2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

# AI-Driven Adaptive Autonomy: Is AI Really Pervasive? Research Gaps From Bibliometric Assessment

Simona Casini*, Andrea Caiti*, Pietro Ducange*, Francesco Marcelloni*, and Lorenzo Pollini*

*Abstract*— **Artificial Intelligence is widely recognised as a driver of adaptive autonomy in robotics. Yet, the extent to which AI techniques truly permeate the functional architecture of autonomous systems is still only partially characterised. Existing bibliometric analyses typically map research themes, keywords or algorithms but provide limited insight into how contributions distribute across the functional logic of autonomous systems. This raises a fundamental question: is AI really pervasive across the functions that enable robots to act adaptively in complex environments? Which areas are mature or under-explored? To achieve this outcome, the paper adopts a functional, control-loop-oriented perspective that avoids the bias of vertical domains or robot-specific applications, focusing instead on a fine-grained architecture with 13 functional modules spanning the perception, planning, execution, knowledge, and diagnosis domains. More than 2500 scientific works, published in the last 25 years, were mapped across the 13 functional modules, using a multi-label neural classification pipeline, and analysed via co-occurrence and structural techniques. This approach allowed to highlight not only areas where AI is already known to be central and consistently confirmed, but also those where its impact would be expected to be significant yet remains surprisingly limited. By combining architectural reasoning with bibliometric evidence, the study provides a broader lens for assessing research gaps and for situating current advances within the long-term agenda of adaptive and human-centred autonomy.**

## I. Introduction and Background

Over the past three decades, the integration of Artificial Intelligence (AI) into robotics, in what is called *intelligent systems* [1], has become one of the most transformative and rapidly expanding research areas. Both academic initiatives and institutional roadmaps (e.g., euRobotics SRIA [2], Stanford HAI [3]) consistently highlight AI as a critical enabler of robotic autonomy. Applications span from mainstream domains such as agriculture and healthcare to niche but strategically important contexts, including underwater and space robotics. Alongside technological advances, the theoretical perspective of Human-Centred Artificial Intelligence (HCAI) [4] has gained traction, emphasising transparency, interpretability, and value alignment problems, relevant when intelligent systems operate in open-ended, human-in-the-loop environments. Two main intellectual currents have historically shaped the debate on how intelligence should be implemented in robots and intelligent systems [5].

*1) Behaviour-based:* With his seminal work *Vehicles: Experiments in Synthetic Psychology* [6], Braitenberg illustrated how complex behaviours can emerge from the simple wiring of sensors and actuators. Building on this intuition, Brooks [7], [8] developed the subsumption architecture, where layered reactive modules connect perception directly to action. This behaviour-based paradigm rejected centralised symbolic planning, summarised in Brooks' claim that "*elephants don't play chess*" [9], and explicitly challenged the classical Sense–Model–Plan–Act paradigm dominant in the 1970s. By emphasising embodied and situated action, it produced systems that were robust in mostly static environments, though limited when scaling to long-term reasoning, semantic representation, dynamic conditions and mission-level decision-making.

*2) Objective-driven:* Recently, more and more researchers, one above all, LeCun [10], [11] (and references therein) have advocated for objective-driven AI, where agents are endowed with models of the internal world and guided by objectives. In this paradigm, intelligence arises from the ability to learn predictive models of the environment through self-supervised learning, simulate future trajectories, and plan actions that maximise long-term outcomes. In robotics, this shift implies moving beyond hand-crafted reactive behaviours to agents capable of generalisation, adaptive planning, and knowledge integration.

Behaviour-based reactivity and objective-driven planning remain central in autonomous robotics, motivating architectural formalisms that integrate reactive and deliberative mechanisms. Over the years, several general-purpose architectural frameworks have been proposed to support the design and evaluation of Robotics and Autonomous Systems (RASs), with the aim of addressing the growing complexity and fragmentation of robotic software development with authoritative contributions as [12], [13], [14], [15], [16], also including the ICRA Workshop on Innovative Robot Control Architectures [17] with vertical transposition in industry. Among existing frameworks, MAPE-K [18] is widely adopted for modelling self-adaptive systems, defining a closed-loop structure with Monitor, Analyse, Plan, Execute, and a shared Knowledge base [19], [20]. While its modular and feedback-driven design suits layered autonomous control, high-level abstractions often overlook key aspects of real-world robotics, including task decomposition, continuous mission execution, and human–robot interaction. Also

[1]Department of Information Engineering, University of Pisa, Largo Lucio Lazzarino 1, 56122 Pisa, Italy. Email: simona.casini@phd.unipi.it, {andrea.caiti, pietro.ducange, francesco.marcelloni, lorenzo.pollini}@unipi.it

This work has been funded by the PNRR - M4C2 - Investimento 1.3, Partenariato Esteso PE00000013 - "FAIR - Future Artificial Intelligence Research" - Spoke 1 "Human-centered AI" under the NextGeneration EU program, and by the Italian Ministry of University and Research (MUR) in the framework of the FoReLab and Cross-Lab projects (Departments of Excellence).

