

# Agentic Edge Intelligence: A Research Agenda

Lauri Lovén  
University of Oulu  
Oulu, Finland  
lauri.loven@oulu.fi

Reza Farahani  
University of Klagenfurt  
Klagenfurt, Austria  
reza.farahani@aau.at

Iilir Murturi  
University of Prishtina  
Prishtina, Kosovo  
ilir.murturi@uni-pr.edu

Stephan Sigg  
Aalto University  
Helsinki, Finland  
stephan.sigg@aalto.fi

Schahram Dustdar  
Technical University of Vienna  
Vienna, Austria  
dustdar@dsg.tuwien.ac.at

## Abstract

Agentic AI is rapidly transforming autonomous decision-making, yet its deployment across the edge-cloud continuum remains poorly understood. This paper introduces the concept of agentic edge intelligence, an emerging paradigm in which autonomous agents operate across the computing continuum to negotiate computational resources, data, and services within dynamic digital marketplaces. We position this concept at the intersection of edge intelligence, multi-agent systems, and computational economics, where distributed decision-making replaces centralized orchestration. The paper outlines key research challenges, including scalability, interoperability, market stability, and ethical governance, and proposes a research agenda addressing theoretical, architectural, and societal dimensions. By integrating mechanism design with trustworthy AI and edge computing, the real-time AI economy envisions a self-organizing infrastructure for efficient, transparent, and equitable resource exchange in future digital ecosystems.

## CCS Concepts

• **Computing methodologies** → **Multi-agent systems; Intelligent agents.**

## Keywords

Agentic AI, Multi-agent Systems, Computing Continuum, Edge Intelligence, Large Language Model (LLM).

## ACM Reference Format:

Lauri Lovén, Reza Farahani, Iilir Murturi, Stephan Sigg, and Schahram Dustdar. 2025. Agentic Edge Intelligence: A Research Agenda. In *2025 IEEE/ACM 18th International Conference on Utility and Cloud Computing (UCC '25)*, December 01–04, 2025, Nantes, France. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3773274.3777421>

## 1 Introduction

Agentic edge intelligence, or language-model driven multi-agent systems [2], distributed in the computing continuum, will transform how digital services are organized and delivered. AI agents have

rapidly advanced to sophisticated planning and negotiation, while the edge intelligence [16] paradigm offers low-latency, privacy-aware infrastructures and distributed AI training and inference frameworks. However, the convergence of these trends raises questions of trust, security, fairness, and scalability when resources are traded ad hoc among heterogeneous stakeholders.

In the '90s and early 2000's, Multi-Agent Systems (MAS) (Fig. 1) provided a structured framework for engineering distributed intelligence. MAS research established models of autonomy and coordination, including the Belief–Desire–Intention (BDI) architecture [18], agent communication languages [12], and organizational frameworks [24]. These systems enable reasoning and negotiation among autonomous entities but fail to scale beyond controlled environments due to unresolved issues of trust, mobility, and interoperability [9, 27]. However, by the mid-2000s, when MAS research dwindled, distributed autonomy remained largely conceptual.

This vision has now re-emerged through Agentic Artificial Intelligence (AAI), which leverages Language Models (LMs), Large and Small (LLMs, SLMs) as general-purpose reasoning engines [2]. LLM-based agents can interpret goals, generate plans, and coordinate with peers through natural language [29]. Frameworks such as ReAct, Toolformer, and Voyager link language-based reasoning to system control via APIs, operationalizing many goals of classical MAS. A critical enabler of this shift is the computing continuum, integrating cloud, edge, and device layers [13, 15]. Large-scale reasoning and orchestration reside in the cloud, while perception and actuation occur near data sources at the edge, allowing agents to act locally and reason globally. This infrastructure provides scalability and trust mechanisms that earlier MAS architectures lacked. Yet, foundational MAS concepts remain relevant: BDI offers a useful deliberation model; FIPA ACL formalizes agent communication; and semantic frameworks such as W3C Web of Things (WoT) and ETSI SAREF now supply the interoperability that early systems lacked.

Yet, long-standing challenges persist [8]. Trust and security issues now manifest as prompt injection, unsafe tool use, and data leakage in LLM-based agents. Verification of reasoning-to-action grounding remains unresolved. Governance and compliance—previously theoretical—are now technical imperatives as agents cross legal and organizational boundaries under frameworks such as the EU AI Act and NIST AI RMF.

