methodologies in Nigeria, including limitations in computational infrastructure, human capacity, access to nuclear data, and research collaboration. The review argues that strategic investment in digital infrastructure, advanced computing, workforce development, and international partnerships would significantly enhance Nigeria's capability to apply modern data science techniques in nuclear research and related applications. Integrating data science with nuclear physics has the potential to strengthen scientific innovation, improve healthcare and environmental monitoring, support evidence-based policymaking, and contribute to sustainable national development.

Keywords:
Data science,
Nuclear physics,
Machine learning,
Artificial intelligence,
Nuclear data,
Nigeria,
Big data,
Predictive modeling.

INTRODUCTION

Data science techniques have been transforming the way scientific experiments are designed and interpreted and have pervaded most inquiries as well as driving decisions in large enterprises. Nuclear physics is no exception. Novel machine-learning techniques have broadened the capacity of experimentalists to handle and analyze data generated in modern-day high-throughput experiments in nuclear and particle physics. The current era of big data has created unprecedented opportunities for research in nuclear science and earth-abundant heavy ion physics. The proper analysis and interpretation of vast amounts of experimental data generated in this field have enormous potential, especially in emerging economies such as those in Africa, which are dramatically expanding their cyberinfrastructure. (Bruce et al., 2020)

The large amount of data from nuclear experiments has implications for cancer research, drug development, and power generation. Data mining and machine learning can enhance current models and lead to discoveries. A new branch of research at the interface of nuclear physics and data science is needed to handle the challenges of the 21st century.

Nuclear physics has deep contemporary relevance in many scientific disciplines and can significantly contribute to the attainment of some SDGs. Thus, the study of nuclear physics is not limited or relegated to nations with nuclear power plants, and there are clear advantages in developing data science techniques and methods for these peripheries that have a direct impact on expanding the role of nuclear science and contributing to several of the UN's Sustainable Development Goals.

From a practical perspective, the increasing complexity of nuclear physics data-driven technologies could provide the solution for nuclear research that is already heavily reliant on large collaborative teams that are interdisciplinary in the analysis of international laboratory data sets often produced by cutting-edge instrumentation. This is an area of benefit to a country such as Nigeria, where the use of data mining techniques in nuclear physics and data science more generally can boost the capacity of local researchers to begin to make advances in the continued analysis of data generated at FAIR. (Bello et al., 2024)

Definition and Scope of Data Science

The term "data science" does not have a unanimous definition yet, but one existing opinion is that data science is not a brand-new science. With its roots in statistics, grounded in mathematical theories, and embracing various techniques and methodologies, data science has been in existence for about half a century. It is an all-encompassing field striving to make sense of very large data collections or sets. Data science is also supported by results and techniques from a wide range of other fields, including computer science and engineering, which power it for the computer-intensive aspects. As a field, data science is vast. This is evident in the way some organizations define the scope of data science in their respective data science courses. (Efron& Hastie, 2021)

The complexity of nuclear physics data is growing due to technological advancements. Nuclear data is voluminous and heterogeneous, making it difficult for data scientists to integrate and work with. Data munging, or cleaning and shaping data, takes up a significant amount of time. Statistical techniques are used to identify and correct errors in the datasets. Visual representations are used to extract information and attract interest.

Fundamental Concepts in Nuclear Physics

The atomic nucleus and its constituents form the central point of interest in any nuclear system. Theoretical descriptions and experimental evidence suggest a hierarchy for the nucleus. These are based on the structure made of nucleons, i.e., protons and neutrons, and the arrangement in the nuclear orbitals. Moreover, the nucleus is known to undergo decay processes that have been established to be the guiding forces for the structure based on experimental evidence. The interaction among the nucleons and their arrangement is shown to be regulated by the nuclear forces. Understanding the various forces that govern the atomic nucleus and nuclear decay becomes the key component of most theories. The formulation of theories is generally the central concern, as these theories advance information on the phenomena governing and setting the rules involved in the description of most experiments. Theoretical models mostly rely on unobserved quantities to perform the description of experiments. The merits of theories are established if they can reproduce new experimental results in agreement with the uncertainties obtained.

