Journal of Artificial Intelligence for Automation
https://www.sciltp.com/journals/jaia

Review

Cross-Layer AI for Cyber-Physical Ecosystems: The Intelligence Continuum

Claudio Savaglio

DIMES Department, University of Calabria, 87036 Rende, Italy; csavaglio@dimes.unical.it

How To Cite: Savaglio, C. Cross-Layer AI for Cyber-Physical Ecosystems: The Intelligence Continuum. *Journal of Artificial Intelligence for Automation* **2026**, *1*(1), 3.

Received: 28 December 2025
Revised: 20 January 2025
Accepted: 26 February 2025
Published: 27 February 2025

Abstract: Automation in Cyber–Physical Systems (CPS) is undergoing a structural transformation, disrupted by recent technologies such as Internet of Things (IoT), Virtual/Augmented Reality (AR/VR), 5/6G, etc. However, an entire paradigm shift is on the horizon, pushing automation beyond traditional borders and towards operation environments where conditions change, uncertainty is high, or data are unstructured. To this end, regardless of the specific application domain, achieving superior responsiveness, autonomy, and resilience will require an Artificial Intelligence (AI) no longer confined to the application or Cloud layer. Instead, intelligence must be natively shared and coordinated across devices, system service, application software, and even the network itself. Therefore, this perspective paper argues that an intriguing direction for intelligent automation is represented from what we term the *Intelligence Continuum*: a seamless integration of Edge AI, Smart Networking, and Cloud AI, jointly operating under the principles of Continuum Computing and Network Intelligence. We first outline the limitations of current architectures and related paradigms in dealing with complex Cyber-Physical Ecosystems (CPeS). Then, we articulate the need for a cross-layer AI to support an application-centric design philosophy with a unified control loop, jointly considering computation and communication across the continuum. Hence, we highlight the enabling and operational role of Digital Twins as the glue and substrate of next-generation intelligent CPeS. Finally, we discuss key research challenges, implications, and future directions for achieving continuum-native intelligent automation.

Keywords: cyber-physical system; internet of things; artificial intelligence; digital twin

1. Introduction

Automation profoundly evolved over the past decades [1]. As Cyber–Physical Systems (CPS) and Internet of Things (IoT) infrastructures rapidly grow in scale, heterogeneity, and criticality [2], early automation systems (predominantly rule-based, with static logic, rigid workflows, and deterministic control schemes) have become increasingly inadequate. Indeed, they required significant manual configuration, lacked adaptability, and operated effectively only in predictable, well-structured environments [3]. However, Industry 4.0, Smart Energy Grids, autonomous vehicles, just to name few examples, demand real-timeness even under resource constraints, resiliency despite uncertainty and dynamicity, context-awareness but also global optimization, coordination and Service Level Agreements (SLAs) distributed among components [4]. In such a scenario, the concept of **Cyber-Physical Ecosystems (CPeS)** has emerged [5,6], outlining open and adaptive ensembles of cyber-physical components whose interactions, control mechanisms, and system boundaries evolve dynamically at runtime and under uncertainty.

The emergence of data-driven AI initially supported CPeS and enabled an early shift of automation toward adaptive models, facilitating pattern recognition, predictive maintenance, and partial context-awareness [7]. However, these AI capabilities have been often siloed and applied only to specific tasks or executed in the Cloud, with additional bottleneck and disadvantages (in terms of responsiveness, autonomy and resilience) caused by inherent limitations of IP-based networking. Recent technology advances and AI derivatives, such as Edge Artificial Intelligence

(Edge AI) [8] and Generative AI (GenAI) [9], along with efficient industrial IoT communication protocols (MODBUS, MQTT, etc.) [10] and revolutionary content-centric networking approaches (ICN, NDN, etc.) [11], have proven successful as standalone solutions: however, they remain poorly integrated, thus preventing the realization of a seamless, end-to-end architecture.

That said, we argue for the need of a paradigm shift toward an **Intelligence Continuum**, where AI permeates orthogonally all the building blocks of the CPeS: the Edge nodes, the Cloud, and, finally, the network. By advancing the *Computing Continuum* [12] vision of an AI considered as an endpoint capability (hence, prerogative of the local/remote computing devices), we advocate the idea of an intelligence that becomes a cross-layer systemic and infrastructural property. The (different) AI pipelines enable components to be dynamic, context-aware, and continuously self-evolving, able to anticipate conditions, self-correct, negotiate resources, and optimize processes in real time; simultaneously, communication infrastructures are not simply aware of the network status but also of the traffic content and intent, as envisioned by the *Network Intelligence* [13] principles. The resulting closed-loop CPeS would be able to autonomously orchestrate computing resources, instantly shifting heavy AI inference between the Cloud and the Edge to maintain zero-latency control. In such a direction, sometimes discussed but never fully developed, the key engineering challenge lies in establishing the operational mechanisms to effectively fuse computational and networking elements. For realizing the Intelligence Continuum, we believe that *Digital Twin* (DT) [14,15] is the right enabling paradigm: indeed, it has now reached the maturity to act as a model for representing, simulating and coordinating the entire system, transforming a heterogeneous set of physical/digital resources and infrastructures into an autonomous and adaptive CPeS [16]. Therefore, in the following, we first briefly analyze the role and state-of-the-art of intelligent automation in Edge, network, and Cloud layer, discussing main limitations. Then, we present the Intelligence Continuum as a solution against siloed layered architectures and its many dimensions. Finally we disclose how DT can play the role of Intelligence Continuum enabler, discussing its distinctive benefits across all the architectural layers. The article concludes by analyzing key open points and future directions for action.

2. Automation Architectures at a Turning Point

CPS and CPeS are typically structured around tiered architectures where automation tasks are rigidly distributed by necessity [17], including:

- time- and safety-critical local responses at the Edge;
- global optimization and massive data processing within the Cloud;
- seamless connectivity between these two layers provided by the network layer.

