Energy-Efficient Edge Computing Scheduling through AI-Enhanced Traffic Prediction in Heterogeneous Wireless Networks

Scatt Rarsean
Department of Computer Science and Engineering, University of Nevada, Reno, Reno, NV, USA.
scott.larsen457@unr.edu

Sanjay D. Sevastava
Department of Computer Science, University of North Texas, Denton, TX, USA.
sanjaysrivastava99@unt.edu

Abstract

The rapid proliferation of mobile devices and the Internet of Things has placed immense pressure on wireless network infrastructures, demanding ultra-low latency, high bandwidth, and energy-efficient computation. Edge computing has emerged as a critical paradigm to alleviate backhaul congestion and provide responsive services by placing computational resources in close proximity to end users. However, the inherent heterogeneity of next-generation wireless networks, encompassing diverse radio access technologies, fluctuating traffic patterns, and resource-limited edge nodes, introduces significant challenges for scheduling tasks in an energy-aware manner. This paper investigates a system-level framework that integrates artificial intelligence-driven traffic prediction with edge scheduling mechanisms to achieve energy-efficient operation across heterogeneous wireless environments. By anticipating spatiotemporal traffic dynamics through advanced predictive models, edge orchestrators can proactively allocate workloads, steer task offloading, and dynamically scale resource usage, thereby minimizing energy consumption while preserving stringent quality-of-service constraints. The discussion focuses on architectural design principles, structural trade-offs between predictive accuracy and computational overhead, robustness against uncertainty, fairness in resource allocation among competing tenants, and the broader governance and sustainability implications of embedding learning-based decision processes into critical communication infrastructures. The analysis draws on cross-domain insights from large-scale data center management, cyber-physical systems, and adaptive control theory to illuminate pathways toward self-optimizing, green edge ecosystems. A governance-aware perspective underscores the need for transparent, accountable mechanisms when AI models influence infrastructure operations that underpin societal functions. The paper avoids mathematical formalization and instead provides deep conceptual exploration of how AI-enhanced prediction can reshape the energy footprint of edge computing in an increasingly heterogeneous networked world.

Keywords

edge computing, energy efficiency, traffic prediction, heterogeneous wireless networks, scheduling, artificial intelligence, sustainability, governance.

1. Introduction

The ongoing expansion of mobile broadband, the Internet of Things, and immersive digital applications is reshaping the landscape of wireless communication infrastructures, forcing a fundamental reexamination of where and how computation is performed. Traditional cloud-centric architectures, while offering virtually unlimited computing capacity, incur high end-to-end latency and strain the core network with massive data backhaul, making them ill-suited for time-critical and context-aware services. Edge computing has consequently risen as a transformative design principle that distributes storage, processing, and intelligence to the network periphery, thereby enabling sub-millisecond response times, localized data handling, and improved resilience [1]. Mobile edge computing, in its many instantiations, positions computational resources at base stations, access points, and aggregation nodes, effectively collapsing the physical and logical distance between data sources and processing engines [2]. However, the promise of edge computing is contingent upon the ability to orchestrate tasks across a highly heterogeneous fabric of wireless technologies, including 5G New Radio, Wi-Fi 6/7, LTE, and emerging spectrum-sharing regimes, each presenting distinct channel characteristics, mobility patterns, and energy constraints [3]. Within such environments, intelligent scheduling that balances latency, throughput, and energy consumption becomes a cornerstone for operational viability.

Energy efficiency is now a first-order design objective rather than an afterthought, driven by escalating electricity costs, carbon footprint concerns, and the thermal limitations of densely deployed small cells and edge devices [4]. Edge nodes, often operating with limited power budgets and passive cooling, cannot rely on the abundant energy resources available in hyperscale data centers. Therefore, reducing unnecessary computation, communication, and cooling energy hinges on aligning workload execution with the actual demand dynamics. Idle periods, over-provisioning, and reactive resource allocation all contribute to energy wastage. Predicting future traffic demands with high spatiotemporal resolution offers a proactive lever to optimize scheduling decisions, allowing edge orchestrators to consolidate workloads, turn off underutilized nodes, or pre-fetch data at energy-favorable times [5]. The confluence of wireless network analytics and advanced artificial intelligence techniques, particularly deep learning models that can capture complex non-linear dependencies and multi-scale temporal patterns, opens new avenues for anticipatory energy management. This paper develops a system-level examination of how AI-enhanced traffic prediction can be embedded into the scheduling fabric of heterogeneous edge networks to drive energy-efficient operation, emphasizing architectural coherence, structural trade-offs, robustness, fairness, and governance considerations that accompany the deployment of learning-driven control loops.

