Human Factors and Simulation, Vol. 83, 2023, 226–235
https://doi.org/10.54941/ahfe1003580

Human Factors for Machine Learning in Astronomy

John Wenskovitch[1,2] and Amruta Jaodand[3]

[1]Pacific Northwest National Laboratory, Richland, WA 99354, USA
[2]Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA
[3]California Institute of Technology, Pasadena, CA 91125, USA

ABSTRACT

In this work, we present a collection of human-centered pitfalls that can occur when using machine learning tools and techniques in modern astronomical research, and we recommend best practices in order to mitigate these pitfalls. Human concerns affect the adoption and evolution of machine learning (ML) techniques in both existing workflows and work cultures. We use current and future surveys such as ZTF and LSST, the data that they collect, and the techniques implemented to process that data as examples of these challenges and the potential application of these best practices, with the ultimate goal of maximizing the discovery potential of these surveys.

Keywords: Astronomy, Machine learning, Human factors, Trust, Reproducibility, Bias

INTRODUCTION

The introduction of machine learning (ML) into the sciences has had an immense impact on the process of knowledge discovery. This is particularly true in astronomy, where researchers must analyze the vast quantities of data that continue to be acquired and cataloged by astronomical surveys (Poudel et al., 2022). The amount of data collected by astronomers continues to increase as a result of many factors, including the construction of larger telescopes which enable the detection of fainter objects, viewing distant objects with greater clarity through advancements such as adaptive optics, and technological discoveries that enable observations across non-visible wavelengths of light (Biewald, 2021). With the petascale era of astronomical data already at hand and the exascale era soon to follow, human-directed computation alone is insufficient (Ivezić et al., 2020). Modern astronomy is therefore increasingly dependent on ML assistance for processing, prioritization, and labeling in order to gain insight into the secrets of the universe.

Exemplar data-filled discovery spaces for rapid astronomical transients are provided by surveys such as the Zwicky Transient Factory (ZTF) and the Large Synoptic Survey Telescope (LSST). In particular, ZTF improves on its predecessor by increasing the number of visits to every field of view and picking out optical transients up to ≤ 18.5 magnitude (Bellm et al., 2019). In ZTF data alone as of the most recent data release (DR 15), there are 4.58 billion unique astronomical sources, with time series data gathered for each source (Zwicky Transient Facility, 2023). Moreover, ZTF also generates about one million alerts per night, far beyond a data scale that can be supported by human-only analysis. Searching for interesting astronomical objects is therefore data-intensive and has required various ML-based approaches such as RealBogus classification to identify spurious sources, periodicity searches for Galactic objects, and anomaly detection algorithms for any photometric variabilities. ZTF in many ways is a predecessor for the upcoming LSST. LSST will take the data-driven discovery potential to further and fainter transients at ≤ 27 magnitude. It is set to conduct trillions of observations of billions of stars and galaxies over its 10-year operation timeline (Ivezić et al., 2019).

However, because ML algorithmic decisions must be paired with human domain expertise, human concerns affect the adoption and refinement of ML techniques into existing processes. Established workflows and work cultures are often resistant to change, impacting the introduction of ML. Even when ML is embraced, mistrust in novel models and slower adoption of novel techniques biases researchers towards selecting known or popular techniques that may not apply to new data. Alternatively, overconfidence in new technology can lead to approval of the latest techniques without thorough validation or guarantees of reproducibility. A constant concern among astronomers is that errors in their understanding of software, and the conclusions that software draws, might lead to errors in their understanding of how the universe works (Biewald, 2021).

In the sections that follow, we discuss the role of ML in modern astronomy research, identifying ways in which the usage of ML connects to a selection of human-centric challenges and suggesting Human Factors (HF) techniques to assist in overcoming these barriers. Each section begins with a challenge related to the use of machine learning in astronomy, which is then followed by related HF knowledge, tools, and techniques. We note that generating a complete list of astronomical machine learning challenges is beyond the scope of this work; instead, we present a selected list of challenges from our own experiences.

