Radar Chart Summary

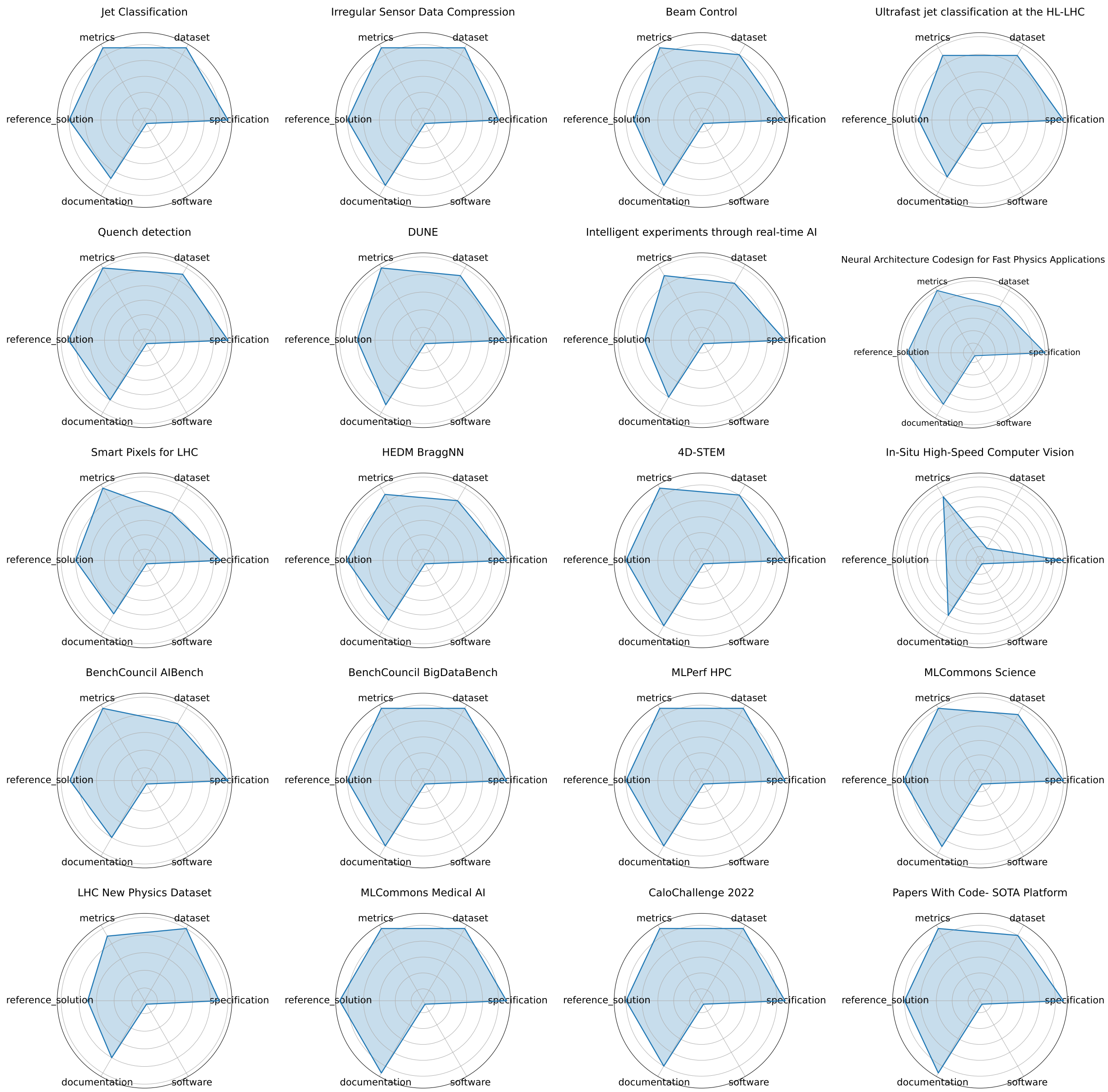

Figure 1: Radar chart overview (page 1)

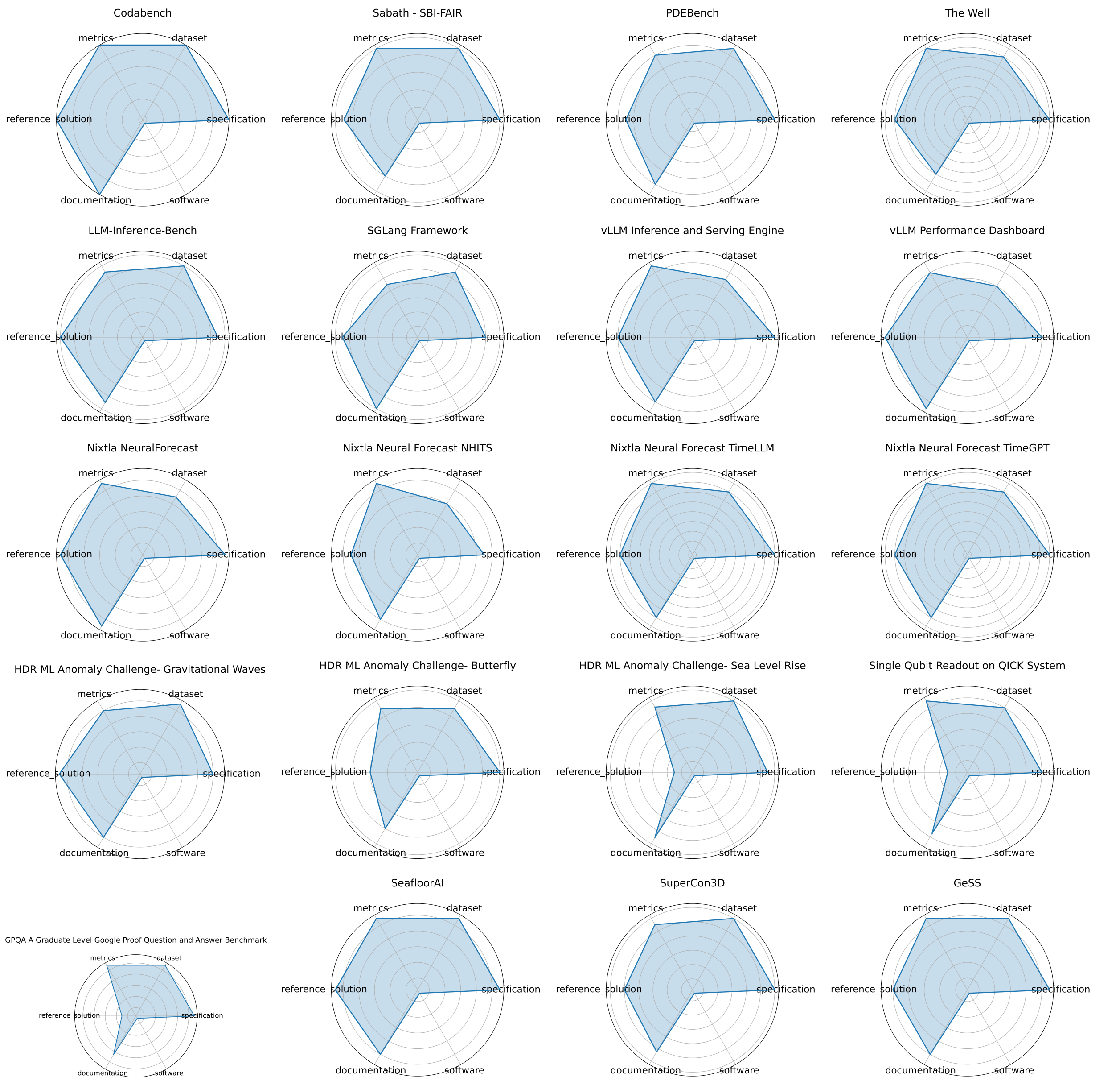

Figure 2: Radar chart overview (page 2)

Figure 3: Radar chart overview (page 3)

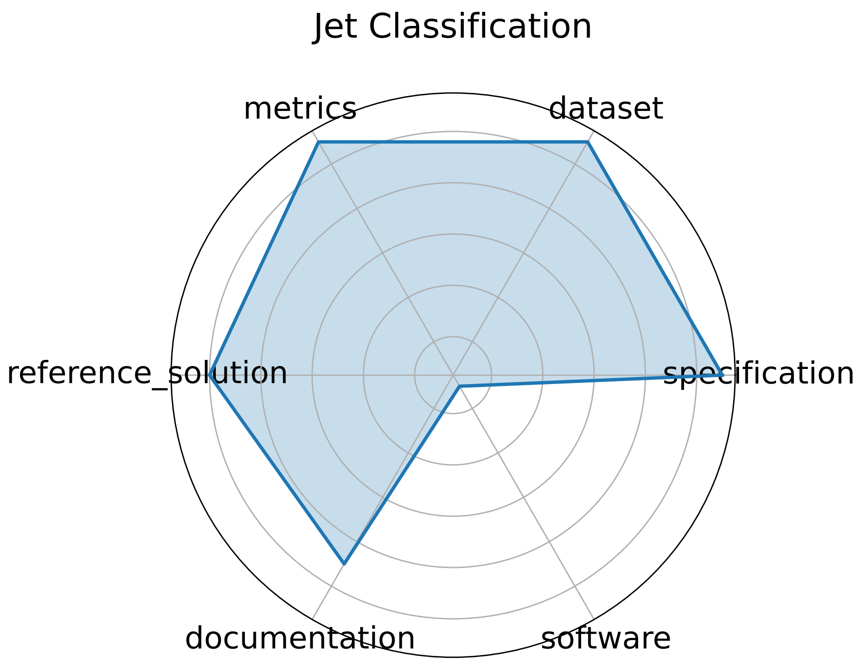

Figure 4: Jet Classification [9]

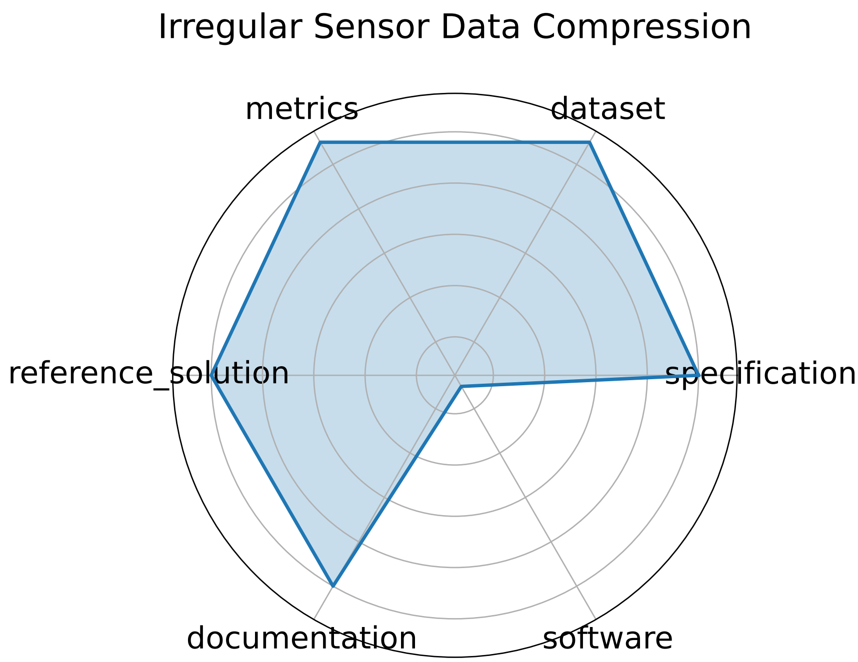

Figure 5: Irregular Sensor Data Compression [10]

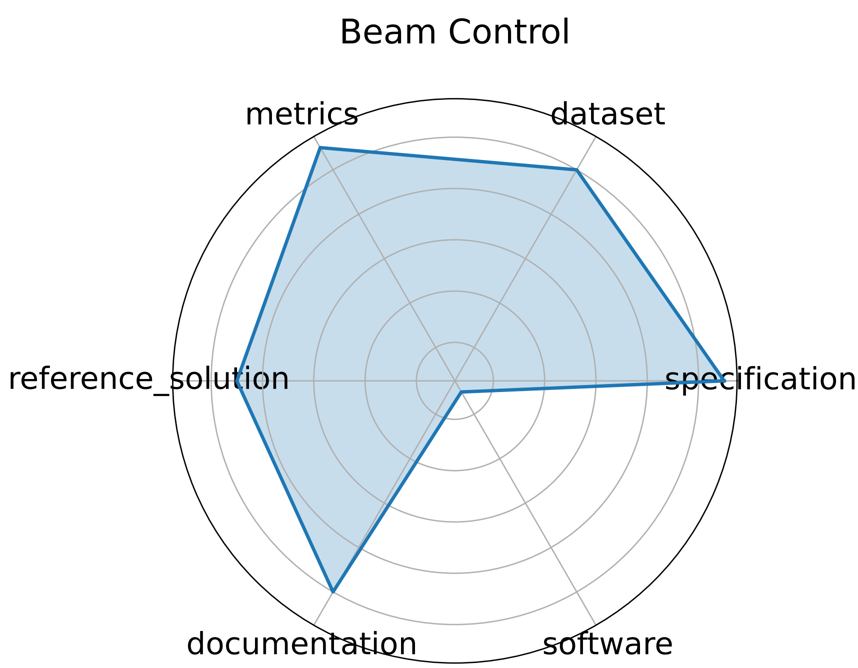

Figure 6: Beam Control [19, 11]

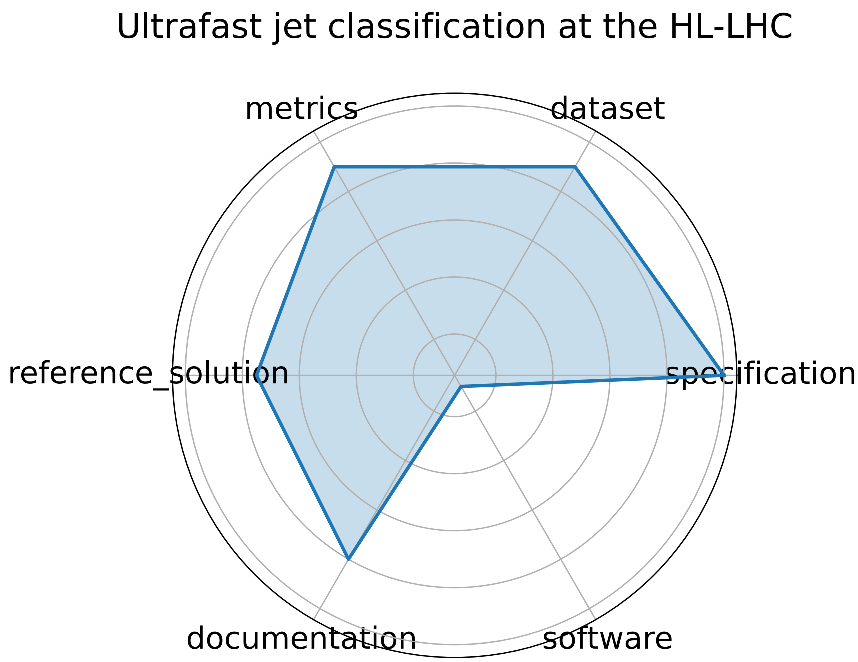

Figure 7: Ultrafast jet classification at the HL-LHC [29]