However, a siloed and static approach to these architectural layers limits system efficiency by creating data bottlenecks and preventing the cross-layer visibility required for truly autonomous, adaptive intelligence, as discussed hereinafter.

2.1. Edge Layer: Perception-Centric and Real-Time Automation

Edge devices equipped with sensors, cameras, and actuators leverage lightweight ML models to perceive local context, extract features, and detect events with minimal latency. Therefore, at the Edge layer, AI-driven automation is primarily perception- and action-oriented [18], focusing on real-time interaction with the physical environment, supporting immediate actions such as anomaly detection, control adaptation, or local task admission [19].

Approaches such as TinyML and on-device inference support localized learning and reasoning under strict resource constraints, enabling fast and autonomous reactions [20,21]. More sophisticated Deep Learning-based perception techniques enable semantic interpretation of signals and images directly at the device level [22,23], but intelligence at the Edge still remains mostly reactive and perception-centric.

In fact, due to the limited resources, transforming Edge automation from simple threshold-based logic into adaptive, self-contained control loops, would require feedback monitoring and periodic model updates coordinated with the Cloud. However, it is not always feasible for the sake of latency, connectivity, computation, energy, or cost: as results, Edge AI models are frequently trained and deployed independently, without coordination or awareness of global system objectives, and such a siloed deployments can lead to myopic decision-making, where local optimization goals may conflict with end-to-end performance requirements [24]. Limited data visibility and constrained model capacity further exacerbate this issue, preventing effective reasoning beyond immediate observations; moreover, the scarcity of collaborative or federated learning mechanisms across Edge nodes increases the risk of model drift and inconsistent behavior [8]. Emblematic is the double definition of “AI for Edge” and “AI on Edge” that highlights the absence of a unified approach [25]: the former focuses on providing optimal solutions to solve key problems in Edge computing (e.g., data offloading, energy management, nodes coordination) with the help of popular AI techniques, while the latter studies how (i.e., which hardware platforms, programming frameworks, methods, and tools) to perform the whole process of AI model building, i.e., training, inference, and optimization, on Edge devices despite their intrinsic resource limitations. Overall, this fragmentation shows that, in most cases, intelligent automation is limited to individual devices or local controllers, thus undermining scalability, robustness, and coherent system-level adaptation [8].

2.2. Network Layer: Connectivity, Robustness, and Policy Enforcement

The network component of CPeS is critical, as it enables the real-time communication, monitoring, and control operations necessary for the tight integration of physical and computational elements. Automation tasks at this layer mainly focus on optimization and robustness, resorting to AI techniques to understand the network state, to learn traffic and workload patterns, and, accordingly, to improve routing, scheduling, resource allocation, and security [26,27].

For example, Zero-Touch Provisioning (ZTP) allows automatically onboarding new sensors, actuators, routers, etc., as they are added to the system [28]; learning-based and optimization-driven approaches, instead, enable network automation to dynamically balance latency, bandwidth, and reliability constraints [29]. Similarly, AI algorithms excel at anomaly detection by monitoring network traffic patterns for signs of cyberattacks and automatically taking defensive actions, such as segmenting the affected device or automatically updating firewall and access control lists across the network [30].

The diversity of goals (management, optimization, and security) and levels (physical, access, transport) bring to a multitude of automation strategies and to a fragmented AI adoption. This typically manifests itself as disjoint learning and optimization components in the routing, scheduling, protection, congestion control, and offloading mechanisms. Independently designed AI network modules can lead to conflicting policies and unstable control behavior [31], thus increasing operational complexity and degrading predictability and performance. In addition, networks are still mostly IP-centric, endpoint-oriented, and agnostic to data semantics, with rigid addressing and routing hierarchies which prevent efficient automation of data dissemination and traffic prioritization [11]. Applying AI as an uppermost "patch" on such a shaky infrastructure and legacy protocol constraints can create additional bottlenecks and delay responses [32], or restrict the benefits of policy refinement, traffic steering, and other network automation tasks.

2.3. Cloud Layer: Learning, Global Reasoning, and Long-Term Optimization

The Cloud layer acts as the learning and reasoning core of the AI-driven automation pipeline. It aggregates large-scale data from Edge and network layers to train complex models and perform long-term reasoning and optimization [33]. ML enables Cloud automation to learn predictive and decision models that guide orchestration, scheduling, and resource management at any system scale [34]. Reasoning mechanisms, encapsulated in advanced paradigms such as software agents, support proactiveness, sociality and goal-driven automation under uncertainty and multi-objective constraints [35]. Actions taken at the Cloud layer include deploying updated models, enforcing QoS policies, and adjusting orchestration strategies across the infrastructure [36]. Continuous improvement is realized through iterative learning cycles and feedback-driven refinement, enabling automation systems to evolve over time [37].

However, centralized intelligent automation often introduces excessive latency for control loops, bottlenecks in data transport, dependence on always-on connectivity, and reduced resilience. In addition, although Cloud platforms provide huge computational and data resources, siloed AI pipelines for learning, orchestration, and optimization often lack awareness of deployment constraints and runtime conditions of lower layers [12]. This disconnect reduces model transferability and leads to inefficient orchestration decisions [8]; also, delayed/incomplete feedback from Edge and network components significantly limits timely retraining and policy updates, weakening the self-improving nature of AI-driven automation and diminishing both orchestration and global optimization capabilities [17].

3. Toward Intelligence Continuum: Aligning Computing and Networking

The fragmented AI exploitation discussed so far confines automation systems to collections of locally intelligent components, preventing the emergence of coherent and adaptive CPeS [17]. Overcoming this limitation lies at the core of the **Intelligence Continuum** vision, which seeks to enable truly integrated, end-to-end intelligent automation. In fact, it envisions a scenario where network, devices and applications jointly co-design decisions, without the separation between computation-centric automation and network-centric control.