2. Related Work and Motivation

A substantial body of literature has addressed the individual pillars upon which this paper rests: edge computing architectures, energy-aware resource allocation, wireless traffic prediction, and AI-driven scheduling. Foundational surveys on mobile edge computing have delineated reference architectures, deployment scenarios, and the communication-theoretic fundamentals that govern performance [2], [3]. Complementary work on energy-efficient offloading has demonstrated that dynamic voltage and frequency scaling, sleep modes, and workload migration can yield significant energy savings when guided by accurate workload forecasts [4], [9]. Separately, the domain of mobile traffic prediction has matured from simple autoregressive models to sophisticated deep spatiotemporal approaches that exploit the grid-like topology of urban base station deployments and the periodicities inherent in human mobility [6], [7]. Graph-based learning methods have further enriched predictive capabilities

by explicitly modeling the non-Euclidean relationships among network elements, capturing correlations that arise from shared user populations, adjacent coverage zones, and backhaul constraints [11], [13]. These advances collectively indicate that proactive resource management, fueled by anticipatory awareness of network conditions, can transcend the limitations of reactive heuristics.

Nevertheless, the integration of predictive intelligence into energy-efficient edge schedulers across heterogeneous wireless networks remains fragmented. Many scheduling frameworks operate under the assumption of quasi-static traffic or employ myopic optimization that fails to account for prediction uncertainty and model staleness. Traffic prediction models are often evaluated in isolation, decoupled from the scheduling objectives that they are meant to serve, leaving open the question of how prediction accuracy translates into tangible energy savings under practical operational constraints. Moreover, the heterogeneity of radio access technologies introduces diverse energy profiles, latency bounds, and reliability requirements that challenge monolithic scheduling policies. Cross-domain lessons from large-scale data center management reveal that even modest improvements in workload anticipation can reduce energy overhead by enabling thermal-aware consolidation and graceful degradation of non-critical services. Applied to the edge, such insights motivate a holistic architecture where prediction and scheduling co-evolve, sharing feedback signals that calibrate model updates and policy adaptations. The need for system-level thinking is further intensified by the rapid proliferation of AI accelerators at the edge, which themselves consume non-negligible power and must be scheduled alongside general-purpose compute and radio resources. The present work contributes to bridging these gaps by articulating a comprehensive system design perspective that weaves together traffic intelligence, multi-access heterogeneity, and energy-centric scheduling as an integrated whole.

3. System Architecture for Heterogeneous Edge Computing

Designing an energy-efficient scheduling system for heterogeneous wireless edge networks necessitates a layered architectural framework that cleanly separates concerns while enabling tight control loops. At the infrastructure layer, the network comprises a multitude of edge nodes with varying computational capabilities, energy budgets, and wireless interfaces. These nodes are interconnected through a mix of wired and wireless backhaul links, forming a multi-tier topology that may include femto-clouds at the far edge, aggregation points at street cabinets, and regional micro data centers. Heterogeneity is not merely a hardware attribute but extends to software stacks, virtualization environments, and ownership domains, as infrastructure may be operated by network operators, neutral hosts, or enterprise tenants [3], [10]. On top of this physical substrate, an orchestration layer abstracts the distributed resources into a logically unified resource pool, employing containerization and lightweight virtualization to enable seamless task migration and replication. This orchestration layer hosts the scheduling engine, which continuously maps incoming computation tasks, content caching requests, and data analytics jobs onto the available nodes while respecting energy budgets and service-level agreements.