Recent advances offer partial solutions. Trusted Execution Environments (TEEs) and zero-trust architectures [23] protect code



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

UCC '25, Nantes, France

© 2025 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-2285-1/25/12

<https://doi.org/10.1145/3773274.3777421>

	Multi-agent Systems	Agentic AI	Agentic Edge Intelligence
Methods	<ul style="list-style-type: none"> <li>- Symbolic reasoning</li> <li>- Logic-based autonomy</li> <li>- BDI models</li> </ul>	<ul style="list-style-type: none"> <li>- LLM-based reasoning</li> <li>- Natural language planning</li> <li>- Tool-use APIs</li> </ul>	<ul style="list-style-type: none"> <li>- Hybrid/neurosymbolic reasoning</li> <li>- Distributed cognition on computing continuum resources</li> </ul>
Challenges	<ul style="list-style-type: none"> <li>- Security of mobile agents</li> <li>- Trust and ontology alignment</li> <li>- Scalability limits</li> </ul>	<ul style="list-style-type: none"> <li>- Prompt-level robustness</li> <li>- Grounding and verification</li> <li>- Behavioural security</li> </ul>	<ul style="list-style-type: none"> <li>- Orchestration of heterogeneous infrastructures</li> <li>- Multi-objective optimisation</li> <li>- Non-cooperative ecosystems, negotiation</li> <li>- Cross-domain governance and compliance</li> </ul>
Computation	<ul style="list-style-type: none"> <li>- Local reasoning on isolated nodes</li> <li>- Limited coordination</li> </ul>	<ul style="list-style-type: none"> <li>- Centralised LLM reasoning and decision generation</li> </ul>	<ul style="list-style-type: none"> <li>- Multi-level/continuum reasoning (IoT-edge-cloud) with localised actuation (IoT/edge)</li> <li>- Adaptive workload offloading, migration and scheduling</li> <li>- Federated adaptation</li> </ul>
Communication	<ul style="list-style-type: none"> <li>- Message passing</li> <li>- Shared ontologies (OWL/RDF)</li> <li>- FIPA ACL Semantics</li> </ul>	<ul style="list-style-type: none"> <li>- API-based interaction</li> <li>- Function calling</li> <li>- Tool orchestration</li> </ul>	<ul style="list-style-type: none"> <li>- Semantic interoperability via WOT, SOSA/SSN, SAREF</li> <li>- Dynamic data exchange via event-driven architectures</li> </ul>
Infrastructure	<ul style="list-style-type: none"> <li>- Ad-hoc middleware</li> <li>- Agent frameworks (JADE/FIPA)</li> </ul>	<ul style="list-style-type: none"> <li>- LangChain</li> <li>- MCP</li> <li>- Semantic Kernel</li> </ul>	<ul style="list-style-type: none"> <li>- Container/WASM-based orchestration</li> <li>- Serverless trusted execution</li> <li>- Federated learning and inference</li> <li>- AI Interconnect</li> </ul>
Evaluation	<ul style="list-style-type: none"> <li>- Task success</li> <li>- Goal satisfaction</li> </ul>	<ul style="list-style-type: none"> <li>- Accuracy</li> <li>- Coherence</li> <li>- Reasoning trace quality</li> </ul>	<ul style="list-style-type: none"> <li>- Performance</li> <li>- Trustworthiness</li> <li>- Energy/carbon footprint</li> <li>- Regulatory compliance (AI act, ISO 42001)</li> </ul>

Figure 1: From Multi-agent systems and Agentic AI to Agentic Edge Intelligence.

and data; Federated learning enables privacy-preserving cooperation; Standardized semantics (WoT, SOSA/SSN, SAREF) enhance machine-to-machine interoperability. Combined with container orchestration and edge AI, these technologies transform distributed autonomy into an operational reality and enable systematic evaluation of energy, cost, and performance trade-offs. The convergence of MAS, Agentic AI, and the computing continuum thus marks a pivotal transition: distributed intelligence has become executable. The next step lies in establishing shared models of trust, interoperability, and performance evaluation to guide the design of scalable, secure, and accountable agentic systems.