In this vision, computational activities are layer-agnostic by design and may be executed at any point across the architectural stack, provided that current (device, network, and application) conditions make such placement optimal or feasible: hence, by deliberately avoiding any a-priori binding between tasks and architectural layers (i.e., device, Edge, or Cloud), the system moves away from static tiering toward context-driven execution placement, treated as a runtime decision problem [38]. In parallel, the communication network evolves into an intelligent and active component of the system: beyond the conventional AI-assisted mechanisms that provide timely awareness, protection, and prediction of the traffic conditions, the network supports content-aware forwarding, the context-aware placement decisions, and the application-aware resource coordination [39]. Moreover, this enables a unified control loop spanning both compute management mechanisms (e.g., Kubernetes, serverless platforms, Edge orchestrators) and AI-aided networking mechanisms (e.g., SDN controllers, traffic analytics engines), allowing coordinated, end-to-end optimization of computation and communication resources. By integrating these formerly independent control loops, the system achieves holistic, closed-loop optimization, allowing workload placement, data flows, and resource management to continuously align with system-wide objectives across the entire continuum. Therefore, the resulting approach is inherently application-centric rather than infrastructure-centric, as execution decisions are guided by both application requirements and context-awareness instead of by rigid, predefined roles imposed by the layers. To emphasize the fundamental architectural shift enabled by the Intelligence Continuum, Table 1 compares conventional siloed layered architectures with the proposed paradigm, highlighting the unification of computation and network control, the move from static tiering to context-driven placement, and the resulting system-level benefits.

Table 1. Comparison between siloed layered architectures and the Intelligence Continuum

Dimension	Siloed Layered Architecture	Intelligence Continuum
Architectural model	Rigid, tier-based separation (device, Edge, Cloud)	Unified continuum spanning devices, Edge, network, and Cloud
Execution placement	Statically assigned to predefined layers	Dynamically determined at runtime across the continuum
Design philosophy	Infrastructure-centric, layers impose fixed roles	Application-centric, driven by goals and constraints
Adaptation mechanism	Limited or reactive, often manual reconfiguration	Continuous, context-aware closed-loop adaptation
Control loops	Separate and independent compute and network control loops	Unified control loop jointly optimizing computation and communication
Network role	Passive data transport layer	Active, intelligent system component
Awareness	Local, layer-specific monitoring	Global, cross-layer situational awareness
Optimization scope	Per-layer or per-domain optimization	End-to-end, system-wide optimization
Handling dynamics	Poor resilience to workload and network variability	Intrinsic support for variability, mobility, and failures
CPeS suitability	Adequate for static or loosely coupled systems	Well-suited for dynamic, latency-sensitive CPeS

Consider, for example, a smart factory populated with autonomous robots generating high-frequency sensor data that require low-latency control. Traditionally, such data would be sent to a fixed Edge server for processing, with control commands issued from there. In contrast, under the Intelligence Continuum, the system dynamically determines whether computation should occur on the robot itself, at a nearby Edge node, or in the Cloud, based on current network congestion, latency, and resource utilization. The task assignment is not a one-shot optimization but a continuous adaptation process where tasks are dynamically reallocated, split into sub-functions across layers, offloaded or pulled back closer to data sources, replicated to increase resilience or reduce latency, thus enabling the system to respond to both short-term fluctuations and long-term trends. Simultaneously, the network prioritizes critical control messages while balancing traffic for analytics workloads. By integrating compute orchestration and Network Intelligence into a single, unified control loop, the system minimizes control delays and ensures high-throughput processing, achieving holistic optimization across the continuum. A similar scenario arises in intelligent transportation systems, where autonomous vehicles continuously exchange sensor data to support cooperative perception and real-time traffic management. Rather than relying on a fixed Cloud or Edge node, computation can dynamically occur on the vehicles, at nearby roadside units, or in regional Cloud data centers depending on network conditions, vehicle density, and the urgency of decisions. The unified control loop jointly orchestrates both computation placement and network behavior, enabling timely dissemination of critical information, reducing latency spikes, and optimizing bandwidth usage. This coordinated adaptation improves both safety and operational efficiency, illustrating the core benefit of the Intelligence Continuum: the seamless, context-aware integration of computing and networking resources across the entire system.

Rooting and Shaping Intelligence Continuum

A well-documented background ensures both theoretical soundness and practical reliability of the proposed Intelligence Continuum. Rooted in an established background of principles and methodologies, the Intelligence Continuum emerges from the convergence into a unified, high-performance framework of two paradigms that exhibit a strong and intrinsic intersection (especially around the aforementioned keywords of context awareness, distributed decision-making, and adaptive control) but that, in practice, have largely evolved along independent trajectories [40]: the Continuum Computing and the Network Intelligence.

Continuum Computing (referred also as Edge–Cloud or Device–Edge–Cloud continuum [17]) frames compute, storage and intelligence as a single, logically continuous fabric that spans local nodes and remote Cloud. The objective is to dynamically place functions (ingestion, preprocessing, inference, training, storage, and orchestration) at the points in the fabric that, in a given moment, best satisfy both goals and constraints, rather than force every task to a single, predefined architectural tier [12, 41]. Particularly attractive for CPeS where sensing, actuation, and control loops must coexist with large-scale analytics and long-term optimization, the Computing Continuum view is no longer purely conceptual: in a few years the research community outlined its architectural blueprint [42], recent state-of-practice surveys synthesize contemporary platforms and deployment experience [24], and multiple national research initiatives and international conferences now consolidate research on emerging trends for the Edge–Cloud fabric [43] and provide reference middleware, pilots and toolchains that explicitly target Cloud–Edge–IoT interoperability (OpenContinuum [44], ENACT [45], SMARTEDGE [46]).