The intelligence layer, which is the focus of this work, ingests real-time telemetry data from radio access network controllers, user equipment measurement reports, and edge node utilization counters, and feeds them into an AI-enhanced traffic prediction pipeline. The predicted spatiotemporal demand maps serve as input to the scheduling decision logic, enabling proactive resource reservation, anticipatory scaling, and selective node deactivation. A critical architectural decision concerns the placement of the prediction and scheduling

intelligence. Centralized approaches, where a global orchestrator aggregates all measurements and computes optimal schedules, benefit from a holistic view but suffer from single points of failure, high communication overhead, and latency in reacting to local fluctuations. Fully decentralized approaches, in which each edge node autonomously predicts its own load and makes independent scheduling choices, enhance resilience and scalability but risk oscillatory behavior and inefficient resource utilization due to lack of coordination. A federated hierarchical design, where regional coordinators synthesize predictions from clusters of edge nodes and negotiate resource allocations with peers, often strikes a favorable balance between global efficiency and local responsiveness. Such an architecture must also incorporate a data governance fabric that enforces privacy policies and access controls, especially when user mobility traces and application usage patterns are leveraged for model training.

Another architectural dimension is the temporal granularity of the control loop. Scheduling decisions for energy efficiency operate on multiple timescales: long-term capacity planning, medium-term server activation and deactivation cycles, and short-term packet-level forwarding and offloading. The prediction pipeline must correspondingly generate forecasts at multiple horizons, from milliseconds for radio resource block scheduling to hours for node power state management. This multi-scale requirement imposes stringent demands on model architecture, as a single monolithic predictor is unlikely to capture both fine-grained burstiness and diurnal trends. Consequently, the system design should accommodate a hierarchy of predictive models that work in concert, with fast models adjusting predictions conditioned on the outputs of slower, more accurate base models. The architectural discussion thus sets the stage for understanding how AI prediction modules seamlessly interleave with scheduling controllers, shaping the trade-offs that will be explored subsequently.

4. AI-Enhanced Traffic Prediction for Network Dynamics

The predictive core of the proposed framework rests on the capacity to model and forecast the spatiotemporal traffic intensity across a heterogeneous wireless landscape. Contemporary mobile traffic exhibits strong regularity arising from human circadian rhythms, weekly work-leisure cycles, and event-driven surges, yet it is also perturbed by unpredictable phenomena such as flash crowds, network failures, and mobility anomalies. Early work on mobile traffic prediction demonstrated the efficacy of recurrent neural networks and convolutional architectures that treat spatial grids of base stations as images, allowing the extraction of local spatial features and temporal evolution patterns [6], [7]. Building upon these foundations, graph neural networks have emerged as a powerful tool for representing the irregular topology of cellular deployments, where adjacency is defined not by Euclidean distance but by radio propagation characteristics, user handover relationships, and backhaul connectivity. By aggregating information from graph neighborhoods, models can learn latent representations that capture interference patterns and collaborative resource dynamics, enabling more accurate inference than grid-based counterparts in environments with non-uniform node distribution [11], [13].

The rise of transformer architectures has further pushed the frontier of traffic prediction by allowing models to attend to long-range dependencies without suffering from the vanishing gradient problems that hamper recurrent models. When combined with graph-structured inputs, transformer-based spatial-temporal predictors can simultaneously model multi-scale temporal correlations and complex spatial interactions, achieving state-of-the-art accuracy in forecasting tasks [12], [14]. Recent advances have introduced large-scale pre-trained models specifically tailored for network intelligence, enabling few-shot adaptation to new

deployment scenarios and reducing the need for extensive historical data at each edge site. The TIDES framework, for example, leverages a DeepSeek architecture to achieve notable spatial-temporal prediction performance, demonstrating that foundation model paradigms can be successfully transplanted to communication network analytics [18]. Such models, however, pose significant challenges in terms of computational footprint and energy consumption during both training and inference, raising a crucial tension: the very mechanism intended to reduce system-wide energy must itself be energy-efficient. This tension necessitates careful model compression, quantization, and hardware-aware deployment strategies, as well as adaptive invocation policies that trigger complex predictions only when simpler heuristics are insufficient.