This paper defines the concept of a “agentic edge intelligence”, outlines a research agenda that integrates multi-agent systems, economics, and edge intelligence, and connects these to trust, ethics, and policy. Section 2 establishes the premises and core arguments. Section 3 clarifies definitions and principles of real-time AI economies. Section 4 highlights technical and social challenges. Section 5 presents our proposed research agenda, and Section 7 concludes.

## 2 Justification

The rapid progress of LLMs and related approaches (e.g., reasoning, multi-modal, and small models) has greatly expanded the capabilities of AI agents, enabling them to reason, plan, and operate autonomously in increasingly complex environments. Meanwhile, ubiquitous sensing and communication infrastructures, supported by communication networks (5G/6G), dense IoT deployments, and the convergence of edge and cloud paradigms, have reshaped how data is generated, shared, and processed. This shift has led to *edge intelligence*, where computational tasks are carried out closer to data sources to reduce latency and safeguard privacy. At the same time, distributed AI techniques have matured, allowing agents to train and infer at the network edge and to coordinate and negotiate among themselves. These advancements allow autonomous agents to manage local decisions and resources across distributed nodes.

Based on these premises, AI agents will soon proliferate throughout the computing continuum, thus forming highly interactive *agent ecosystems*. Because these agents operate in diverse administrative and operational domains, they require mechanisms to interact, compete, and cooperate effectively. We posit that such interactions will increasingly pivot around dynamic resource exchange (e.g., computational power, sensor data, specialized AI models), creating a new, fast-paced economy at the edge [11]. In this scenario, conventional static or pre-negotiated contracts prove insufficient when agents must adapt in real time to changing network conditions, user needs, and service availability. Consequently, we anticipate a move away from fixed, *a priori* agreements toward *ad hoc*, on-demand resource and service trading, all backed by new methods ensuring trust, fairness, and optimal utilization within decentralized environments.

In view of these technological and economic transitions, our central claim is that agentic edge intelligence will foster *real-time AI economies* as autonomous edge agents become primary drivers of data exchange and decision-making. These economies will involve real-time negotiation, resource allocation, and collaboration across multi-tenant networks, with agents making operational decisions. To enable this vision, a robust and interoperable *edge ecosystem* is necessary – capable of spanning administrative boundaries, supporting fluid resource trading, and maintaining reliability and security at scale. The shift to dynamic, *ad hoc* agreements and on-the-fly bartering, signals a fundamental change in how digital services are organized and delivered, with profound implications not only for technical architectures but also for governance, policy, and the social contract underlying trustworthy data ecosystems.

## 3 Definitions and Principles

In a real-time AI economy, autonomous agents act as self-governing economic actors that perceive, reason, and negotiate within digital marketplaces [11]. An agent may represent an individual, an enterprise, or an independent AI system with specific objectives and

budgets. Agents manage computations, data, or model resources and allocate them dynamically to optimize performance or economic utility. For example, an industrial maintenance agent may temporarily acquire GPU capacity from a nearby edge node during peak demand. At the same time, healthcare analytics agents could lease anonymized datasets to enhance diagnostic models while ensuring privacy. These interactions show how agents function as both computational and economic entities, engaging in transactions governed by trust and ethical standards.

Economic exchanges in a real-time AI economy span the computing continuum, where latency, locality, and trust determine value. Agents engage in continuous negotiation and micro-transactions involving computational power, bandwidth, models, or data streams. These exchanges occur under stringent temporal constraints, often requiring sub-second decision cycles to maintain system responsiveness. Dynamic market mechanisms, such as real-time auctions, peer-to-peer contracts, and bartering schemes, enable efficient resource allocation under heterogeneous objectives. The required autonomy of these agents necessitates their awareness of internal and external states, as well as a value system that guides their actions. For example, autonomous vehicles may bid for priority access to bandwidth during congestion, while underutilized edge nodes may offer surplus computing resources through decentralized marketplaces. These mechanisms improve scalability while raising the challenges of fairness, verifiability, and protection against adversarial actions.

The structure of the real-time AI economy thus derives from three intersecting theoretical foundations: (1) distributed systems and edge intelligence, which emphasize locality, autonomy, resilience, and awareness [15, 22]; (2) mechanism design and market-based coordination in multi-agent systems [25, 26]; and (3) AI governance and ethics, which provide normative and societal grounding for autonomous economic behavior [3, 4] (Fig. 2). These foundations collectively motivate the following principles.

