

Appendix N Bibliography

- [1] United States Government Accountability Office (GAO), "Artificial Intelligence: Status of Developing and Acquiring Capabilities for Weapons Systems," Washington, DC, 2022.
- [2] A. Brown, N. Forsgren, J. Humble, N. Kersten and G. Kim, "2016 State of DevOps Report, Puppet + DORA, 2016. [Online]. Available: <https://services.google.com/fh/files/misc/state-of-devops-2016.pdf>.
- [3] V. Kaulgud, A. Saxena, S. Podder, V. S. Sharma and C. Dinakar, "Shifting testing beyond the deployment boundary. In Proceedings of the International Workshop on Continuous Software Evolution and Delivery (pp. 30-33).," May 2016. [Online]. Available: <https://doi.org/10.1145/2896941.2896954>.
- [4] C. Collins, "DRAFT T&E as a Continuum," *ITEA Publication*, vol. 2, no. 44, 2023.
- [5] S. J. Hutchison, "Shift Left!: Test Earlier in the Life Cycle," *Defense Acquisition, Technology and Logistics Magazine*, vol. 42, no. 234, pp. 35-39, 2013.
- [6] M. Korolov, "6 reasons why AI projects fail," 6 August 2019. [Online]. Available: <https://www.cio.com/article/220445/6-reasons-why-ai-projects-fail.html>.
- [7] DoD, "DoDI 5000.87 Operation of the Software Acquisition Pathway," October 2, 2020.
- [8] P. Laura Freeman, P. Justin Kauffman, D. Sobien, P. Tyler Cody and P. Erin Lanus, "Best Practices for Addressing New Challenges in Testing and Evaluating Artificial Intelligence Enabled Systems," *The ITEA Journal of Test and Evaluation*, vol. (3), no. 43, pp. 166-172, 2022.
- [9] MITRE, "A Sensible Regulatory Framework for AI Security," 2023. [Online]. Available: <https://www.mitre.org/news-insights/publication/sensible-regulatory-framework-ai-security>.
- [10] F. Reeder, D. Kotras, J. Lockett, C. Pomales and R. Lukas, "The Future State of Test and Evaluation of Artificial Intelligence - Enabled Systems in the Department of Defense," Mclean, Virginia, 2022.
- [11] J. Jordan, "Effective testing for machine learning systems," 2020 August 2020. [Online]. Available: <https://www.jeremyjordan.me/testing-ml/>.
- [12] R. M. Lakshmanan, *Machine Learning Design Patterns*, Sebastopol, CA: O'Reilly, 2021.
- [13] R. Black, J. Davenport, J. R. Olszewska, A. L. Smith and J. Wright, *Artificial Intelligence and Software Testing: Building Systems You Can Trust*, Swindon, UK: BCS, 2022.
- [14] K. Gereb, "Best Practices for MLOps and the Machine Learning Lifecycle," 12 May 2020. [Online]. Available: <https://www.contino.io/insights/ml-ops-and-the-machine-learning-lifecycle>.
- [15] DOD, "Department of Defense Responsible Artificial Intelligence Strategy and Implementation Pathway," June 2022. [Online]. Available: <https://media.defense.gov/2022/Jun/22/2003022604/-1/-1/0/Department-of-Defense-Responsible-Artificial-Intelligence-Strategy-and-Implementation-Pathway.PDF>.

- [16] H. Belani, M. Vukovic and Ž. Car, "Requirements engineering challenges in building AI-based complex systems," *IEEE 27th International Requirements Engineering Conference Workshops*, pp. 252-255, 2019.
- [17] Data Bricks, "What is ML Ops?," 2023. [Online]. Available: <https://www.databricks.com/glossary/mlops>.
- [18] DoD, "Instruction 5000.80, Operation of the Middle Tier of Acquisition (MTA)," 30 December 2019. [Online]. Available: https://static.e-publishing.af.mil/production/1/saf_aq/publication/dodi5000.80_dafi63-146/dodi5000.80_dafi63-146.pdf.
- [19] DoD, 19 November 2020. [Online]. Available: <https://acqnotes.com/wp-content/uploads/2021/03/DoD-Instruction-5000.89-Test-and-Evaluation-19-Nov-2020.pdf>.
- [20] A. Ghahrai, "DevOps Foundations and Concepts," 1 May 2020. [Online]. Available: <https://devqa.io/what-is-devops/>.
- [21] Executive Order of the President, "Maintaining American Leadership in Artificial Intelligence," [Online]. Available: <https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence>.
- [22] N. Shadab, A. U. Kulkarni and A. Salado, "Shifting Paradigms in Verification and Validation of AIES: A Systems Theoretic Perspective," in *Systems Engineering and Artificial Intelligence*, Springer, 2021, pp. 363-378.
- [23] Towards AI, "Top 10 reasons Why 87% of Machine learning projects fail?," 17 September 2020. [Online]. Available: <https://towardsai.net/p/machine-learning/top-10-reasons-why-87-of-machine-learning-projects-fail>.
- [24] P. McDermott, C. Dominguez, N. Kasdaglis, M. Ryan, I. Trhan and A. Nelson, "Human Machine Teaming Systems Engineering Guide," MITRE, Bedford, MA, 2018.
- [25] B. T. Molloy, "Project Governance for Defense Applications of Artificial Intelligence," *PRISM*, vol. 9, no. 3, pp. 106-121, 1 January 2021.
- [26] T. Temin, "Too few DoD program managers using agile methodologies, GAO cautions," 13 October 2021. [Online]. Available: <https://federalnewsnetwork.com/defense-main/2021/10/too-few-dod-program-managers-using-agile-methodologies-gao-cautions/>.
- [27] J. Palat, "Agile Development Applied to Machine Learning Projects," 20 February 2021. [Online]. Available: <https://www.infoq.com/articles/machine-learning-agile/>. [Accessed 21 July 2022].
- [28] K. Marko, "Infrastructure for machine learning, AI requirements, examples," 20 October 20. [Online]. Available: <https://www.techtarget.com/searchdatacenter/feature/Infrastructure-for-machine-learning-AI-requirements-examples>.
- [29] Department of the Air Force, Massachusetts Institute of Technology, "Artificial Intelligence Acquisition Guidebook," 14 February 2022. [Online]. Available: https://aia.mit.edu/wp-content/uploads/2022/02/AI-Acquisition-Guidebook_CAO-14-Feb-2022.pdf.

