Human Factors in Energy: Oil, Gas, Nuclear and Electric Power, Vol. 54, 2022, 56–64
https://doi.org/10.54941/ahfe1002221

Humans with/as Big Data in Nuclear Energy

Ronald L. Boring[1], Torrey J. Mortenson[1], Thomas A. Ulrich[1], and Roger Lew[2]

[1]Idaho National Laboratory, Idaho Falls, ID 83415, USA
[2]University of Idaho, Moscow, ID 83844, USA

ABSTRACT

Data are considered big when they contain significant variety, volume, and velocity in comparison with standard datasets. Big data are used by humans, but humans are also potential sources of big data. In nuclear power, much of the current research and application of big data principles focuses on instrumenting additional sensors or analyzing and visualizing the resulting data in useful ways. Gathering data on hardware enables enhanced diagnoses, planning, and prognoses, leading to greater efficiency through reduced maintenance costs and optimized power production. In this paper, we explore human interactions with big data, as well as potential ways that human performance data can inform nuclear power operations. This latter case considers humans as sources of big data.

Keywords: Big data, Nuclear power plant, Reactor operators, Data sources

INTRODUCTION

The push for big data methods in nuclear power plants (NPPs) is driven by advancements in sensor and diagnostic capabilities and the ready, off-the-shelf availability of deep learning algorithms and hardware systems. In this sense, more data mean more information about plant states and processes, ensuring plant safety by detecting current trends or predicting emerging anomalies (Hu et al., 2021). However, while computational and data handling methods and tools continue to adapt and grow, the capability of human operators does not. One critical aspect of the trend toward increased big data usage is that, for big data to be successful, the eventual users and consumers must consider such data to be both meaningful and useful.

Another aspect of big data is that users may themselves serve as sources of operational data. Data are acted upon by humans, and those actions form the basis for an added data dimension. Human performance measures in nuclear power remain largely within the realm of human factors research or qualification evaluations, and are often used in assessing the risk of human error or designing systems and interfaces for human operators. So far, such human operational data have not been used to inform algorithmic approaches to monitor and control NPPs. In this paper, we explore the relationship between big data and human operators at NPPs, viewing those humans as both users and sources of big data.

A challenge in advanced computational approaches such as artificial intelligence (AI) and machine learning (ML) resides in providing transparency and communication between the machine systems and human users. Many of these computational techniques rely on complex algorithms and advanced mathematics, potentially making it extremely difficult to explain and interpret the outputs to general user groups, especially ones lacking in relevant background knowledge and experience. The field of explainable AI seeks to make AI processes and outputs more transparent to various types of end users (Wickramasinghe et al., 2021). Human factors plays an important role in understanding user needs and designing AI interfaces or systems that closely align with human cognitive performance and abilities (Liao and Varshney, in press). This is the first element of big data in nuclear—the opportunity to design AI systems that truly benefit reactor operators and other plant personnel.

The second area of opportunity explained herein frames humans as data sources for the advanced computational techniques mentioned above. While measuring human performance can be intrusive at times, nonintrusive measures of performance, such as reactor operators' path on procedures or the relationship between operator actions and plant parameters, can be collected without necessitating additional or prespecified operator interactions with a data logging system. Such data, which should not represent a privacy risk to the individual or be used punitively, may enable predictions of future plant states or the enhancement of current human-system interactions. For example, logging computerized procedures in the main control room could anticipate field worker tasking and ready them for future actions. Frequent checking of specific systems or components by a reactor operator may indicate an anomaly that requires mitigation. Operators are often aware of when the plant is behaving unusually due to a suspected degraded component, but such insights are rarely codified as part of standard operations (Boring et al., 2019a). Yet, this expertise in monitoring emergent plant behaviors may ultimately prove as insightful as prognostic monitoring of components. Tracking how operators interact with systems unlocks a new source of data to feed into prognostic analyses. Similarly, outside of the main control room, nonintrusive technologies such as computer vision may be used in field work to track concentrations of gathered workers to prevent bottlenecks or inefficiencies (Sun et al., 2020).