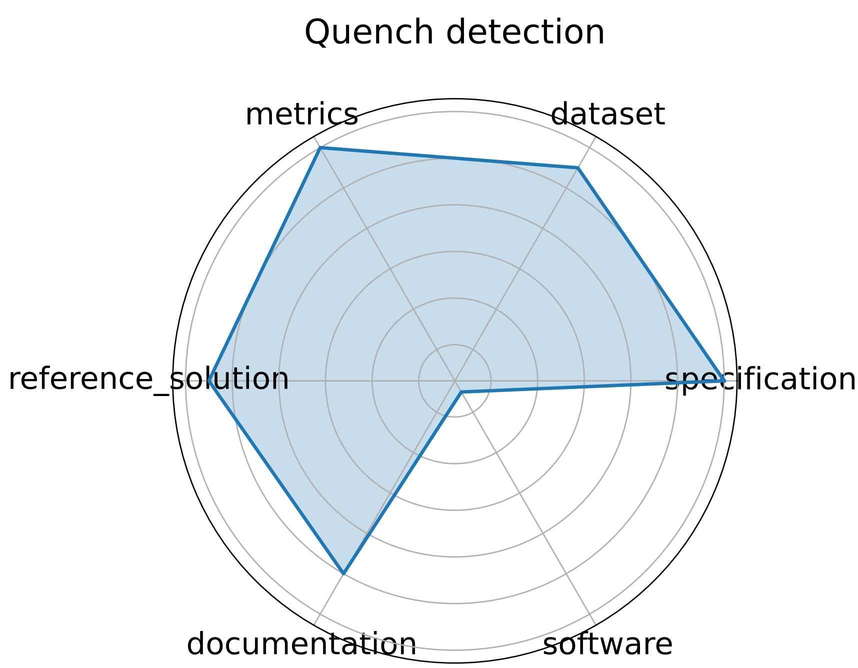

Figure 8: Quench detection

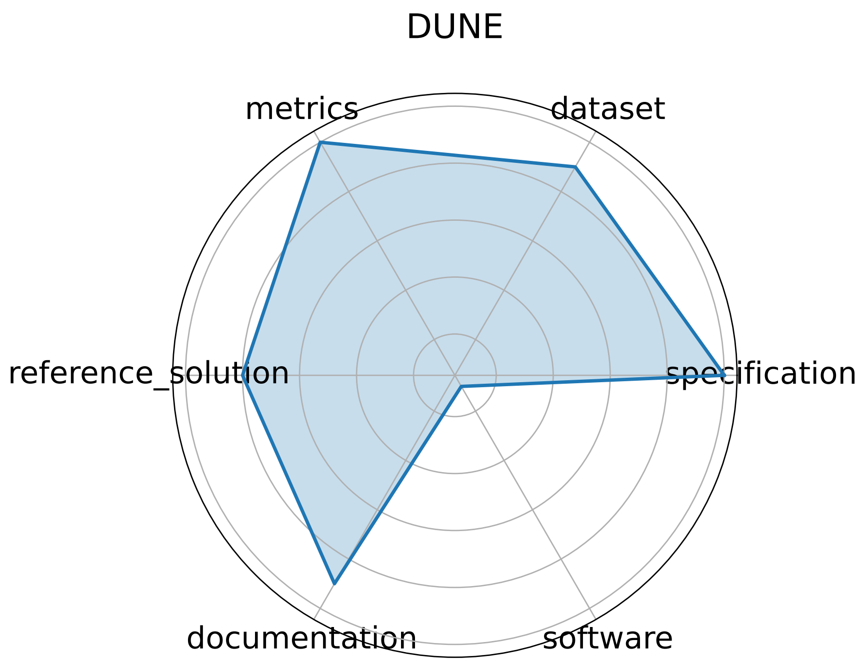

Figure 9: DUNE

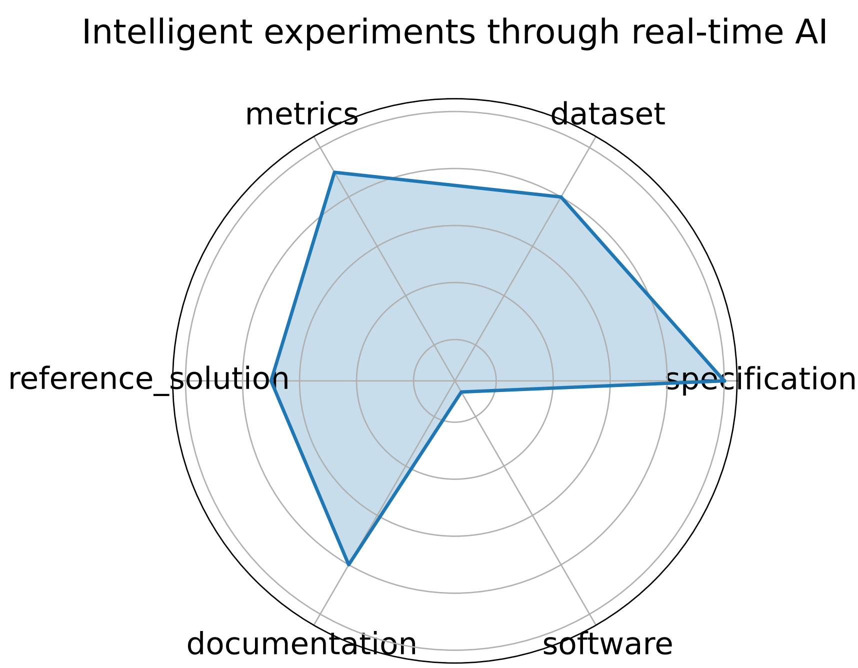

Figure 10: Intelligent experiments through real-time AI [22]

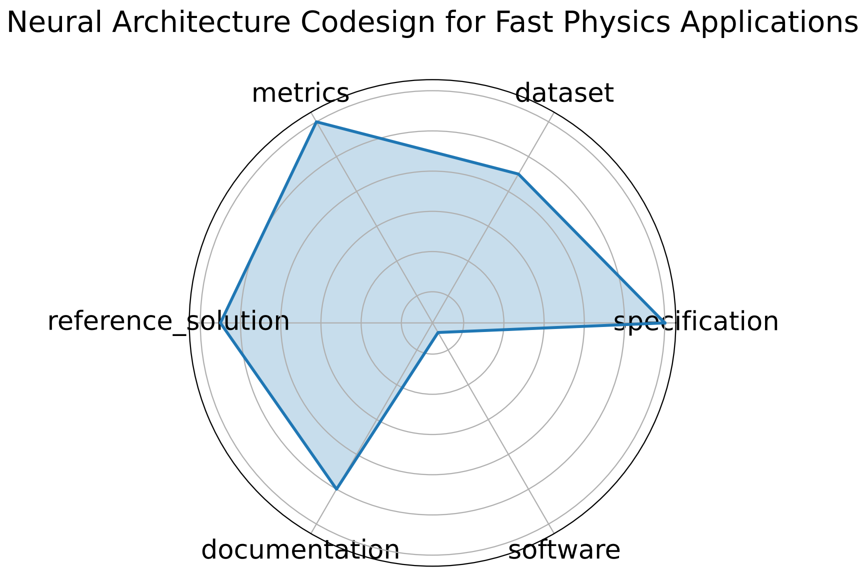

Figure 11: Neural Architecture Codesign for Fast Physics Applications [40]

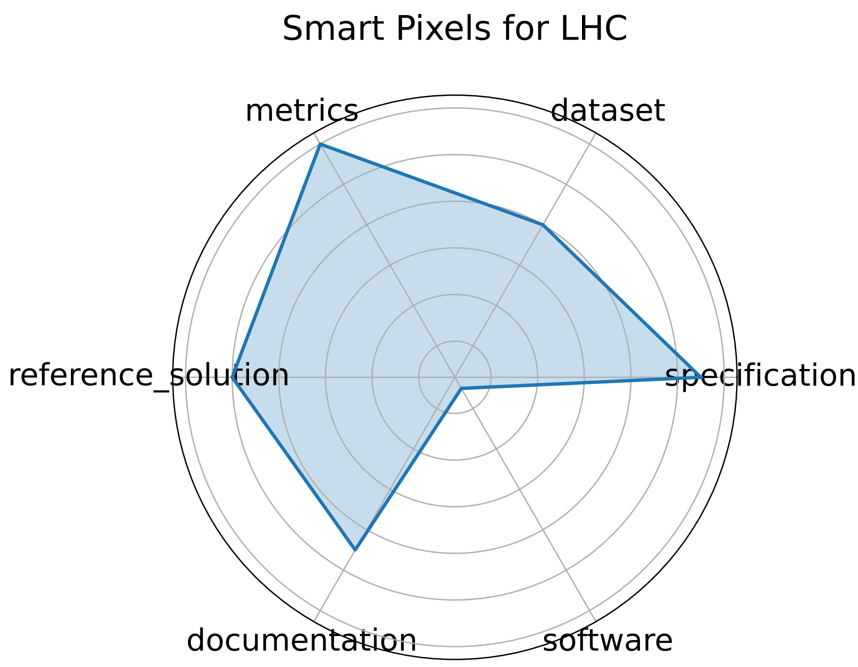

Figure 12: Smart Pixels for LHC [32]

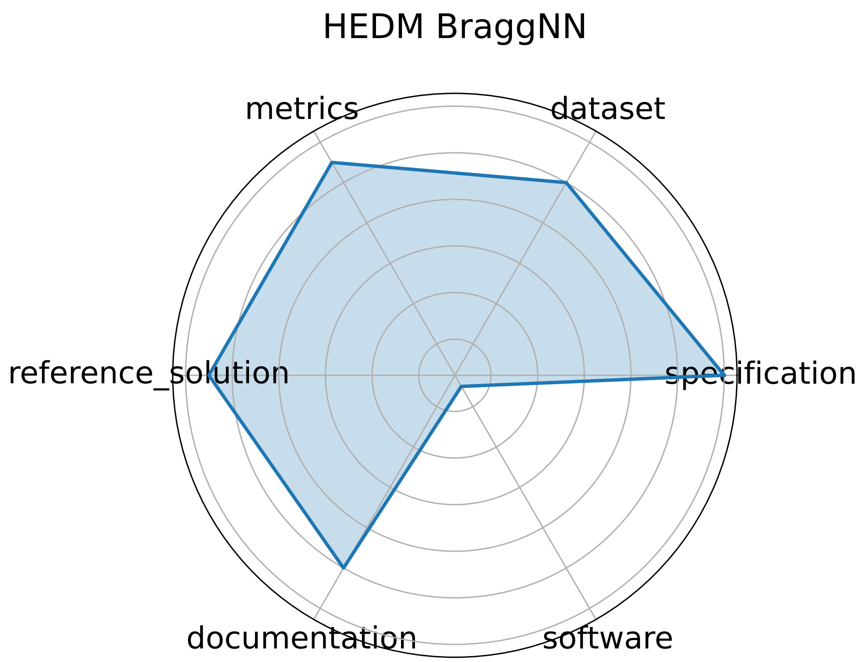

Figure 13: HEDM BraggNN [25]

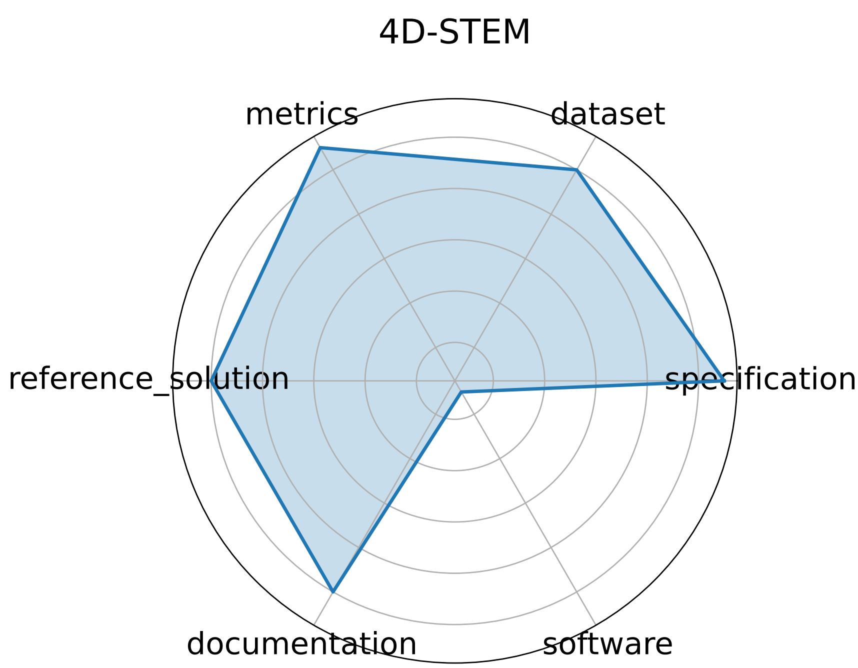

Figure 14: 4D-STEM [35]

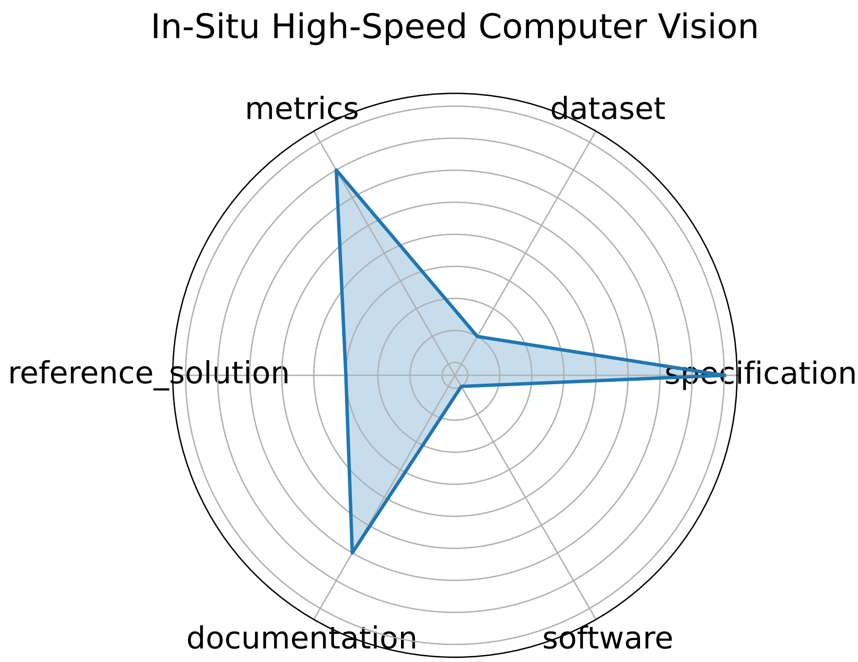

Figure 15: In-Situ High-Speed Computer Vision [39]