- [30] Acquisition.gov, "Federal Acquisition Regulation," 30 12 2022. [Online]. Available: <https://www.acquisition.gov/browse/index/far> .
- [31] National Archives, "Defense Federal Acquisition Regulation Supplement (DFARS)," [Online]. Available: <https://www.federalregister.gov/defense-federal-acquisition-regulation-supplement-dfars->.
- [32] CDAO, "Fundamentals Common Service Draft Document," 2021.
- [33] EGNYTE, "EGNYTE - Data Ownership," 2 June 2021. [Online]. Available: <https://www.egnyte.com/guides/governance/data-ownership>. [Accessed 27 July 2022].
- [34] J. Rotner, R. Hodge and L. Danley, "AI Fails and How We Can Learn from Them," MITRE Corporation, 2020.
- [35] D. L. Visengeriyeva, A. Kammer, I. Bär, A. Kniesz, M. Plöd (DDD Advisor) and S. Eberstaller, "MLOps Principles," March 2020. [Online]. Available: <https://ml-ops.org/content/mlops-principles>. [Accessed 21 July 2022].
- [36] DoD, "Test and Evaluation Enterprise Guidebook," August 2022. [Online]. Available: <https://www.test-evaluation.osd.mil/Portals/120/Documents/TE%20Enterprise%20Guidebook/TE%20Enterprise%20Guidebook%208.02.pdf?ver=uqBTY9tLbZTH1oSOQgRmUg%3D%3D>.
- [37] DOT&E/DTE&A, "T&E Strategy/TEMP Chapter," 2022.
- [38] JAIC (transitioned to CDAO), "TEMP Template for a Program with AI," Jan 14, 2022.
- [39] JAIC (transitioned to CDAO), "T&E Workshop," October 2, 2021.
- [40] V. Turri, R. Dzombak, E. Heim, N. VanHoudnos, J. Palat and A. Sinha, "Cornell University - arXiv," 7 April 2022. [Online]. Available: <https://arxiv.org/pdf/2204.04211.pdf>. [Accessed 7 October 2022].
- [41] T. Gebru, J. Morgenstern, B. Vecchione, J. W. Vaughan, H. Wallach, H. Daumé III and K. Crawford, "Cornell University - arXiv," 14 April 2019. [Online]. Available: <https://arxiv.org/abs/1803.09010.pdf>. [Accessed 22 June 2020].
- [42] T. Gebru, J. Morgenstern, B. Vecchione, J. W. Vaughan, H. Wallach, H. Daume and K. Crawford, "Datasheets for Datasets," 19 March 2020. [Online]. Available: <https://arxiv.org/abs/1803.09010>. [Accessed 19 March 2021].
- [43] M. Mitchell, S. Wu, A. Zaldivar, P. Barnes, L. Vasserman, B. Hutchinson, E. Spitzer, I. D. Raji and T. Gebru, "Model cards for model reporting," in *Conference on Fairness, Accountability, and Transparency*, 2019.
- [44] S. Schelter, D. Lange, P. Schmidt, M. Celikel, F. Biessmann and A. Grafberger, "Automating large-scale data quality verification," in *Proceedings of the VLDB Endowment*, 2018.
- [45] M. Arnold, J. Boston, M. Desmond, E. Duesterwald, B. Elder, A. Murthi, J. Navratil and D. Reimer, "Towards Automating the AI Operations Lifecycle," in *MLOps Workshop at MLSys 2020*, 2020.

- [46] X. Wang, Q. Kang, M. Zhou, L. Pan and A. Abusorrah, "Multiscale drift detection test to enable fast learning in nonstationary environments," *IEEE Transactions on Cybernetics*, vol. 51, no. 7, pp. 3483-3495, 2020.
- [47] H. van Merode, "Operate and Monitor," in *Continuous Integration (CI) and Continuous Delivery (CD)*, Berkeley, CA, Apress, 2023, pp. 331-357.
- [48] A. Ravichandran, K. Taylor and P. Waterhouse, "Deploy: Building an Agile, Resilient, and Scalable Continuous Delivery Pipeline," in *DevOps for Digital Leaders*, Berkeley, CA, Apress, 2016, pp. 87-103.
- [49] E. Laukkanen, J. Itkonen and C. Lassenius, "Problems, causes and solutions when adopting continuous delivery — A systematic literature review," *Information and Software Technology*, vol. 82, pp. 55-79, February 2017.
- [50] P. d. Hamer, P. Krensky, A. Linden, C. Idoine, E. Brethenoux, J. Hare, S. Sicular and F. Choudhary, "Critical Capabilities for Data Science and Machine Learning Platforms," 2 March 2020. [Online]. Available: <https://b2bsalescafe.files.wordpress.com/2021/11/gartner-critical-capabilities-for-data-science-and-machine-learning-platforms-march-2020.pdf>. [Accessed July 27 2022].
- [51] MITRE Corporation, "MITRE AI Maturity Model and Assessment Tool Guide, DRAFT 20210128.pdf," 2021.
- [52] Y. Mehta, "CIO - Best practices to embrace an ‘MLOps’ mindset," 28 June 2021. [Online]. Available: <https://www.cio.com/article/188887/best-practices-to-embrace-an-mlops-mindset.html>. [Accessed 21 July 2022].
- [53] A. Gupta, "Gartner - Data Fabric Architecture is Key to Modernizing Data Management and Integration," 11 March 2021. [Online]. Available: <https://www.gartner.com/smarterwithgartner/data-fabric-architecture-is-key-to-modernizing-data-management-and-integration>. [Accessed 27 July 2022].
- [54] D. Groombridge, "Gartner - Top Strategic Technology Trends for 2022," Gartner, 10 January 2022. [Online]. Available: <https://www.gartner.ca/en/information-technology/insights/top-technology-trends>. [Accessed 27 July 2022].
- [55] DoD, "DOD Adopts Ethical Principles for Artificial Intelligence," 4 February 2020. [Online]. Available: <https://www.defense.gov/News/Releases/Release/Article/2091996/dod-adopts-ethical-principles-for-artificial-intelligence/>.
- [56] D. Bickson, "The 10 Recommender System Metrics You Should Know About," 3 October 2012. [Online]. Available: <http://bickson.blogspot.com/2012/10/the-10-recommender-system-metrics-you.html>. [Accessed 25 January 2023].
- [57] P. N. Stuart Russell, *Artificial Intelligence a Modern Approach*, Fourth Edition, Hoboken, NJ: Pearson, 2021.
- [58] O. Gunderson and S. Kjensmo, "State of the Art: Reproducibility in Artificial Intelligence," 2017. [Online]. Available: <https://www.aaai.org/GuideBook2018/17248-73943-GB.pdf>. [Accessed 30 November 2022].