**Autonomy and Decentralization:** Rooted in distributed AI and multi-agent systems theory [26], autonomy and decentralization assert that a decision-making authority should be distributed among agents rather than centralized. Decentralization enhances scalability, reduces single points of failure, and enables local adaptation under heterogeneous and dynamic conditions. This design philosophy follows the edge intelligence paradigm [15], where computation and decision-making occur near data sources. Finally, autonomy demands awareness and hence, agents need to implement a value system which is at the core of computational awareness [6].

**Trust and Verifiability:** Drawing from blockchain research and distributed ledger technologies [30], verifiability is essential to maintain accountability in non-cooperative agent ecosystems. As agents transact autonomously, cryptographic proofs, verifiable credentials, and immutable audit trails ensure the integrity of exchanges and deter fraudulent or manipulative behaviors. This principle aligns with trust frameworks developed for decentralized IoT and federated edge systems, where reputational and technical verification jointly sustain reliability.

**Latency-Aware Value Creation:** The notion that temporal proximity affects economic value is supported by real-time systems and edge economics literature [15, 22]. In latency-critical applications, such as autonomous driving or industrial control, the utility

of data and compute diminishes rapidly with delay. Hence, the market must internalize time sensitivity as a determinant of price and allocation efficiency—an extension of dynamic pricing theory adapted to spatiotemporal constraints in the computing continuum.

**Interoperability and Standardization:** This principle follows from the need for cross-domain interoperability identified in service-oriented architectures and semantic web research [19]. Shared ontologies, open APIs, and protocol standards enable seamless negotiation among agents operating under diverse governance, ownership, and technology stacks. Interoperability fosters market liquidity and prevents fragmentation, ensuring that economic exchanges remain composable across layers of the continuum.

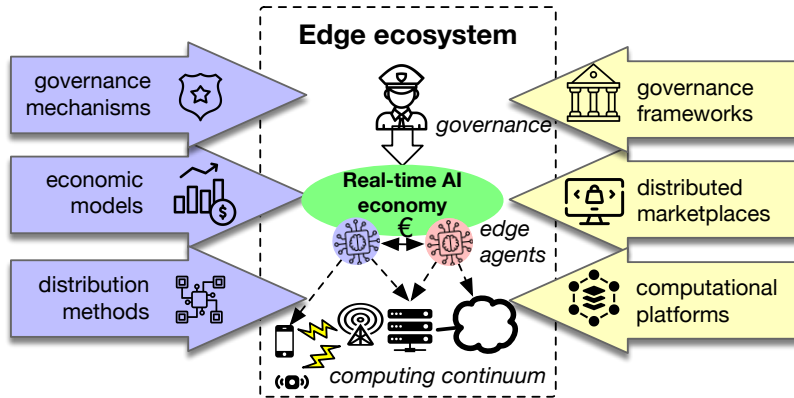
**Ethical and Sustainable Operation:** Originating from AI ethics and sustainable computing discourses [3, 4], this principle establishes that optimization objectives in autonomous markets must be constrained by fairness, transparency, and ecological responsibility. Embedding ethical norms within economic algorithms ensures alignment with human values, regulatory compliance, and long-term societal trust. Such integration is particularly critical in real-time economies where algorithmic decisions may have immediate social or environmental consequences. These principles provide a framework for real-time AI economies by integrating insights from distributed intelligence, economic coordination, and ethical AI. They create a self-organizing system for continuous, value-driven negotiation, emphasizing that efficiency and autonomy must develop alongside verifiability, interoperability, and normative oversight for sustainable societal adoption.

## 4 Key Technical and Social Challenges

**Scalability and Interoperability:** A central challenge for real-time AI economies is scalability across heterogeneous infrastructures. Systems may comprise millions of agents operating across edge, fog, and cloud domains under different administrative controls. Achieving interoperability requires standardized communication protocols and shared ontologies for data, resources, and services [19]. To prevent fragmentation and vendor dependence, we need open frameworks and standards like those that ensure Internet interoperability. Scalable coordination mechanisms are essential to balance computational costs and responsiveness, enabling local decision-making while maintaining system coherence. Without these mechanisms, agent ecosystems could face congestion, inconsistent operations, and economic inefficiencies.