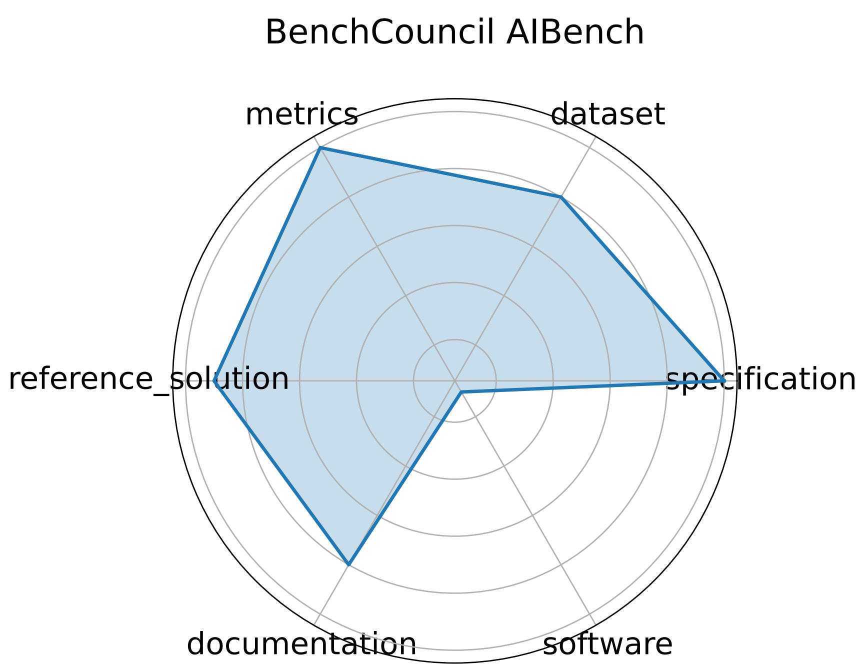

Figure 16: BenchCouncil AIBench [13]

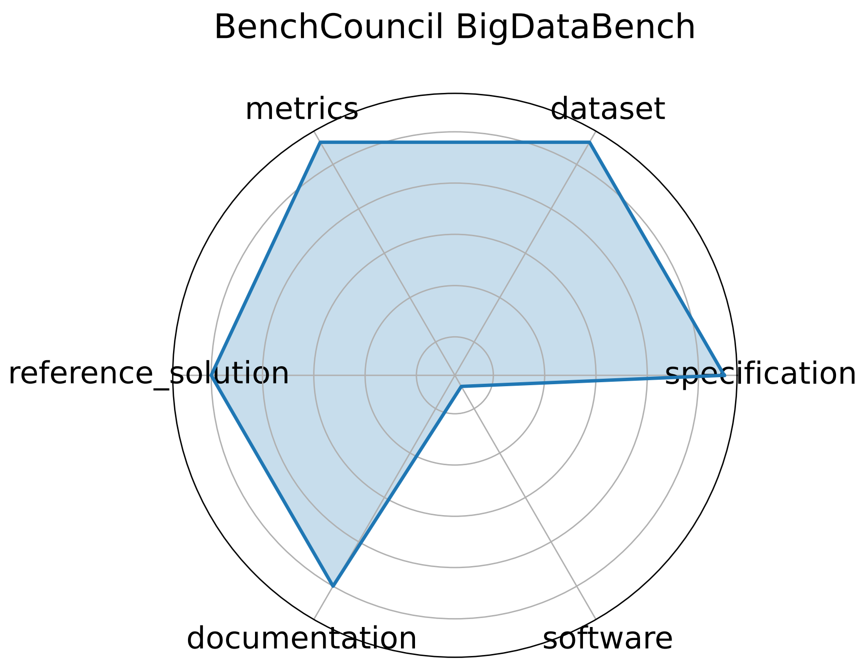

Figure 17: BenchCouncil BigDataBench [14]

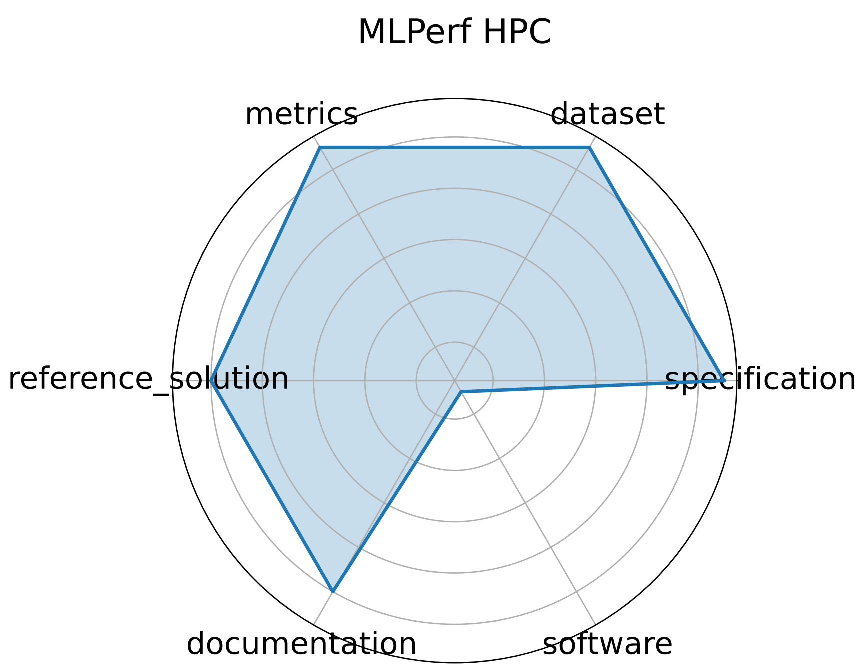

Figure 18: MLPerf HPC [12]

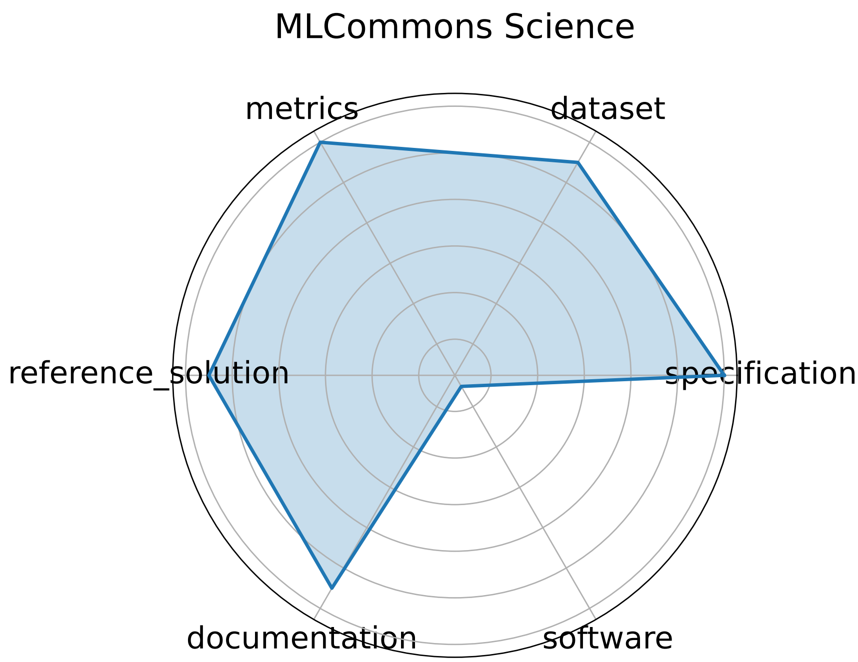

Figure 19: MLCommons Science [16]

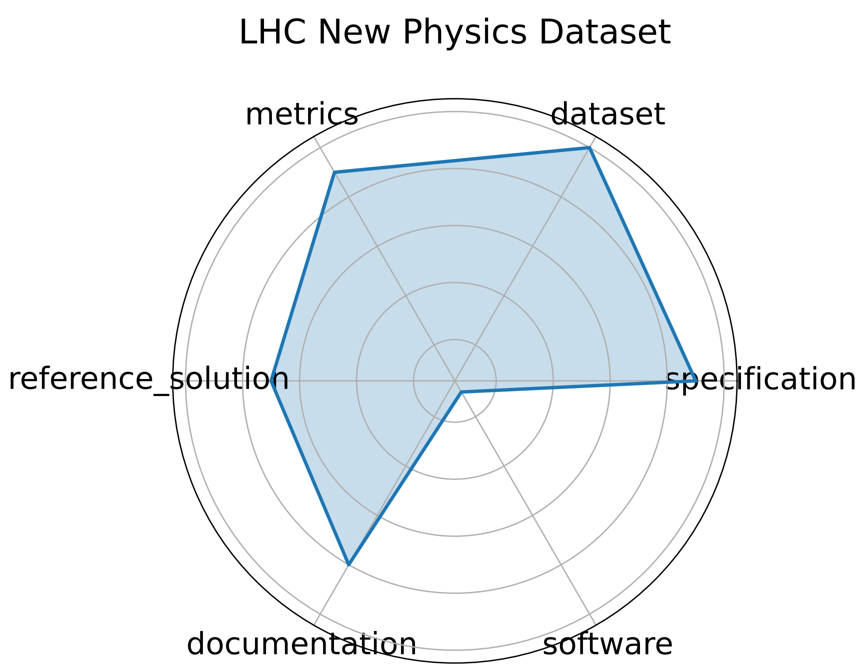

Figure 20: LHC New Physics Dataset [1]

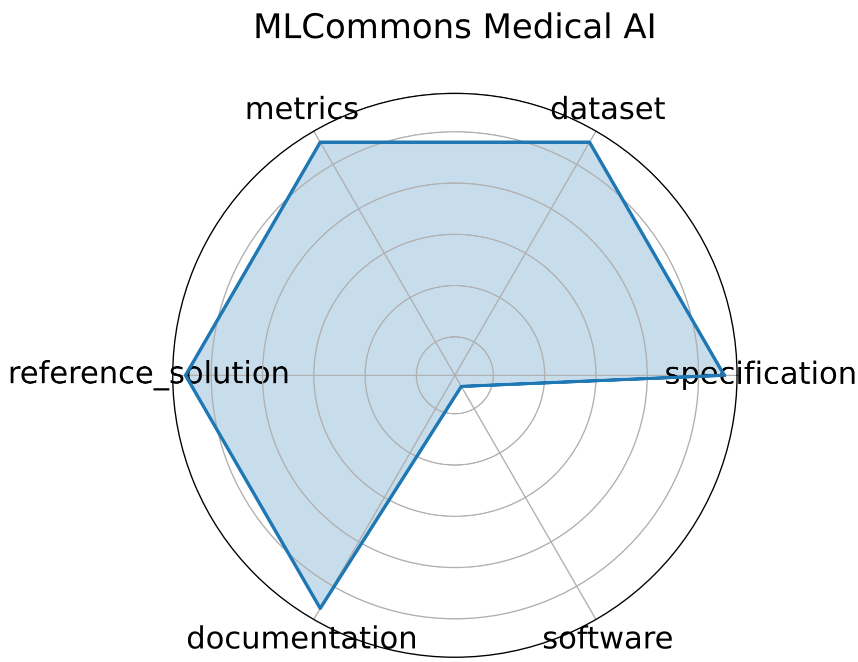

Figure 21: MLCommons Medical AI [20]

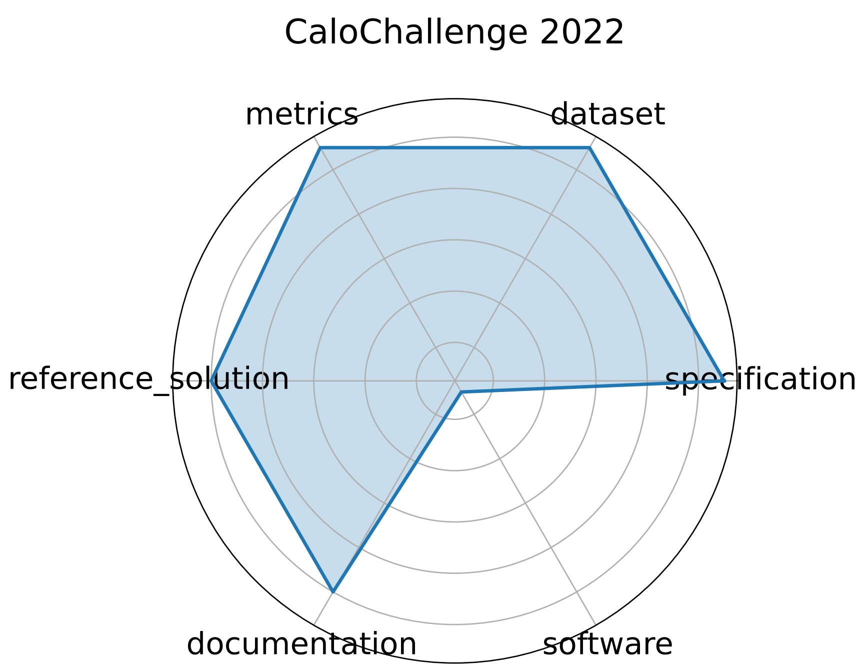

Figure 22: CaloChallenge 2022 [21]