A further critical aspect of the prediction subsystem is its ability to quantify uncertainty. Point forecasts, while useful, provide no indication of confidence, leaving the scheduler vulnerable to misinformed decisions when predictions deviate from reality. Probabilistic forecasting methods, including Bayesian neural networks, Monte Carlo dropout, and deep ensembles, produce predictive distributions that can be fed into risk-aware scheduling policies. For instance, if the predictive uncertainty surrounding a surge in traffic is high, the scheduler might opt for a conservative activation of additional edge nodes rather than relying on aggressive sleep modes. The feedback loop between scheduling outcomes and model retraining also deserves architectural attention; prediction errors that result in quality-of-service violations or energy overruns should be systematically logged and used to fine-tune the model, closing the loop between planning and execution. In heterogeneous networks, the diversity of traffic patterns across different radio access technologies implies that prediction models may need to be technology-specific or conditioned on radio access type embeddings, further complicating the model management lifecycle. Addressing these challenges requires a robust MLOps pipeline integrated into the edge orchestrator, capable of continuous evaluation, drift detection, and safe model updates without disrupting live traffic.

5. Energy-Efficient Scheduling Design and Trade-offs

The scheduling engine that consumes the predicted traffic maps must reconcile multiple interacting objectives: minimizing total energy consumption, satisfying per-application latency and reliability constraints, maintaining fairness among tenants, and ensuring system stability. Energy efficiency at the edge can be pursued through several complementary mechanisms. Workload consolidation concentrates active tasks onto a minimal set of nodes, allowing idle servers to transition into low-power sleep states; the timing and selection of which nodes to deactivate critically depend on accurate predictions of how long the low-demand period will last, because frequent state transitions incur energy penalties and wear-and-tear. Task offloading decisions determine whether a computation is executed locally on the user device, at a nearby edge node, or forwarded to a distant cloud data center, each option carrying distinct energy implications for both the device and the infrastructure. Dynamic frequency and voltage scaling adapts processor performance to the instantaneous workload, but must be coordinated with radio resource allocation to avoid creating bottlenecks in the wireless segment.

Integrating AI-driven traffic prediction transforms these mechanisms from reactive to proactive. With forecasts of upcoming demand, the scheduler can pre-emptively spin up additional edge nodes in anticipation of load spikes, avoiding last-second scrambling that often leads to over-provisioning and energy waste. Similarly, content caching decisions can be optimized by predicting which multimedia objects will experience high request volumes,

pre-loading them onto edge caches during off-peak energy tariff periods. The quality of the prediction fundamentally governs the efficiency gains achievable; overly conservative predictions lead to under-utilization and idle energy draw, while overly aggressive predictions cause performance degradation and potential service disruptions that might incur even higher energy costs through retransmissions and user dissatisfaction. Therefore, the scheduling design must embed a stochastic optimization perspective that treats predictions as imperfect probability distributions, balancing expected energy savings against risk of violation.

The structural trade-off between predictive accuracy and computational overhead is paramount. Complex transformer-based models or deep graph networks consume substantial processing cycles and memory, which, if not carefully accounted for, can erode the energy savings they enable. A practical system might adopt a cascaded architecture where a lightweight, low-power model handles routine predictions with high confidence, while a more sophisticated model is invoked for ambiguous or high-stakes scenarios. Such an approach aligns with system-level energy proportionality principles, ensuring that the intelligence layer does not become a dominant energy consumer itself. Furthermore, scheduling policies must be designed with an awareness of heterogeneity: a 5G millimeter-wave small cell may have radically different energy consumption patterns and traffic profiles compared to a Wi-Fi access point or a satellite backhaul node. Granularity of the scheduling decisions, spanning individual containers, virtual network functions, and physical resource blocks, introduces additional complexity, as the combinatorial search space explodes. Heuristic, metaheuristic, or learning-based schedulers, such as those using deep reinforcement learning, can approximate near-optimal policies in such expansive state-action spaces, but they require careful curriculum design and reward shaping that properly weights energy, latency, and fairness [15], [16]. Robustness to model drift and environmental non-stationarity further demands online learning capabilities, so that the scheduler continuously adapts to shifts in user behavior, application mix, and network topology without human intervention.