**Resource Constraints and Market Failures:** Edge environments face limitations in computation, bandwidth, and energy. Market inefficiencies arise when agents behave strategically, hoard resources, or exploit pricing mechanisms [1, 17]. Classical equilibrium models do not capture the temporal and informational constraints of such systems. Mechanism design should ensure stability and fairness through anti-collusion protocols, reputation systems, and adaptive pricing. Context-aware policies that prioritize safety-critical tasks can align resource use with reliability and the public interest.

**Governance Models and Institutional Design:** Trust and accountability require formal governance structures defining rules,



**Figure 2: Overview of the agentic edge ecosystem.** The computing continuum serves as a platform for the real-time AI economy, comprising distributed AI agents and their transactions in the continuum marketplaces. Governance mechanisms and frameworks ensure trust, fairness, verifiability, and protection against adversarial actions. Advances in both theory and methods (left, in blue) as well as tools (right, in yellow) are needed to realize this vision.

sanctions, and enforcement [5, 7]. Centralized oversight is impractical in decentralized networks; governance must arise from distributed institutions that ensure norm compliance while fostering innovation. Normative programming [28], distributed ledgers, and meta-governance architectures provide tools for specifying and enforcing behavioral constraints. Smart contracts automate rewards and penalties, but strict enforcement can limit adaptability, while excessive autonomy may cause instability. Effective governance requires collaboration across computing, economics, and law to align autonomous operations with societal standards.

## 5 Research Agenda

**Theoretical Advances:** Future research should extend the formal foundations for real-time AI economies by integrating game theory, algorithmic economics, distributed optimization [1, 25], as well as an improved understanding of computational awareness [6]. Central challenges include modeling agent incentives under dynamic and resource-constrained conditions, designing robust mechanisms for market equilibrium, and ensuring fairness in non-cooperative interactions. Existing equilibrium models must be adapted to account for latency-sensitive transactions, incomplete information, and the bounded rationality of autonomous agents. New incentive-compatible protocols are needed to prevent manipulation and ensure stable cooperation in the computing continuum.

**Architecture and Implementation:** At the architectural level, research should explore scalable infrastructures that couple data-plane orchestration with agent-level negotiation [15, 19]. Promising directions involve AI interconnect architectures for low-latency coordination among agents, partitioning strategies for distributed intelligence, and neural pub-sub paradigms for adaptive communication. A unified control and negotiation fabric across the edge–cloud continuum would allow agents to dynamically allocate computational, network, and data resources in response to real-time market signals. Reference architectures should emphasize modularity, verifiability, and semantic interoperability to facilitate adoption in industrial and civic applications.

**Empirical Evaluations and Pilots:** Empirical validation is essential to assess the feasibility and societal impact of real-time AI economies. Experimental testbeds such as “Follow-Me AI” [20], manufacturing environments, or industrial IoT deployments provide natural contexts for evaluation [10]. Key performance indicators include cooperation efficiency, fairness, market stability, resilience to adversarial behaviors, and user-perceived trust. Longitudinal studies combining quantitative data with qualitative user feedback can uncover phenomena that challenge theoretical premises and guide improvements in governance and design principles.

**Cross-Disciplinary Collaboration:** The realization of a real-time AI economy requires sustained collaboration among computer scientists, economists, ethicists, and policymakers [3, 4]. Technical innovations must be complemented by ethical frameworks and regulatory guidance addressing accountability, transparency, and social welfare. Interdisciplinary dialogue can also inform the development of meta-governance structures that balance innovation with safeguards against systemic risks. Ultimately, understanding and steering these agentic markets demands a convergence of expertise across computational, economic, and human-centric domains.

## 6 Discussion

Agentic intelligence reframes the computing continuum as a socio-technical arena in which autonomous agents negotiate value, resources, and responsibilities. This perspective extends prior work on distributed AI and multi-agent systems by emphasizing that agents are not merely computational units but institutional actors whose interactions generate economic, organizational, and ethical consequences. The resulting real-time AI economies challenge traditional assumptions about orchestration, replacing centralized coordination with market-based negotiation under constraints of latency, locality, and heterogeneous trust requirements [14, 15].