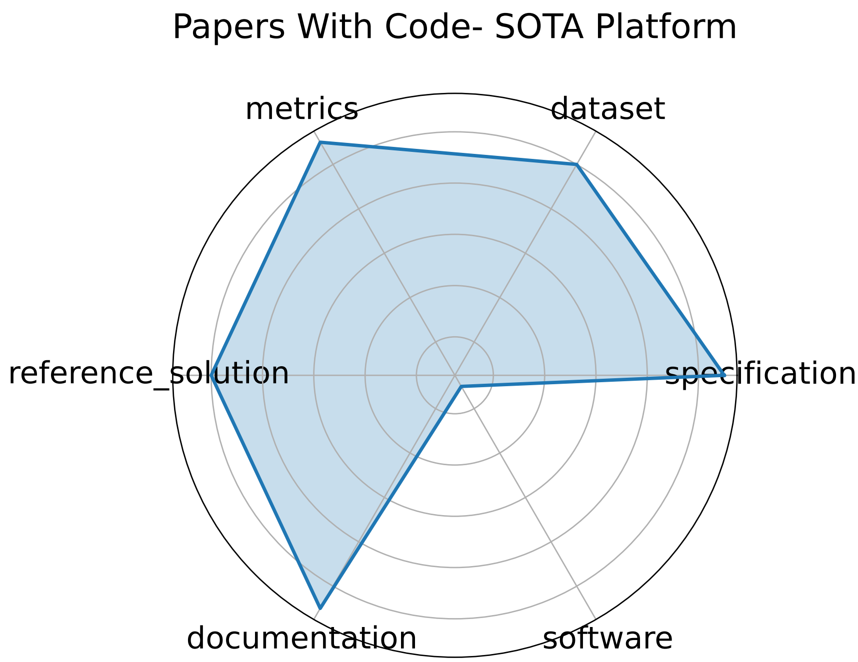

Figure 23: Papers With Code- SOTA Platform [8]

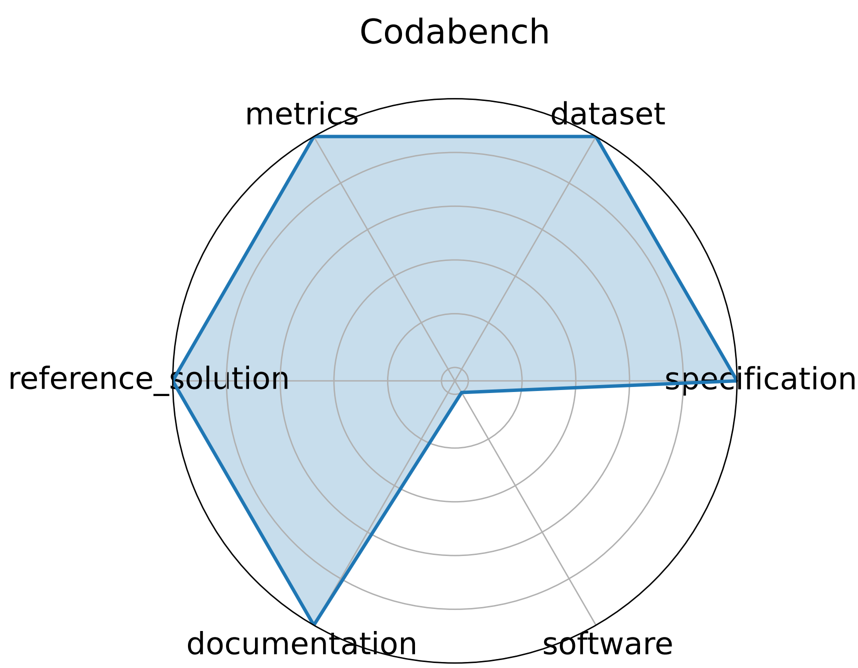

Figure 24: Codabench [41]

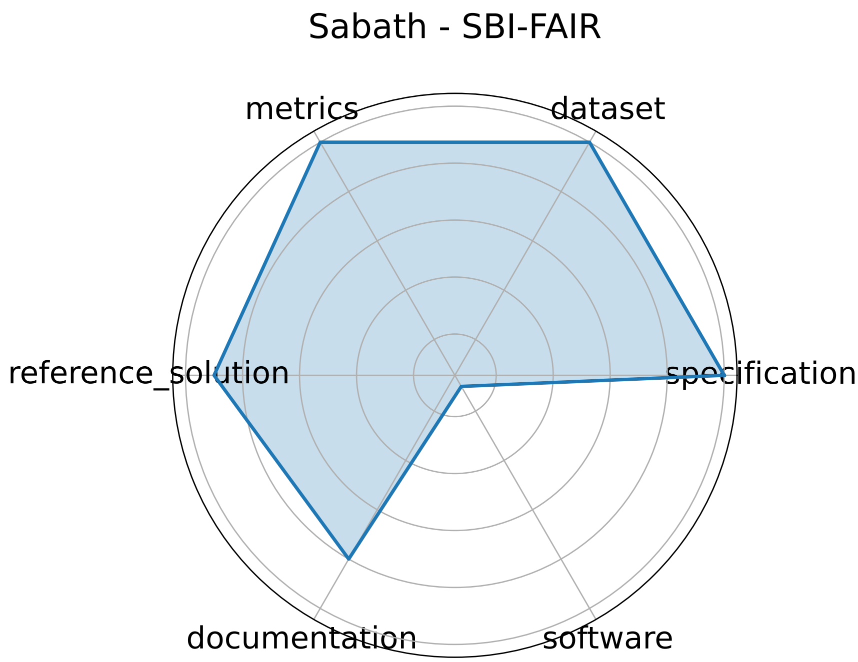

Figure 25: Sabath - SBI-FAIR [26]

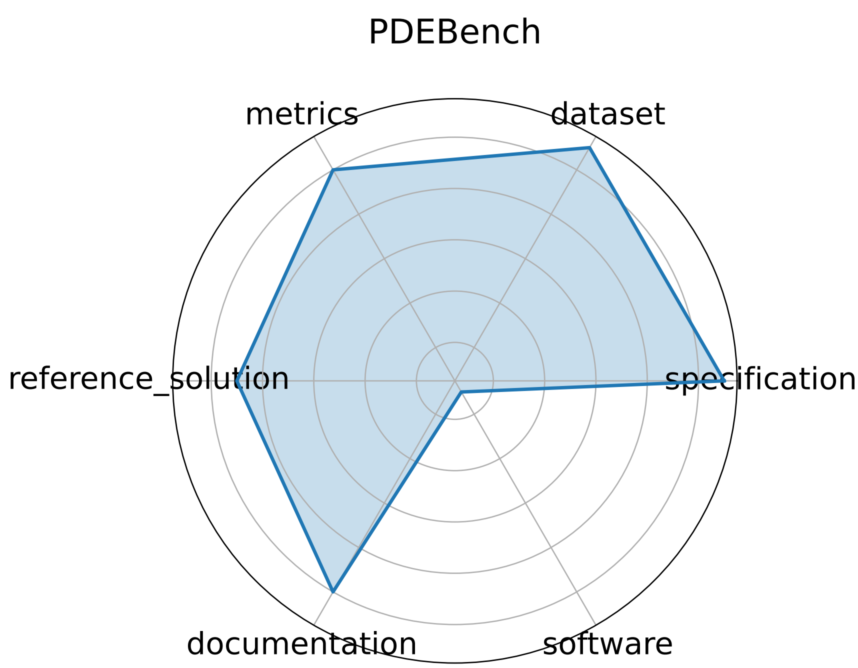

Figure 26: PDEBench [37]

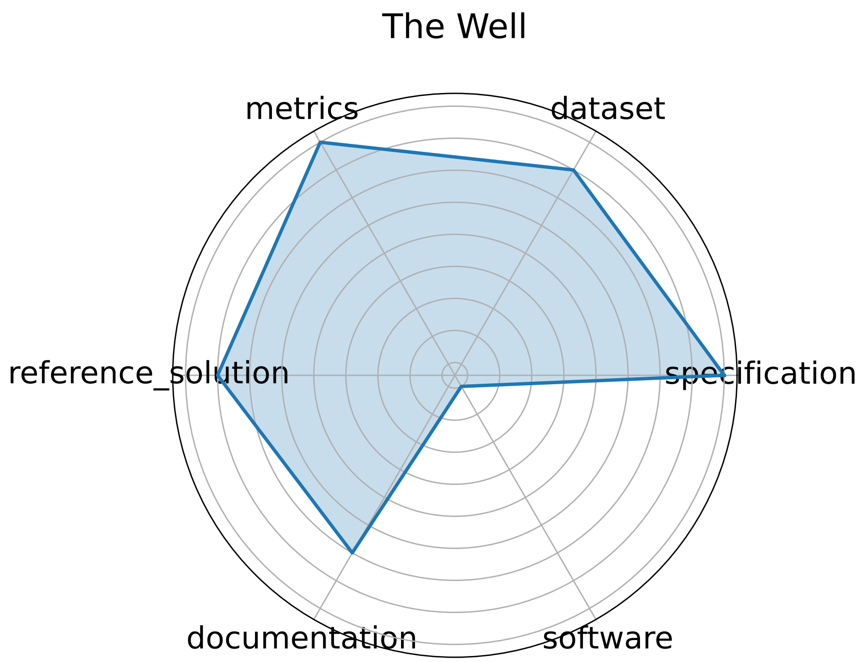

Figure 27: The Well [30]

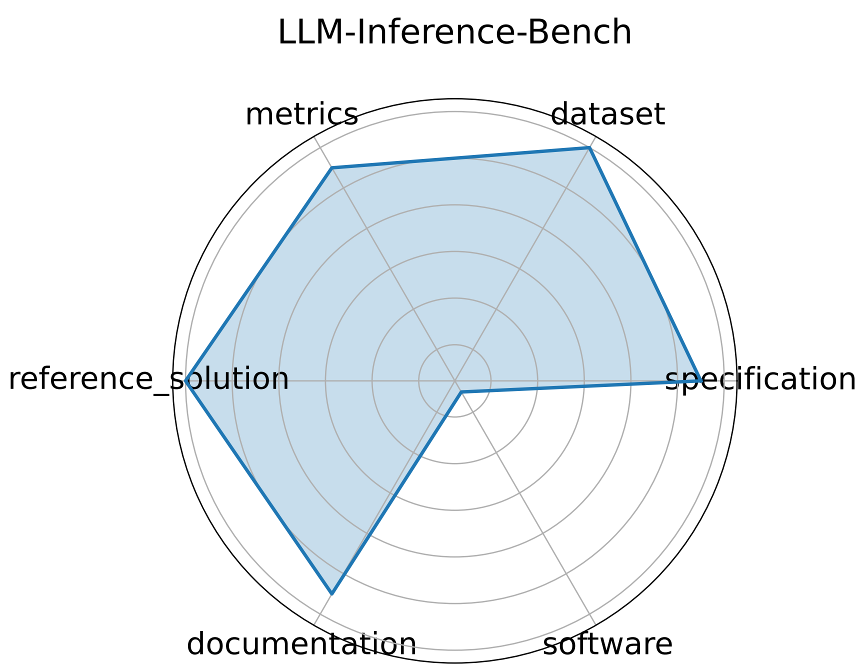

Figure 28: LLM-Inference-Bench [7]

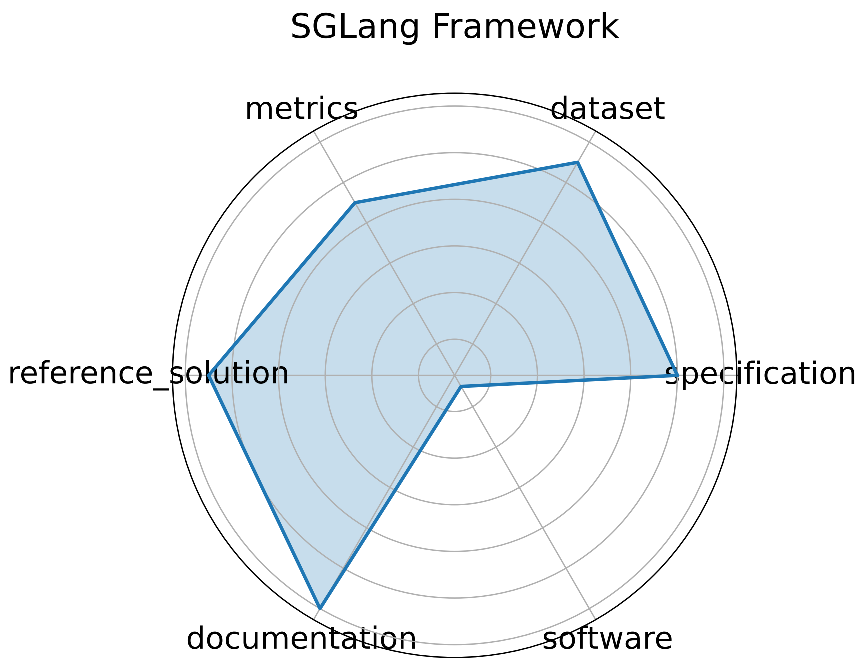

Figure 29: SGLang Framework [42]

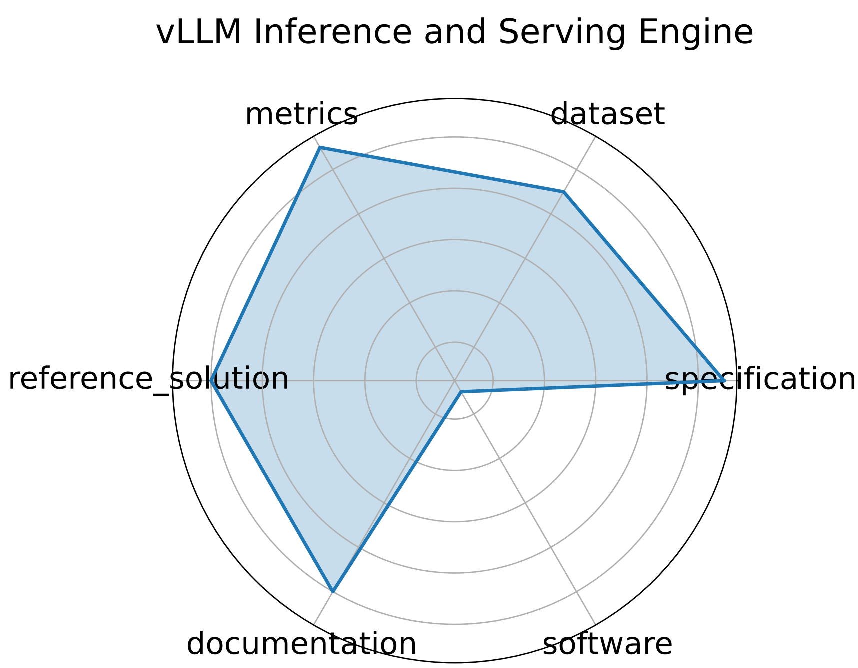

Figure 30: vLLM Inference and Serving Engine [23]

