AI Toolbox Concept for the Arrowhead Framework

Dániel Ficzere†, Gergely Hollósi‡, Attila Frankó†&*, Pál Varga‡

†Dept. of Telecommunications and Media Informatics, Faculty of Electrical Engineering and Informatics Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary
‡IoT Division, AITIA International Inc., 48-50 Czetz J. Str. H-1039 Budapest, Hungary
corr.: ficzure.daniel@vik.bme.hu

Abstract—Artificial Intelligence (AI) has become a game-changer across numerous industrial areas, revolutionizing the way businesses operate and enhancing their competitiveness. The Arrowhead Framework, renowned for its service-oriented architecture and interconnectivity principles, presents an ideal platform for the development and deployment of AI-driven solutions for industrial Cyber-physical System of Systems (CPSoS). This paper delves into the formulation of an AI Toolbox that enhances the capabilities of the Arrowhead Framework, aiming to harness the synergies between AI and a robust architectural foundation. The paper presents the main objectives and requirements for the AI Toolbox, and also describes its concept, operation, and deployment principles. For a better understanding, the paper demonstrates how the AI Toolbox works through a generic industrial safety use case. In conclusion, this paper contributes a comprehensive perspective on the formulation of the Arrowhead AI Toolbox, demonstrating how the Arrowhead Framework can offer AI-based services for industrial use cases.

Index Terms—AI, machine learning, artificial intelligence, Arrowhead, industrial AI, edge AI, industrial automation, SOA, intelligent services

I. INTRODUCTION

AI has been present in the industry for several decades now. The AI industry evolved from a few million dollars in the last century to billions of dollars nowadays, including thousands of companies building expert systems, vision systems, robots, software, and hardware specialized for these purposes. Unveiling the early prospects of AI exposed a multitude of obstacles, with the most prominent being the lack of computational power. Computers proved incapable of storing sufficient information or processing it rapidly, presenting a significant challenge to overcome. The development of machine learning and deep learning, along with the emergence of big data, sparked a renewed interest in AI across various domains, captivating companies, investors, governments, the media, and the general public [1].

Even though by 2023 AI assisted industrial applications are spreading rapidly, there are still several challenges ahead with their deployment and integration. The Arrowhead Framework is helping the industrial systems to be designed loosely coupled, however, integrating AI tools into the framework promises a shorter time-to-market in intelligent industrial services.

This paper introduces the vision of an AI Toolbox for the Arrowhead Framework. The overall goal is to improve the design, operation, and maintenance for Cyber-Physical System of Systems (CPSoS) with the help of AI-related technologies. We describe the necessary objectives and requirements for the AI Toolbox based on the state-of-the-art while also considering the smooth integration into the Arrowhead Framework. The paper addresses the possible implementation approaches, including service granularity, lifecycle, and Edge AI questions. The introduced concepts will be explained through a safety-related use case, where the operation of the AI Toolbox will be presented in interaction with the Arrowhead Framework.

This paper is structured as follows. Section II introduces the Arrowhead framework and its main principles. Besides, the role of AI in the industry is highlighted in several application areas. The Arrowhead AI Toolbox vision is introduced in Section III, including the requirements, objectives, architecture and deployment. Section IV guides you through the concepts presented in Section III via a safety use case study, while Section V concludes the paper.

II. RELATED WORKS

A. Arrowhead framework

The Arrowhead architecture incorporates systems, services, and service-oriented architecture principles [2]. The Eclipse Arrowhead framework aims to facilitate the development, deployment, and operation of interconnected, cooperative systems. It is built upon the philosophy of Service Oriented Architecture and consists of a set of mandatory core systems that provide essential service-oriented features such as service registration, discovery, authentication, and authorization. The framework’s building blocks are systems that can both provide and consume services while collaborating as part of larger system of systems. The architectural objective is to facilitate the creation and dynamic operation of self-contained local automation clouds for CPSoS. Every local Arrowhead cloud shall provide the mandatory core services besides the application services which implement the actual business logic [3]. Secure inter-cloud information exchange is also ensured [4].

The services of the core systems – namely the Service Registry, Orchestration, and Authorization (see Figure 1) – are commonly utilized by application systems that adhere to the Arrowhead Framework’s guidelines. The mandatory services ensure the main objectives of the Arrowhead local cloud, namely
Figure 3 illustrates the AI adoption across industries and AI functions in 2022. The highest adoption rate was observed in the risk domain for high tech/telecom, accounting for 38%. It was closely followed by service operations in the consumer goods/retail sector at 31% and product and/or service development in the financial services industry, also at 31%.

Figure 1. High-level overview of Arrowhead framework.

- loose coupling,
- late binding,
- service discovery and
- information assurance.