INITIAL MACHINE LEARNING ADOPTION

The use of computers and computation in astronomical research is not a recent development. The necessity to store and manipulate vast quantities of survey data has led to collaborations between astronomers and computer scientists that have been ongoing for decades, and this includes the use of ML for astronomical research. However, the rate at which ML (and deep learning in particular) is finding a role in astronomy projects and publications is accelerating rapidly (Smith and Geach, 2022). Successful research on cutting-edge astronomical challenges nearly always requires the usage of advanced computation. And yet while academic changes are gradually being introduced, the current curriculum used to instruct astronomers does not provide sufficient training in machine learning, statistics, or even software engineering (Biewald, 2021). In addition to recent graduates not receiving the education and support necessary for success in modern astronomical research, established researchers also struggle with limited resources for continuing education, lacking both training resources and a thorough primer on current best practices for the use of ML (National Academies of Sciences, Engineering, and Medicine, 2021). Overall, the introduction of technology into existing workflows presents a well-known challenge (Wickens, 2004), and even identifying locations in a workflow that will benefit from new technology can prove to be challenging (Liebowitz, 2000).

Researchers in HF have long studied impediments to the adoption of technology, dating back at least to Rogers' *Diffusion of Innovation* (Rogers, 1962). Determining the underlying source of the barrier is the first step necessary to overcome the challenge. Ertmer proposes a pair of adoption categories, with the first group related to external factors such as a lack of institutional support or time to learn new technologies, and the second group related to internal perceptions and attitudes such as resistance to change and the perception of the new technology (Ertmer, 1999). Noting whether an adoption obstacle is due to external constraints or internal human bias plays a role in building a strategy to move forward.

When introducing machine intelligence into an established workflow, a further concern is the reduction in human decision-making autonomy that accompanies sharing a workload with untrusted technology (BenMessaoud et al., 2011). One of the outstanding concerns for the use of machine learning in astronomy and generally any field that deals with copious amounts of data is the black box nature of it. While we save the discussion of ML trust and interpretability for another section, it is noteworthy to identify this common point that can lead to a lack of confidence in the results of the research and the quality of the process.

ADAPTING TO NEW TECHNOLOGY

While introducing ML into an established workflow is difficult, it may be even more challenging to update that existing ML with novel techniques. In addition to adopting techniques that are more efficient or more accurate, this can also include introducing new techniques that are more appropriate for the current data under investigation. The costs that are involved in switching to new models can range from financial to temporal. Common concerns that are voiced when transitioning to a new technology include concerns about loss of productivity while integrating a new tool into a workflow (Snoeyink and Ertmer, 2001), information flow disruption and data continuity, and issues with data standardization and interoperability. Additionally, when experts adopt new paradigms, their workload often drastically increases during the adjustment process (Ludwick and Doucette, 2009).

Another important issue with astronomical data is the high dimensionality and large volume, along with data gaps leading to information loss (Goodman, 2012). Hence, changes in ML-based approaches while analyzing the data can lead to vastly different outcomes. As a result, astronomers often find themselves working with approaches which are well-known, well-tested, and with high signal-to-noise ratio datasets. However, these approaches may not be efficient or accurate for datasets with large uncertainties. A common example of this is the use of random forest classification in astronomy. While random forest classification algorithms are fast to train and seem to outperform other classifications, they are inept at handling labeling uncertainties (Reis et al., 2018), and their performance lags behind more modern deep learning techniques (Carrasco-Davis et al., 2019).

Task analysis represents one HF method that can assist in overcoming such issues. The goal of task analysis is to understand and provide insight into the processes involved in the completion of a task (the astronomical research challenge under investigation), in order to identify and optimize the overarching workflow (Stanton et al., 2017). Given a thorough understanding of the process, task analysis can help to identify the advantages and disadvantages of altering that workflow. Indeed, the workflow itself represents a human-centric process of which the choice of ML technology is only one component, and so a thorough understanding and analysis of the full process is necessary when determining the best data science models and ML technology that can work collaboratively with researchers.