979-8-3315-8160-2/26/$31.00 ©2026 IEEE 

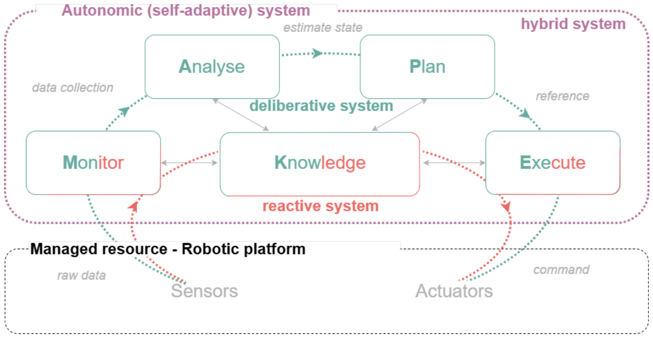


Fig. 1. MAPE-K-based control framework adapted for RAS. The red loop represents the reactive subsystem, where perception is directly coupled to action for rapid responses. The green loop depicts the deliberative subsystem, in which perception informs reasoning and planning before execution. The outer purple layer shows their hybrid integration, illustrating how autonomous systems combine fast reactivity with higher-level decision-making to balance safety and performance.

considering bibliometric analysis, over the years, multiple contributions have come out (e.g. [21], [22]). A recent study [23] presented a classical bibliometric analysis of AI-driven robotic systems from 2000 to 2024, using science mapping tools such as SciMAT and VOSviewer. That work identified historical trends, key thematic clusters, and the emergence of HCAI as a guiding paradigm. Building on these insights, the present study enriches that perspective by embedding the analysis within a fine-grained functional, MAPE-K-oriented, framework. This framework reorganises AI contributions into functional blocks such as perception, planning, knowledge management, and execution, providing an architectural lens that connects bibliometric evidence to the operational logic of autonomous robots. Unlike traditional bibliometric studies based on keyword clustering, the proposed approach offers three main key contributions:

- introduction of a functional, control-loop-oriented framework adapted from MAPE-K that situates AI contributions within the operational architecture of robotic autonomy, effectively framing the relationship between reactive and objective-driven paradigms, regardless of the nature of the robot (underwater, aerial, legged, wheeled, etc.).
- a functional map of AI in autonomous robotics that helps researchers place their work, identify underexplored modules, and understand how contributions relate to system-level design.
- a reproducible methodological pipeline for applying this framework at scale, enabling both cross-sectional and longitudinal analyses of the literature.

The remainder of the paper is structured as follows: Section II introduces the functional control-loop architecture, Section III describes the bibliometric methodology and classification pipeline, Section IV presents the main findings and structural analyses, Section V discusses their implications for AI-driven robotics, and Section VI concludes with final remarks and future research directions.

## II. Robotic and Autonomous System Architecture

Refining and extending the original MAPE-K model involves a two-level architectural adaptation process, aimed at aligning the framework with the structural and functional characteristics of RASs. The first corresponds to a high-level control-loop integration of well-established robotic primitives and control paradigms, which have historically provided the foundation for the generation of robot behaviour [7], [24]. The Sense, Plan, and Act primitives map onto the fundamental stages of autonomy: perception of the environment, reasoning over internal state and objectives, and execution of physical actions. Consequently, control paradigms, namely deliberative, reactive, and hybrid architectures, further shape the temporal and structural characteristics of robotic decision-making. Deliberative systems emphasize internal planning and long-term reasoning, while reactive systems rely on immediate sensor-to-actuator coupling for fast adaptation. Hybrid models aim to balance these approaches by layering reactive responsiveness with mission-level deliberation. These paradigms influence both system design and algorithmic deployment and are essential when adapting a software-oriented framework (like MAPE-K) to the robotics domain (also demonstrated in [12], [14]). Figure 1 visually illustrates the MAPE-K alignment with robotics-specific control logic and paradigmatic requirements, accommodating both low-latency reactivity and long-term deliberation.

The second level of refinement focuses on the internal decomposition of each MAPE-K component, with the aim of capturing the functional richness and specific operational demands of real-world robotic systems. In particular, each module is expanded into sub-modules that reflect both reactive capabilities (such as real-time control, low-latency monitoring, and sensor fusion) and deliberative mechanisms (including mission decomposition, planning under uncertainty, and strategy adaptation). This granular architecture addresses the needs of long-term, self-adaptive robotic systems operating in real-world environments, where complex missions are decomposed into evolving sequences of tasks. Its emphasis on modularity and internal specialization ensures applicability across heterogeneous platforms and supports AI contributions at both functional and architectural levels. Considering Figure 2, the grey blocks on the background represent the MAPE-K architecture structure. The relationship between sub-blocks has been maintained, also guaranteeing paradigms match. On the left side of the figure, the framework incorporates human demands and vision, which articulate the purpose and expected role of the robot within its environment. These inputs are processed by the **vision manager**, a high-level interface responsible for interpreting long-term user intent and bridging the human-machine division. The interpreted vision feeds into the knowledge module, which comprises a **learning repository** containing models of the system, environment, missions, and past experiences. This repository also encodes strategy policies, control rules, and