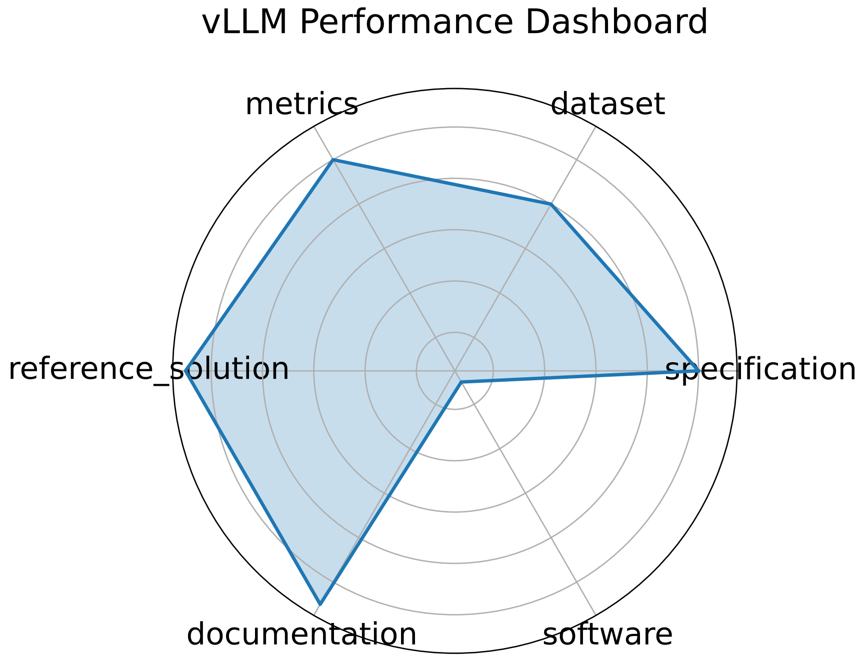

Figure 31: vLLM Performance Dashboard [27]

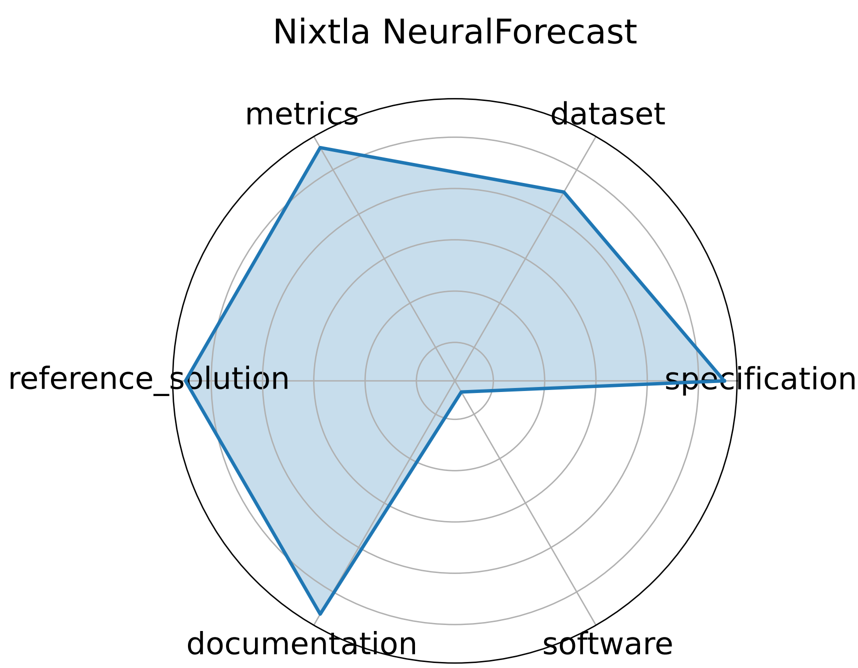

Figure 32: Nixtla NeuralForecast [31]

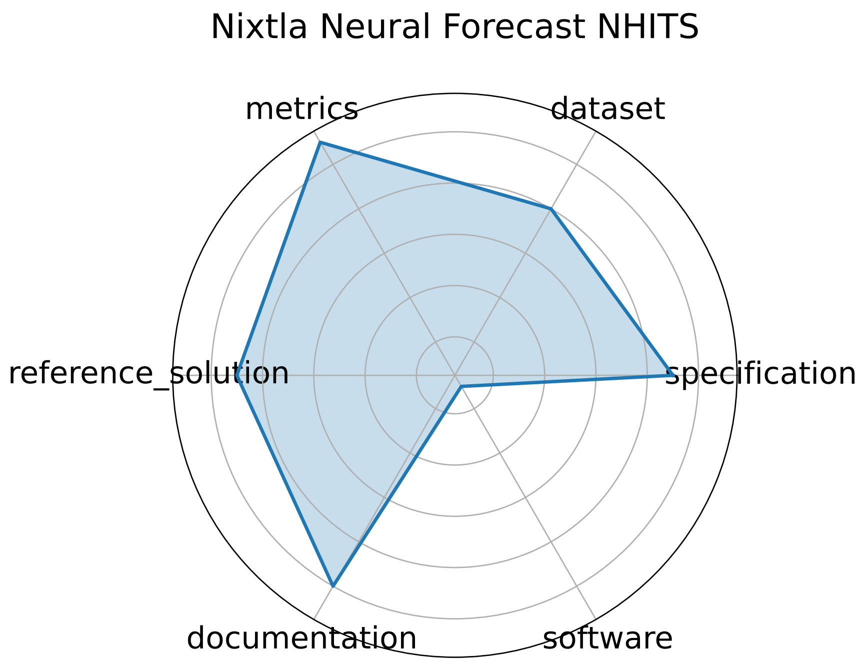

Figure 33: Nixtla Neural Forecast NHITS [6]

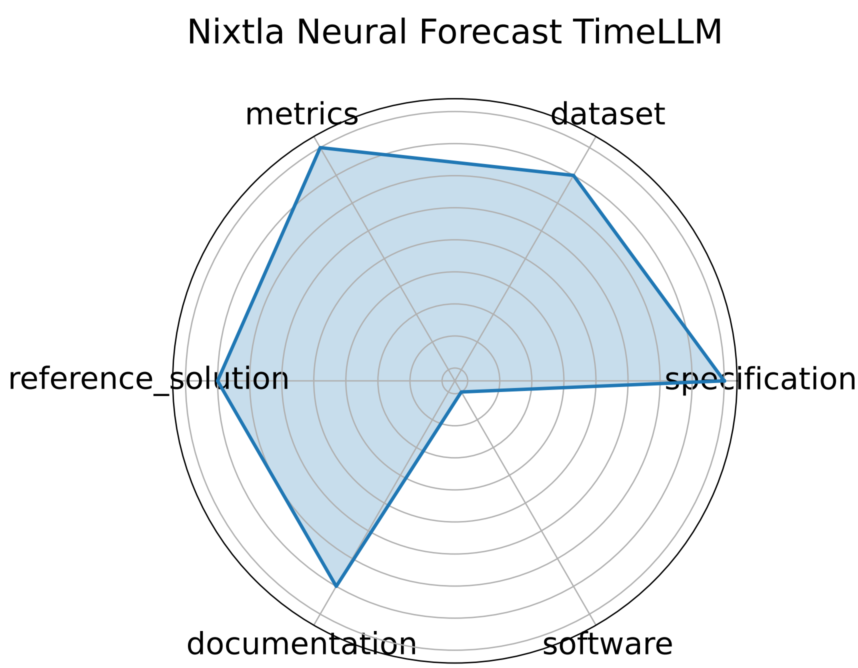

Figure 34: Nixtla Neural Forecast TimeLLM [18]

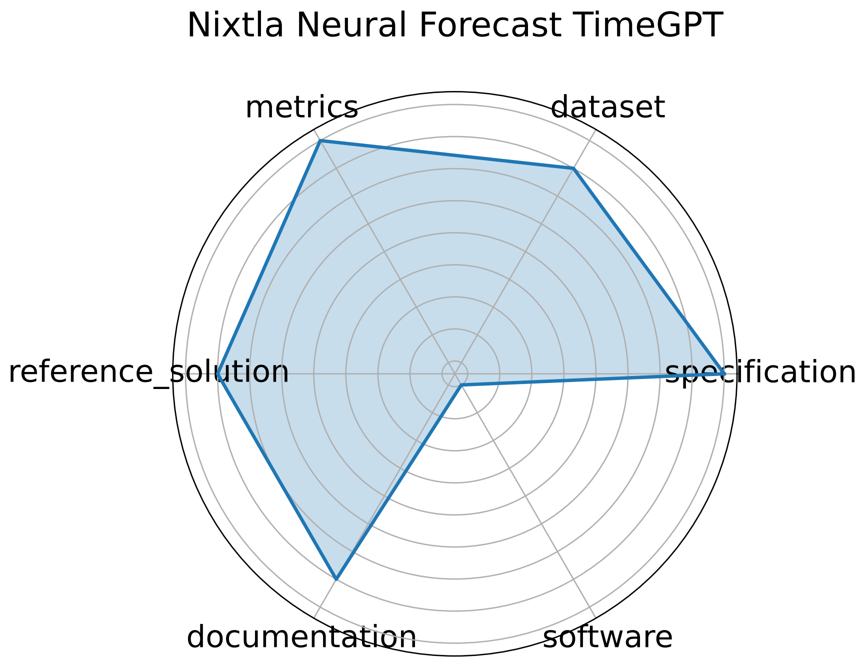

Figure 35: Nixtla Neural Forecast TimeGPT [15]

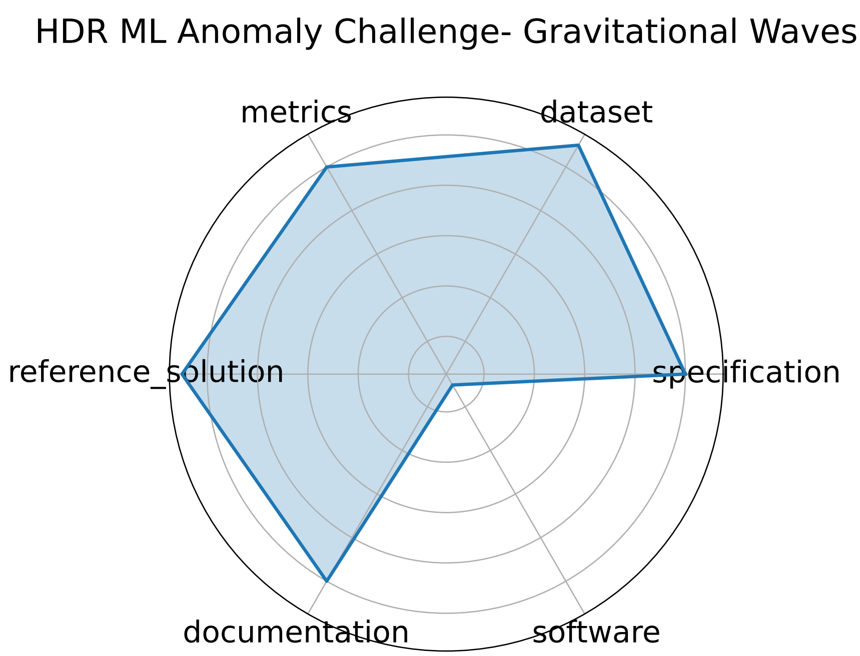

Figure 36: HDR ML Anomaly Challenge- Gravitational Waves [3]

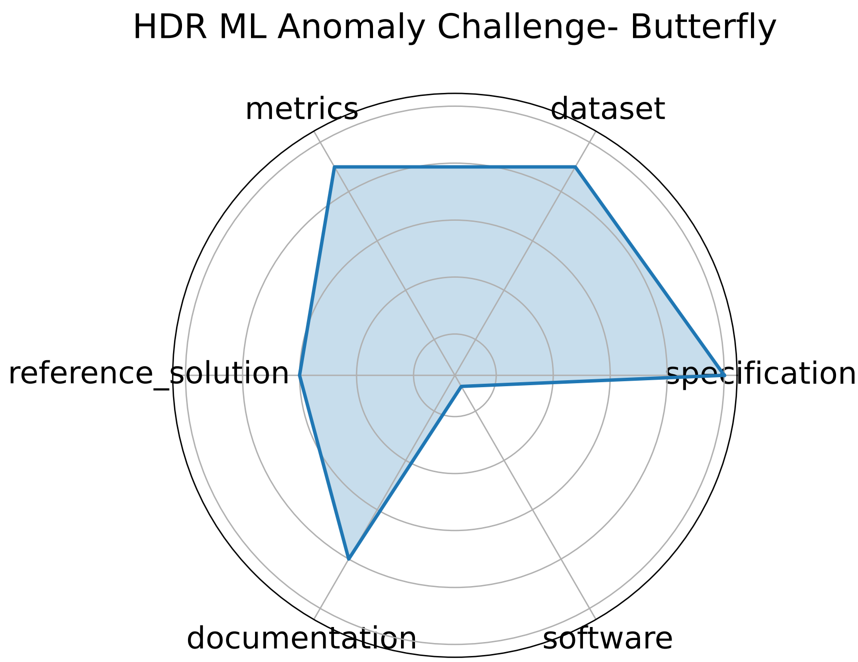

Figure 37: HDR ML Anomaly Challenge- Butterfly [4]

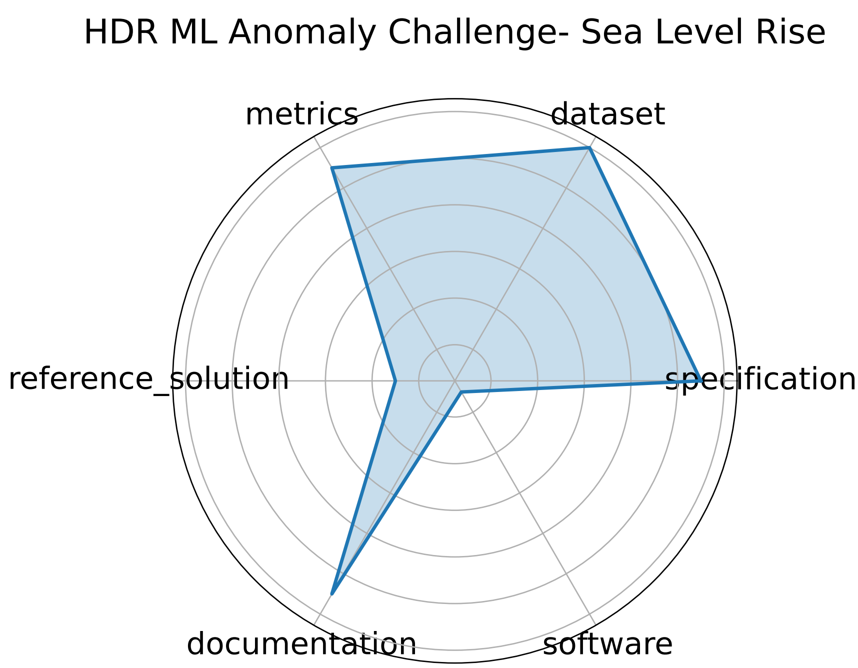

Figure 38: HDR ML Anomaly Challenge- Sea Level Rise [5]