As with task analysis, Cognitive Work Analysis (CWA) represents a similar approach towards understanding the cognitive skills and strategies used by humans in their workflows, providing a collection of methods to understand tasks, actors, strategy, and cooperation in complex systems (Vicente, 1999). Charting techniques such as process charts and event tree analysis also support understanding these workflows and interactions in a visual medium, clarifying actions and events that both humans and ML own (Kirwan and Ainsworth, 1992). As technology continues to evolve and machine intelligence continues to develop, these approaches can also assist in modifying the current research paradigm in which ML is a tool into a new model where ML is a partner, an area of research known as human-machine teaming (Lyons et al., 2019).

TRUST IN MACHINE LEARNING PROCESSES AND RESULTS

While many shallow-learning ML models have somewhat interpretable properties, many of the frequently-used deep learning models in current astronomical research are much more opaque. It is therefore difficult to understand the decisions made by complex ML models, yet alone to understand the processes undertaken to reach those decisions (Goebel et al., 2018). Further, uncertainty in interactions between black box models and humans can be extended by ML that is unaware of human goals or processes (Wenskovitch and North, 2020). This lack of understanding leads to a variety of issues in astronomical research, including issues with trust in the research outcomes, questions about the accuracy of the computations, and concerns in detecting errors in the process.

Establishing trust is further complicated by data, as astronomical datasets are heteroskedastic, non-homogenous, and contain gaps. It therefore gets complicated to quantify the underlying statistical noise uncertainties using machine learning language of aleatoric and epistemic uncertainties (Caldeira and Nord 2020, Dvorkin et al., 2022). Therefore, while the ML methods are being widely adopted by the astronomy community, confidence in these methods is lacking. As Caldeira and Nord suggest, in deep learning we can reduce the gap with aleatoric and statistical uncertainties by ensuring a wide range of noise realizations. However, for epistemic uncertainties originating from the data, the problem is still tough to resolve and requires specific understanding of the scientific problem (Chen et al., 2022). For example, the distribution of a number of parameters including orbital periodicity, eccentricity, amplitude, and more require estimation in advance of data collection. Because of data gaps, noise, and limitations in astronomers' understanding of how uncertainty cascades through the parameter space, the actual detection of transients in large-scale surveys such as ZTF may not map well to expected detection rates (Feindt et al., 2019).

The nascent field of Explainable Artificial Intelligence (often styled as XAI) provides a human-centered vision for the future interpretability of ML models, seeking to explain how or why a model has generated its output (Longo et al., 2020). Indeed, effectively opening the "black box" of deep learning represents one of the most significant opportunities for the impact of HF on ML in the sciences. Providing interpretations and evidence for the decision-making process of a complex model in a way that astronomers can easily comprehend can lead to the identification of issues in the model that may require correction, as well as providing new insight into relationships that underlie the data. Increased trust in the computation and the results represents a significant benefit to astronomers, while the ML models can also benefit from improved robustness and algorithmic performance when their latent properties are better understood (Phillips et al., 2020).

While explainability serves as a first step towards trusting an unfamiliar ML model, a number of other factors including stability and usability, accountability, security, and transparency play a role in developing trustworthy ML (Smith, 2019). Fundamentally, a usable and trustworthy technology must be ready for operational use by humans rather than experimental. The Human Readiness Level (HRL) scale provides a framework for capturing the maturity of a technology with respect to usability (Salazar et al., 2020), paralleling the earlier Technology Readiness Levels (TRL) and ranging from a Level 1 system (relevant human capabilities are identified) to a Level 9 system (the system was successfully used in operations with measured performance benefits). Trust itself can be measured by a variety of metrics (Damacharla et al., 2018), and scales such as the NASA Task Load Index provide a mental workload assessment to understand how humans are responding to working with technology (Hart and Staveland, 1988).