While the Arrowhead framework provides examples and best practices on how to implement application services, while the Arrowhead implementation gives great freedom to design application services. Also, AI functions in intelligent services can be formed freely and in numerous ways, resulting in heterogeneous services not suitable for cooperation or unified deployment. To ease the implementation of intelligent Arrowhead services and decrease the time-to-market while providing great scalability and reliability, the design of an integrated Arrowhead AI Toolbox is beneficial for both the Arrowhead framework and the several different AI-supported scenarios and use-cases.

B. AI in the industry

In recent years, AI has rapidly transformed the way businesses operate, revolutionizing processes and driving unprecedented levels of efficiency and innovation. The AI Index Report [5] gathers and presents data on artificial intelligence. Its goal is to offer reliable and globally sourced information – besides many others – about AI-related economic and industrial trends. AI integration in the industry has rapidly transformed various sectors, enabling advanced automation, data analysis, and predictive capabilities. Regarding the AI capabilities integrated into at least one function or business unit, as shown in Figure 2, robotic process automation demonstrated the highest embedding rates within high tech/telecom, financial services and business, and legal and professional services industries, reaching 48%, 47%, and 46%, respectively. Across all industries, the most prevalent embedded AI technologies included robotic process automation (39%), computer vision (34%), natural language text understanding (33%), and virtual agents (33%).

Besides the economic interest, there are several researches and surveys which are covering the role of AI in industrial domains and use cases. The study by [6] presents an analysis of AI applications in real manufacturing environments, highlighting key enabling technologies and design principles. They formulate crucial challenges and opportunities for future research, proposing a conceptual framework to facilitate the transition towards a digitized and data-driven culture, encouraging industrial adoption. This paper represents one of the pioneering works, providing a comprehensive understanding of Industrial Artificial Intelligence within the Industry 4.0 landscape, examining its foundational components and emerging trends. Additionally, [7] systematically explores the generation, definition, characteristics, classification, technical system, and current state of Industrial Artificial Intelligence.
(I-AI). Based on existing research and industrial projects, the paper presents a detailed framework and reference model for I-AI implementation in various industries. Besides, there are many other comprehensive studies focusing on the challenges of AI and implementation issues in the industry [8], [9], [10].

AI’s integration in the industry has significantly impacted how businesses interact with clients, partners, and manage supply chains [11], [12], [13]. Besides, AI researches are focusing on inventory management [14], [15], customer feedback [16], [17], [18] and customization [19], [20] also. Leveraging AI-powered solutions has led to enhanced customer experiences, streamlined collaborations with partners, and optimized supply chain operations, ultimately driving overall business success.

Other prominent segment of the industry that utilizes AI are design process [21], [22], planning, and manufacturing [23], [24], [25], [26] processes has ushered in a new era of innovation and efficiency. From concept ideation to production optimization [27], AI-powered solutions are reshaping how industries approach product development, planning, and manufacturing, resulting in improved design quality, streamlined operations, and enhanced productivity. Within the manufacturing sector, AI-driven robotics and automation [28], [29] systems are optimizing production lines, enhancing product quality with AI-supported quality assurance[30], [31] and process monitoring [32], [33], [34].

Furthermore, AI’s integration in maintenance, operation, and recycling processes has revolutionized how industries manage their assets, optimize operations, and promote sustainability. From predictive maintenance [35], [36], [37] to life-cycle management [38], [39], AI-powered solutions are transforming the way businesses approach asset management, reduce downtime, and waste contamination [40], [41]. A relevant survey machine learning applied for smart maintenance reduce downtime, and waste contamination [42]. Network and service optimization is essentially supported by AI-based solutions in this field.

III. ARROWHEAD AI TOOLBOX IN INDUSTRY

A. Objectives & Requirements

While the application of AI has widespread, Arrowhead Framework still lacks of AI services. The paper presents the requirements and different design aspects of the so called AI Toolbox for the Arrowhead Framework. The main purpose of the AI Toolbox is to provide interchangeable, re-usable AI services to facilitate the implementation of intelligent services in the Arrowhead Framework. The objectives of the AI Toolbox can be summarised as follows:

- **Decrease the time-to-market** Support fast development of intelligent services. This requires re-usable services and easy deployment.
- **Reliability and security** Since AI tools handle potentially sensible data, the AI Toolbox needs to provide means to prevent unauthorized access to data.
- **Scalability** The computing needs of AI algorithms are so high that scalability of the algorithms is a great concern.
- **Upgradeability** Modern AI algorithms change rapidly, or the existing algorithms learn from new data, so the state-of-the-art models must be continuously deployed, especially in the Edge AI paradigm.

While Arrowhead Framework specifies a couple of lightweight requirements against application services, like loose coupling, and late binding [2], it states no rigid bounds to implement application services. However, for AI services, some additional requirements can be stated based on the objectives of the AI Toolbox presented earlier.