Practical examples of applications can be found in the description of radiation and the release of energy through fission and fusion events by the atomic nucleus. These events are known to cause severe damage to human health and endanger the surroundings. Significant amounts of data can be generated, and these will promote reliability. Most nuclear systems exhibit some level of complexity, and these warrant adequate representation of uncertainties. Understanding the aforementioned nuclear phenomena is central to the development of new and alternative power sources attributed to nuclear processes such as fission and fusion. Nonetheless, large and complex nuclear data can mask and hinder the physical interpretation of the state of nuclear activities. This can render a shift in attention in nuclear model developments to adopt data science methodologies, hence promoting the analyses and interpretation of nuclear systems. In line with our awareness, we seek methodologies that can promote data representation through descriptions of experiments. (Terrani et al. 2020).

MATERIALS AND METHODS

Applications of Data Science in Nuclear Physics

Numerous applications of data science in the field of nuclear physics could drive fresh perspectives and introduce innovative methods leading to the transformation of the field. Some of the applications applicable to present-day nuclear physics may be highlighted. With the advent of exaflop computing systems powered by petabytes of data, the use of data science in carrying out full-scale, full-complexity simulations of entire nuclear power systems has been proposed. Such “digital twins” would be capable of predicting the power plant performance and dynamics over a wide range of parameters, including material parameters.

In another example, in nuclear particle interactions < 1 GeV, hybridizing machine learning models with traditional nuclear physics phenomenology provides a compact representation of intra-nuclear cascade calculations for the NHL muon tomography facility. Data science tools aid in increasing discovery potential in high-energy physics experiments due to rare events. Data science can also predict non-standard analysis outcomes and neutron emission events in the nonproliferation industry and reactor safety analysis. Data science-assisted algorithms optimize nuclear processes by tuning model inputs. AI and machine learning techniques predict the composition of fission sources and identify man-made disturbances affecting reactor operation. (Kahn et al., 2022) (Zharenov, 2023)

Given the above-mentioned applications of data science in nuclear physics, especially in model development, performance prediction, and reactor safety, its integration into the Nigerian nuclear industry could foster and stimulate a new direction in nuclear research aimed at optimizing reactor power/efficiency and extending nuclear facility life. Furthermore, data science can be used to predict system behavior under unknown input parameters, provide meaningful insights that could help data and nuclear dataand experiments, and advance the field of nuclear medicine. There has been an advance in the incorporation of AI/ML methods directly as a nuclear response to a problem, such as identifying optimal simulation strategies and selecting sub-physics parameters that improve solver accuracy, reduce calculation time, and reduce data waste, without the use of an external AI/ML tool. As this area continues to grow, new ideas and techniques are being introduced, such as quantum AI methods, reinforcement learning, and other advanced statistical methods. Furthermore, the adoption of cutting-edge computing practices like neural networks for scientific computing and the exploration of quantum machine learning-assisted scientific computing are new trends. (Gohel et al. 2020)

Nuclear Reactor Monitoring and Safety

Nuclear reactors have complex systems with many potential operational risks. They have hundreds of sensors that gather large amounts of data during operation. These sensors use various techniques, such as thermocouples for hot leg recirculation and ex-core detectors for core flux distribution. The data collected helps detect conditions that may lead to engine burnouts, like loss of flow accidents.

Given the sometimes harsh working conditions in some plants, data on engine performance is transmitted to data centers in real-time, where advanced analytics such as digital twins of the plants using neural networks and other machine learning algorithms, genetic algorithms, and other state-of-the-art optimization techniques are applied to predict future engine behavior based on current trends. Emergency response plans are also developed. Real-time monitoring and predictive modeling are important to the operational safety of many nuclear facilities. Advanced analytics can visualize trends and discern subtle indications of anomalous behavior, thus shedding light on hidden correlations and indicators. Existing predictive maintenance models in real operating reactors have data accuracy above 95%. Such capabilities give operators the necessary time to act to forestall or mitigate these catastrophic occurrences. Implemented following a disaster, a series of improvements were made, and protective actions were taken by staff at a nuclear power plant as part of drills against such events. While the emergency offsite plan has long existed in the vicinity of nuclear power facilities, improvements may abound, with near real-time questions over partnerships entered into between regulatory bodies and meteorological agencies. Practical case studies and lessons learned are available for various sectors. (Bae et al., 2021)

It is worthy of note to state that both the Nigerian Nuclear Regulatory Authority and relevant international agencies require that records of reactor performance, such as staff doses, operation and maintenance activities, and all incidents, are kept. In Nigeria's ongoing use of small modular reactors, reactor safety will be paramount in the licensing guidelines. Hence, reactor safety could receive a real boost in Nigeria through implementing the proposed nuclear data techniques for accurate and reliable monitoring of reactors. Ensuring the accuracy and reliability of data will be important to avoid loss of confidence by the public in the nuclear safety authority, as demonstrated in the direct and indirect effects of a disaster. Data science techniques for robust nuclear operation in Nigeria will ensure that a measure of trust is developed, regulator protection from slander suits, investment of resources in assets with the most value-added potential, and a well-advertised supply chain, which are significant added values of the operational decision-making, can be guaranteed.