In this paper, we first explore how humans interact with big data tools in a nuclear power context. We then identify possibilities for using human performance data as part of plant analytics. We conclude by outlining key considerations for future research aimed at using operational big data to design better plant interfaces.

HUMANS WITH BIG DATA

Oracle defines big data as "data that contain greater variety, arriving in increasing volumes, and with more velocity" (2022). The scale of the data is certainly bigger than with conventional sensor data, but the data are also more varied and frequently gathered. Processing such data can be particularly challenging without robust computational and data science infrastructure, and humans will often find difficulty in attempting to make sense of so many competing information points. The latter is a key concern in nuclear power, as the burden for decision explainability is as high as for decision reliability.

Today, a proliferation of research exists on big data in NPPs. Such research often focuses on the data tools, such as those employed for capturing new and advanced sensor data (Agarwal et al., 2017) or applying ML methods to plant prognostics and health management (Zhao et al., 2021). One factor missing from this conceptualization of big data is the consideration of humans. Where do humans fit into this approach to data collection and processing? And, what do they need to properly process the final product(s) of big data analytics? The most important consideration is that big data do not exist solely for their own benefit but are ultimately used by humans for a specific purpose, usually one related to decision making. As such, big data implementations that fail to consider the human in the loop risk becoming impediments rather than aids to human data users. Failure to consider the human user undermines the foundational reason behind big data.

It can be argued that in the context of NPPs, big data are not entirely new. The main control rooms of NPPs were conceived to be able to capture all the plant's indicators and controls in one location (Boring et al., 2016). Incorporated within the main control room of current-generation plants are thousands of data points, some representing controls, and others representing the indicators. The human is the point of synthesis for those various data sources. The human operators control and monitor the plant, ensuring safe and efficient operation of a complex process.

The impetus for control rooms was the need to consolidate large volumes of information sources and controls to support complex and safety-critical systems. The fundamental nature of control rooms has remained largely unchanged, even amid the transition of control room technology from analog to digital systems. While there is great variability, a typical main control room in an NPP consists of approximately 3,000 indicators and controls. With the advent of digital technology now comes the opportunity to improve and increase the various sensors installed. However, any increase in the number of sensors introduces the need for more sensor instrumentation to convey the data to the reactor operators in the control room. Analog instrumentation usually represents a one-to-one mapping between sensors and indicators. In most plants, it is simply infeasible to add more indicators to the existing panel real estate in control rooms. Digital displays may provide a solution to this problem, as they often nest information across multiple screens or windows rather than displaying it all at once. On the other hand, operators must navigate to the nested information screens, creating the potential for important sensor information to become buried on screens not currently being viewed.

Therein resides a quandary for big data implementation. Vendors and plants have the opportunity to add sensors, but these new sensors may exceed the threshold of operator perception or may overload operators with too much information (Medema et al., 2019). At the same time, one of the biggest economic challenges (especially in new reactor designs) has been finding ways to reduce the staffing levels required to operate NPPs. Simply adding more sensors without considering the human users of those sensors results in increased workload, undermining the move toward optimized staffing levels.

Is the solution to add advanced analytics to encompass the increased number of sensors? Though a better approach, it can still cause operators to be overwhelmed with additional information if other information is not removed first. Developers of big data analytics cannot create more data and hope to maintain the complexity of that environment at a steady level for humans simply by adding operator tools. Ideally, big data analytics should decrease the operator burden. This can be accomplished by consolidating multiple sensors into a single indicator or by automating certain functions performed by the human operators.

The question then becomes, do we remove the human in the loop via automation? Taking the operator out of the sensor-and-indicator loop risks eliminating those tools operators need to make effective decisions. In other words, improperly implemented automation threatens to remove the contextual and situation awareness elements on which operators rely to make informed decisions on controlling the plant. Automation is not a full surrogate for humans in terms of decision making. A proper balance must be reached between adding information, distilling it to operators in reasonable amounts, and maintaining the operators' key role in making decisions and controlling the plant.