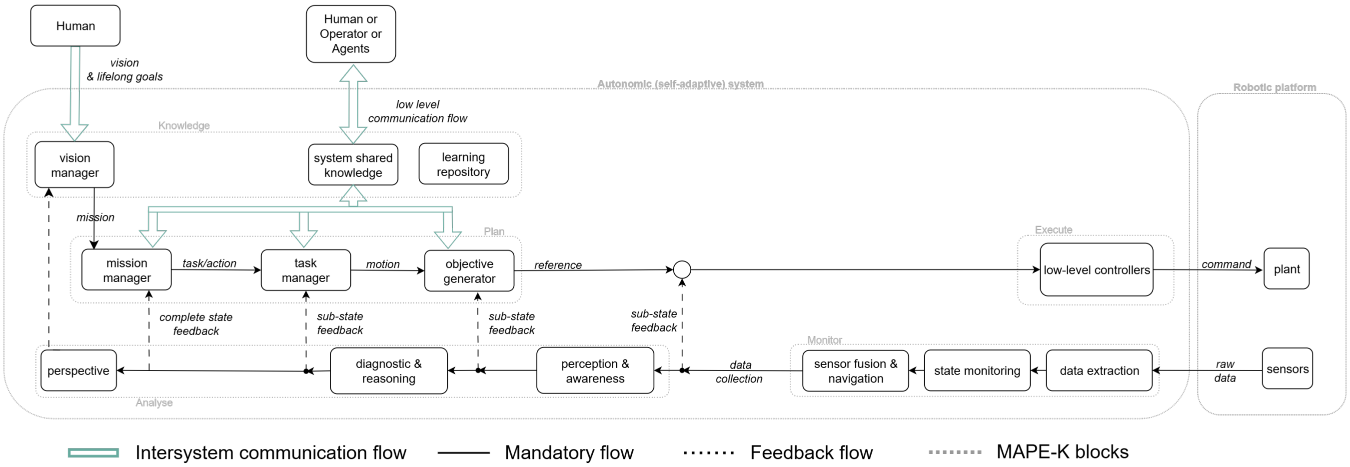


Fig. 2. Proposed control loop scheme with reference (in grey) to the MAPE-K architecture. This structure enables fine-grained classification of scientific contributions according to their functional role in autonomous and self-adaptive robotic systems.

interpretations of human instructions, enabling persistent and shared knowledge across modules. In the planning module, the system decomposes goals into structured mission components. A **mission manager** orchestrates high-level objectives, while a **task manager** translates these into concrete actions. An **objective generator** then formulates executable commands based on current goals and system state. The planning module communicates extensively with the analysis module, which is tasked with assessing system state and reasoning over diagnostic and perceptual data. Modules such as **diagnostic and reasoning** and **perception and awareness** ensure that the planning process is grounded in an up-to-date understanding of the robot's environment and internal status. The execution module takes the refined motion objectives and translates them into commands through a **low-level controller**, responsible for real-time actuation. Feedback loops from this module provide continuous sub-state information back to the planning and analysis blocks, ensuring reactivity and adaptability in dynamic environments.

Monitoring occurs at multiple levels. The monitoring module integrates raw data from sensors using a **sensor fusion** submodule, filters and interprets the data via **state monitoring**, and performs feature extraction and diagnostics within the **data extraction** component. These outputs inform both the analysis and execution processes. The system interfaces directly with the robotic platform, composed of a plant (i.e., actuators) and a variety of sensors (i.e., vision, inertial, position, and environmental sensors) through which it perceives and interacts with the physical world.

A brief description of the scheme blocks is also reported in Table I. To assess the plausibility and coverage of the proposed 13-module framework, a qualitative validation was conducted. First, a targeted qualitative assessment was conducted on a representative subset of robotics literature, selected to cover heterogeneous functional roles, including perception, planning, control, and knowledge-related processes. In addition, the framework was confronted with a widely adopted real-world robotic software architecture, namely the ROS 2 Navigation Stack (Nav2 [25]). The main

TABLE I

FUNCTIONAL MODULES OF THE PROPOSED FRAMEWORK, GROUPED BY MAPE-K BLOCKS.

| Module | Function |
|---|---|
| *Monitor (M)* | |
| Data Extraction | Acquires raw information from sensors |
| State Monitoring | Tracks internal robot states and variables |
| Sensor Fusion & nav. | Merges heterogeneous perceptions into navigational goals, linking sensing to motion control |
| *Analyse (A)* | |
| Perception & Awareness | Combines sensing with semantic context to support decision-making and human interaction. |
| Diagnostic & Reasoning | Detects anomalies, evaluates system health, supports recovery |
| Perspective | Provides higher-level context for vision implementation |
| *Plan (P)* | |
| Mission Manager | Oversees mission-level task planning |
| Task Manager | Task decomposition in multiple motion objectives |
| Motion Objective Generator | Produces intermediate control reference goals |
| *Execute (E)* | |
| Low-level Controller | Sends commands to actuators, ensures stability |
| *Knowledge (K)* | |
| Vision Manager | Collects and interprets lifelong goals from the Operator/Human to define missions |
| Shared Knowledge | Stores and shares semantic/contextual data, intersystem communication flow block |
| Learning Repository | Manages training datasets and learned models |

Nav2 components, including perception and state estimation (e.g., costmaps and localization), planning, control, and behaviour orchestration, can be consistently associated with the corresponding functional modules of the proposed architecture. Overall, the architecture integrates multiple closed feedback loops, supporting both top-down mission planning and bottom-up responsiveness, and provides a functional scaffold for mapping AI contributions in robotics. From a human-centred perspective, it accommodates different au-

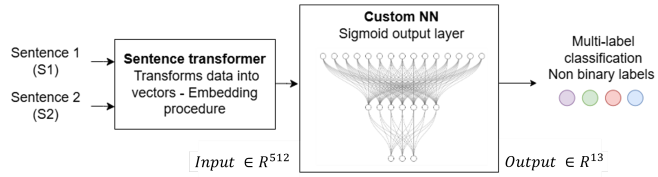


Fig. 3. Multi-label classification pipeline. Abstracts and keywords are embedded into text vectors and processed by a neural network to predict probabilities for 13 functional modules.

tonomy levels—from vision-driven autonomy to supervised decision-making and full human responsibility—supported by the Knowledge module through continuous interaction with humans, peer robots, and supervisory systems.