Particle Physics Experiments

Particle physics experiments generate vast amounts of data, particularly from high-granularity detectors. The data ranges from kilobytes to megabytes per collision and can reach terabytes to petabytes over a year-long experiment. This data complexity is just the beginning of the experimentation process.

Algorithms are developed to take the disorderliness of raw data and process it into an organized summary that is as close to what one would have estimated before actual data collection. Pertinently, this means that not all of the data collected is essential for answering our primary physical question. It is peculiar that this final processed data is what gets made accessible to other research parts following a standard walk-through; this axis of data then proceeds to basic or detailed simulation to make the data more relevant and adequate for different types of physical analysis. Generically, the term machine learning encompasses a lot of varied concepts. In terms of the uses envisioned for machines, one common approach taken in designing learning in machines or algorithms is to make machines possess some generalized form of intelligence, called artificial intelligence. Albeit the overlapping designs, the choice of a focus on artificial intelligence or machine learning impacts cross-disciplinary research implications. It is therefore expedient to differentiate the two, even if only loosely. Development in the ability to classify objects in bi-directional focus has nearly been the same in image and non-image data. Nonetheless, the data science algorithms needed may not meet if this classification skill set is to be transposed into non-image equally good usability or learning. For one, clustering is easier to accomplish with non-image data than with image data, although clustering is a derivative concept of classification. Therefore, such difficulties inform the necessity for focused data science research to detect classification principles that can be made enormously relevant and more practically impactful on the physical data-taking experiments. It is known that an intrusion detection classification algorithm for such a dataset may clamor for the in-depth study of the two or more fundamentally take on an out-of-the-box approach to solving the classification problem. Machine learning has improved its performance in pattern recognition to detect particles among the debris of proton-proton collisions observed. (Lee et al. 2020)

RESULTS AND DISCUSSION

This section examines the issues, limitations, and problems confronting the implementation of data science methodologies in the Nigerian nuclear physics sector. One of the key requirements for data science analysis is the availability of modern computing facilities. Besides, the non-availability of certain facilities may also limit the availability of useful data required for scientific reckoning with the contemporary world. The solutions to these challenges entail the refurbishment and development of technology needed to conduct cutting-edge technological and scientific research in Nigeria. Capacity building in information technology and big data implementation capabilities is required to achieve this end. Well-developed education and training curricula are needed to produce a workforce that has the necessary capabilities to carry out data science analysis.

The gap between the activities in nuclear research in Nigeria, on the one hand, and data science analysis potential, particularly in nuclear science, points to opportunities rather than limitations. Could there be neglected benefits in overlooking data science methodologies in scientific research in Nigeria? The combination of data science capabilities in nuclear data measurement and the assessment of quality parameters and standardization may open new frontiers in the efforts to re-establish radioactivity monitoring capabilities after the closure of the Siting and Design Facility. Advanced data science techniques could also be brought into nuclear physics research currently being conducted by colleagues to further delineate the longevity of the derived results and processes in neutron detection efficiency, Fermi wavevector, fractional integral derivative models, etc. In characterizing the modern paradigm shift in nuclear activity data collection, representation, and utilization in nuclear physics, other synergistic ideas could be generated, especially in radiation protection and human health optimization that are of interest to relevant stakeholders. (Battistoni et al., 2021)(Deiana et al., 2022)

Infrastructure and Resources

Current State At the moment, no neutron source in Nigeria is operational for research purposes. At several levels, Nigeria also does not have the required infrastructure to support data science. Because the instrumentation required to undertake data-driven initiatives in nuclear physics is typically sensitive and technologically advanced, investments must be made in facilities and infrastructure to acquire credible data for quality science. On the computational aspects of data science, there are currently basic computational facilities in various tertiary institutions across the country. Most of these facilities are basic and outdated, offering little or no support for massive data analytics efforts and artificial intelligence/machine learning studies. Exclusive and privileged access to databases, including neutron, charged particle, and spectrometry databases accessible as needed, would be of immense advantage. Reduction software is not available, and hence Nigerian researchers would not be able to do real-time data reduction on experiments. (Asuku et al., 2023)

As discovered at the data science workshop, infrastructure is vital for research. Monetized services at the center will attract partnerships. Collaboration in creating a repository will enhance international appeal. Training in Python, R, Julia, Git, etc. will establish Nigeria as a data excellence center. Established facilities will offer community-wide strategic plans for data sharing. Collaborative efforts will ensure the delivery of quality data. The refinement of data can be done at the neutron source. Nigeria and its partners will wait for the test facility to become operational. Metadata will be agreed upon online and documented.