REPRODUCIBILITY

Reproducibility is not an issue tied solely to the use of machine learning, as research teams can come to differing conclusions from the same data and when testing the same hypothesis in many contexts (Breznau et al., 2022). That said, ML has had an outsized effect on reproducibility issues in the sciences in general (Kapoor and Narayanan, 2022) and in astronomy specifically. Reproducibility issues with ML can also include related technology concerns, such as documentation or human-centric misunderstandings about the capabilities of the ML. For example, while searching for periodic objects through millions of time series objects in optical surveys such as ZTF, although training, test and validation datasets might agree the absolute number of objects belonging to each of the periodic sub-class might differ widely with difference in periodicity search algorithms or detectability thresholds used by different collaborations.

Concerns with addressing reproducibility can be found in several HF techniques and approaches, such as establishing inter- and intra-rater reliability in task analysis and in assessing situational awareness (Stanton et al., 2017). As such, a number of methods exist to mitigate (but notably not to eliminate) reproducibility concerns when humans and technology intersect. Team performance analysis techniques such as team communication analysis (Jentsch and Bowers, 2004) represent one such methodology. Other techniques that are designed to assess and reduce human errors can be adapted to technology, including those designed for both retrospective and predictive analysis of errors (Shorrock and Kirwan, 2002). Additionally, charting techniques can assist in visually clarifying the roles of humans and technology in scientific workflows, and in the case of Decision Action Diagrams can also depict decision points and options available to both humans and ML (Kirwan and Ainsworth, 1992).

SAMPLE SELECTION BIAS

In addition to challenges resulting from a lack of clarity in ML decision-making, the inappropriate selection of data used to train astronomical ML models presents a mechanism for introducing error and bias into scientific processes. As detailed by Scaife (2020), the increasing sensitivity of astronomical surveys to fainter and more distant objects can cause issues with the use of previous knowledge for training models for these new surveys, since sample distributions in the previous less-sensitive surveys will be biased towards brighter object classes. As a result, new astronomical sources identified by these new surveys will be probabilistically more likely to be classified as objects that are more common in existing datasets, reducing both model accuracy and the chances of discovering new object classes and distributions (Clarke et al., 2020).

As this challenge is primarily a data collection issue, a natural tendency would be to look towards HF data collection techniques for a solution. However, most HF techniques focus on the collection of data from humans using interviews, questionnaires, and observations (Stanton et al., 2017). Therefore, assistance from HF to address this challenge can best be found in other classes of techniques. For example, some design techniques can be used or modified to the ML use case, incorporating special attention in the design stages of a new ML pipeline and/or human-ML workflow to assist in the development of software, equipment, and data selection. In particular, mission analysis (initially designed to analyze operational procedures and requirements for cockpit design) can assist in breaking down the goals of learning from new data into mission phases and operation modes, isolating task function requirements and understanding the necessary inputs for both identifying the required information for ML and for monitoring the accuracy of ML (Wilkinson, 1992). Similarly, processes from task-centered system design

(Greenberg, 2003) can be adapted for evaluating new system design concepts based upon existing systems and tasks. While not obviously as applicable, some situation awareness techniques can be modified to support an understanding of the effects of using existing data in new ML models, notably rating the quantity and information (Taylor, 1990) and in understanding the human's ability to locate and understand relevant information regarding the current state of the ML (Durso et al., 1998).

CONCLUSION

Looking to the future, we envision a new goal of human-centered machine learning for astronomy, in which the capabilities and strengths of both human and machine are optimized and balanced. Approaching the best use of the creativity of the human and the computational power of the machine necessitates both human insight into the processes of the machine and well as machine insight into the goals of the human (Wenskovitch and North, 2020). As astronomers look to emerging technology for integration into ongoing workflows, practices, and collaborations, a significant role exists for human factors experts to assist in mediating and enhancing the communication between researcher and automation.