## III. Functional Mapping - Methodologies

This section presents a functionally oriented methodology that combines large-scale literature analysis with supervised classification to map AI contributions within a control-oriented architecture for robotic autonomy.

The database search query was designed to capture AI methods within the entire autonomous robotics domain; no application-specific labels (e.g. *fault detection*, *obstacle avoidance*, etc.) were employed to keep the query as general and comprehensive as possible. Although the inclusion of *autonomous robot** keyword restricts the search space, it ensures a consistent focus on robotic systems and does not compromise the structural insights derived from the analysis.

---

*(TITLE-ABS-KEY("artificial intelligence" OR "machine learning" OR deep OR reinforcement OR "data driven") AND TITLE-ABS-KEY ("autonomous robot*") AND NOT TITLE-ABS-KEY (review*)) AND PUBYEAR>1999 AND PUBYEAR<2026 AND (LIMIT-TO(LANGUAGE, "English"))*

---

The dataset considered comprises over 2,500 peer-reviewed publications retrieved from the Scopus database, spanning the years 2000 to 2025 and authored by a wide range of researchers across multiple subdomains of robotics and artificial intelligence. As a result, the textual content is inherently heterogeneous, reflecting diverse writing styles, terminologies, and methodological focuses. This introduces a non-negligible level of semantic noise, as well as class imbalance across the functional categories, making the classification task particularly challenging from a modelling perspective. For each document, the abstract and indexed keywords were extracted, then concatenated into a single textual input. Titles were excluded to reduce redundancy, as they often overlap with abstracts or keywords and provide less detailed methodological information, which is more effectively captured by abstracts and indexed keywords.

Each document is then processed through a neural network, trained on a manually annotated/labelled subset of 800 publications. Annotations were independently performed by the 5 authors, acting as domain experts in robotics and AI (including explainability, trustworthiness, and human-centred AI); disagreements were resolved through collective discussion. Given (as mentioned) the inherent heterogeneity and imbalance of the dataset, comprising papers from diverse subdomains of robotics and AI, a random sampling stratified strategy (from PyTorch library [26]) was adopted to ensure a representative and unbiased training set. [1]

The output of this process is a functional transposition of the literature, in which each publication is associated with one or more components of the architecture. This enables both synchronic and diachronic analyses: identifying dominant clusters, monitoring the evolution of specific research directions over time, and highlighting underexplored or emergent areas. Ultimately, this methodology provides a system-level lens through which the role of AI in robotics can be interpreted, not only thematically, but also operationally.

### A. Neural Network Model for Functional Classification

TABLE II
MultiLayer Perceptron details

| Perceptron Design Parameters | Value |
|---|---|
| Embedder | distiluse-base-multilingual |
| Number of hidden layer | 2 |
| Number of inputs | 512 |
| Number of outputs | 13 |
| Hidden layer activation function | ReLU |
| Output layer function | Sigmoid |
| Training Function | BCE loss |
| Performance | Mean squared error |
| K-fold number | 10 |
| Epochs | 100 |
| Training ratio | 80% |
| Test ratio | 20% |
| Hyper-parameter optimization | Random Search |
| **Results** | **Value** |
| Weighted Precision - Test set | 0.91 |
| Weighted Recall - Test set | 0.907 |
| Weighted F1 - Test set | 0.907 |

Text classification is a well-established task in natural language processing, with models ranging from rule-based methods to deep neural networks. For this study, the objective was not to advance classification techniques but to provide a reliable tool to support functional mapping. For this reason, a MultiLayer Perceptron (MLP) was selected as a simple, interpretable, and reproducible solution, offering solid performance at relatively low computational cost [27]. While Large Language Models (LLMs) represent promising alternatives for richer semantic analysis, the focus here is on methodological clarity, leaving LLM-based extensions for future work.

The pipeline is illustrated in Figure 3. Abstracts and indexed keywords were concate-

[1]Some threats nonetheless may exist: the dataset relies on Scopus indexing, which may miss very recent publications; class imbalance across modules may bias results toward well-represented areas; and the use of random splits, while ensuring robustness, does not capture temporal concept drift. We believe these limitations do not undermine the methodological contribution but rather outline natural extensions for future work.

nated, preprocessed, and embedded using the `distiluse-base-multilingual-cased-v1` sentence transformer [28], chosen for its balance between semantic richness and efficiency. The resulting vectors were fed into a two-layer MLP with dropout regularisation, producing multi-label outputs aligned with the 13 functional sub-modules of the proposed architecture. This enables each paper to be associated with multiple functions (e.g., a reinforcement learning navigation paper may contribute simultaneously to perception, planning, and strategy management).