6. Robustness, Fairness, and Policy Implications

Deploying AI-driven prediction and scheduling in production heterogeneous edge networks brings robustness to the forefront. Adversarial perturbations, data pipeline failures, and concept drifts can degrade prediction quality, potentially leading to cascading energy mismanagement and service outages. A robust system must incorporate graceful degradation pathways, such as fallback to simpler prediction or rule-based scheduling when confidence metrics dip below thresholds, and must include anomaly detection modules that identify when the prediction-scheduling loop is operating outside its validated envelope. Architectural redundancy, including geo-distributed model replicas and leader election for centralized scheduling components, can mitigate single points of failure. Testing and validation frameworks that go beyond offline accuracy metrics to include closed-loop energy efficiency and stability under realistic network simulators or digital twins are essential for building system trustworthiness [19]. The integration of formal verification approaches for neural network components, though still nascent, offers a longer-term avenue to certify safety properties.

Fairness emerges as a multidimensional requirement in multi-tenant edge environments. Infrastructure providers must allocate computational and radio resources among competing service providers, each with different traffic characteristics and energy priorities, while avoiding discrimination or resource starvation. Predictive scheduling can inadvertently amplify inequities if models are trained predominantly on data from densely populated urban

areas, leading to under-prediction and consequent resource undersupply in rural or underserved regions. Counterfactual fairness metrics and inclusive training data curation become necessary governance mechanisms to ensure that AI-driven optimization does not perpetuate digital divides. Energy-aware scheduling also raises questions of intergenerational and geographical equity, as the carbon footprint of edge infrastructure affects communities differently depending on the local energy mix and climate. Thus, the design of scheduling policies must be sensitive to the environmental justice dimension, potentially incorporating carbon-intensity signals from regional power grids into the objective function, incentivizing workload migration to times and places where renewable energy is abundant [20].

From a broader policy and governance standpoint, embedding AI into critical communication infrastructures demands transparency, accountability, and contestability. When an automated scheduler decides to throttle a video stream or delay a vehicular safety message to save energy, affected stakeholders must have access to explainable decision rationales, and regulatory bodies need auditing capabilities to verify compliance with net neutrality and public safety regulations. The use of end-user behavioral data for traffic prediction triggers privacy concerns that intersect with data protection frameworks such as the GDPR, requiring federated learning, differential privacy, or on-device inference techniques to minimize data exposure while retaining predictive utility [21]. The governance of AI-enhanced edge scheduling should be viewed as part of a broader socio-technical system, where technical design choices are entangled with institutional norms, market structures, and public values. Multi-stakeholder governance bodies, independent algorithmic auditors, and open benchmarking standards can foster a responsible innovation ecosystem that balances energy efficiency with fundamental rights and societal welfare. Ultimately, the transition toward green, intelligent edge networks is not a purely technical undertaking; it is a governance challenge that will determine the legitimacy and long-term sustainability of pervasive AI-managed infrastructures.

7. Conclusion

This paper has presented an interdisciplinary system-level exploration of energy-efficient edge computing scheduling through AI-enhanced traffic prediction in heterogeneous wireless networks. By weaving together perspectives from distributed systems, machine learning, wireless communications, and technology governance, the discussion has illuminated the architectural imperatives, structural trade-offs, and societal dimensions that accompany the insertion of predictive intelligence into the scheduling fabric of the edge. The analysis emphasized that the promise of proactive energy management hinges on a delicate interplay between prediction fidelity, computational overhead, scheduling granularity, and architectural coherence. Heterogeneity, far from being an obstacle to be abstracted away, is a fundamental design parameter that must be explicitly modeled and leveraged to achieve true energy proportionality and resilience. Robustness considerations dictate that AI components be integrated with fallback mechanisms, uncertainty quantification, and continuous validation to avoid brittle automation failures. Fairness and governance concerns remind us that energy-efficient scheduling must be aligned with ethical principles, regulatory mandates, and the equitable distribution of digital and environmental benefits. Looking ahead, the convergence of foundation models for network intelligence, open radio access network interfaces, and sustainability regulations will likely accelerate the deployment of self-optimizing edge ecosystems. Future research should target the co-design of lightweight predictive architectures and scalable scheduling algorithms that jointly optimize energy, performance,

and trustworthiness across the entire edge-to-cloud continuum, while engaging policy scholars and civil society to ensure that the resulting infrastructures serve the public good. The path toward green edge computing is thus as much a socio-technical undertaking as an engineering one, demanding sustained interdisciplinary dialogue and responsible innovation.