Human factors for automation is based on an information processing conceptualization of the human, interface, and machine (Card et al., 1983). Humans have sensory inputs through which they perceive information. This information is processed cognitively, and decisions or strategies are developed. Some of the decision responses are translated into physical actions taken by the operators, which are the control actions that serve as the inputs to the control system in this conceptualization.

A counterpart to human information processing exists in the form of the control system. Human actuations of controls serve as the inputs to the system. The control system of course has its own processing and reactions. Certain functions such as the reactor trip may be automated, using sensor data without human intervention, although manual trips should also remain an option. The control room contains various outputs that manifest as indicators. These indicators are actually the inputs to the operator's sensory system. Most system outputs to the operator are visual in nature and take the form of instrumented gauges, status indicators, and alarms, but may also include audible alarms. This interplay forms a control loop between the human and the control system. Of course, the human does not typically operate in isolation with the control system. The human may interact with other humans, with multiple subsystems, or with operator aids (e.g., operating procedures) that guide them in interfacing with the system. This overall feedback loop forms the current concept of operations for NPPs.

Increasing automation changes this concept of operations, not always in an effective manner. Boring (2011) considered information foraging theory in the context of nuclear power plant control rooms, which can illustrate how automation may disrupt optimal operator performance. In this exploration of information foraging, operators are classified symbolically as either wolves or spiders, which reflects the way that they interact with information. A wolf hunts for its prey. In terms of an information analogy, a wolf is someone who seeks information; it actively pulls the information it needs. By analogy, this lupine operator actively seeks specific indicators needed to make decisions and take actions. The counterpart to the wolf is the spider. The spider builds a web and waits for its prey to come to it. In the information foraging analogy, the symbolically arachnid operator subscribes to information or monitors information that is relevant to him or her. They monitor indicators and await information such as alarms to be pushed to them. The control room interface is built around these basic roles of pushing and pulling information. In the wolf role, the operators actively seek information in the control room to support the diagnosis of plant states. In the spider role, the operators passively receive information in the control room from alarms or other indicators. The challenge of automating analytics is that it potentially forces the operators into the passive spider role. A spider may miss the active process of capturing operational information, which supports its situation awareness and vigilance. Eliminating the seeking activity also potentially reduces the ability of the operator to make timely decisions in response to changing conditions at the plant.

Humans and automation have been the subject of formal study since at least the time of Fitts' List (1951). This list sought to catalogue what humans do well vs. what machines do well. In many areas, technologies such as ML are narrowing the distinction between machine and human performance, and there is opportunity to automate many functions in new control systems. To allocate human activities to AI systems, two key types of automation (Boring et al., 2019b) should be considered. Control automation takes actions that would normally be performed manually by human operators. In turn, information automation consolidates indicators so as to give operators the information they need to make decisions at a glance. Big data analytics may be used to support control automation, but a large portion of big data is informational in nature and primarily serves to guide operator decision making. One takeaway is the necessity for big data analytics to keep the operator in the loop for activities (e.g., decision making) in which human input is necessary. If plant safety is a function of key decisions made by human operators, those operators must retain that responsibility, and automation should support rather than replace their role.

HUMANS AS BIG DATA

The previous section highlighted ways in which big data can assist human operators. However, humans can also serve as sources of big data, and human factors research can help access and extract these data so as to turn them into meaningful information. The actions of operators leave traces that are capturable by the digital data historian in the control system. For example, controls that are manually actuated by human operators are logged. These logs can be considered alongside plant parameters to understand the context of operations. Operators in NPPs make frequent use of operating procedures, whether paper-based or computerized. Steps taken through procedures can be tracked and logged for informational purposes. Humans may also leave traces in other, more novel ways. For example, humans equipped with physiological monitoring may provide information pertaining to how they interface with the system. A classic example of this is eye tracking. While it is not standard to equip reactor operators with eye tracking devices, their use in experimental research has proven valuable in determining what operators look at and for how long (Kovesdi et al., 2018). Such insight into an operator's attention can generate extremely useful information on plant operations and interface usage. For example, if an advanced visualization is displayed with a variety of data points but operators do not spend much time focusing on it, the human factors team can infer that the visualization is not of use to the operators at that moment. Other operator performance data may be analyzed to understand which paths operators took and with which systems they interfaced. These may be cataloged to develop patterns of operation. Beyond the research realm, human-centered performance data can support analytic tools at the plant, informing plant systems as to the nature of various human activities, anticipating plant changes, and dynamically adjusting information and control automation.