Capacity Building and Education

The successful integration of data science into nuclear physics depends not only on the infrastructure, but also on the availability of a well-trained workforce that can apply computational techniques to complex nuclear problems. The recent literature has pointed out the lack of expertise in artificial intelligence, machine learning, scientific computing and advanced statistical analysis as important obstacles in the adoption of data-driven methodologies in developing countries. These challenges are especially difficult in the area of nuclear science, which requires knowledge that crosses disciplines such as physics, mathematics, computer science and engineering. The increase in data science education in Nigeria has raised awareness about computational methods but there is a lack of specialized training that includes nuclear physics and advanced analytics. Current university courses tend to cover the fundamentals of either physics or computer science, but rarely combine the two in a single curriculum. As a consequence, graduates often need further professional training before applying machine learning and data analytics to nuclear applications. The reviewed literature emphasizes that strengthening national capacity requires sustained investment in postgraduate education, interdisciplinary curricula, continuing professional development, and international collaboration. Partnerships with established nuclear research centres can facilitate knowledge transfer, access to advanced computational tools, and exposure to emerging analytical techniques. Furthermore, mentorship programmes and collaborative research networks can help develop the next generation of Nigerian nuclear scientists while improving institutional research capacity. Beyond formal education, continuous professional development is essential because data science technologies evolve rapidly. Regular training in programming languages such as Python and R, scientific computing, cloud-based analysis, high-performance computing, and reproducible research practices would strengthen the ability of Nigerian researchers to participate effectively in international collaborations involving data-intensive nuclear science.

Case Studies and Success Stories

Published studies demonstrate that data science has become an important component of nuclear research across several application areas, including reactor analysis, radiation transport modeling and particle physics, nuclear medicine, and radiation protection. Rather than replacing conventional nuclear physics methodologies, data-driven techniques enhance the efficiency, accuracy, and interpretability of experimental and computational investigations. These examples collectively illustrate the increasing importance of interdisciplinary collaboration between nuclear physicists, computer scientists, statisticians, and engineers in addressing complex scientific problems. The studied literature also underscores the need of comprehensive data management throughout the study lifecycle. Effective collection, storage, curation, sharing, and preservation of scientific datasets are important needs for reproducible research and long-term scientific development. Institutions that have embraced structured data management systems claim improved collaboration, enhanced research transparency, and greater options for secondary data analysis. These experiences provide significant insights for countries looking to boost their national research ecosystems through digital transformation. For Nigeria, published examples demonstrate that collaboration among universities, research institutes, regulatory agencies, and foreign organizations can greatly increase national competence in nuclear science. Such agreements provide access to specialist knowledge, computer infrastructure, and shared research facilities while fostering the creation of sustainable scientific networks. Consequently, joint data-sharing activities should be considered as strategic expenditures that enhance both scientific excellence and national development.

Data-Driven Innovations in Nigerian Nuclear Research

Recent studies reveal rising interest in applying computational tools and data-driven approaches to nuclear research in Nigeria. Published work has addressed reactor physics analysis, neutron transport calculations, radiation measurement, computational modeling and simulation-based research that help nuclear engineering and radiation science. These papers show the growing importance of computational tools in solving scientific concerns, where experimental research may be limited by infrastructure or resource constraints. Computational modelling has been especially useful in reactor physics. Simulation tools allow researchers to study reactor behavior, optimize operating parameters and analyze nuclear interactions across multiple situations. Such methods decrease the necessity for large-scale experimental campaigns and supply accurate data for reactor analysis and safety evaluation. As computer resources continue to grow, prospects for the integration of machine learning with conventional reactor physics models are projected to significantly increase. The literature also implies that collaboration between Nigerian universities and international partners in research activities is an effective tool for improving national expertise in computational nuclear physics. These agreements promote information exchange, provide greater access to modern software and research infrastructure, and improve the quality and visibility of scientific products. This implies that further investment in collaborative research programmes, computational infrastructure and human capacity development will likely promote the adoption of data science approaches in the Nigerian nuclear research community.