To ensure robustness, the dataset was randomly partitioned into training and test sets. This choice reflects the main goal—functional classification—rather than temporal forecasting, as a chronological split would be biased by concept drift and the absence of recent terms in the training data. Performance was evaluated with 10-fold cross-validation, achieving weighted precision, recall, and F1-scores above 90% (Table II, Figure 4). These results confirm that the model provides sufficient accuracy to support the architectural-level analysis while maintaining simplicity and interpretability.

### B. Multi-Label Classification Structural Analysis

To explore interdependencies among functional modules and extract structural insights from the classified literature, a post-classification analysis was conducted based on the output of the multi-label model. For each publication, only the most relevant functional labels were retained, limiting the analysis to the top three predicted classes. This choice reflects a trade-off between preserving the multi-domain nature of research contributions and controlling noise in the co-occurrence analysis. These class co-occurrences were aggregated across the corpus to construct a co-occurrence matrix, where each entry indicates how often two architectural modules were jointly assigned to the same publication. This matrix was used to generate an undirected, weighted graph in which nodes represent functional components and edges encode the strength of their co-occurrence. Edge thickness reflects the number of times two components co-occurred, while node size is proportional to the number of times each component appeared in these co-assignments, referred to as node occurrences. To quantify the structural relevance of each module within the overall architecture, degree centrality was computed for every node. This metric captures how many direct connections a node has with others in the graph, offering a measure of its transversality. Importantly, high centrality does not necessarily correspond to a high number of prediction mass. While prediction mass refers to the cumulative labels attributed to a component across all documents, node occurrences reflect how often that component appears in the final co-assignment set. This distinction is crucial: a component may have moderate overall attribution but still play a central, integrative role within the network. The resulting co-occurrence graph complements the overall distribution of class probabilities. Together, these visualizations enable a dual-layered analysis: (i) a quantitative overview

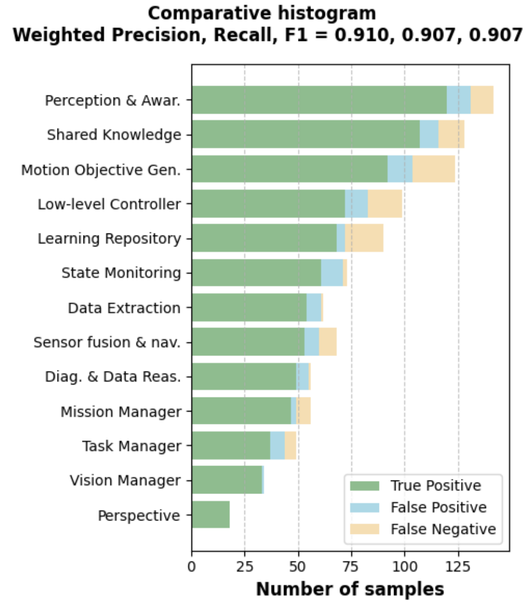


Fig. 4. Comparative histogram of True Positives, False Positives, and False Negatives across functional blocks in the test set. The results show consistent model performance across all 13 classes.

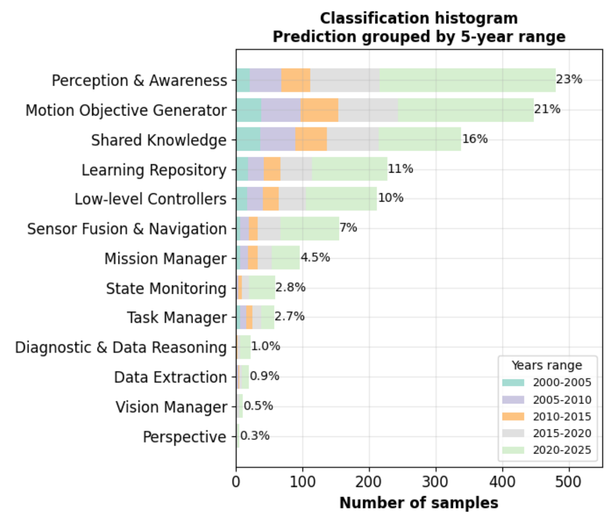


Fig. 5. Distribution of classification labels across all publications. Bars represent the cumulative occurrences assigned to each functional module, showing strong concentration in *Perception & awareness* and *Motion Objective Generator*, with other modules less represented. The colour map shows the distribution over the years grouped by a 5-year range.

of how frequently functional components are emphasized across the literature, and (ii) a structural perspective on how components are related, clustered, and interconnected, revealing thematic convergence and functional overlap within the AI-robotics research landscape. These analyses set the foundation for the interpretation of results discussed in Section IV.