- [59] DoD, "Cybersecurity Test and Evaluation Guidebook," 1 July 2015. [Online]. Available: <http://www.acqnotes.com/wp-content/uploads/2014/09/Cybersecurity-Test-and-Evaluation-Guidebook-27-Oct-2015.pdf>.
- [60] A. Hur and R. Yeonseung, "Developing Anti-tamper Functionalities through Continuous Integration," in *Proceedings of the 2020 5th International Conference on Intelligent Information Technology*, 2020.
- [61] D. Agrawal, B. Archambeault, J. R. Rao and P. Rohatgi, "The EM side—channel(s)," in *Cryptographic Hardware and Embedded Systems-CHES 2002: 4th International Workshop*, Redwood Shores, CA, 2003.
- [62] AcqNotes, "Program Management Tool for Test & Evaluation Operational Test & Evaluation (OT&E)," 21 April 2022. [Online]. Available: <https://acqnotes.com/acqnote/careerfields/operational-test-and-evaluation-ote>.
- [63] "OECD iLibrary," [Online]. Available: https://www.oecd-ilibrary.org/sites/543e84ed-en/1/3/4/index.html?itemId=/content/publication/543e84ed-en&_csp_=535d2f2a848b7727d35502d7f36e4885&itemIGO=oecd&itemContentType=book. [Accessed 11 May 2023].
- [64] AcqNotes, "Program Management Tool for Aerospace Test & Evaluation Evaluation of Survivability," 5 June 2018. [Online]. Available: <https://acqnotes.com/acqnote/careerfields/evaluation-of-survivability>.
- [65] J. P. C. S. P. T. N. Kufro, "Machine Learning: Trust and Trustworthiness Metrics," *MITRE Internal Report*, 26 June 2020.
- [66] R. R. Hoffman, S. T. Mueller, G. Klein and J. Litman, "Metrics for explainable AI: Challenges and prospects," 2018. [Online]. Available: <https://arxiv.org/vc/arxiv/papers/1812/1812.04608v1.pdf>. [Accessed 13 February 2023].
- [67] DoD, "DoDI 5000.75 Business Systems Requirements and Acquisition," 24 January 2020. [Online]. Available: <https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/500075p.PDF>.
- [68] Department of Defense (DoD), "DODI 5000.2, Operation of the Defense Acquisition System," 2003.
- [69] M. A. Lapham, "Sustaining Software-Intensive Systems," May 2006. [Online]. Available: https://resources.sei.cmu.edu/asset_files/TechnicalNote/2006_004_001_14669.pdf.
- [70] D. W. McKeown, "Office of the Secretary of Defense: Continuous Authorization To Operate (cATO)," 3 February 2022. [Online]. Available: <https://media.defense.gov/2022/Feb/03/2002932852/-1/-1/0/CONTINUOUS-AUTHORIZATION-TO-OPERATE.PDF>.
- [71] C. Banbury, V. Reddi, M. Lam, W. Fu, A. Fazel, J. Holleman, X. Huang, R. Hurtado, D. Kanter, A. Likhmotov and D. Patterson, "Benchmarking TinyML Systems: Challenges and Direction," in *Proceedings of the 3rd MLSys Conference*, Austin, TX, 2020.
- [72] N. Andresky and A. Taliaferro, "Operationalizing Artificial Intelligence for Multi-Domain Operations," US Army Futures and Concepts Center Future Warfare Division, 2019.