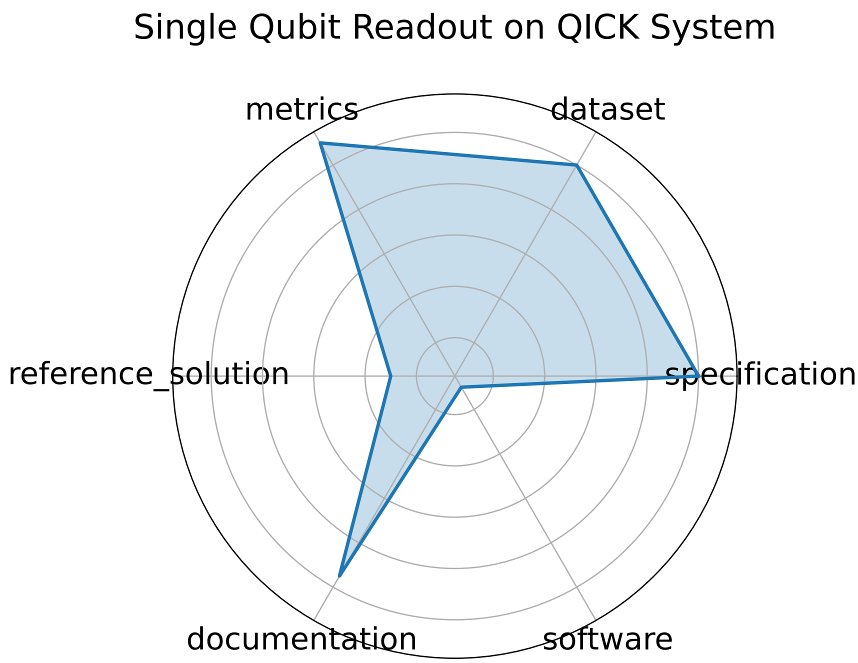

Figure 39: Single Qubit Readout on QICK System [17]

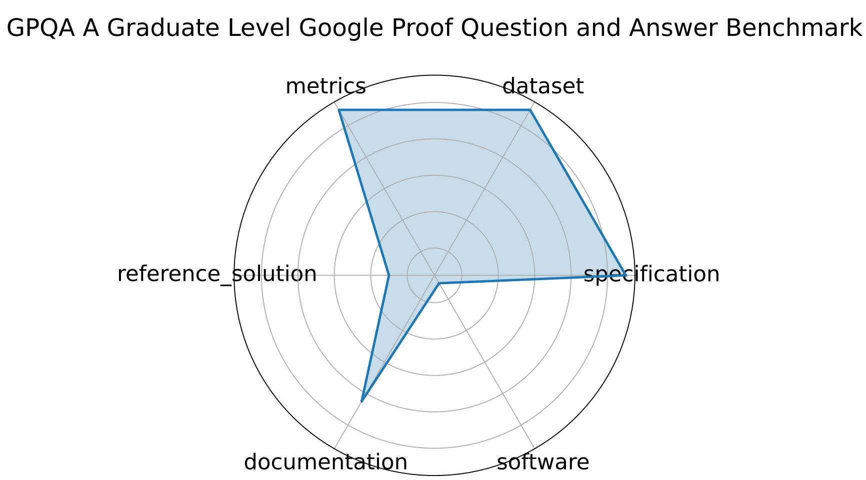

Figure 40: GPQA A Graduate Level Google Proof Question and Answer Benchmark [36]

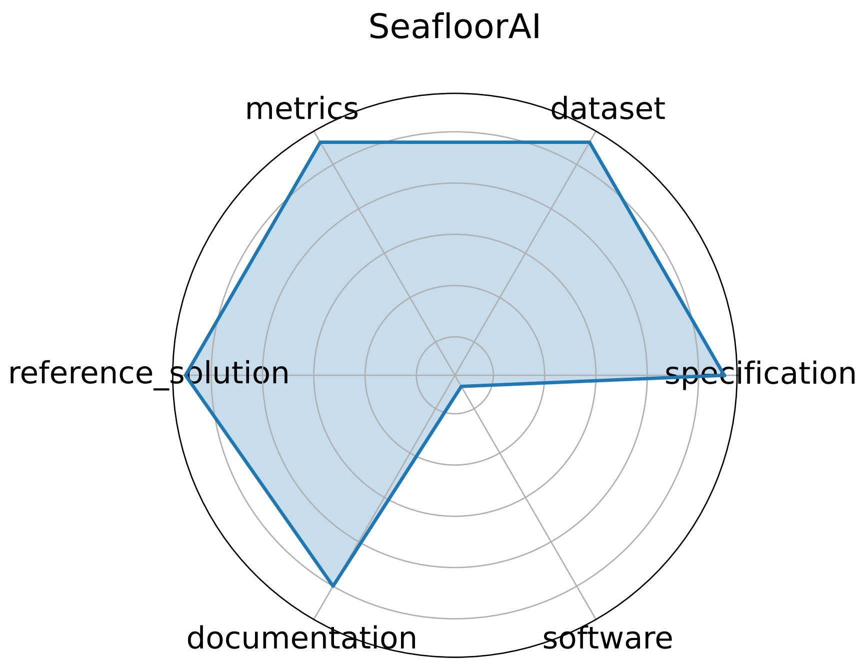

Figure 41: SeafloorAI [28]

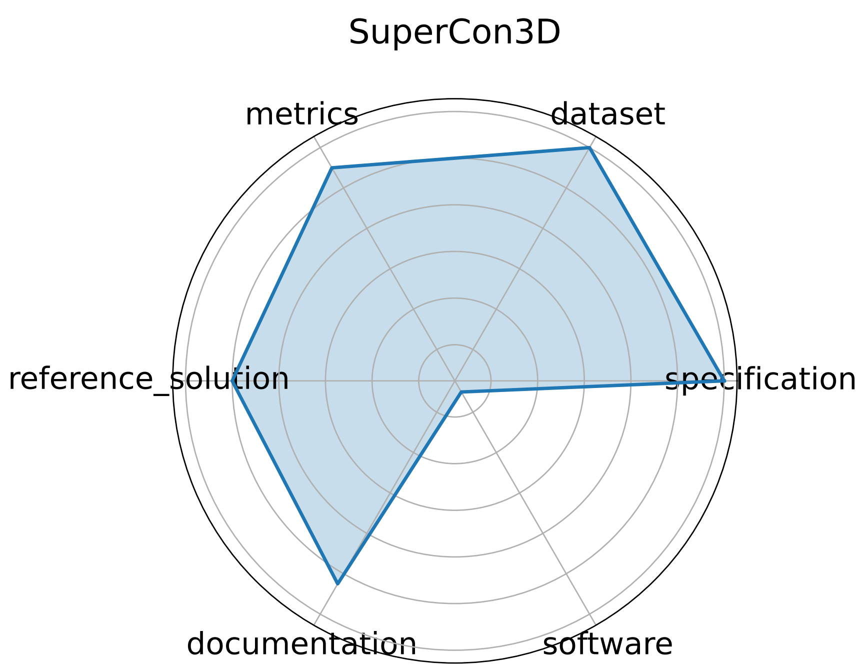

Figure 42: SuperCon3D [45]

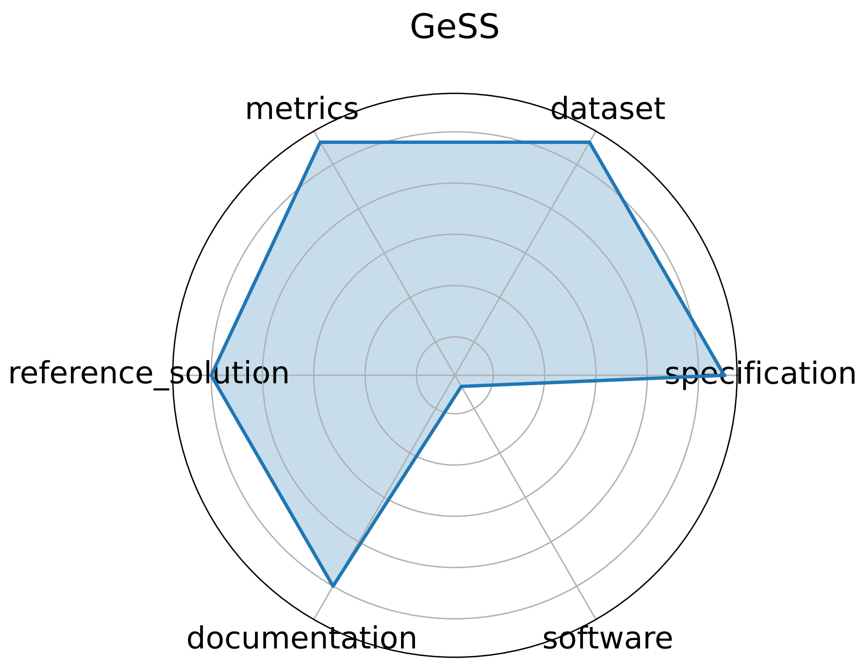

Figure 43: GeSS [44]

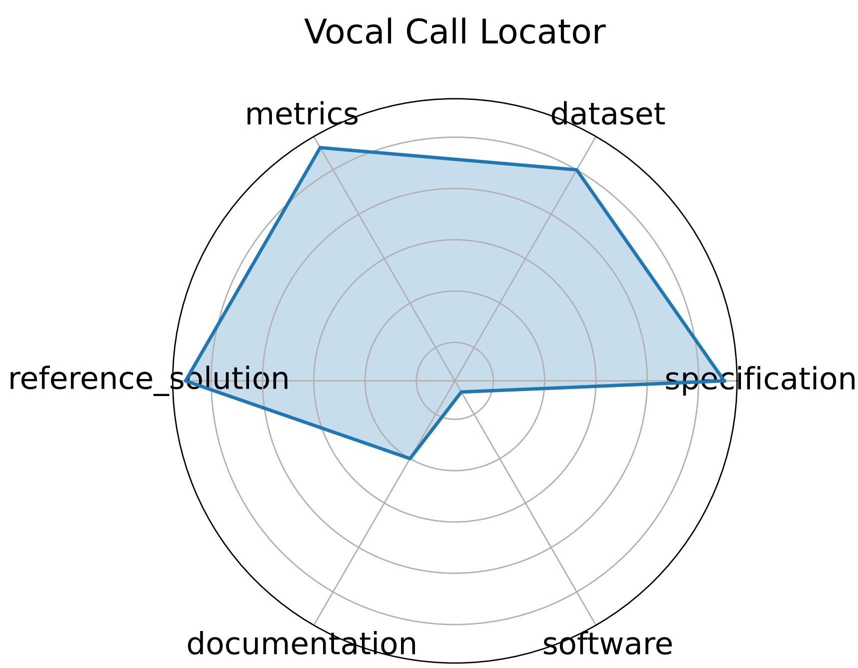

Figure 44: Vocal Call Locator [34]

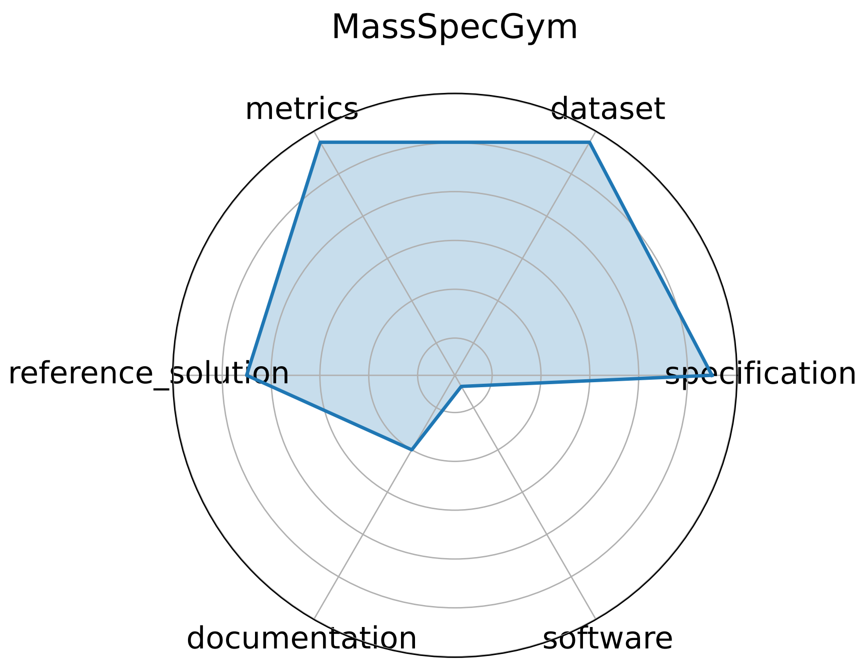

Figure 45: MassSpecGym [2]

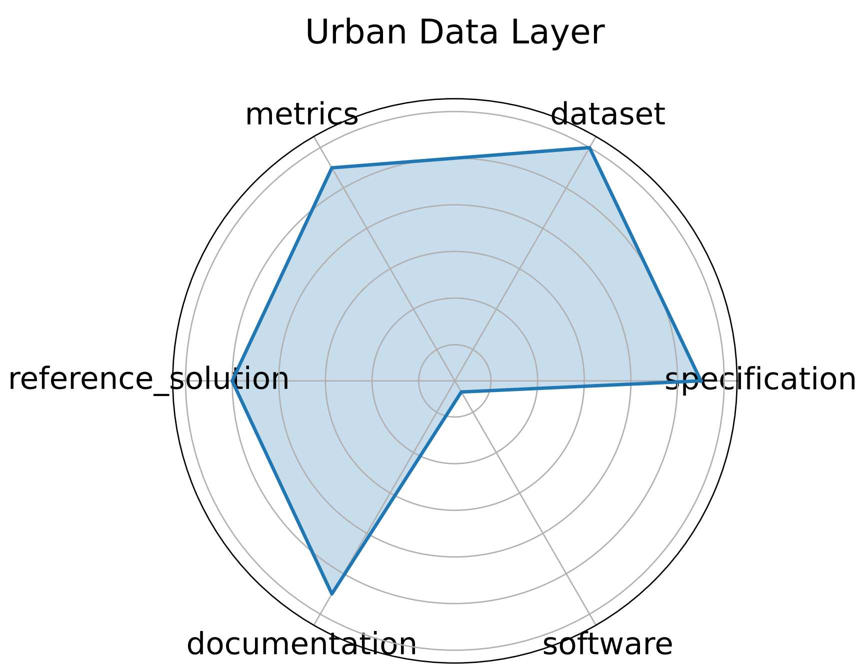

Figure 46: Urban Data Layer [38]