Network Intelligence [13], originated within the telecommunications community, initially focused on traffic monitoring, analytics, and performance optimization, and later incorporated ML techniques to support predictive and adaptive network control [29]. By elevating the communication network from a passive data transport layer to an active, cognitive component of the system, Network Intelligence couples network state and automation logic, thus allowing CPeS to maintain control stability, meet real-time deadlines and improve resilience under dynamic and uncertain traffic situation. Especially in conjunction with ML, 5-6G, and SDN technologies, CPeS leverage intelligent networking to dynamically coordinate robots, sensors, and machines under strict latency and reliability constraints [47], to optimize network slicing and configuration [48], to enable data-driven automated orchestration and resource adaptation in industrial CPS [49,50].

Over the years, the relationship between the computing and the network has been contentious and revealing, and these two paradigms have never been developed in a tightly coupled manner. Core Continuum Computing problems such as task offloading, service placement, and end-to-end QoS enforcement inherently depend on the accurate and timely knowledge of network conditions, as well as on the predictive capabilities traditionally associated with smart communication: however, rather than an enabler for effective orchestration, the network is often abstracted as a static or externally predicted parameter (e.g., a constraint and a source of telemetry used by orchestration engines [24,51]), while Network Intelligence solutions typically treat computation as an external workload endpoint rather than as a first-class optimization dimension. This practical separation is reinforced by differences in research communities (software and computer engineers vs. telco engineers and providers), system abstractions (service/task/function vs. flow/packet/session), and evaluation methodologies, thus confining compute intelligence, network control and application logic in separate planes with weak semantic bindings [52]. From here, the fragmented AI exploitation discussed so far and the obstacles to the evolution of next-generation intelligent CPeS.

Nevertheless, research and early pilots recently started pointing to convergence: network-aware and network-adaptive orchestration, in-network computing, intent-aware fabrics and AI-assisted network controllers are active research directions [40,53], both in industry and academia. Therefore, we conclude that the time is ripe to pursue a convergence of Computing Continuum and Network Intelligence principles under the umbrella of the Intelligence Continuum. This convergence can be enabled by the DT paradigm, which has reached a significant level of maturity and flexibility, as discussed in the following.

4. Operationalizing Intelligence Continuum: The Digital Twins

Said that research on Network Intelligence and on the Computing Continuum has largely evolved along independent trajectories, DT emerges as a promising enabling paradigm to finally align them and realize the Intelligence Continuum, as highlighted in Table 2. Originally introduced as high-fidelity digital representations of physical assets and processes [14], and recently refined around the pillars of Entanglement, Reflection, Synchronization, Augmentation, and Servitization [15], DTs have evolved from simple data shadows to increasingly autonomous, cognitive entities capable of reasoning, prediction, and adaptation [54]. Importantly, they provide the perfect sandboxes for assessing learning-based approaches, evaluating policies, and refining them over real data but before the actual implementation in the live system, thereby reducing operational risk. As the latest step of their evolution, DTs have been also hybridized by GenAI [16,55] to conciliate model- and data-driven approaches, so as to further accommodate the particular needs and characteristics of IoT scenarios in terms of uncertainty management, data accuracy, information processing, and system reliability [9, 56]. Interpreted beyond asset-level modeling, recently DTs have been extended to represent also networking resources and infrastructures [57]. In fact, rather than treating the network as a black-box provider of metrics, a network DT can model topology, traffic patterns, protocol behavior, and performance dynamics, enabling AI-driven reasoning about communication alongside computation and control [58].

Table 2. Enabling role of Digital Twins for the Intelligence Continuum

Continuum Layer	Scope of Intelligence	Limitations without Digital Twins	Enabling Role of Digital Twins
Edge AI	Local inference, task execution, and latency-sensitive adaptation close to data sources	Limited global context, reactive decisions, heuristic offloading strategies, and fragmented awareness of network conditions	Digital Twins model Edge resources and access networks, enabling predictive offloading, local what-if analysis, and safe adaptive control
Network Intelligence	Traffic inference, routing, and QoS/QoE optimization across heterogeneous IP and non-IP networks	Flow-centric abstractions, weak linkage to service and task semantics, and limited support for non-IP protocols	Digital Twins provide protocol-agnostic and semantic representations of traffic and topology, enabling service- and task-aware network optimization
Cloud Intelligence	Global analytics, large-scale learning, and long-term system optimization	Delayed feedback, limited operational realism, and weak coupling with Edge and network dynamics	Digital Twins aggregate multi-domain state and historical data, supporting system-wide learning, SLA synthesis, and capacity planning
Intelligence Continuum	End-to-end orchestration and joint optimization across Edge, network, and Cloud	Fragmented control loops, inconsistent system models, and limited cross-domain reasoning	Digital Twins act as a shared semantic and control substrate, enabling closed-loop co-optimization of networking and computation across the continuum

Exactly thanks to such a versatility, DT can be engineered to explicitly and simultaneously encode multi-layer semantics and multi-objective optimization goals, spanning network characteristics (e.g., topology, traffic flows, latency distributions), computational resources (e.g., heterogeneous nodes, accelerators, energy profiles), service models (e.g., tasks, workflows, microservices), and contextual dimensions (e.g., mobility, workload dynamics, user intent), thereby addressing the long-standing representational mismatch between task-centric Continuum Computing abstractions and flow-centric Network Intelligence models. In fact, by naturally integrating network telemetry (e.g., delays, congestion, failures), compute telemetry (e.g., CPU, memory, GPU, queues), service-level metrics (e.g., throughput, deadlines) with simulation and AI capabilities of different size [59], DT can be continuously updated with the CPeS status, thus ensuring that decisions are based on an accurate, timely, and global view of the contexts. Hence, coordinated actions (rather than sequential or isolated ones) derived from joint intelligence, such as re-routing flows, migrating services, data augmentation, or offloading task, can be first enacted on the virtual representation, validated, and eventually reflected back to the physical realm. By maintaining such a bi-directional and continuously synchronized virtual representation of the CPeS (including network topology and dynamics, computational resources, service structures, and contextual information), DTs provide a unified and semantically rich abstraction upon which Network Intelligence and Continuum Computing can coherently operate [16], as depicted in the illustrative example of Figure 1. Moreover, by integrating model-driven and data-driven representations within a single framework, DTs reconcile heterogeneous control and optimization processes across networking [60] and computation domains [61], so that reactions to transient conditions smoothly coexist with longer-term strategies. Finally, by abstracting over protocol-specific details and exposing semantic representations of both traffic and computation, DTs extend the reach of Network Intelligence beyond traditional IP-centric designs, paving the way towards Future Internet scenarios [11]. Indeed, ICN/NDN research for IoT and CPeS demonstrates how data-centric networking could simplify automation logic and improve robustness, particularly in distributed and mobile scenarios [62]. Summarizing, the unified semantic representation provided by the DT paradigm enables the joint observability of the CPeS through the fusion of network-, compute-, and service-level telemetry, supports safe policy evaluation and learning via what-if analysis, and facilitates closed-loop control across networking and computation layers. Therefore, as highlighted in Table 2, the DT can establish a unified representational and operational framework for the Intelligence Continuum, enabling several key functionalities at each different layer and across them.