- [73] S. Soro, "TinyML for Ubiquitous Edge AI," MITRE Technical Report #200519, 2020.
- [74] T. S. A. T. A. a. S. V. Li, "Federated learning: Challenges, methods, and future directions," *IEEE Signal Processing*, vol. 37, no. 3, pp. 50-60, 2020.
- [75] E. P. T. C. C. K. W. L. H. B. D. a. A. T. Blasch, "Machine learning/artificial intelligence for sensor data fusion—opportunities and challenges," *IEEE Aerospace and Electronic Systems*, vol. 36, no. 7, pp. 80-93, 2021.
- [76] DOD, "Cybersecurity Test and Evaluation Guidebook," 10 February 2020. [Online]. Available: <https://www.dau.edu/cop/test/DAU%20Sponsored%20Documents/Cybersecurity-Test-and-Evaluation-Guidebook-Version2-change-1.pdf>. [Accessed January 2023].
- [77] L. a. J. D. Aguilera, "DECA: DoD Enterprise Cloud Architecture Concept for Cloud-Based Cross Domain Solutions," *2021 4th Artificial Intelligence and Cloud Computing Conference*, pp. 165-171, 2021.
- [78] S. D. S. R. B. P. R. G. B. A. a. S. K. Thielemans, "Experiences with on-premise open source cloud infrastructure with network performance validation," *IECON 2022—48th Annual Conference of the IEEE Industrial Electronics Society*, pp. 1-6, 2022.
- [79] E. L. C. F. I. C. J. a. M. A. Sturzinger, "Improving the performance of AI models in tactical environments using a hybrid cloud architecture," *Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications III*, vol. 11746, pp. 18-32, 2021.
- [80] P. F. L. a. M. A. Bellavista, "Decentralised learning in federated deployment environments: A system-level survey," *ACM Computing Surveys (CSUR)*, vol. 54, no. 1, pp. 1-38, 2021.
- [81] M. Baxter, "Centrally Pretrained Federated Fine-Tuning: Enabling a Secure and Accurate Military Security Application on Embedded Hardware," Naval Postgraduate School, 2020.
- [82] G. a. V. D. Cirincione, "Federated machine learning for multi-domain operations at the tactical edge," in *Proceedings of SPIE Defense + Commercial Sensing*, Baltimore, MD, 2019.
- [83] A. Taïk and S. Cherkaoui, "Federated edge learning: Design issues and challenges," *IEEE Network*, vol. 35, no. 2, pp. 252-258, 2020.
- [84] A. Taïk, Z. Mlika and S. Cherkaoui, "Data-aware device scheduling for federated edge learning," *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 1, pp. 408-21, 28 July 2021.
- [85] H. Zhu, H. Zhang and Y. Jin, "From federated learning to federated neural architecture search: a survey," *Complex & Intelligent Systems*, vol. 7, pp. 639-657, 2021.
- [86] C. a. S. C. Shi, "Federated multi-armed bandits," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 11, pp. 9603-9611, 2021.
- [87] A. G.-L. C. V. A. G. S. N. D. G. S. L. D. B. A. S. A. G. A. a. G. Y. Lavin, "Technology readiness levels for machine learning systems," arXiv preprint arXiv:2101.03989, 2021.
- [88] A. D. M. B. S. R. Debbiche, "Challenges When Adopting Continuous Integration: A Case Study," in *Product-Focused Software Process Improvement. PROFES 2014. Lecture Notes in Computer Science*, vol. 8892, Springer, 2014.

- [89] W. M. V. R. T. D. P. E. M. K. M. A. a. O. P. Hummer, "Modelops: Cloud-based lifecycle management for reliable and trusted AI," in *2019 IEEE International Conference on Cloud Engineering (IC2E)*, 2019.
- [90] S. Deng, H. Zhao, W. Fang, J. Yin, S. Dustdar and A. Y. Zomaya, "Edge intelligence: The confluence of edge computing and artificial intelligence," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 7457-7469, 2020.
- [91] H. Raza, G. Prasad and Y. Li, "Adaptive learning with covariate shift-detection for non-stationary environments," in *2014 14th UK Workshop on Computational Intelligence (UKCI)*, 2014.
- [92] B. Karlaš, M. Interlandi, C. Renggli, W. Wu, C. Zhang, D. M. I. Babu, J. Edwards, C. Lauren, A. Xu and M. Weimer, "Building continuous integration services for machine learning," in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2020.
- [93] E. Blasch and J. Sung, "Standards for the artificial intelligence community," *Automatic Target Recognition XXXI*, vol. 11729, pp. 286-297, 2021.
- [94] M. R. Zafar and N. Khan, "Deterministic local interpretable model-agnostic explanations for stable explainability," *Machine Learning and Knowledge Extraction*, vol. 3, no. 3, pp. 525-541, 2021.
- [95] L. Antwarg, R. M. Miller, B. Shapira and L. Rokach, "Explaining anomalies detected by autoencoders using Shapley Additive Explanations," *Expert Systems with Applications*, vol. 186, 2021.
- [96] J. Nguyen, M. Kshitiz, H. Zhan, A. Yousefpour, M. Rabbat, M. Malek and D. Huba, "Federated learning with buffered asynchronous aggregation," in *International Conference on Artificial Intelligence and Statistics*, 2022.
- [97] J. Lu, A. Liu, F. Dong, F. Gu, J. Gama and G. Zhang, "Learning under Concept Drift: A Review," *IEEE Transactions on Knowledge and Data Engineering*, vol. 31, no. 12, pp. 2346-2363, 2019.
- [98] E. Minkov, R. Wang, A. Tomasec and W. Cohen, "NER Systems that Suit User's Preferences: Adjusting the Recall-Precision Trade-off for Entity Extraction," in *NAACL/HLT*, 2006.
- [99] K. Darwish, "Named entity recognition using cross-lingual resources: Arabic as an example," in *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, 2013.
- [100] C. Helwe and S. Elbassuoni, "Arabic named entity recognition via deep co-learning," *Artificial Intelligence Review*, vol. 52, no. 1, pp. 197-215, 2019.
- [101] J. Ni, G. Dinu and R. Florian, "Weakly supervised cross-lingual named entity recognition via effective annotation and representation projection," arXiv preprint: 1707.02483, 2017.
- [102] R. Grishman and B. Sundheim, "Message understanding conference-6: A brief history," in *16th Conference on Computational Linguistics (COLING)*, 1996.