At the same time, this paradigm exposes a series of tensions. Autonomy and decentralization improve scalability and resilience but introduce risks related to adversarial behavior, collusion, and misalignment, requiring governance structures capable of enforcing

norms and ensuring accountability across administrative boundaries. The integration of mechanism design into edge contexts demands models that address incomplete information, bounded rationality, and spatiotemporal constraints, while also maintaining fairness and societal legitimacy. These conceptual challenges underscore that real-time AI economies are not simply technical systems but emergent institutions requiring interdisciplinary analysis and careful normative design.

In addition to solving current issues in agentic systems such as distributed memory coordination [21], negotiation [20], and mission drift [8], addressing these challenges requires coordinated advances across infrastructure, economics, and governance (Fig. 2). Emerging architectures for edge-native AI—such as distributed orchestration fabrics, trusted execution, and neural pub/sub communication—offer mechanisms for securing autonomous interactions while preserving low-latency responsiveness [15]. Standardization efforts around semantic interoperability (e.g., WoT, SAREF) mitigate fragmentation by enabling shared resource descriptions and machine-readable negotiation protocols, thereby improving liquidity and reducing coordination overhead in decentralized markets. From an economic perspective, adaptive mechanism design and incentive-compatible protocols can counteract strategic manipulation and promote stable cooperation under resource constraints, extending existing insights on algorithmic economics and cooperative-competitive multi-agent dynamics [26]. Finally, governance models drawing on artificial institutions, normative programming, and verifiable transaction records provide a pathway for embedding accountability and norm compliance into agent interactions without reverting to centralized control [5, 7]. Together, these developments suggest that real-time AI economies can remain both resilient and aligned with societal expectations, provided that technical design choices co-evolve with institutional and ethical considerations.

## 7 Conclusion

This paper introduced agentic edge intelligence as a foundation for real-time AI economies in which autonomous agents coordinate and compete across the computing continuum. We argued that realizing this vision requires advances in distributed architectures, incentive mechanisms, and governance models capable of ensuring scalability, trustworthiness, and fairness. By outlining the technological, economic, and organizational implications of this emerging paradigm, we aim to establish a coherent agenda that supports the development of accountable and resilient agentic systems for future digital environments.

## Acknowledgments

This work was supported by the Research Council of Finland through the 6G Flagship program (grant 318927) and the CO2CREATION SRC project (grant 372355), by Business Finland through the Neural pub/sub research project (diary number 8754/31/2022), and by the ERDF (project numbers A81568, A91867).