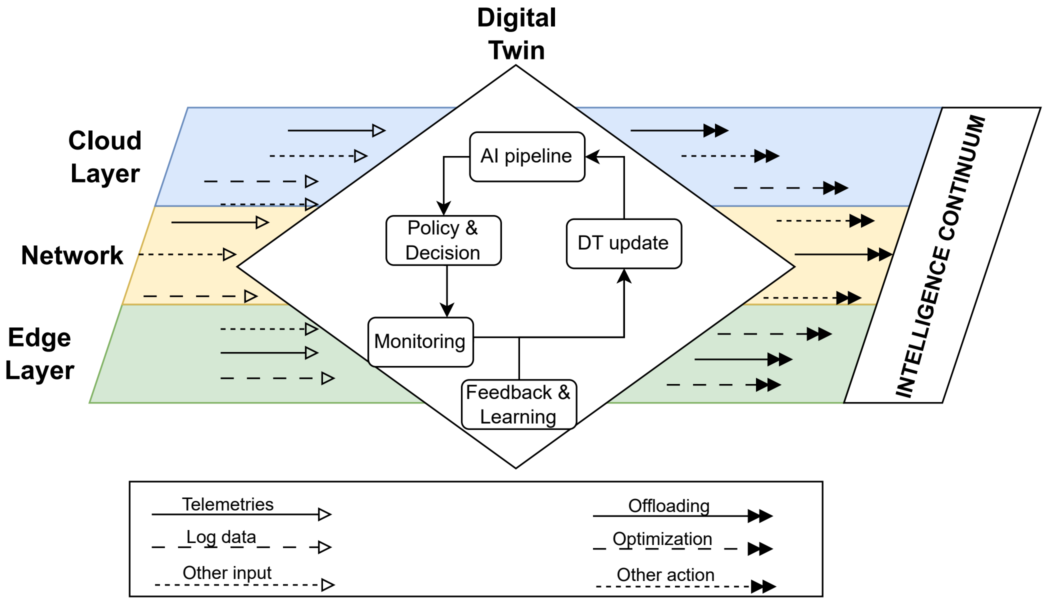

Figure 1. DT implementing the Intelligence Continuum: it gathers different kinds of inputs from each architectural layer, and orchestrates the coordinated actions dictated by the unified control loop.

5. Paving the Way for the Intelligence Continuum

Despite significant advances in both Continuum Computing and Network Intelligence, their tight integration into a unified Intelligence Continuum is still an articulated process with technical, operational, and conceptual open points. Some relevant ones include:

- *Managing operational complexity and fostering trust in autonomous network control.* The operational complexity and stringent reliability requirements of communication infrastructures naturally make operators cautious about delegating autonomy to learning-based control mechanisms. As telecommunication systems are mission-critical, the adoption of AI-driven decisions calls for increased trust through explainability, predictability, and formal guarantees. Recent advances in explainable AI and safe learning for networking indicate that trustworthy autonomous control is an achievable and impactful goal [37,63].
- *Harmonizing heterogeneous control-loop timescales across the continuum.* Control-loop timescales vary widely across system layers: network control often operates at millisecond granularity, orchestration and resource management at seconds to minutes, and learning and model adaptation at minutes to hours. While this temporal diversity introduces coordination challenges, it also motivates hierarchical and multi-timescale control frameworks capable of aligning fast network reactions with slower optimization and learning processes [26,64].
- *Evolving legacy architectures toward cross-layer knowledge sharing.* The long-standing separation between telecom and Cloud systems has resulted in legacy architectures that limit tight cross-layer knowledge exchange. However, ongoing efforts toward network softwarization and ZTP provide a strong foundation for improved observability and integrated decision-making across network, compute, and application layers [65,66].
- *Establishing semantically interoperable abstractions of network state.* The absence of a dominant abstraction that represents network state in a form semantically interoperable with compute and application models remains a key gap. While Cloud-native abstractions for services and workloads are well established, emerging intent-based and model-driven networking paradigms offer promising directions for unifying representations across the continuum [66,67].
- *Advancing integrated Network Intelligence-Continuum Computing DTs.* DTs are increasingly recognized as a powerful enabler of intelligent systems, yet current implementations predominantly focus on either networking

or computing domains in isolation. Extending DTs toward integrated, continuum-scale, multi-layer, and multi-timescale models represents a particularly promising opportunity to bridge the historical separation between Network Intelligence and Continuum Computing, supporting simulation-assisted learning, proactive optimization, and safe closed-loop control [68,69].