- [103] E. Sang and F. De Meulder, "Introduction to the CoNLL-2003 shared task: Language-independent named entity recognition," arXiv preprint cs/0306050, 2003.
- [104] G. Doddington, A. Mitchell, M. Przybocki, L. Ramshaw, S. Strassel and R. Weischedel, "The automatic content extraction (ACE) program-tasks, data, and evaluation," in *LREC*, 2004.
- [105] NIST, "Automatic Content Extraction 2008 Evaluation Plan (ACE08) – Assessment of Detection and Recognition of Entities and Relations within and across Documents," NIST/ACE, 2007.
- [106] Q. Wang, H. Wang, C. Gupta and S. Serita, "Regularized Operating Envelope with Interpretability and Implementability Constraints," <https://arxiv.org/pdf/1912.10158.pdf>, Santa Clara, 21 Dec 2019.
- [107] I. K. Chen, "JAIC (transitioned to CDAO) T&E Whitepaper: Data Traceability and Explainable Model Development.pdf," 2021.
- [108] I. K. Chen, "JAIC (transitioned to CDAO) T&E: Maritime Search & Rescue_Data Split Approach and Methods.pdf," 2021.
- [109] MITRE, "MITRE ATLAS™ (Adversarial Threat Landscape for Artificial-Intelligence Systems)," [Online]. Available: <https://atlas.mitre.org>.
- [110] EAGLES, "EAGLES Evaluation of Natural Language Processing Systems. Final Report," Center for Sprogteknologi, Copenhagen, 1996.
- [111] National Virtual Translation Center, "Translation Memory Metrics Development Experiment Series," NVTC, Washington, DC, 2011.
- [112] J. Vardaro, M. Schaeffer and S. Hansen-Schirra, "Translation Quality and Error Recognition in Professional Neural Machine Translation Post-Editing," *Informatics*, vol. 6, no. 41, 2019.
- [113] Office of the Director of National Intelligence, "Intelligence Community Human Language Technology Roadmap: The Way Forward," 2015.
- [114] K. Yost, L. Booker, D. Gerber, C. Howell, C. Elsaesser, T. Levitt, A. McLeod, F. Reeder and M. Tanner, "Threats Posed by Advances in Artificial Intelligence," MITRE Technical Report to OSD Office of Net Assessment, McLean, VA, 2016.
- [115] University of Illinois, "Artificial Intelligence and Autonomous Systems," 2019. [Online]. Available: <https://ece.illinois.edu/research/aias.asp>. [Accessed 18 August 2019].
- [116] J. Kleinberg, S. Mullainathan and M. Raghavan, "Inherent Trade-Offs in the Fair Determination of Risk Scores," in *Proceedings of Innovations in Theoretical Computer Science*, 2017.
- [117] B. Marr, "What is the Difference Between Artificial Intelligence and Machine Learning?," *Forbes*, 6 December 2016.
- [118] Institute of Electrical and Electronics Engineers (IEEE), "2017 IEEE AI Taxonomy," 2017. [Online]. Available: http://ieeauthorcenter.ieee.org/wp-content/uploads/taxonomy_v101.pdf. [Accessed 2018].

- [119] Association for Computing Machinery (ACM), "ACM Computing Classification System," 2012. [Online]. Available: <https://www.acm.org/publications/class-2012/>. [Accessed 2018].
- [120] D. Rumelhart and J. McLelland, *Parallel Distributed Processing*, Volume 1, Bradford, 1987.
- [121] T. Klosowski, "Facial Recognition Is Everywhere. Here's What We Can Do About It.," *New York Times*, 15 July 2020.
- [122] PEO-I EW&S, "Machine Foreign Language Translation Systems Bridge The Gap Between Soldiers, Locals," US Army, 10 May 2011. [Online]. Available: <https://peoiews.army.mil/machine-foreign-language-translation-systems-bridge-the-gap-between-soldiers-locals/>. [Accessed 26 January 2022].
- [123] A. Greenberg, "How to Steal an AI," *WIRED*, September 30, 2016.
- [124] Office of the Deputy Assistant Secretary of Defense, "Department of Defense Digital Engineering Strategy," DOD, USA, Washington, DC, June 2018.
- [125] N. Bajema, "AI's 6 Worst-Case Scenarios," *IEEE Spectrum*, January 2022.
- [126] C. Williams, "A Brief Introduction To Artificial Intelligence," *IEEE*, 1983.
- [127] J. A. Tirpak, "AI, Networks, Hypersonics Are the Pentagon's Top Research Priorities," *Air Force Magazine*, Arlington, VA, January 2022.
- [128] DODI, "Operation of the Adaptive Aquisition Framework," DOD, January 2020.
- [129] R. Shultz and R. Clarke, "Big Data at War: Special Operations Forces, Project Maven, and Twenty-First-Century Warfare," *Modern War Institute at West Point*, 25 August 2020.
- [130] National Security Commission , "National Security Commission on Artificial Intelligence," DOD, March, 2021.
- [131] A. M. Church, "The Science of Avoidance," *Air Force Magazine*, no. 99, pp. 34-38, 2016.
- [132] E. Tabassi, "NISTIR 8269 (Draf)," NIST, 2019.
- [133] DoD, "MIL-STD-882E, Standard Practice for System Safety," 2000.
- [134] DoD, "DoDD 3000.09, November 21, 2012, Incorporating Change 1," 2017.
- [135] Defense Innovation Board, "AI Principles: Recommendations on the Ethical Use of AI by the Department of Defense," 2020.
- [136] Y. Pinelis, "Systems Engineering Research Center Talks: "Progress in T&E of AI-Enabled Systems in the DoD," 2021.
- [137] F. Lucini, "The Real Deal About Synthetic Data," *MIT Sloan Management Review*, 20 October 2021.
- [138] M. Mitchell, S. Wu, A. Zalivar, P. Barnes, L. Vasserman, B. Hutchinson, S. E. D. Raji and T. Gebru, "Model Cards for Model Reporting," in *FAT 19*, Atlanta, GA, 2019.
- [139] C. Pomales, N. Sydney and C. Johns, "Artificial Intelligence Test and Evaluation: Current State Survey," MITRE, McLean, VA, June 2021.
- [140] Defense Innovation Board, "AI Principles: Recommendations on the Ethical Use of AI by the Department of Defense," 2020.