In short, we assert that astronomers should consider utilizing the expertise of human factors researchers to the same degree that they look to machine learning researchers. By understanding and supporting human-centric processes, HF strives to identify challenges, formulate solutions, and generally create more refined processes for scientists to collaborate with technology (de Winter and Hancock, 2021). Notably, the LSST Collaboration has begun to fund fellowships in social science prior to the beginning of scientific operations as the Vera C. Rubin Observatory, supporting research into the ways that astronomers collaborate and analyze the substantial quantities of data that will result from this survey (LSST Corporation, 2022). In looking to the future of technological and computational support for data discovery, human factors will play a significant and critical role for successful projects and collaborations.

ACKNOWLEDGMENT

This work was supported by Chandra Observational Grant Award GO0-21067X, as well as a NuSTAR mission grant at Caltech.

REFERENCES

Bellm, E. C., Kulkarni, S. R., Graham, M. J., Dekany, R., Smith, R. M., Riddle, R., Masci, F. J., Helou, G., Prince, T. A., Adams, S. M. and Barbarino, C., 2018. The Zwicky Transient Facility: System Overview, Performance, and First Results. *Publications of the Astronomical Society of the Pacific*, *131*(995), p. 018002.

BenMessaoud, C., Kharrazi, H. and MacDorman, K. F., 2011. Facilitators and Barriers to Adopting Robotic-Assisted Surgery: Contextualizing the Unified Theory of Acceptance and Use of Technology. *PloS One*, 6(1), p. e16395.

Biewald, L. (2021). "Josh Bloom – The Link Between Astronomy and ML." [Podcast] *Gradient Dissent*, Weights & Biases. August 20. Available at https://wandb.ai/wandb_fc/gradient-dissent/reports/Josh-Bloom-The-Link-Between-Astronomy-and-ML--Vmlldzo4NDk5NjA [Accessed 22 January, 2023].

Breznau, N., Rinke, E. M., Wuttke, A., Nguyen, H. H., Adem, M., Adriaans, J., Alvarez-Benjumea, A., Andersen, H. K., Auer, D., Azevedo, F. and Bahnsen, O., 2022. Observing Many Researchers Using the Same Data and Hypothesis Reveals a Hidden Universe of Uncertainty. *Proceedings of the National Academy of Sciences*, 119(44), p. e2203150119.

Caldeira, J. and Nord, B., 2020. Deeply Uncertain: Comparing Methods of Uncertainty Quantification in Deep Learning Algorithms. *Machine Learning: Science and Technology*, 2(1), p. 015002.

Carrasco-Davis, R., Cabrera-Vives, G., Förster, F., Estévez, P. A., Huijse, P., Protopapas, P., Reyes, I., Martínez-Palomera, J. and Donoso, C., 2019. Deep Learning for Image Sequence Classification of Astronomical Events. *Publications of the Astronomical Society of the Pacific*, *131*(1004), p. 108006.

Chen, T. Y., Dey, B., Ghosh, A., Kagan, M., Nord, B. and Ramachandra, N., 2022. Interpretable Uncertainty Quantification in AI for HEP [online]. [Preprint]. [Viewed 13 February, 2023] Available from doi: https://doi.org/10.2172/1886020.

Clarke, A. O., Scaife, A. M. M., Greenhalgh, R. and Griguta, V., 2020. Identifying Galaxies, Quasars, and Stars with Machine Learning: A New Catalogue of Classifications for 111 Million SDSS Sources without Spectra. Astronomy & Astrophysics, 639, p. A84.

Damacharla, P., Javaid, A. Y., Gallimore, J. J. and Devabhaktuni, V. K., 2018. Common Metrics to Benchmark Human-Machine Teams (HMT): A Review. *IEEE Access*, 6, pp. 38637–38655.

De Winter, J. C. F. and Hancock, P. A., 2021. Why Human Factors Science is Demonstrably Necessary: Historical and Evolutionary Foundations. *Ergonomics*, 64(9), pp. 1115–1131.