6. Conclusions

In this perspective paper, the idea of an Intelligence Continuum has been outlined as a unifying conceptual framework for CPeS, that brings together Continuum Computing and Network Intelligence. Indeed, Continuum Computing is inherently network-aware and fundamentally network-dependent; however, Network Intelligence (intended as an active, native component of the continuum) remains fragmented and insufficiently integrated. Closing that gap remains a central research and engineering priority for enabling autonomy and SLA-driven operation across heterogeneous entities and infrastructures. Therefore, rather than treating intelligence as a collection of isolated, stand-alone solutions, the Intelligence Continuum frames it as a cross-layer property emerging from the coordinated interaction of networking, computing, control, and application components. Within this framework, DTs have been discussed as a promising enabling paradigm for mitigating the semantic, observational, and control-plane fragmentation that historically separated computation and networking domains. By providing shared representations of application requirements, system behavior, network state, and context across layers, DTs can support more coherent reasoning and coordination across the CPeS, thus providing a viable pathway toward the realization of the Intelligence Continuum.

Looking ahead, progress toward realizing the Intelligence Continuum will require sustained, interdisciplinary engagement across AI engineering, networking research, and CPeS design. Key areas of focus are diverse, and include the formulation of standardized abstractions integrating networking, computing, and application domains; the design of trustworthy AI-driven control mechanisms capable of operating across heterogeneous timescales; and the development of continuum-aware DTs, to name a few. However, while these efforts encompass significant technical complexity, the development of the Intelligence Continuum can be viewed as an essential step towards future truly autonomous, resilient, and adaptive CPeS.

Funding

The Author acknowledges financial support under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.1, Call for tender No. 1409 published on 14.9.2022 by the Italian Ministry of University and Research (MUR), funded by the European Union–NextGenerationEU–Project ”Entrust: usEr ceNtric plaTform foR continoUS healThcare”—CUP H53D23008110001—Grant Assignment Decree No. 1382 adopted on 01-09-2023 by the Italian MUR.

Data Availability Statement

Not applicable.

Acknowledgments

The author gratefully acknowledges the coauthors of the previously published works referenced in this article for their valuable contributions and collaborations, which have significantly influenced and informed the perspectives presented herein.

Conflicts of Interest

The author declares no conflict of interest.

Use of AI and AI-Assisted Technologies

During the preparation of this work, the author used ChatGPT with minimal intervention to enhance grammatical correctness and readability. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the published article.

References

1. Singh, N.; Panigrahi, P.K.; Zhang, Z.; et al. Cyber-Physical Systems: A Bibliometric Analysis of Literature. *J. Intell. Manuf.* **2025**, *36*, 2335–2371.

2. Derler, P.; Lee, E.A.; Sangiovanni-Vincentelli, A.L. *Addressing Modeling Challenges in Cyber-Physical Systems*; University of California at Berkeley: Berkeley, CA, USA, 2011.
3. Attaran, M. The Impact of 5G on the Evolution of Intelligent Automation and Industry Digitization. *J. Ambient Intell. Humaniz. Comput.* **2023**, *14*, 5977–5993.
4. Fortino, G.; Savaglio, C.; Spezzano, G.; et al. Internet of Things as System of Systems: A Review of Methodologies, Frameworks, Platforms, and Tools. *IEEE Trans. Syst. Man Cybern. Syst.* **2020**, *51*, 223–236.
5. Bartelt, C.; Rausch, A.; Rehfeldt, K. Quo Vadis Cyber-Physical Systems: Research Areas of Cyber-Physical Ecosystems: A Position Paper. In Proceedings of the 1st International Workshop on Control Theory for Software Engineering, Bergamo, Italy, 31 August 2015; pp. 22–25.
6. Casadei, R.; Aguzzi, G.; Audrito, G.; et al. Software Engineering for Collective Cyber-Physical Ecosystems. *ACM Trans. Softw. Eng. Methodol.* **2025**, *34*, 1–40.
7. Adekunle, B.I.; Chukwuma-Eke, E.C.; Balogun, E.D.; et al. Machine Learning for Automation: Developing Data-Driven Solutions for Process Optimization and Accuracy Improvement. *Mach. Learn.* **2021**, *2*, 800–808.
8. Barbuto, V.; Savaglio, C.; Chen, M.; et al. Disclosing Edge Intelligence: A Systematic Meta-Survey. *Big Data Cogn. Comput.* **2023**, *7*, 44.
9. Mangione, F.; Savaglio, C.; Fortino, G. Generative Artificial Intelligence for Internet of Things Computing: A Systematic Survey. *arXiv* **2025**, arXiv:2504.07635.
10. Mao, W.; Zhao, Z.; Chang, Z.; et al. Energy-Efficient Industrial Internet of Things: Overview and Open Issues. *IEEE Trans. Ind. Inform.* **2021**, *17*, 7225–7237.
11. Islam, R.U.; Savaglio, C.; Fortino, G. Leading Smart Environments towards the Future Internet through Name Data Networking: A Survey. *Future Gener. Comput. Syst.* **2025**, *167*, 107754.
12. Dustdar, S.; Pujol, V.C.; Donta, P.K. On Distributed Computing Continuum Systems. *IEEE Trans. Knowl. Data Eng.* **2022**, *35*, 4092–4105.
13. Shen, X.; Gao, J.; Wu, W.; et al. Holistic Network Virtualization and Pervasive Network Intelligence for 6G. *IEEE Commun. Surv. Tutorials* **2021**, *24*, 1–30.
14. Grieves, M.W. *Digital Twins: Past, Present, and Future*; Springer International Publishing: Cham, Switzerland, 2023; pp. 97–121. https://doi.org/10.1007/978-3-031-21343-4_4.
15. Minerva, R.; Lee, G.M.; Crespi, N. Digital Twin in the IoT Context: A Survey on Technical Features, Scenarios, and Architectural Models. *Proc. IEEE* **2020**, *108*, 1785–1824.
16. Savaglio, C.; Barbuto, V.; Mangione, F.; et al. Generative Digital Twins: A Novel Approach in the IoT Edge-Cloud Continuum. *IEEE Internet Things Mag.* **2024**, *8*, 42–48.
17. Savaglio, C.; Fortino, G.; Ma, J. *Device-Edge-Cloud Continuum*; Springer: Cham, Switzerland, 2023.
18. Li, S.; Xu, L.D.; Zhao, S. The Internet of Things: A Survey. *Inf. Syst. Front.* **2015**, *17*, 243–259.
19. Stankovski, S.; Ostojić, G.; Baranovski, I.; et al. The Impact of Edge Computing on Industrial Automation. In Proceedings of the 2020 19th International Symposium Infoteh-Jahorina (Infoteh), East Sarajevo, Bosnia and Herzegovina, 18–20 March 2020; pp. 1–4.
20. Lane, N.D.; Bhattacharya, S.; Mathur, A.; et al. Squeezing Deep Learning into Mobile and Embedded Devices. *IEEE Pervasive Comput.* **2017**, *16*, 82–88.
21. Tundo, A.; Mobilio, M.; Ilager, S.; et al. An Energy-Aware Approach to Design Self-Adaptive AI-Based Applications on the Edge. In Proceedings of the 2023 38th IEEE/ACM International Conference on Automated Software Engineering (ASE), Luxembourg, 11–15 September 2023; pp. 281–293.
22. LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521*, 436–444.
23. Wang, J.; Chen, Y.; Hao, S.; et al. Deep Learning for Sensor-Based Activity Recognition: A Survey. *Pattern Recognit. Lett.* **2019**, *119*, 3–11.
24. Belcastro, L.; Marozzo, F.; Orsino, A.; et al. Navigating the Edge-Cloud Continuum: A State-of-Practice Survey. *arXiv* **2025**, arXiv:2506.02003.
25. Deng, S.; Zhao, H.; Fang, W.; et al. Edge Intelligence: The Confluence of Edge Computing and Artificial Intelligence. *IEEE Internet Things J.* **2020**, *7*, 7457–7469.
26. Chiang, M.; Zhang, T. Fog and IoT: An Overview of Research Opportunities. *IEEE Internet Things J.* **2016**, *3*, 854–864.
27. Chen, Y.; Liu, Z.; Zhang, Y.; et al. Deep Reinforcement Learning-Based Dynamic Resource Management for Mobile Edge Computing in Industrial Internet of Things. *IEEE Trans. Ind. Inform.* **2020**, *17*, 4925–4934.