References

1. Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. IEEE Internet of Things Journal, 3(5), 637–646.
2. Mao, Y., You, C., Zhang, J., Huang, K., & Letaief, K. B. (2017). A survey on mobile edge computing: The communication perspective. IEEE Communications Surveys & Tutorials, 19(4), 2322–2358.
3. Abbas, N., Zhang, Y., Taherkordi, A., & Skeie, T. (2018). Mobile edge computing: A survey. IEEE Internet of Things Journal, 5(1), 450–465.
4. Chen, Y., Zhang, N., Zhang, Y., Chen, X., Wu, W., & Shen, X. S. (2020). Energy efficient dynamic offloading in mobile edge computing via deep reinforcement learning. IEEE Transactions on Mobile Computing, 20(5), 1772–1784.
5. Agiwal, M., Roy, A., & Saxena, N. (2016). Next generation 5G wireless networks: A comprehensive survey. IEEE Communications Surveys & Tutorials, 18(3), 1617–1655.
6. Wang, J., Tang, J., Xu, Z., Wang, Y., Xue, G., Xing, Y., & Yang, D. (2017). Spatiotemporal modeling and prediction in cellular networks: A big data enabled deep learning approach. IEEE Journal on Selected Areas in Communications, 35(9), 1924–1937.
7. Zhang, C., & Patras, P. (2018). Long-term mobile traffic forecasting using deep spatio-temporal neural networks. In Proceedings of the Eighteenth ACM International Symposium on Mobile Ad Hoc Networking and Computing (pp. 231–240). ACM.
8. Li, Y., Yu, R., Shahabi, C., & Liu, Y. (2018). Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. In International Conference on Learning Representations.
9. Wang, C., Liang, C., Yu, F. R., Chen, Q., & Tang, L. (2017). Computation offloading and resource allocation in wireless cellular networks with mobile edge computing. IEEE Transactions on Wireless Communications, 16(8), 4924–4938.
10. Zhang, J., Xia, W., Yan, F., & Shen, L. (2018). Joint computation offloading and resource allocation optimization in heterogeneous networks with mobile edge computing. IEEE Access, 6, 54752–54763.
11. Wang, S., Wang, T., & Wang, X. (2022). Graph neural networks for wireless communications: From basics to applications. IEEE Communications Surveys & Tutorials, 24(1), 686–723.
12. Zerveas, G., Jayaraman, S., Patel, D., Bhamidipaty, A., & Eickhoff, C. (2021). A transformer-based framework for multivariate time series representation learning. In Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (pp. 2114–2124). ACM.

13. Bai, L., Yao, L., Li, C., Wang, X., & Wang, C. (2020). Adaptive graph convolutional recurrent network for traffic forecasting. In Advances in Neural Information Processing Systems, 33.

14. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. In Advances in Neural Information Processing Systems, 30.

15. Huang, L., Bi, S., & Zhang, Y. J. (2020). Deep reinforcement learning for online computation offloading in wireless powered mobile-edge computing networks. IEEE Transactions on Mobile Computing, 19(11), 2581–2593.

16. Mao, Y., Zhang, J., & Letaief, K. B. (2016). Dynamic computation offloading for mobile-edge computing with energy harvesting devices. IEEE Journal on Selected Areas in Communications, 34(12), 3590–3605.

17. Zhang, C., Zhang, H., Qiao, J., Li, Z., & Alouini, M. S. (2025). TIDES: Traffic Intelligence with DeepSeek Enhanced Spatial Temporal Prediction. IEEE Journal on Selected Areas in Communications.

18. Deng, R., Lu, R., Lai, C., Luan, T. H., & Liang, H. (2016). Optimal workload allocation in fog-cloud computing toward balanced delay and power consumption. IEEE Internet of Things Journal, 3(6), 1171–1181.

19. Beloglazov, A., & Buyya, R. (2012). Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data centers. Concurrency and Computation: Practice and Experience, 24(13), 1397–1420.

20. Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, 28(4), 689–707.