- [141] A. Chandrasekaran, D. Cearley and S. Alaybeyi, "Top 10 Strategic Technology Trends for 2020: AI Security," Gartner ID G00450643, 10 March, 2020.
- [142] E. Tabassi, K. Burns, M. Hadjimichael, A. Molina-Markham and J. Sexton, "A Taxonomy and Terminology of Adversarial Machine Learning," NIST, October 2019.
- [143] L. Pons and I. Ozkaya, "Priority Quality Attributes for Engineering AI-enabled Systems," in *Association for the Advancement of Artificial Intelligence AI in Public Sector Workshop*, Washington, DC, November 7-9, 2019.
- [144] Loucks, Hupfer and Jarvis, "Future in the Balance? How Countries are pursuing and AI advantage - Insights from Deloitte's State of AI in the Enterprise, 2nd Edition Survey," Deloitte, May 2019.
- [145] C. Hundhausen, S. Trent, A. Balkar and M. Nuur, "The design and experimental evaluation of a tool to support the construction and wizard-of-oz testing of low fidelity prototypes," *IEEE Symposium on Visual Languages and Human-Centric Computing*, pp. 86-90, September 2008.
- [146] S. Amershi, A. Begel, C. Bird, R. DeLine, H. Gall, E. Kamar, N. Nagappan, B. Nushi and T. Zimmermann, "Software Engineering for Machine Learning: A Case Study," Microsoft, 2018.
- [147] E. T. Chancey, J. P. Bliss, Y. Yamani and H. A. Handley, "Trust and the compliance–reliance paradigm: The effects of risk, error bias, and reliability on trust and dependence," vol. 59, no. 3, pp. 333-345, 2017.
- [148] G. Falco, B. Schneiderman, J. Badger, R. Carrier, A. Dahbura, D. Danks, M. Eling, A. Goodloe, J. Gupta, C. Hart, M. Jirotko, H. Johnson, C. LaPointe, A. Llorens, A. Mackworth, C. Maple, S. Palsson, F. Pasquale, A. Winfield and Z. K. Yeong, "Governing AI Safety through Independent Audits," *Nature*, July 2021.
- [149] M. Flournoy, A. Haines and G. Chevitz, "Building Trust through Testing," WestExec Advisors, 2020.
- [150] D. Gunning and D. Aha, "DARPA's explainable artificial intelligence (XAI) program," *AI Magazine*, vol. 40, no. 2, pp. 44-58, 2019.
- [151] A. Horneman, A. Mellinger and I. Ozkayam, "AI Engineering: 11 Foundational Practices," Carnegie Mellon University Software Engineering Institute, Pittsburgh, 2019.
- [152] B. Stanton and T. Jensen, "Trust and Artificial Intelligence," Department of Commerce, USA, December 2020.
- [153] H. M. Wojton, D. T. Porter, S. Lane, C. Bieber and P. Madhavan, "Initial validation of the trust of automated systems test (TOAST)," *Journal of Social Psychology*, vol. 160, no. 6, pp. 735-750, 2020.
- [154] N. L. L. Freeman, S. Welsh, S. Adams, P. Beling and K. J. Meimandi, "Metrics for the Evaluation of Human-Agent Teaming," Hume Center for National Security and Technology, Virginia Tech Link Lab, School of Engineering.

- [155] J. Jian, A. Bisantz and C. Drury, "Foundations for an empirically determined scale of trust in automated systems," *International Journal of Cognitive Ergonomics*, vol. 4, no. 1, p. 53–71, 2000.
- [156] S. M. Merritt, D. Lee, J. L. Unnerstall and K. Huber, "Are well-calibrated users effective users? Associations between calibration of trust and performance on an automation-aided task.," *Human Factors*, vol. 57, no. 1, pp. 34-47, 2015.
- [157] S. M. Merritt and D. R. Illgen, "Not all trust is created equal: Dispositional and history based trust in human automation interactions," *Human Factors*, vol. 50, no. 2, pp. 194-210., 2008.
- [158] D. J. Niedober, N. T. Ho, G. Masequesmay, K. Koltai, M. Skoog, A. Cacanindin and J. B. Lyons, "Influence of cultural, organizational and automation factors on human-automation trust: A case study of Auto-GCAS engineers and developmental history.," in *International Conference on Human-Computer Interaction*, 2014.
- [159] IEEE, "IEEE Std 1012-2016, IEEE Standard for System, Software, and Hardware Verification and Validation," 2017.
- [160] C. Rudin, "Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead," in *NIPS 2018 Workshop on Critiquing and Correcting Trends in Machine Learning*, 2019.
- [161] A. Thierer, A. Castillo O'Sullivan and R. Russell, "Artificial Intelligence and Public Policy," Mercatus Center at George Mason University, Arlington, VA, 2017.
- [162] T. Dutton, "AI Policy 101: An Introduction," *Medium*, 5 July 2018.
- [163] D. Gelhaus, L. Koslosky, K. Goode and C. Perkins, "U.S. AI Workforce: Policy Recommendations," Center for Security and Emerging Technology, Georgetown University, 2021.
- [164] Defense Innovation Unit, "Responsible AI Guidelines: Operationalizing DoD's Ethical Principles for AI," 2021.
- [165] Stanford University, "AI Policy, Now and in the Future | One Hundred Year Study on Artificial Intelligence (AI100)," 2020.
- [166] Office of the Director of National Intelligence, "Intelligence Community Directive (ICD) 503, Intelligence Community Information Technology Systems Security Risk Management," 2015.
- [167] Committee on National Security Systems, "Instruction Number 1253, Security Categorization and Control Selection for National Security Systems, Version 1," 2014.
- [168] Department of Homeland Security (DHS), "DHS National Security Systems Policy Directive 4300B," 2018.
- [169] Department of Homeland Security (DHS) / Office of Intelligence and Analysis (I&A), "Rapid Requirement Development and Data Environment (R2D2E) Assessment and Recommendation Letter," 2021.