28. Coronado, E.; Behravesh, R.; Subramanya, T.; et al. Zero Touch Management: A Survey of Network Automation Solutions for 5G and 6G Networks. *IEEE Commun. Surv. Tutorials* **2022**, *24*, 2535–2578.
29. Savaglio, C.; Pace, P.; Aloi, G.; et al. Lightweight Reinforcement Learning for Energy Efficient Communications in Wireless Sensor Networks. *IEEE Access* **2019**, *7*, 29355–29364.
30. Bringhenti, D.; Marchetto, G.; Sisto, R.; et al. Automation for Network Security Configuration: State of the Art and Research Trends. *ACM Comput. Surv.* **2023**, *56*, 1–37.
31. Samdanis, K.; Abbou, A.N.; Song, J.; et al. AI/ML Service Enablers and Model Maintenance for Beyond 5G Networks. *IEEE Netw.* **2023**, *37*, 162–172.
32. Dong, Q.; Xu, J.; Pillai, P.; et al. Does Accurate Real-Time AI Need Edge Offload? In Proceedings of the Tenth ACM/IEEE Symposium on Edge Computing, Arlington, VA, USA, 3–6 December 2025; pp. 1–18.
33. Gupta, A.K.; Govindankutty, S.; Nandyala, A.R.; et al. AI-Driven Test Automation for Big Data Systems in Cloud Environments. In Proceedings of the 2025 First International Conference on Advances in Computer Science, Electrical, Electronics, and Communication Technologies (CE2CT), Nainital, India, 21–22 February 2025; pp. 1354–1359.
34. Zhang, Q.; Cheng, L.; Boutaba, R. Cloud Computing: State-of-the-Art and Research Challenges. *J. Internet Serv. Appl.* **2010**, *1*, 7–18.
35. Karaduman, B.; Tezel, B.T.; Challenger, M. Rational Software Agents with the BDI Reasoning Model for Cyber-Physical Systems. *Eng. Appl. Artif. Intell.* **2023**, *123*, 106478.
36. Pournaropoulos, F.; Patras, A.; Antonopoulos, C.D.; et al. Fluidity: Providing Flexible Deployment and Adaptation Policy Experimentation for Serverless and Distributed Applications Spanning Cloud-Edge-Mobile Environments. *Future Gener. Comput. Syst.* **2024**, *157*, 210–225.
37. Sutton, R.S.; Barto, A.G. *Reinforcement Learning: An Introduction*, 2nd Ed.; MIT Press: London, UK, 2018.
38. D'Aniello, G.; Greco, L.; Pupo, F.; et al. Entrust: User Centric Platform for Continuous Healthcare. In Proceedings of the 2024 20th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT), Abu Dhabi, United Arab Emirates, 29 April–1 May 2024; pp. 595–600.
39. Yu, F.R. From Information Networking to Intelligence Networking: Motivations, Scenarios, and Challenges. *IEEE Netw.* **2021**, *35*, 209–216.
40. Wu, H.; Xiang, Z.; Nguyen, G.T.; et al. Computing Meets Network: Coin-Aware Offloading for Data-Intensive Blind Source Separation. *IEEE Netw.* **2021**, *35*, 21–27.
41. Gkonis, P.; Giannopoulos, A.; Trakadas, P.; et al. A Survey on IoT-Edge-Cloud Continuum Systems: Status, Challenges, Use Cases, and Open Issues. *Future Internet* **2023**, *15*, 383.
42. IEEE Computer Society. *Edge-Cloud Continuum*; IEEE Computer Society Whitepaper; IEEE Computer Society: Washington, DC, USA, 2021.
43. In-Network Computing: Emerging Trends for the Edge-Cloud Continuum. IEEE Network CFP, 2021. Available online: https://www.comsoc.org/publications/magazines/ieee-network/cfp/network-computing-emerging-trends-edge-cloud-continuum (accessed on 20 February 2026).
44. Opencontinuum Project. CORDIS Project Page. 2024. Available online: https://cordis.europa.eu/project/id/101070030 (accessed on 20 February 2026).
45. Enact—Cognitive Computing Continuum. CORDIS Project Page. 2025. Available online: https://cordis.europa.eu/project/id/101135423 (accessed on 20 February 2026).
46. Smartedge Project. Project Website. 2025. Available online: https://www.smart-edge.eu/ (accessed on 20 February 2026).
47. Jiang, Z.; Fu, S.; Zhou, S.; et al. AI-Assisted Low Information Latency Wireless Networking. *IEEE Wirel. Commun.* **2020**, *27*, 108–115.
48. Rustamov, K. 5G-Enabled Internet of Things: Latency Optimization through AI-Assisted Network Slicing. *Qubahan Techno J.* **2023**, *2*, 1–10.
49. Chi, H.R.; Wu, C.K.; Huang, N.-F.; et al. A Survey of Network Automation for Industrial Internet-of-Things toward Industry 5.0. *IEEE Trans. Ind. Inform.* **2022**, *19*, 2065–2077.
50. Salau, B.A.; Rawal, A.; Rawat, D.B. Recent Advances in Artificial Intelligence for Wireless Internet of Things and Cyber-Physical Systems: A Comprehensive Survey. *IEEE Internet Things J.* **2022**, *9*, 12916–12930.
51. Arulraj, J.; Chatterjee, A.; Daglis, A.; et al. eCloud: A Vision for the Evolution of the Edge-Cloud Continuum. *Computer* **2021**, *54*, 24–33.
52. Baresi, L.; Mendonça, D.F.; Garriga, M.; et al. A Unified Model for the Mobile-Edge-Cloud Continuum. *ACM Trans. Internet Technol.* **2019**, *19*, 1–21.