## IV. Results

The following analysis examines a functional overview of the literature. It highlights how contributions are allocated among the modules of the autonomy-oriented architecture. Figure 5 reports the cumulative distribution of classifica-

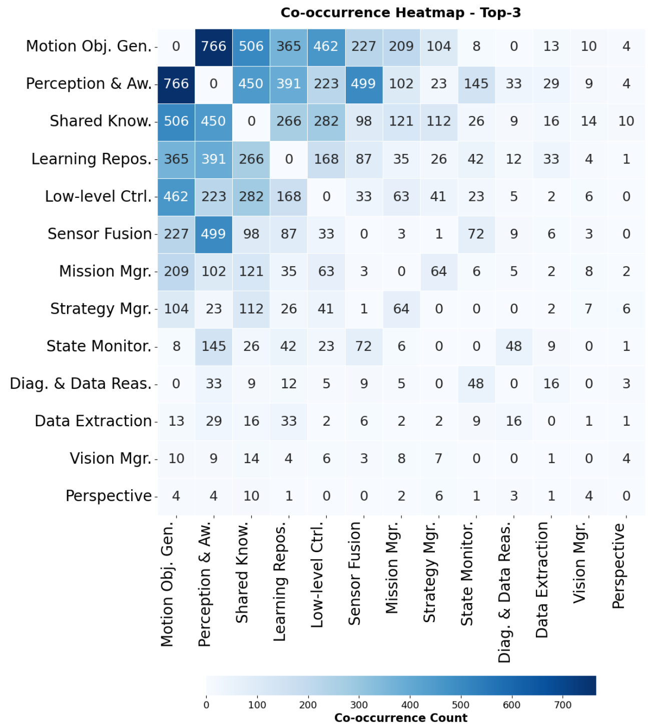


Fig. 6. Co-occurrence heatmap of functional sub-blocks. Colour intensity reflects the frequency of co-assigned modules, highlighting strong functional couplings. Node centrality corresponds to the number of non-zero co-occurrences per row.

tion predictions for each module. To provide a diachronic perspective, it also reports the distribution of classification outputs grouped into five-year intervals. The results show that the dominance of *Perception & Awareness* and *Motion Objective Generation* is consistent across all time windows, while knowledge-oriented components (e.g., Shared Knowledge and Learning Repository) exhibit a clearer growth in the most recent period (2020–2025). In wider terms, this concentration reflects a strong research focus on real-time environmental awareness, motion planning, and adaptive behaviour; critical capabilities for autonomy in dynamic and unstructured settings. Other modules also attract significant attention. *Shared Knowledge* and *Learning Repository* show frequent attribution, confirming the growing interest in experience reuse, semantic understanding, and long-term system adaptation. In contrast, several components remain underrepresented. *Perspective*, *Vision Manager*, and *Data Extraction* occupy only a small share of the classification output. This suggests areas that are either still underexplored or functionalities embedded implicitly in broader modules such as perception or learning. Particularly notable is *Diagnostic & Reasoning*. Despite its importance for system introspection and fault detection, it accounts for less than 1% of the predictions, exposing a potential research gap.

Beyond the quantitative distribution, the structural analysis highlights interdependencies among components. Figure 6, derived from the co-occurrence matrix introduced in Section III - Figure 7, visualises these relations in a network. Nodes represent architectural blocks, edges indicate the frequency with which pairs of modules co-occur in the same publication. Node size reflects total co-attribution, edge thickness co-occurrence strength, and node colour degree centrality. This analysis reveals several tightly interconnected clusters. *Perception & Awareness* and *Motion Objective Generator* frequently co-occur, highlighting the established coupling between environment modelling, goal formulation, and motion planning. Similarly, *Sensor Fusion* and *Low-Level Controller* show strong alignment, emphasising the integration of perception and control in real-time systems but also the rise of end-to-end AI solutions. Interestingly, transversal roles emerge. *Shared Knowledge*, *Learning Repository*, and *Strategy Manager* appear as high-centrality nodes, linking perception, planning, and execution. Their prominence suggests a shift from narrowly scoped, reactive systems toward architectures that prioritise knowledge integration, experience accumulation, and coordination across modules. Equally relevant are components with low probability mass but high connectivity. For instance, *Data Extraction* is closely associated with higher-level modules such as *Mission Manager* and *Strategy Manager*. Although underrepresented in volume, its position in the network underscores the transversal role it plays within the autonomy architecture.

## V. Discussion

The results presented in this work extend previous bibliometric efforts by shifting the focus from thematic clustering to a functionally grounded perspective. Earlier studies (e.g. [23]) mainly highlighted correspondences between broad research areas (e.g. vision, planning, control) and algorithms (e.g. supervised, unsupervised, reinforcement learning), whereas our architecture-based mapping provides a structured view of AI contributions within robotic autonomy. By aligning literature classification with the control logic of autonomous systems, the analysis highlights both dominant functional trends and underexplored capability gaps, clarifying which components of an autonomy architecture have been extensively studied and which remain neglected. This perspective also reinforces the conceptual debate introduced in Section I. The concentration of contributions in *perception* and *motion planning* reflects the behaviour-based tradition of reactive architectures, where sensing and acting are tightly coupled, while the growing attention to knowledge-oriented modules (e.g. *Shared Knowledge Base*, *Learning Repository*) points toward convergence with objective-driven paradigms centred on long-term reasoning and world modelling. Our findings suggest thus that contemporary research increasingly integrates elements of both, combining reactive perception–action loops with deliberative, knowledge-rich components. From these observations, we can distil some research trajectories that emerge as particularly relevant for future developments in autonomous robotics:

- **Self-Diagnosis**: contributions explicitly addressing the system's introspective functions are relatively scarce. While in structured and supervised tasks, also in industry, this module is well known with a lot of literature in anomaly detection for example, this could be a structural gap in current dynamic multi-tasks autonomous systems. Without such capabilities, robots struggle to achieve endurance and resilience in long-term deployments, making this an urgent area for further research and innovation.