There is no shortage of data to be collected on human operators in control rooms. The *Guideline for Operational Nuclear Usability and Knowledge Elicitation* (GONUKE; Boring et al., 2015) offers a method of understanding and classifying the types of collectible information from humans in NPPs. GONUKE is used primarily as an evaluation tool during the development of advanced control rooms. However, the human performance measures it identifies can also be used for real-time operational monitoring. GONUKE delineates early- and late-stage development activities. Early-stage (i.e., formative-stage) development activities tend to benefit more from qualitative data such as operator preferences for particular designs. Late-stage evaluations (i.e., summative evaluations) tend to be more quantitative. The goal of a summative evaluation is to ascertain whether the completed design functions as intended and whether operators can perform their tasks safely. Quantitative measures can include things such as time-on-task and error rates. GONUKE emphasizes the importance of both qualitative and quantitative data. Qualitative data represent the perceived experience of the operator and can be an extremely valuable source of insight (Ulrich et al., 2018). While no magic tool exists to peer into the black box of human cognition, some measures can be obtained in real time to ensure that the operator understands the necessary information and is using it in a meaningful way. Qualitative measures may require some level of intrusiveness, such as periodic surveying of operators. Quantitative measures are generally easier to obtain in a less intrusive manner, without needing to interrupt the operator. However, they may lack depth, as they do not always indicate the reasoning behind the operators' actions.

Through their actions, operators indirectly communicate several types of data. For example, they are conveying the problem on which they are focused, their understanding of it (or lack thereof), their situation awareness, their stress levels, their context-specific knowledge, and even their performance. Such information is part of the overall puzzle of plant operations and can serve as big data to complement the other plant data already being gathered. Operator performance data can tell us how to tailor our big data to operators by understanding the context surrounding what they are doing and anticipating what plant data they need. Such tailoring could enable context-dependent information visualizations and even respond dynamically with different levels of automation. Thus, human operators represent an untapped source of information and big data.

CONCLUSION

Big data should be seen as a two-way street. Humans—particularly operators—are the users of big data. Big data should thus be tailored to provide information that truly supports operators in making decisions and conducting accurate, safe, and timely responses. Conversely, operators are also potential sources of data for the system. Control room operations currently fail to capture many of the details regarding operator actions. Such information could provide the specific operational context that enables a smart system to appropriately tailor its information to the operator.

There remains a large opportunity to improve the diversity and depth of our data and to inform considerations of big data and ML applications. Explainable AI is not just about conveying the outcomes of big data analytics to operators; it is also about explaining how the operator interacts or intends to interact with the ML system. Big data must consider the full interplay between the human input and output with a dynamic control system.

To conclude, this paper offers three key takeaways that should be considered in future research and development conducted on big data applications for NPPs. First, innate human limitations mean that operators cannot process infinite amounts of data, and data must be distilled to be meaningful and useful. Second, there remain important research challenges to better understand how human operators interface with analytic systems and team with AI or automation, not to mention how best to design these interactions in a manner that maximizes the system's overall effectiveness and usability. Lastly, human actions can be a source of data to deepen our understanding and coordination of system characteristics, which are not included in current sensor strategies or datasets. Thus, it is a necessary and worthwhile to explore a framework for collecting and using humans as sources of big data.

ACKNOWLEDGMENT

This work of authorship was prepared as an account of work sponsored by Idaho National Laboratory (under Contract No. DE-AC07-05ID14517), an agency of the U.S. Government. Neither the U.S. Government, nor any

agency thereof, nor any of their employees makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe on privately owned rights.