Durso, F. T., Hackworth, C. A., Truitt, T. R., Crutchfield, J., Nikolic, D. and Manning, C. A., 1998. Situation Awareness as a Predictor of Performance for En Route Air Traffic Controllers. *Air Traffic Control Quarterly*, 6(1), pp. 1–20.

Dvorkin, C., Mishra-Sharma, S., Nord, B., Villar, V. A., Avestruz, C., Bechtol, K., Ćiprijanović, A., Connolly, A. J., Garrison, L. H., Narayan, G. and Villaescusa-Navarro, F., 2022. Machine Learning and Cosmology [online]. [Preprint]. [Viewed 10 February, 2023]. Available from doi: https://doi.org/10.48550/arXiv.2203.08056.

Ertmer, P. A., 1999. Addressing First- and Second-Order Barriers to Change: Strategies for Technology Integration. *Educational Technology Research and Development*, *47*(4), pp. 47–61.

Feindt, U., Nordin, J., Rigault, M., Brinnel, V., Dhawan, S., Goobar, A. and Kowalski, M., 2019. simsurvey: Estimating Transient Discovery Rates for the Zwicky Transient Facility. *Journal of Cosmology and Astroparticle Physics*, 2019(10), p. 005.

Goebel, R., Chander, A., Holzinger, K., Lecue, F., Akata, Z., Stumpf, S., Kieseberg, P. and Holzinger, A., 2018, August. Explainable AI: The New 42?. In *International Cross-Domain Conference for Machine Learning and Knowledge Extraction* (pp. 295–303). Springer.

Goodman, A. A., 2012. Principles of High-Dimensional Data Visualization in Astronomy. *Astronomische Nachrichten*, *333*(5-6), pp. 505–514.

Greenberg, S., 2003. Working Through Task-Centred System Design. In *The Handbook of Task Analysis for Human Computer Interaction*, London: Lawrence Erlbaum Associates Inc.

Hart, S. G. and Staveland, L. E., 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology* (Vol. 52, pp. 139–183). North-Holland.

Ivezić, Ž., Kahn, S. M., Tyson, J. A., Abel, B., Acosta, E., Allsman, R., Alonso, D., AlSayyad, Y., Anderson, S. F., Andrew, J. and Angel, J. R. P., 2019. LSST: From Science Drivers to Reference Design and Anticipated Data Products. *The Astrophysical Journal*, *873*(2), p. 111.

Ivezić, Ž., Connolly, A. J., VanderPlas, J. T. and Gray, A., 2020. "Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data, Updated Edition." Princeton University Press.

Jentsch, F. and Bowers, C. A., 2004. Team Communications Analysis. In *Handbook of Human Factors and Ergonomics Methods* (pp. 505–509). CRC Press.

Kapoor, S. and Narayanan, A., 2022. Leakage and the Reproducibility Crisis in ML-Based Science [online]. [Preprint]. [Viewed 05 February, 2023]. Available from doi: https://doi.org/10.48550/arXiv.2207.07048.

Kirwan, B. and Ainsworth, L. K. eds., 1992. *A Guide to Task Analysis: The Task Analysis Working Group*. CRC Press.

Liebowitz, J., Rubenstein-Montano, B., McCaw, D., Buchwalter, J., Browning, C., Newman, B. and Rebeck, K., 2000. The Knowledge Audit. *Knowledge and Process Management*, 7(1), pp. 3–10.

Longo, L., Goebel, R., Lecue, F., Kieseberg, P. and Holzinger, A., 2020, August. Explainable Artificial Intelligence: Concepts, Applications, Research Challenges and Visions. In *International Cross-Domain Conference for Machine Learning and Knowledge Extraction* (pp. 1–16). Springer.

LSST Corporation, 2022. Press Release: First Astrophysics Fellowship to Include a Social-Science Dimension Announces Awardees [online]. [Viewed 05 February, 2023]. Available from https://www.lsstcorporation.org/node/353.