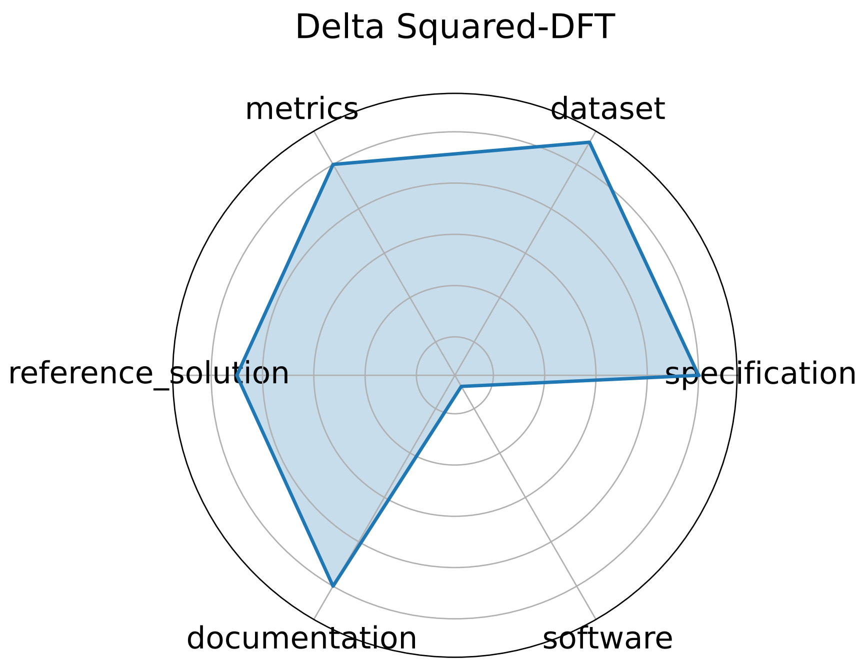

Figure 47: Delta Squared-DFT [24]

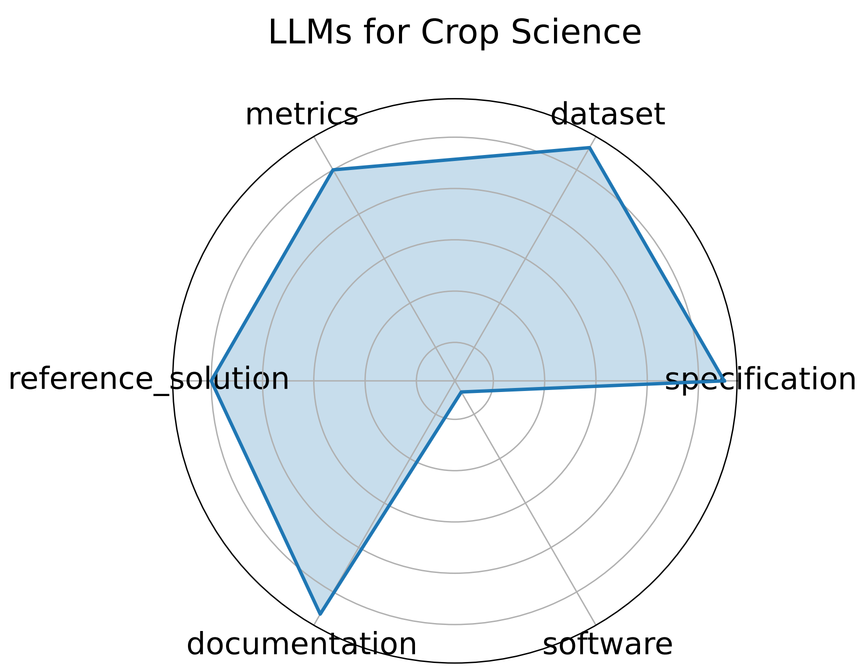

Figure 48: LLMs for Crop Science [33]

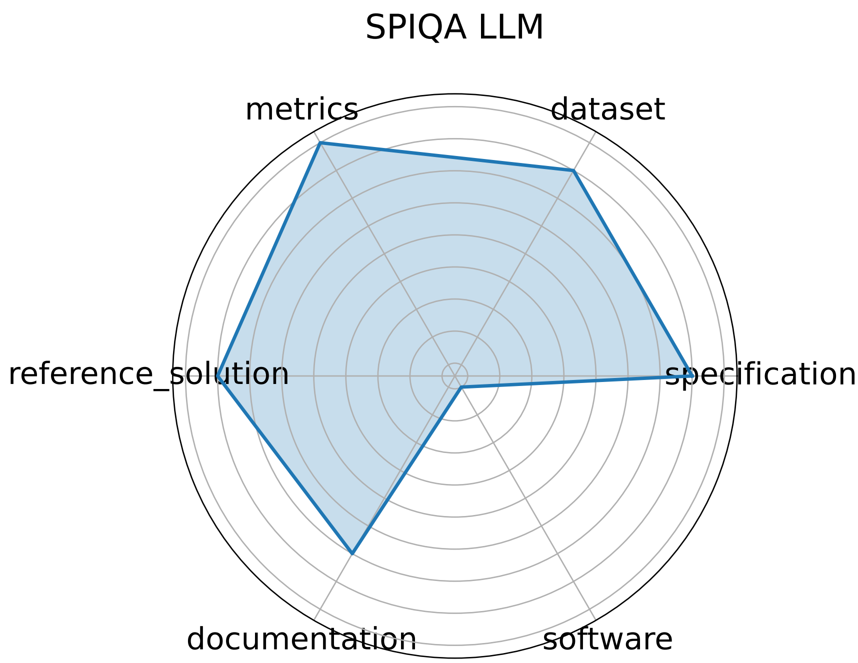

Figure 49: SPIQA LLM [43]

References

[1] Thea Aarrestad, Ekaterina Govorkova, Jennifer Ngadiuba, Ema Puljak, Maurizio Pierini, and Kinga Anna Wozniak. Unsupervised new physics detection at 40 mhz: Training dataset, June 2021.

[2] Roman Bushuiev, Anton Bushuiev, et al. Massspecgym: A benchmark for the discovery and identification of molecules, 2024. NeurIPS Spotlight Poster.

[3] Elizabeth G. Campolongo, Yuan-Tang Chou, Ekaterina Govorkova, Wahid Bhimji, Wei-Lun Chao, Chris Harris, Shih-Chieh Hsu, Hilmar Lapp, Mark S. Neubauer, Josephine Namayanja, Aneesh Subramanian, Philip Harris, Advaith Anand, David E. Carlyn, Subhankar Ghosh, Christopher Lawrence, Eric Moreno, Ryan Raikman, Jiaman Wu, Ziheng Zhang, Bayu Adhi, Mohammad Ahmadi Gharehtoragh, Saúl Alonso Monsalve, Marta Babicz, Furqan Baig, Namrata Banerji, William Bardon, Tyler Barna, Tanya Berger-Wolf, Adji Bousso Dieng, Micah Brachman, Quentin Buat, David C. Y. Hui, Phuong Cao, Franco Cerino, Yi-Chun Chang, Shivaji Chaulagain, An-Kai Chen, Deming Chen, Eric Chen, Chia-Jui Chou, Zih-Chen Ciou, Miles Cochran-Branson, Artur Cordeiro Oudot Choi, Michael Coughlin, Matteo Cremonesi, Maria Dadarlat, Peter Darch, Malina Desai, Daniel Diaz, Steven Dillmann, Javier Duarte, Isla Duporge, Urbas Ekka, Saba Entezari Heravi, Hao Fang, Rian Flynn, Geoffrey Fox, Emily Freed, Hang Gao, Jing Gao, Julia Gonski, Matthew Graham, Abolfazl Hashemi, Scott Hauck, James Hazelden, Joshua Henry Peterson, Duc Hoang, Wei Hu, Mirco Huennefeld, David Hyde, Vandana Janeja, Nattapon Jaroenchai, Haoyi Jia, Yunfan Kang, Maksim Kholiavchenko, Elham E. Khoda, Sangin Kim, Aditya Kumar, Bo-Cheng Lai, Trung Le, Chi-Wei Lee, JangHyeon Lee, Shaocheng Lee, Suzan van der Lee, Charles Lewis, Haitong Li, Haoyang Li, Henry Liao, Mia Liu, Xiaolin Liu, Xiulong Liu, Vladimir Loncar, Fangzheng Lyu, Ilya Makarov, Abhishikth Mallampalli Chen-Yu Mao, Alexander Michels, Alexander Migala, Farouk Mokhtar, Mathieu Morlighem, Min Namgung, Andrzej Novak, Andrew Novick, Amy Orsborn, Anand Padmanabhan, Jia-Cheng Pan, Sneh Pandya, Zhiyuan Pei, Ana Peixoto, George Percivall, Alex Po Leung, Sanjay Purushotham, Zhiqiang Que, Melissa Quinnan, Arghya Ranjan, Dylan Rankin, Christina Reissel, Benedikt Riedel, Dan Rubenstein, Argyro Sasli, Eli Shlizerman, Arushi Singh, Kim Singh, Eric R. Sokol, Arturo Sorensen, Yu Su, Mitra Taheri, Vaibhav Thakkar, Ann Mariam Thomas, Eric Toberer, Chenghan Tsai, Rebecca Vandewalle, Arjun Verma, Ricco C. Venterea, He Wang, Jianwu Wang, Sam Wang, Shaowen Wang, Gordon Watts, Jason Weitz, Andrew Wildridge, Rebecca Williams, Scott Wolf, Yue Xu, Jianqi Yan, Jai Yu, Yulei Zhang, Haoran Zhao, Ying Zhao, and Yibo Zhong. Building machine learning challenges for anomaly detection in science, 2025.

[4] Elizabeth G. Campolongo, Yuan-Tang Chou, Ekaterina Govorkova, Wahid Bhimji, Wei-Lun Chao, Chris Harris, Shih-Chieh Hsu, Hilmar Lapp, Mark S. Neubauer, Josephine Namayanja, Aneesh Subramanian, Philip Harris, Advaith Anand, David E. Carlyn, Subhankar Ghosh, Christopher Lawrence, Eric Moreno, Ryan Raikman, Jiaman Wu, Ziheng Zhang, Bayu Adhi, Mohammad Ahmadi Gharehtoragh, Saúl Alonso Monsalve, Marta Babicz, Furqan Baig, Namrata Banerji, William Bardon, Tyler Barna, Tanya Berger-Wolf, Adji Bousso Dieng, Micah Brachman, Quentin Buat, David C. Y. Hui, Phuong Cao, Franco Cerino, Yi-Chun Chang, Shivaji Chaulagain, An-Kai Chen, Deming Chen, Eric Chen, Chia-Jui Chou, Zih-Chen Ciou, Miles Cochran-Branson, Artur Cordeiro Oudot Choi, Michael Coughlin, Matteo Cremonesi, Maria Dadarlat, Peter Darch, Malina Desai, Daniel Diaz, Steven Dillmann, Javier Duarte, Isla Duporge, Urbas Ekka, Saba Entezari Heravi, Hao Fang, Rian Flynn, Geoffrey Fox, Emily Freed, Hang Gao, Jing Gao, Julia Gonski, Matthew Graham, Abolfazl Hashemi, Scott Hauck, James Hazelden, Joshua Henry Peterson, Duc Hoang, Wei Hu, Mirco Huennefeld, David Hyde, Vandana Janeja, Nattapon Jaroenchai, Haoyi Jia, Yunfan Kang, Maksim Kholiavchenko, Elham E. Khoda, Sangin Kim, Aditya Kumar, Bo-Cheng Lai, Trung Le, Chi-Wei Lee, JangHyeon Lee, Shaocheng Lee, Suzan van der Lee, Charles Lewis, Haitong Li, Haoyang Li, Henry Liao, Mia Liu, Xiaolin Liu, Xiulong Liu, Vladimir Loncar, Fangzheng Lyu, Ilya Makarov, Abhishikth Mallampalli Chen-Yu Mao, Alexander Michels, Alexander Migala, Farouk Mokhtar, Mathieu Morlighem, Min Namgung, Andrzej Novak, Andrew Novick, Amy Orsborn, Anand Padmanabhan, Jia-Cheng Pan, Sneh Pandya, Zhiyuan

Pei, Ana Peixoto, George Percivall, Alex Po Leung, Sanjay Purushotham, Zhiqiang Que, Melissa Quinnan, Arghya Ranjan, Dylan Rankin, Christina Reissel, Benedikt Riedel, Dan Rubenstein, Argyro Sasli, Eli Shlizerman, Arushi Singh, Kim Singh, Eric R. Sokol, Arturo Sorensen, Yu Su, Mitra Taheri, Vaibhav Thakkar, Ann Mariam Thomas, Eric Toberer, Chenghan Tsai, Rebecca Vandewalle, Arjun Verma, Ricco C. Venterea, He Wang, Jianwu Wang, Sam Wang, Shaowen Wang, Gordon Watts, Jason Weitz, Andrew Wildridge, Rebecca Williams, Scott Wolf, Yue Xu, Jianqi Yan, Jai Yu, Yulei Zhang, Haoran Zhao, Ying Zhao, and Yibo Zhong. Building machine learning challenges for anomaly detection in science, 2025.