- **Knowledge Sharing and Semantic Awareness**: the significant presence of modules like the *Shared Knowledge Base* and *Learning Repository* in our mapping suggests an emphasise knowledge reuse, semantic reasoning, and contextual awareness. These modules enable robots to integrate information over time, share insights with other agents, and interpret the environment at a higher semantic level. Such capabilities are foundational for multi-robot coordination, transfer learning between tasks or domains, and sustained autonomy beyond narrow, pre-programmed behaviours. The fact that many recent contributions incorporate shared knowledge or world-model components indicates a community-wide recognition that scalable autonomy requires not just reactive skills, but also rich semantic understanding of the world.
- **HCAI in Robotics**: confirms that modules tied to transparency, explainability, and human–robot collaboration have been found. These principles are increasingly being built into robotic systems. Beyond performing tasks autonomously, robots are now expected to justify their decisions and adapt their behaviour in alignment with human intent and oversight. For example, components that provide understandable system state feedback or allow humans to intervene (human-in-the-loop decision modules) appeared across various architectures in the literature.
- **Perspective and Task Management**: intended to provide an external, reflective view of the robot's operation. This module relates to the previous item, ensuring that system decisions can be contextualised and communicated in human-understandable terms. Closely connected to this, the absence of a dedicated task manager highlights the importance of task decomposition, prioritisation, and alignment with human values. These areas, still only marginally addressed in the literature, open promising directions for integrating introspection, explainability, and value-sensitive design into robotic autonomy.

Moreover interesting is that these insigths confirm and align with priorities outlined in the 2024 SRI Agenda of euRobotics [2], which emphasises autonomous systems that are capable and adaptive while also safe, explainable, and socially acceptable. The prominence of human-centred AI in our mapping mirrors this vision, grounded in transparency, accountability, and user-centric design. Moreover, the identified gaps — from self-diagnosis and mission-level reasoning to human-in-the-loop control — echo the SRIA's focus on semantic understanding, introspection, and long-term dependability.

## VI. Conclusions

This paper introduced a functional control-loop-oriented framework to map the contributions of AI to robotic autonomy as a refinement of the MAPE-K model. The study provided a system-level perspective of AI use in robotics that complements traditional bibliometric approaches by applying supervised classification to more than 2,500 publications. The analysis confirmed the predominance of AI in perception and motion planning, highlighted the transversal role of knowledge integration, and exposed critical gaps in diagnostic reasoning and human-in-the-loop decision-making. These findings suggest a progressive shift from component-level intelligence to modular, integrative, and knowledge-driven architectures, in line with the long-term vision of the euRobotics SRIA. Beyond offering a reproducible pipeline for large-scale literature mapping, the framework also contributes with a conceptual lens to interpret past developments, highlight current imbalances, and guide future directions. It provides researchers and practitioners with a reference map to situate their work, identify under-explored areas, and orient the design of the next generation of autonomous and human-aligned robotic systems. Future research may build on these results to characterise the features of self-adaptive robots in real-world environments and to extend the methodology to specific subdomains such as human–robot interaction, multi-agent coordination, and ethical decision-making in human-centred autonomy.

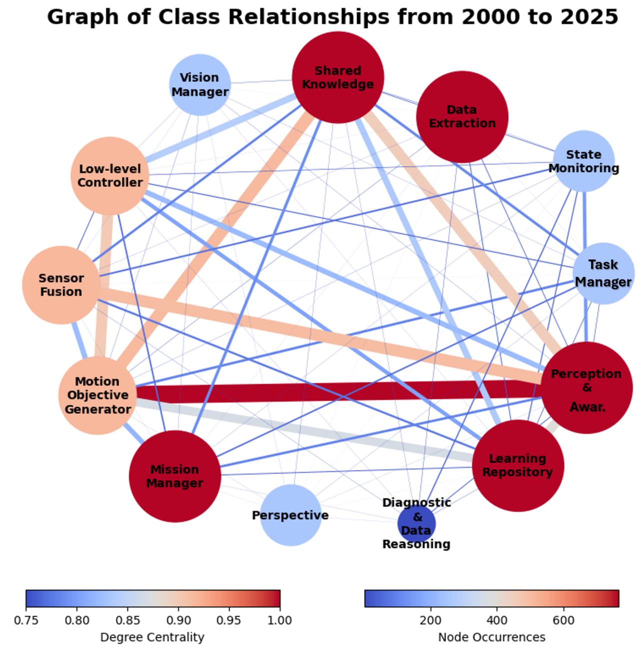


Fig. 7. Co-occurrence graph of functional components (2000–2025). Nodes represent sub-modules sized by occurrence and coloured by degree centrality, with edges weighted by co-occurrence strength. High-centrality nodes highlight transversal components that frequently connect across modules, independent of their overall frequency.

## References


[1] Rolf Pfeifer, Fumiya Iida, and Josh Bongard. "New robotics: Design principles for intelligent systems". In: *Artificial life* 11.1-2 (2005), pp. 99–120.

[2] euRobotics. *Robotics Strategic Research and Innovation Agenda (2024)*. https://eu-robotics.net/strategy/. Accessed: 2025-03-26. 2024.

[3] Stanford University Human-Centered Artificial Intelligence. *Artificial Intelligence Index Report 2025*. https://hai.stanford.edu/ai-index/2025-ai-index-report. 2025.