Ludwick, D. A. and Doucette, J., 2009. Adopting Electronic Medical Records in Primary Care: Lessons Learned from Health Information Systems Implementation Experience in Seven Countries. *International Journal of Medical Informatics*, 78(1), pp. 22–31.

Lyons, J. B., Wynne, K. T., Mahoney, S. and Roebke, M. A., 2019. Trust and Human-Machine Teaming: A Qualitative Study. In *Artificial Intelligence for the Internet of Everything* (pp. 101–116). Academic Press.

National Academies of Sciences, Engineering, and Medicine, 2021. *Pathways to Discovery in Astronomy and Astrophysics for the 2020s*.

Phillips, P. J., Hahn, C. A., Fontana, P. C., Broniatowski, D. A., and Przybocki, M. A., 2020. Four Principles of Explainable Artificial Intelligence. National Institute of Standards and Technology: United States Department of Commerce. NISTIR 8312.

Poudel, M., Sarode, R. P., Watanobe, Y., Mozgovoy, M. and Bhalla, S., 2022. A Survey of Big Data Archives in Time-Domain Astronomy. Applied Sciences, 12(12), p. 6202.

Reis, I., Baron, D. and Shahaf, S., 2018. Probabilistic Random Forest: A Machine Learning Algorithm for Noisy Data Sets. *The Astronomical Journal*, *157*(1), p. 16.

Rogers, E. M., 1962. *Diffusion of Innovations*. The Free Press, New York.

Salazar, G., See, J. E., Handley, H. A. and Craft, R., 2020, December. Understanding Human Readiness Levels. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 64, No. 1, pp. 1765–1769). Sage CA: Los Angeles, CA: SAGE Publications.

Scaife, A. 2020. DARA: Development in Africa with Radio Astronomy: AI and Bias in Radio Astronomy. Available at: https://youtu.be/26BY635ZbC0 [Accessed 20 January, 2023].

Shorrock, S. T. and Kirwan, B., 2002. Development and Application of a Human Error Identification Tool for Air Traffic Control. *Applied Ergonomics*, 33(4), pp. 319–336.

Smith, C. J., 2019. *Designing Trustworthy AI: A Human-Machine Teaming Framework to Guide Development*. Carnegie Mellon University Software Engineering Institute Pittsburgh United States.

Smith, M. J. and Geach, J. E., 2022. Astronomia Ex Machina: A History, Primer, and Outlook on Neural Networks in Astronomy [online]. [Preprint]. [Viewed 23 January, 2023]. Available from doi: https://doi.org/10.48550/arXiv.2211.03796.

Snoeyink, R. and Ertmer, P. A., 2001. Thrust into Technology: How Veteran Teachers Respond. *Journal of Educational Technology Systems*, 30(1), pp. 85–111.

Stanton, N. A., Salmon, P. M., Rafferty, L. A., Walker, G. H., Baber, C. and Jenkins, D. P., 2017. *Human Factors Methods: A Practical Guide for Engineering and Design*. CRC Press.

Taylor, R. M., 2017. Situational Awareness Rating Technique (SART): The Development of a Tool for Aircrew Systems Design. In *Situational Awareness* (pp. 111–128). Routledge.

Vicente, K. J., 1999. *Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work*. CRC Press.

Wenskovitch, J. and North, C., 2020. Interactive Artificial Intelligence: Designing for the "Two Black Boxes" Problem. *Computer*, 53(8), pp. 29–39.

Wickens, C. D., Gordon, S. E., Liu, Y. and Lee, J., 2004. *An Introduction to Human Factors Engineering* (Vol. 2). Upper Saddle River, NJ: Pearson Prentice Hall.

Wilkinson, P. R., 1992. The Integration of Advanced Cockpit and Systems Design. *ALGARD*, 18, p. 18.

Zwicky Transient Facility (ZTF), 2023. Public Data Release 15 [online]. Available at: https://www.ztf.caltech.edu/ztf-public-releases.html [Accessed 6 February, 2023].