- [170] Department of Homeland Security (DHS) / Office of Intelligence and Analysis (I&A), "Rapid Requirement Development and Data Environment (R2D2E) Plan of Action & Milestones (POA&M)," 2021.
- [171] Department of Homeland Security (DHS) / Office of Intelligence and Analysis (I&A), "Rapid Requirement Development and Data Environment (R2D2E) Security Assessment Report (SAR)," 2021.
- [172] Department of Homeland Security (DHS) / Office of Intelligence and Analysis (I&A), "Rapid Requirement Development and Data Environment (R2D2E) System Security Plan (SSP), Version 1, Revision 3," Archer, 2021.
- [173] Deloitte, "CIO Journal - The Wall Street Journal," 31 August 2021. [Online]. Available: <https://deloitte.wsj.com/articles/how-public-policy-can-enable-trustworthy-ai-0163043398>. [Accessed 08 April 2022].
- [174] DOD, "Cyber DT Cross-Service Working Group (held Jan 2021)," 2021.
- [175] TechArcis, "What is the role of AI in Test Automation," 5 February 2021. [Online]. Available: <https://www.techarcis.com/what-is-the-role-of-artificial-intelligence-ai-in-test-automation/>. [Accessed 4 October 2022].
- [176] Application Lifecycle Management Body of Knowledge, "DevOps Toolchain," 17 August 2017. [Online]. Available: <https://almbok.com/devops/devops>.
- [177] S. C. Gupta, "MLOps: Machine Learning Lifecycle," 1 September 2022. [Online]. Available: <https://www.ml4devs.com/articles/mlops-machine-learning-life-cycle/>.
- [178] Z. C. Hrvoje Belani & Marin Vuković, "Requirements Engineering Challenges in Building AI-Based Complex Systems," 2019. [Online]. Available: <https://arxiv.org/pdf/1908.11791.pdf>.
- [179] CompTIA, "CompTIA - AI in Business: Top Considerations before Implementing AI," 21 April 2021. [Online]. Available: <https://connect.comptia.org/content/guides/business-considerations-before-implementing-ai>. [Accessed 25 July 2022].
- [180] J. Martel, "Gartner - Comparing Platforms and Capabilities for Data Science and AI," Gartner, Inc., 2021.
- [181] P. Bhandari, "MLOps: Simplify collaboration between Data Scientists and Ops," 30 May 2020. [Online]. Available: <https://medium.com/@bhandaripallavi.98/mlops-ease-collaboration-between-data-scientists-and-it-ops-e7e3ccac9eed>.
- [182] A. L. M. B. o. K. (ALMBok), "DevOps Toolchain," 17 August 2022. [Online]. Available: <https://almbok.com/devops/devops>.
- [183] S. Shellien-Walker, "Why AI Testing Is Different to Traditional Software Testing," 19 March 2019. [Online]. Available: <servian.dev/why-testing-a-dialogflow-agent-is-different-to-traditional-software-testing-9e81a71de016>.
- [184] K2View, "What is Data Fabric? The Complete Guide," 2021. [Online]. Available: <https://www.k2view.com/what-is-data-fabric>. [Accessed 27 July 2022].
- [185] Alegion, "How to Measure Quality of Labeled Data for Machine Learning Applications.pdf," 2021.

- [186] B. Downs, "JAIC (transitioned to CDAO) T&E Whitepaper: K-Fold Cross Validation.pdf," 2021.
- [187] W. Novak, "Artificial Intelligence and Machine Learning Acquisition and Policy Implications," February 2021. [Online]. Available: <https://apps.dtic.mil/sti/pdfs/AD1122292.pdf>.
- [188] M. Billa, "Testers guide for Testing Machine Learning Models," 12 October 2019. [Online]. Available: <https://medium.com/analytics-vidhya/testers-guide-for-testing-machine-learning-models-e7e5cea81264>.
- [189] Defense Acquisition University, "Acquisition Guidebooks & References," [Online]. Available: <https://aaf.dau.edu/guidebooks/>.
- [190] DoD, "DoD Directive 5000.01 Defense Acquisition System," 28 July 2022. [Online]. Available: <https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodd/500001p.pdf>.
- [191] Department of the Army, "Army Regulation 73-1 Test and Evaluation Policy," 8 June 2018. [Online]. Available: https://armypubs.army.mil/epubs/DR_pubs/DR_a/pdf/web/ARN7727_AR_73-1_WEB_Final.pdf.
- [192] Defense Acquisition University, "DAG Test and Evaluation Master Plan Template," [Online]. Available: <http://www.acqnotes.com/Attachments/DAG Test and Evaluation Master Plan Template.doc>.
- [193] DHS, "Template: Test and Evaluation Master Plan," 7 November 2008. [Online]. Available: <http://www.acqnotes.com/Attachments/Test & Evaluation Master Plan Template.doc>.
- [194] Defense Acquisition University, "Test and Evaluation Guide Sixth Edition," December 2012. [Online]. Available: <https://www.acqnotes.com/Attachments/DoD%20Test%20and%20Evaluation%20Management%20Guide%20Dec%202012.pdf>.
- [195] U.S. Air Force, "Air Force Test & Evaluation Guidebook," December 2004. [Online]. Available: <http://www.acqnotes.com/Attachments/Air Force Test & Evaluation Guidebook Dec 2004.doc>.
- [196] DoD, "DOT&E and TEMP Guidebook," 28 May 2013. [Online]. Available: <https://www.acqnotes.com/Attachments/DOT&E%20and%20TEMP%20Guidebook%20-%2028%20Mar%202013.pdf>.
- [197] U.S. Army, "Army TEMP Procedures and Guidelines," 11 October 1996. [Online]. Available: <http://www.acqnotes.com/Attachments/Army TEMP Procedures and Guidelines.pdf>.
- [198] U.S. Government, "Title 10 USC 2366 Survivability testing and lethality testing required before full-scale production," [Online]. Available: <https://law.justia.com/codes/us/2017/title-10/subtitle-a/part-iv/chapter-139/sec.-2366/>.