53. Toma, I.; Serles, U. Intend: Intent-Based Data Operation in the Computing Continuum. In Proceedings of the ESWC 2025: Project Networking Track, Portoroz, Slovenia, 1–5 June 2025. https://doi.org/10.5281/zenodo.15609773.
54. Tao, F.; Qi, Q.; Wang, L.; et al. Digital Twins and Cyber-Physical Systems toward Smart Manufacturing and Industry 4.0: Correlation and Comparison. *Engineering* **2019**, *5*, 653–661.
55. Lou, S.; Tan, R.; Zhou, M.; et al. A Cognitive Digital Twin for Industry 5.0 Based on a Large Language Model Agent. In Proceedings of the 2025 IEEE 5th International Conference on Human-Machine Systems (ICHMS), Abu Dhabi, United Arab Emirates, 26–28 May 2025; pp. 223–228.
56. Fortino, G.; Savaglio, C. Integration of Digital Twins & Internet of Things. In *The Digital Twin*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 205–225.
57. Almasan, P.; Ferriol-Galmés, M.; Paillisse, J.; et al. Network Digital Twin: Context, Enabling Technologies, and Opportunities. *IEEE Commun. Mag.* **2022**, *60*, 22–27.
58. Nguyen, H.X.; Trestian, R.; To, D.; et al. Digital Twin for 5G and Beyond. *IEEE Commun. Mag.* **2021**, *59*, 10–15.
59. Savaglio, C.; Barbuto, V.; Awan, F.M.; et al. Opportunistic Digital Twin: An Edge Intelligence Enabler for Smart City. *ACM Trans. Sens. Netw.* **2023**, *accepted*. https://doi.org/10.1145/3616014.
60. Liu, N.; Li, C.; Chen, L.; et al. Hybrid Data-Driven and Model-Based Distribution Network Reconfiguration with Lossless Model Reduction. *IEEE Trans. Ind. Inform.* **2021**, *18*, 2943–2954.
61. Bariah, L.; Debbah, M. The Interplay of AI and Digital Twin: Bridging the Gap between Data-Driven and Model-Driven Approaches. *IEEE Wirel. Commun.* **2024**, *31*, 219–225.
62. Amadeo, M.; Campolo, C.; Quevedo, J.; et al. Information-Centric Networking for the Internet of Things: Challenges and Opportunities. *IEEE Netw.* **2016**, *30*, 92–100.
63. Sun, H.; Liu, Y.; Al-Tahmeesschi, A.; et al. Advancing 6G: Survey for Explainable AI on Communications and Network Slicing. *IEEE Open J. Commun. Soc.* **2025**, *6*, 1372–1412.
64. Taleb, T.; Samdanis, K.; Mada, B.; et al. On Multi-Access Edge Computing: A Survey of the Emerging 5G Network Edge Cloud Architecture and Orchestration. *IEEE Commun. Surv. Tutorials* **2017**, *19*, 1657–1681.
65. Benzaid, C.; Taleb, T. AI-Driven Zero Touch Network and Service Management in 5G and Beyond: Challenges and Research Directions. *IEEE Network* **2020**, *34*, 186–194.
66. Afolabi, I.; Taleb, T.; Samdanis, K.; et al. Network Slicing and Softwarization: A Survey on Principles, Enabling Technologies, and Solutions. *IEEE Commun. Surv. Tutorials* **2018**, *20*, 2429–2453.
67. Leivadeas, A.; Falkner, M. A Survey on Intent-Based Networking. *IEEE Commun. Surv. Tutorials* **2022**, *25*, 625–655.
68. Tao, F.; Qi, Q.; Liu, A.; et al. Data-Driven Smart Manufacturing. *J. Manuf. Syst.* **2018**, *48*, 157–169.
69. Khan, L.U.; Han, Z.; Saad, W.; et al. Digital Twin of Wireless Systems: Overview, Taxonomy, Challenges, and Opportunities. *IEEE Commun. Surv. Tutorials* **2022**, *24*, 2230–2254.