CONCLUSION

With the increasing integration of data science into nuclear physics, new opportunities are emerging to improve the analysis, interpretation, and use of nuclear data across a wide range of scientific and technical fields. This review synthesizes recent material showing that machine learning, artificial intelligence, statistical modelling and sophisticated computational approaches have become useful complements to traditional theoretical and experimental methodologies. Applications include reactor monitoring, radiation protection, particle physics, nuclear medicine, environment monitoring and predictive modelling where they improve analytical accuracy, computational efficiency and decision making. Furthermore, the examined literature indicates that there is a significant opportunity in the use of data-driven approaches to enhance nuclear research in Nigeria. However, achieving these benefits depends on continued investment in computational infrastructure, sophisticated data management systems, high-performance computing facilities and specialized human resources. Furthermore, strengthening multidisciplinary education and developing international collaborations would be vital to enable Nigerian researchers to engage more effectively in global nuclear science projects. Future work in this area should be directed at the integration of machine learning and traditional nuclear physics approaches, better access to high-quality nuclear data sets, the development of explainable artificial intelligence for nuclear applications, and fostering reproducible computational research techniques. Such initiatives would promote scientific innovation and improve the application of nuclear technologies in health care, environmental protection, agriculture, industry and national energy development. Overall, the examined literature shows that data science is becoming an integral part of current nuclear physics and has a great potential to speed up scientific progress and sustainable development in Nigeria and other growing nuclear programmes.

REFERENCES

Asuku, A., Ibrahim, Y. V., Jonah, S. A., Umar, S., & Simon, J. . (2023). Application of SCALE code for reactor physics analysis of the as built Nigerian Miniature Neutron Source Reactor running with low enriched uranium fuel. Nuclear Engineering and Design, 411, 112416. https://doi.org/10.1016/j.nucengdes.2023.112416

Bae, J., Kim, G., & Lee, S. J. (2021). Real-time prediction of trends of nuclear power plant parameters following operator operations. Expert Systems with Applications.

Battistoni, G., Toppi, M., & Patera, V. (2021). Measuring the impact of nuclear interaction in particle therapy and radioprotection in space: the FOOT experiment. Frontiers in Physics. frontiersin.org

Bello, S., Simon, J., Aliyu, A. S., Abdulqadir, A., Rilwan, U., & El-Taher, A. (2024). The use of nuclear energy to solve Nigeria's energy crisis and help the country achieve its SDGs. J Rad Nucl Appl, 9, 77-28. researchgate.net

Bruce, P., Bruce, A., &Gedeck, P. (2020). Practical statistics for data scientists: 50+ essential concepts using R and Python. STM-marseille.com

Deiana, A. M., Tran, N., Agar, J., Blott, M., Di Guglielmo, G., Duarte, J., & Warburton, T. K. (2022). Applications and techniques for fast machine learning in science. Frontiers in Big Data, 5, 787421. frontiersin.org

Efron, B. & Hastie, T. (2021). Computer age statistical inference, student edition: algorithms, evidence, and data science. academia.edu

Gohel, H. A., Upadhyay, H., Lagos, L., Cooper, K. & Sanzetenea, A. (2020). Development of machine learning predictive maintenance framework for nuclear facilities. Nuclear Engineering and Technology, 52(7), 1436-1442. https://doi.org/10.1016/j.net.2019.11.023

Kahn, Y., Monzani, M. E., Palladino, K. J., Anderson, T., Bard, D., Baxter, D., &Westerdale, S. (2022). Snowmass2021 Cosmic Frontier: Modeling, statistics, simulations, and computing needs for direct dark matter detection. arXiv preprint arXiv:2203.07700. [PDF]

Lee, B., Yoon, S., Lee, J. W., Kim, Y., Chang, J., Yun, J., & Lee, J. H. (2020). Statistical characterization of the morphologies of nanoparticles through machine learning-based electron microscopy image analysis. ACS Nano, 14(12), 17125-17133. acs.org

Sarker, I. H. (2021). Data science and analytics: an overview from data-driven smart computing, decision-making, and applications perspective. SN Computer Science. springer.com

Terrani, K. A., Capps, N. A., Kerr, M. J., Back, C. A., Nelson, A. T., Wirth, B. D., & Stanek, C. R. (2020). Accelerating nuclear fuel development and qualification: Modeling and simulation integrated with separate-effects testing. Journal of Nuclear Materials, 539, 152267. sciencedirect.com

Zharenov, D. (2023). Interaction of antiprotons with nuclear matter within the INCL model. hal.science