REFERENCES

Agarwal, V., Buttles, J.W., Beaty, L.H., Naser, J., and Hallbert, B.P. (2017) Wireless online position monitoring of manual valve types for plant configuration management in nuclear power plants, IEEE Sensors Journal, Volume 17, No. 2, pp. 311–322.

Boring, R.L. (2011) Information foraging in nuclear power plant control rooms. Proceedings of 2011 European Safety and Reliability (ESREL) Conference: Advances in Safety, Reliability, and Risk Management, pp. 654–660.

Boring, R.L., Ulrich, T.A., Joe, J.C., and Lew, R.T. (2015) Guideline for operational nuclear usability and knowledge elicitation (GONUKE), Procedia Manufacturing, Volume 3, pp. 1327–1334.

Boring, R., Ulrich, T., and Lew, R. (2016) RevealFlow: A process control visualization framework. Lecture Notes in Artificial Intelligence, Volume 9744, pp. 145–156.

Boring, R.L., Ulrich, T.A., Medema, H.M., and Lew, R. (2019a) Operator resilience to cyber interdictions in nuclear power plants. Proceedings of IEEE Resilience Week 2019, pp. 247–251.

Boring, R.L., Ulrich, T.A., and Mortenson, T.J. (2019b) Level-of-automation considerations for advanced reactor control rooms. Proceedings of the 11th Nuclear Plant Instrumentation, Control and Human-Machine Interface Technologies (NPIC&HMIT 2019), pp. 1210–1221.

Card, S.K., Moran, T.P., and Newell, A. (1983) The Psychology of Human-Computer Interaction. Mahwah, NJ: Lawrence Erlbaum Associates.

Fitts, P.M. (1951) Human engineering for an effective air navigation and traffic control system. Washington, DC: National Research Council.

Hu, G., Zhou, T., and Liu, Q. (2021) Data-driven machine learning for fault detection and diagnosis in nuclear power plants: A review. Frontiers of Energy Research, Volume 9, Article 663296.

Kovesdi, C., Spielman, Z., Le Blanc, K., and Rice, B. (2018) Application of eye tracking for measurement and evaluation in human factors studies in control room modernization. Nuclear Technology, Volume 202, pp. 220–229.

Liao, Q.V., and Varshney, K.R. (In press) Human-centered explainable AI (XAI): From algorithms to user experiences. Book chapter.

Medema, H., Savchenko, K., Boring, R., Ulrich, T., and Park, J. (2019) Human reliability considerations for the transition from analog to digital control technology in nuclear power plants. Proceedings of the 11th Nuclear Plant Instrumentation, Control and Human-Machine Interface Technologies (NPIC&HMIT 2019), pp. 132–141.

Oracle Corporation. (Retrieved March 8, 2022) What is Big Data? https://www.oracle.com/big-data/what-is-big-data

Sun, Z., Zhang, C., Chen, J., Tang, P., and Yilmax, A. (2020) Predictive nuclear power plant outage control through computer vision and data-driven simulation. Progress in Nuclear Energy, 127, Article 103448.

Ulrich, T., Boring, R., and Lew, R. (2018) Qualitative or quantitative data for nuclear control room usability studies? A pragmatic approach to data collection and presentation. Proceedings of the 62nd Human Factors and Ergonomics Society, pp. 1674–1678.

Wickramasinghe, C.S., Amarasinghe, K., Marino, D.L., Rieger, C., and Manic, M. (2021) Explainable unsupervised machine learning for cyber-physical systems, IEEE Access, Volume 9, pp. 131824–131843.

Zhao, X., Kim, J., Warns, K., Wang, X., Ramuhalli, P., Cetiner, S., Kang, H.G., and Golay, M. (2021) Prognostics and health management in nuclear power plants: An updated method-centric review with special focus on data-driven methods. Frontiers of Energy Research, Volume 9, Article 696785.