- [199] U.S. Government, "10 USC 2399 Operational Test and Evaluation of Defense Programs," [Online]. Available: <https://law.justia.com/codes/us/1998/title10/subtitlea/partiv/chap141/sec2399/>.
- [200] U.S. Air Force, "Air Force Operational Text and Evaluation Center (AFOTEC)," [Online]. Available: <https://www.afotec.af.mil>.
- [201] U.S. Army, "Army Test and Evaluation Command," [Online]. Available: <https://www.army.mil/atec>.
- [202] U.S. Navy, "Navy Commander Operational Test and Evaluation Force," [Online]. Available: <https://www.cotf.navy.mil/test-and-evaluation/>.
- [203] IBM, "The Adversarial Robustness Toolbox: Securing AI Against Adversarial Threats," 17 April 2018. [Online]. Available: <https://www.ibm.com/blogs/research/2018/04/ai-adversarial-robustness-toolbox/>. [Accessed 20 August 2022].
- [204] Government Accounting Office, "Artificial Intelligence: Status of Developing and Acquiring Capabilities for Weapon Systems," GAO, Washington, DC, 2022.
- [205] A. B. J. a. O. H. Munappy, "Data pipeline management in practice: Challenges and opportunities," in *International Conference on Product-Focused Software Process Improvement*, 2020.
- [206] J. G. Y. Z. D. L. Q. a. Z. Y. Ren, "Distributed and efficient object detection in edge computing: Challenges and solutions.," *IEEE Network*, vol. 32, no. 6, pp. 137-143, 2018.
- [207] A. O'Brien, "Core Security Principles for Edge Computing in Government," Amazon Web Services (AWS), 2022.
- [208] Amazon Web Services (AWS), "Bringing cloud capability to the Air Force at the “speed of mission need”, " 7 May 2021. [Online]. Available: <https://aws.amazon.com/blogs/publicsector/bringing-cloud-air-force-speed-of-mission-need/>. [Accessed 9 December 2022].
- [209] United States Air Force (USAF), "ABMS signs more companies post onramp," 24 September 2020. [Online]. Available: <https://www.af.mil/News/Article-Display/Article/2359938/abms-signs-more-companies-post-onramp/>. [Accessed 9 December 2022].
- [210] M. Berube, T. Giannelia and G. Vial, "Barriers to the Implementation of AI in Organizations: Findings from a Delphi Study," in *Hawaii International Conference on System Sciences*, Manoa, Jan 2021.
- [211] I. Franz and G. Fogerty, "Agile Product Management: How To Build Weapons Faster & Better," 5 March 2021. [Online]. Available: <https://breakingdefense.com/2021/03/agile-product-management-how-to-build-weapons-faster-better/>.
- [212] S. Saurabh, "Top 5 Data Fabric Takeaways from 2021 Gartner D&A Summit," Nexla, 2021. [Online]. Available: https://www.nexla.com/top-5-data-fabric-takeaways/?utm_term=&utm_campaign=DSA&utm_source=adwords&utm_medium=ppc&hsa_acc=7917802207&hsa_cam=16866137031&hsa_grp=137135942122&hsa_ad=592595

979287&hsa_src=g&hsa_tgt=dsa-19959388920&hsa_kw=&hsa_mt=&hsa_net=adwor.
[Accessed 2022 July 2022].

- [213] D. T. Moore, "Commercial Surveys: Interviews on AI Development and Testing," MITRE, 2022.
- [214] Department of Defense, "DoD Ethical Principles for AI," Department of Defense, Washington, DC, 2020.
- [215] General Accounting Office, "AI – Status of Developing and Acquiring Capabilities for Weapon Systems," 17 2 2022. [Online]. Available: <https://www.gao.gov/products/gao-22-104765>.
- [216] U. D. o. D. Inspector General, "Audit of Department of Defense Middle Tier of Acquisition Rapid Prototyping and Rapid Fielding Programs, Report No. DODIG-2021-131," DoD, 2021.
- [217] Q. Wang, H. Wang, C. Gupta and S. Serita, "Regularized Operating Envelope with Interpretability and Implementability Constraints," Santa Clara, 21 Dec 2019.
- [218] M. Rawat, "Top AI Tools/Platforms To Perform Machine Learning ML Model Monitoring," 30 October 2022. [Online]. Available: <https://www.marktechpost.com/2022/10/30/top-ai-tools-platforms-to-perform-machine-learning-ml-model-monitoring/>.
- [219] NIST, "AI Risk Management Framework," January 2023. [Online]. Available: <https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf>.
- [220] L. Freeman, "Test and Evaluation for Artificial Intelligence," *Insight*, vol. 23, pp. 27-30, 2020.
- [221] J. Sprockhoff, B. Lukic, V. Janson, A. Ahlbrecht, U. Durak, S. Gupta and T. Krueger, "Model-Based Systems Engineering for AI-Based Systems," in *AIAA SCITECH 2023 Forum*, National Harbor, MD, 2023.
- [222] P. Slingerland and a. et, "Adapting a Trusted AI Framework to Space Mission Autonomy," in *2022 IEEE Aerospace Conference (AERO)*, 2022.
- [223] U. S. Department of Health and Human Services, "Artificial Intelligence (AI) Strategy," 2021.