[4] Mark O. Riedl. “Human-centered artificial intelligence and machine learning”. In: *Human Behavior and Emerging Technologies* 1.1 (2019), pp. 33–36. DOI: https://doi.org/10.1002/hbe2.117.

[5] Stuart J. Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach, Global Edition*. 4th. Pearson Education Limited, 2021. ISBN: 9781292401133.

[6] Valentino Braitenberg. *Vehicles: Experiments in synthetic psychology*. MIT press, 1986.

[7] Rodney Brooks. “A robust layered control system for a mobile robot”. In: *IEEE journal on robotics and automation* 2.1 (2003), pp. 14–23.

[8] Rodney A Brooks. “Intelligence without reason”. In: *The artificial life route to artificial intelligence*. Routledge, 2018, pp. 25–81.

[9] Rodney A. Brooks. “Elephants don’t play chess”. In: *Robotics and Autonomous Systems* 6.1 (1990). Designing Autonomous Agents, pp. 3–15. ISSN: 0921-8890.

[10] Yutaka Matsuo et al. “Deep learning, reinforcement learning, and world models”. In: *Neural Networks* 152 (2022), pp. 267–275. ISSN: 0893-6080.

[11] Yann LeCun. *Objective-Driven AI: A Path Towards Autonomous Intelligence*. Harvard CMSA Distinguished Lecture, Center of Mathematical Sciences and Applications. Mar. 2024. URL: https://cmsa.fas.harvard.edu/media/lecun-20240328-harvard_reduced.pdf.

[12] Davide Brugali et al. “Trends in Robotic Software Frameworks”. In: vol. 30. Apr. 2007, pp. 259–266. ISBN: 978-3-540-68949-2.

[13] David Kortenkamp, Reid Simmons, and Davide Brugali. “Robotic Systems Architectures and Programming”. In: *Springer Handbook of Robotics*. Ed. by Bruno Siciliano and Oussama Khatib. Cham: Springer International Publishing, 2016, pp. 283–306. ISBN: 978-3-319-32552-1.

[14] Konstantina Christakopoulou, Shibl Mourad, and Maja Matarić. *Agents Thinking Fast and Slow: A Talker-Reasoner Architecture*. 2024. arXiv: 2410.08328 [cs.AI]. URL: https://arxiv.org/abs/2410.08328.

[15] Rodrigo Bernardo, João M.C. Sousa, and Paulo J.S. Gonçalves. “Ontological framework for high-level task replanning for autonomous robotic systems”. In: *Robotics and Autonomous Systems* 184 (2025), p. 104861. ISSN: 0921-8890.

[16] Julius Schöning, Adrian Riechmann, and Hans-Jürgen Pfisterer. “AI for closed-loop control systems: New opportunities for modeling, designing, and tuning control systems”. In: *Proceedings of the 2022 14th International Conference on Machine Learning and Computing*. 2022, pp. 318–323.

[17] “Innovative Robot Control Architectures: How to Modify and Enhance Industrial Controllers”. In: *Proceedings of the ICRA 2010 Workshop on Innovative Robot Control Architectures*. Workshop Proceedings. Anchorage, Alaska, USA, 2010.

[18] S.R. White et al. “An architectural approach to autonomic computing”. In: *International Conference on Autonomic Computing, 2004. Proceedings.* 2004, pp. 2–9.

[19] Imane Ettahiri and Karim Doumi. “A novel dynamic enterprise architecture model: leveraging MAPE-K loop and case-based reasoning for context awareness”. In: *International Journal of Electrical and Computer Engineering (IJECE)* 14 (Feb. 2024), pp. 1875–1889.

[20] Eric Rutten, Nicolas Marchand, and Daniel Simon. “Feedback Control as MAPE-K Loop in Autonomic Computing”. In: ed. by Rogério de Lemos et al. Springer International Publishing, 2017, pp. 349–373. ISBN: 978-3-319-74183-3.

[21] Dahlia Fernandez et al. “Robotic process automation: bibliometric reflection and future opportunities”. In: *European Journal of Innovation Management* 27.2 (2024), pp. 692–712.

[22] Pavitra Dhamija and Surajit Bag. “Role of artificial intelligence in operations environment: a review and bibliometric analysis”. In: *The TQM Journal* 32.4 (2020), pp. 869–896.

[23] Simona Casini et al. “Human-Centered AI and Autonomy in Robotics: Insights from a Bibliometric Study”. In: *2025 International Joint Conference on Neural Networks (IJCNN)*. 2025, pp. 1–8. DOI: 10.1109/IJCNN64981.2025.11227890.

[24] Erann Gat. “On Three-Layer Architectures”. In: 1997. URL: https://api.semanticscholar.org/CorpusID:7980015.

[25] Steven Macenski et al. “The Marathon 2: A Navigation System”. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2020.

[26] Adam Paszke et al. “PyTorch: An Imperative Style, High-Performance Deep Learning Library”. In: *Advances in Neural Information Processing Systems*. Ed. by H. Wallach et al. Vol. 32. Curran Associates, Inc., 2019.

[27] J. C. Blandón Andrade et al. “Computational Methods for Information Processing from Natural Language Complaint Processes—A Systematic Review”. In: *Computers* 14.1 (2025). ISSN: 2073-431X.

[28] Nils Reimers and Iryna Gurevych. “Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks”. In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Nov. 2019.