[5] Elizabeth G. Campolongo, Yuan-Tang Chou, Ekaterina Govorkova, Wahid Bhimji, Wei-Lun Chao, Chris Harris, Shih-Chieh Hsu, Hilmar Lapp, Mark S. Neubauer, Josephine Namayanja, Aneesh Subramanian, Philip Harris, Advaith Anand, David E. Carlyn, Subhankar Ghosh, Christopher Lawrence, Eric Moreno, Ryan Raikman, Jiaman Wu, Ziheng Zhang, Bayu Adhi, Mohammad Ahmadi Gharehtoragh, Saúl Alonso Monsalve, Marta Babicz, Furqan Baig, Namrata Banerji, William Bardon, Tyler Barna, Tanya Berger-Wolf, Adji Bousso Dieng, Micah Brachman, Quentin Buat, David C. Y. Hui, Phuong Cao, Franco Cerino, Yi-Chun Chang, Shivaji Chaulagain, An-Kai Chen, Deming Chen, Eric Chen, Chia-Jui Chou, Zih-Chen Ciou, Miles Cochran-Branson, Artur Cordeiro Oudot Choi, Michael Coughlin, Matteo Cremonesi, Maria Dadarlat, Peter Darch, Malina Desai, Daniel Diaz, Steven Dillmann, Javier Duarte, Isla Duporge, Urbas Ekka, Saba Entezari Heravi, Hao Fang, Rian Flynn, Geoffrey Fox, Emily Freed, Hang Gao, Jing Gao, Julia Gonski, Matthew Graham, Abolfazl Hashemi, Scott Hauck, James Hazelden, Joshua Henry Peterson, Duc Hoang, Wei Hu, Mirco Huennefeld, David Hyde, Vandana Janeja, Nattapon Jaroenchai, Haoyi Jia, Yunfan Kang, Maksim Kholiavchenko, Elham E. Khoda, Sangin Kim, Aditya Kumar, Bo-Cheng Lai, Trung Le, Chi-Wei Lee, JangHyeon Lee, Shaocheng Lee, Suzan van der Lee, Charles Lewis, Haitong Li, Haoyang Li, Henry Liao, Mia Liu, Xiaolin Liu, Xiulong Liu, Vladimir Loncar, Fangzheng Lyu, Ilya Makarov, Abhishikth Mallampalli Chen-Yu Mao, Alexander Michels, Alexander Migala, Farouk Mokhtar, Mathieu Morlighem, Min Namgung, Andrzej Novak, Andrew Novick, Amy Orsborn, Anand Padmanabhan, Jia-Cheng Pan, Sneh Pandya, Zhiyuan Pei, Ana Peixoto, George Percivall, Alex Po Leung, Sanjay Purushotham, Zhiqiang Que, Melissa Quinnan, Arghya Ranjan, Dylan Rankin, Christina Reissel, Benedikt Riedel, Dan Rubenstein, Argyro Sasli, Eli Shlizerman, Arushi Singh, Kim Singh, Eric R. Sokol, Arturo Sorensen, Yu Su, Mitra Taheri, Vaibhav Thakkar, Ann Mariam Thomas, Eric Toberer, Chenghan Tsai, Rebecca Vandewalle, Arjun Verma, Ricco C. Venterea, He Wang, Jianwu Wang, Sam Wang, Shaowen Wang, Gordon Watts, Jason Weitz, Andrew Wildridge, Rebecca Williams, Scott Wolf, Yue Xu, Jianqi Yan, Jai Yu, Yulei Zhang, Haoran Zhao, Ying Zhao, and Yibo Zhong. Building machine learning challenges for anomaly detection in science, 2025.

[6] Cristian Challu, Kin G. Olivares, et al. Nhits: Neural hierarchical interpolation for time series forecasting. In *AAAI 2023*, 2023.

[7] Krishna Teja Chitty-Venkata, Siddhisanket Raskar, Bharat Kale, Farah Ferdaus, Aditya Tanikanti, Ken Raffenetti, Valerie Taylor, Murali Emani, and Venkatram Vishwanath. Llm-inference-bench: Inference benchmarking of large language models on ai accelerators, 2024.

[8] Papers With Code. Papers with code: Open machine learning benchmarks and leaderboards, 2025.

[9] Javier Duarte, Nhan Tran, Ben Hawks, Christian Herwig, Jules Muhizi, Shvetank Prakash, and Vijay Janapa Reddi. Fastml science benchmarks: Accelerating real-time scientific edge machine learning, 2022.

[10] Javier Duarte, Nhan Tran, Ben Hawks, Christian Herwig, Jules Muhizi, Shvetank Prakash, and Vijay Janapa Reddi. Fastml science benchmarks: Accelerating real-time scientific edge machine learning, 2022.

[11] Javier Duarte, Nhan Tran, Ben Hawks, Christian Herwig, Jules Muhizi, Shvetank Prakash, and Vijay Janapa Reddi. Fastml science benchmarks: Accelerating real-time scientific edge machine learning, 2022.

[12] Steven Farrell, Murali Emani, Jacob Balma, Lukas Drescher, Aleksandr Drozd, Andreas Fink, Geoffrey Fox, David Kanter, Thorsten Kurth, Peter Mattson, Dawei Mu, Amit Ruhela, Kento Sato, Koichi Shirahata, Tsuguchika Tabaru, Aristeidis Tsaris, Jan Balewski, Ben Cumming, Takumi Danjo, Jens Domke, Takaaki Fukai, Naoto Fukumoto, Tatsuya Fukushi, Balazs Gerofi, Takumi Honda, Toshiyuki Imamura, Akihiko Kasagi, Kentaro Kawakami, Shuhei Kudo, Akiyoshi Kuroda, Maxime Martinasso, Satoshi Matsuoka, Henrique Mendonça, Kazuki Minami, Prabhat Ram, Takashi Sawada, Mallikarjun Shankar, Tom St. John, Akihiro Tabuchi, Venkatram Vishwanath, Mohamed Wahib, Masafumi Yamazaki, and Junqi Yin. Mlperf hpc: A holistic benchmark suite for scientific machine learning on hpc systems, 2021.

[13] Wanling Gao, Fei Tang, Lei Wang, Jianfeng Zhan, Chunxin Lan, Chunjie Luo, Yunyou Huang, Chen Zheng, Jiahui Dai, Zheng Cao, Daoyi Zheng, Haoning Tang, Kunlin Zhan, Biao Wang, Defei Kong, Tong Wu, Minghe Yu, Chongkang Tan, Huan Li, Xinhui Tian, Yatao Li, Junchao Shao, Zhenyu Wang, Xiaoyu Wang, and Hainan Ye. Aibench: An industry standard internet service ai benchmark suite, 2019.

[14] Wanling Gao, Jianfeng Zhan, Lei Wang, Chunjie Luo, Daoyi Zheng, Xu Wen, Rui Ren, Chen Zheng, Xiwen He, Hainan Ye, Haoning Tang, Zheng Cao, Shujie Zhang, and Jiahui Dai. Bigdatabench: A scalable and unified big data and ai benchmark suite, 2018.

[15] Azul Garza, Cristian Challu, and Max Mergenthaler-Canseco. Timegpt-1, 2024.

[16] MLCommons Science Working Group. Mlcommons science working group benchmarks, 2023.

[17] Giuseppe Di Guglielmo, Botao Du, Javier Campos, Alexandra Boltasseva, Akash V. Dixit, Farah Fahim, Zhaxylyk Kudyshev, Santiago Lopez, Ruichao Ma, Gabriel N. Perdue, Nhan Tran, Omer Yesilyurt, and Daniel Bowring. End-to-end workflow for machine learning-based qubit readout with qick and hls4ml, 2025.

[18] Ming Jin, Shiyu Wang, Lintao Ma, Zhixuan Chu, James Y. Zhang, Xiaoming Shi, Pin-Yu Chen, Yuxuan Liang, Yuan-Fang Li, Shirui Pan, and Qingsong Wen. Time-llm: Time series forecasting by reprogramming large language models, 2024.

[19] Diana Kafkes and Jason St. John. Boostr: A dataset for accelerator control systems, 2021.

[20] Alex Karargyris, Micah J Sheller, et al. Federated benchmarking of medical artificial intelligence with medperf. *Nature Machine Intelligence*, 2023.

[21] Claudius Krause, Michele Faucci Giannelli, Gregor Kasieczka, Benjamin Nachman, Dalila Salamani, David Shih, Anna Zaborowska, Oz Amram, Kerstin Borras, Matthew R. Buckley, Erik Buhmann, Thorsten Buss, Renato Paulo Da Costa Cardoso, Anthony L. Caterini, Nadezda Chernyavskaya, Federico A. G. Corchia, Jesse C. Cresswell, Sascha Diefenbacher, Etienne Dreyer, Vijay Ekambaram, Engin Eren, Florian Ernst, Luigi Favaro, Matteo Franchini, Frank Gaede, Eilam Gross, Shih-Chieh Hsu, Kristina Jaruskova, Benno Käch, Jayant Kalagnanam, Raghav Kansal, Taewoo Kim, Dmitrii Kobylianskii, Anatolii Korol, William Korcari, Dirk Krücker, Katja Krüger, Marco Letizia, Shu Li, Qibin Liu, Xiulong Liu, Gabriel Loaiza-Ganem, Thandikire Madula, Peter McKeown, Isabell-A. Melzer-Pellmann, Vinicius Mikuni, Nam Nguyen, Ayodele Ore, Sofia Palacios Schweitzer, Ian Pang, Kevin Pedro, Tilman Plehn, Witold Pokorski, Huilin Qu, Piyush Raikwar, John A. Raine, Humberto Reyes-Gonzalez, Lorenzo Rinaldi, Brendan Leigh Ross, Moritz A. W. Scham, Simon Schnake, Chase Shimmin, Eli Shlizerman, Nathalie Soybelman, Mudhakar Srivatsa, Kalliopi Tsolaki, Sofia Vallecorsa, Kyongmin Yeo, and Rui Zhang. Calochallenge 2022: A community challenge for fast calorimeter simulation, 2024.

[22] J. Kvapil, G. Borca-Tasciuc, H. Bossi, K. Chen, Y. Chen, Y. Corrales Morales, H. Da Costa, C. Da Silva, C. Dean, J. Durham, S. Fu, C. Hao, P. Harris, O. Hen, H. Jheng, Y. Lee, P. Li, X. Li, Y. Lin, M. X. Liu, V. Loncar, J. P. Mitrevski, A. Olvera, M. L. Purschke, J. S. Renck, G. Roland, J. Schambach, Z. Shi, N. Tran, N. Wuerfel, B. Xu, D. Yu, and H. Zhang. Intelligent experiments through real-time ai: Fast data processing and autonomous detector control for sphenix and future eic detectors, 2025.

[23] Woosuk Kwon et al. Efficient memory management for large language model serving with pagedattention. In *SOSP 2023*, 2023.

[24] Wei Liu, Rong Chen, et al. Delta squared-dft: Machine-learning corrected density functional theory for reaction energetics, 2024. NeurIPS Poster.

[25] Zhengchun Liu, Hemant Sharma, Jun-Sang Park, Peter Kenesei, Antonino Miceli, Jonathan Almer, Rajkumar Kettimuthu, and Ian Foster. Braggnn: Fast x-ray bragg peak analysis using deep learning, 2021.

[26] Piotr Luszczek et al. Sabath: Fair metadata technology for surrogate benchmarks. Technical report, University of Tennessee, 2021.

[27] Simon Mo. vllm performance dashboard, 2024.

[28] Kien X. Nguyen, Fengchun Qiao, Arthur Trembanis, and Xi Peng. Seafloorai: A large-scale vision-language dataset for seafloor geological survey, 2024.

[29] Patrick Odagiu, Zhiqiang Que, Javier Duarte, Johannes Haller, Gregor Kasieczka, Artur Lobanov, Vladimir Loncar, Wayne Luk, Jennifer Ngadiuba, Maurizio Pierini, Philipp Rincke, Arpita Seksaria, Sioni Summers, Andre Sznajder, Alexander Tapper, and Thea K Årrestad. Ultrafast jet classification at the hl-lhc. *Machine Learning: Science and Technology*, 5(3):035017, July 2024.

[30] Ruben Ohana, Michael McCabe, Lucas Meyer, et al. The well: a large-scale collection of diverse physics simulations for machine learning. *NeurIPS*, 37:44989–45037, 2024.

[31] Kin G. Olivares, Cristian Challú, et al. Neuralforecast: User friendly state-of-the-art neural forecasting models. PyCon US, 2022.

[32] Benjamin Parpillon, Chinar Syal, Jieun Yoo, Jennet Dickinson, Morris Swartz, Giuseppe Di Guglielmo, Alice Bean, Douglas Berry, Manuel Blanco Valentin, Karri DiPetrillo, Anthony Badea, Lindsey Gray, Petar Maksimovic, Corrinne Mills, Mark S. Neubauer, Gauri Pradhan, Nhan Tran, Dahai Wen, and Farah Fahim. Smart pixels: In-pixel ai for on-sensor data filtering, 2024.

[33] Deepak Patel, Lan Zhao, et al. Large language models for crop science: Benchmarking domain reasoning and qa, 2024. NeurIPS Poster.

[34] Ralph Peterson, Aramis Tanelus, et al. Vocal call locator benchmark for localizing rodent vocalizations, 2024. NeurIPS Poster.

[35] Shuyu Qin, Joshua Agar, and Nhan Tran. Extremely noisy 4d-tem strain mapping using cycle consistent spatial transforming autoencoders. In *AI for Accelerated Materials Design - NeurIPS 2023 Workshop*, 2023.

[36] David Rein, Betty Li Hou, Asa Cooper Stickland, Jackson Petty, Richard Yuanzhe Pang, Julien Dirani, Julian Michael, and Samuel R. Bowman. Gpqa: A graduate-level google-proof q&a benchmark, 2023.

[37] Makoto Takamoto, Timothy Praditia, Raphael Leiteritz, Dan MacKinlay, Francesco Alesiani, Dirk Pflüger, and Mathias Niepert. Pdebench: An extensive benchmark for scientific machine learning, 2024.

[38] Yiheng Wang, Tianyu Wang, et al. Urbandatalayer: A unified data pipeline for urban science, 2024. NeurIPS Poster.

[39] Y. Wei, R. F. Forelli, C. Hansen, J. P. Levesque, N. Tran, J. C. Agar, G. Di Guglielmo, M. E. Mauel, and G. A. Navratil. Low latency optical-based mode tracking with machine learning deployed on fpgas on a tokamak. *Review of Scientific Instruments*, 95(7), July 2024.

[40] Jason Weitz, Dmitri Demler, Luke McDermott, Nhan Tran, and Javier Duarte. Neural architecture codesign for fast physics applications, 2025.

[41] Zhen Xu, Sergio Escalera, et al. Codabench: Flexible, easy-to-use, and reproducible meta-benchmark platform. *Patterns*, 3(7):100543, 2022.

[42] Lianmin Zheng, Liangsheng Yin, Zhiqiang Xie, Chuyue Sun, Jeff Huang, Cody Hao Yu, Shiyi Cao, Christos Kozyrakis, Ion Stoica, Joseph E. Gonzalez, Clark Barrett, and Ying Sheng. Sglang: Efficient execution of structured language model programs, 2024.

[43] Xiaoyan Zhong, Yijian Gao, et al. Spiqa-llm: Evaluating llm adapters on scientific figure qa, 2024. NeurIPS Poster.

[44] Deyu Zou, Shikun Liu, et al. Gess: Benchmarking geometric deep learning under scientific applications with distribution shifts, 2024. NeurIPS Poster.

[45] Zhong Zuo et al. Supercon3d: Learning superconductivity from ordered and disordered material structures, 2024. NeurIPS Poster.