## References

- [1] Rediet Abebe and Kira Goldner. 2018. Mechanism design for social good. *AI Matters* 4, 3 (2018), 27–34.
- [2] Peter Belcak et al. 2025. Small Language Models are the Future of Agentic AI. *arXiv preprint arXiv:2506.02153* (2025).
- [3] Emily M Bender et al. 2021. On the dangers of stochastic parrots: Can language models be too big?. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. 610–623.
- [4] Allan Dafoe et al. 2020. Open problems in cooperative ai. *arXiv preprint arXiv:2012.08630* (2020).
- [5] Maiquel de Brito et al. 2018. Situated artificial institutions: stability, consistency, and flexibility in the regulation of agent societies. *Autonomous Agents and Multi-Agent Systems* 32 (2018), 219–251.
- [6] Stanislas Dehaene, Hakwan Lau, and Sid Kouider. 2021. What is consciousness, and could machines have it? *Robotics, AI, and humanity: Science, ethics, and policy* (2021), 43–56.
- [7] Nicoletta Fornara and Marco Colombetti. 2008. Specifying and enforcing norms in artificial institutions. In *International Workshop on Declarative Agent Languages and Technologies*. Springer, 1–17.
- [8] Taicheng Guo, Xiuying Chen, Yaqi Wang, Ruidi Chang, Shichao Pei, Nitesh V Chawla, Olaf Wiest, and Xiangliang Zhang. 2024. Large language model based multi-agents: A survey of progress and challenges. *arXiv preprint arXiv:2402.01680* (2024).
- [9] Fritz Hohl. 1998. Time limited blackbox security: Protecting mobile agents from malicious hosts. In *Mobile agents and security*. Springer, 92–113.
- [10] Henna Kokkonen et al. 2022. Autonomy and intelligence in the computing continuum: Challenges, enablers, and future directions for orchestration. *arXiv preprint arXiv:2205.01423* (2022).
- [11] Henna Kokkonen, Susanna Pirttikangas, and Lauri Lovén. 2023. EISim: A Platform for Simulating Intelligent Edge Orchestration Solutions. *arXiv preprint arXiv:2311.01224* (2023).
- [12] Shujun Li and Mieczyslaw M Kocar. 2012. Agent communication language. In *Flexible Adaptation in Cognitive Radios*. Springer, 37–44.
- [13] Lauri Lovén et al. 2019. EdgeAI: A vision for distributed, edge-native artificial intelligence in future 6G networks. In *The 1st 6G Wireless Summit*. Levi, Finland, 1–2.
- [14] Lauri Lovén, Miguel Bordallo López, Roberto Morabito, Jaakko Sauvola, and Sasu Tarkoma (Eds.). 2025. Large Language Models in the 6G-Enabled Computing Continuum: a White Paper. *6G Research Visions* 14 (2025).
- [15] Tobias Meuser, Lauri Lovén, et al. 2024. Revisiting Edge AI: Opportunities and Challenges. *IEEE Internet Computing* 28, 4 (2024), 49–59.
- [16] Victor Casamayor Pujol, Praveen Kumar Donta, Andrea Morichetta, Ilir Murturi, and Schahram Dustdar. 2023. Edge intelligence—research opportunities for distributed computing continuum systems. *IEEE Internet Computing* 27, 4 (2023), 53–74.
- [17] Houming Qiu et al. 2022. Applications of auction and mechanism design in edge computing: A survey. *IEEE Transactions on Cognitive Communications and Networking* 8, 2 (2022), 1034–1058.
- [18] Anand S Rao, Michael P Georgeff, et al. 1995. BDI agents: From theory to practice.. In *Icmas*, Vol. 95. 312–319.
- [19] Alaa Saleh et al. 2025. Towards message brokers for generative AI: Survey, challenges, and opportunities. *Comput. Surveys* 58, 1 (2025), 1–37.
- [20] Alaa Saleh, Praveen Kumar Donta, Roberto Morabito, Naser Hossein Motlagh, Sasu Tarkoma, and Lauri Lovén. 2025. Follow-me AI: Energy-efficient user interaction with smart environments. *IEEE Pervasive Computing* (2025).
- [21] Alaa Saleh, Sasu Tarkoma, Anders Lindgren, Praveen Kumar Donta, Schahram Dustdar, Susanna Pirttikangas, and Lauri Lovén. 2025. MemIndex: Agentic Event-based Distributed Memory Management for Multi-agent Systems. *ACM Transactions on Autonomous and Adaptive Systems* (2025).
- [22] Mahadev Satyanarayanan. 2017. The emergence of edge computing. *Computer* 50, 1 (2017), 30–39.
- [23] V Stafford. 2020. Zero trust architecture. *NIST special publication* 800, 207 (2020), 800–207.
- [24] Peter Stone and Manuela Veloso. 2000. Multiagent systems: A survey from a machine learning perspective. *Autonomous Robots* 8 (2000), 345–383.
- [25] Toby Walsh. 2024. Approximate mechanism design for facility location with multiple objectives. In *ECAI 2024*. IOS Press, 3533–3540.
- [26] Jianrui Wang et al. 2022. Cooperative and competitive multi-agent systems: From optimization to games. *IEEE/CAA Journal of Automatica Sinica* 9, 5 (2022), 763–783.
- [27] Michael Wooldridge and Nicholas R Jennings. 1995. Intelligent agents: Theory and practice. *The knowledge engineering review* 10, 2 (1995), 115–152.
- [28] Elena Yan et al. 2024. An Agent-Centric Perspective on Norm Enforcement and Sanctions. *arXiv preprint arXiv:2403.15128* (2024).
- [29] Shunyu Yao et al. 2022. React: Synergizing reasoning and acting in language models. *arXiv preprint arXiv:2210.03629* (2022).
- [30] Anastasios Zafeiropoulos et al. 2024. AI-Assisted Synergetic Orchestration Mechanisms for Autoscaling in Computing Continuum Systems. *IEEE Communications Magazine* (2024).