- **Re-usability** To decrease development time, AI services should be implemented in a specific granularity which makes the services general enough to be utilized in different use cases but specific enough to implement complex tasks.
- **Composeability** Rapid development of industrial use cases requires that the AI services use common data models and interfaces to combine them to implement more complex tasks easily.
- **Upgradeability** Changes in objectives or learning from new data result in new models, which should be deployed seamlessly and easily.
- **Heterogeneity** Heterogeneity should be required in multiple senses against AI services. Heterogeneity in platforms, implementations, and all other factors related to reliability shall be considered while designing AI services.

B. **AI Service scale levels**

The implementation of AI services to extend the Eclipse Arrowhead framework has different scales or entry levels. There are a couple of trade-offs between different levels, regarding scalability, granularity, re-usability, and development time. E.g. a specific use-case can be implemented using different existing AI solutions – e.g., Tensorflow models – as an Arrowhead service. However, this service will provide only functions to fulfill a specific use case, preventing re-usability. Also, providing clear and standard services ease the problem of vendor lock-in phenomena by providing interchangeable services. The different AI Toolbox scale levels can be seen in Figure 4, where the concept of the Arrowhead AI Toolbox is depicted. There are four levels of AI Service scale, differing primarily in granularity. The three implementation based levels are supported by the fourth – mostly theoretic – methodology scale. In the following, the four levels are presented more deeply and compared to the objectives and requirements.

1) **Application scale:** Application scale is the highest scale level of AI services. Application scale means that AI services provide solutions for complete use cases, e.g., a whole predictive maintenance service specifically for railway switches. The service is implemented as an Arrowhead application service and can be accessed through a specified, custom interface. However, services at this scale can hardly fulfill the requirements for re-usability since every use case needs its own service to be implemented using standard tools. Of course, there can be quite general use cases – such as geofencing –
where it is easier to transform the input data to according to the application service interface requirements and use the Application scale implementation rather to implement the whole application.

2) Tools scale: Tools scale is a compromise between algorithmic scale and app scale. Tools basically provide applied AI methods, which are commonly one or more AI models combined with some inner logic to solve a recurring problem. E.g., localization of assets in a factory consists of a couple of small steps, but the problem arises in numerous use cases. Tools scale can provide excellent reusable scale and support rapid development. Also, careful design of service interfaces can provide good composability.

3) Algorithmic scale: The level of Algorithmic scale is granular since well-known AI and ML algorithms are implemented as AI services. Algorithmic scale is mostly about choosing the best service interfaces to provide composable services. However, using pure algorithms – while being highly reusable – requires great effort to implement complex use cases. Also, using pure AI algorithms requires great knowledge of AI methods and techniques, which is against rapid development.

4) Methodology scale: Being on a theoretical scale, the methodology provides specifications, requirements, and best practices for implementing AI services. However, while the Methodology scale can also be considered an independent level of AI service scales, it is best to imagine it as a basis of all other service scale levels since it can provide means to fulfill all the objectives and requirements of AI Toolbox services.

C. AI Toolbox lifecycle

The lifecycle of AI services shares the very same four steps (see Figure 4). To provide re-usable services, the customization step makes it possible to tailor the service to special needs. This includes e.g., the settings of the parameters. The next step is optional, however, most algorithm requires training to be able to perform specific tasks, e.g., detecting uncommon objects in an image. Training can be complex and shall support widely spread frameworks to efficiently accomplish fine-tuned models. Deployment is a crucial point in the lifecycle, since models require variable resources or perhaps massive parallelization. Also, real-time upgradeability needs the necessary
tools to replace old models with new models on-the-fly. After deployment, the AI service is assessed to prove it fulfills its function, meeting the expected metrics in the application context. The lifecycle might start from the beginning, or it can be periodic, as models are fine-tuned from time to time.

D. AI Toolbox implementation

While having different AI Toolbox scale levels, it does not mean that one of them must be chosen. In the implementation of the AI Toolbox, all scale levels can be implemented side by side to provide different services for different needs. Also, methodology shall be provided at all scale levels to provide integrated and unified interfaces and promote heterogeneity while implementing AI services. However, hierarchical implementation (i.e., task scale uses algorithm scale services) is not always recommended because of performance issues.

The most promising implementation scale is the tools scale: it provides the best re-usability while requiring little knowledge of AI methods. However, to satisfy composability, standardized and clean service interfaces shall be designed. Here, the methodology can provide guidance to best practices of interface development. Also, unified and independently designed interfaces make it possible to create interchangeable services with different implementations – helping to avoid vendor lock-in and ensure reliability.

E. Edge AI concerns

In our discussion Edge AI refers to AI models running on edge infrastructure. State-of-the-art studies show how to create, deploy and operate these AI models, investigating algorithm performance, cost-effectiveness, privacy, reliability, and efficiency challenges [44], [45], [46]. Also, there are recent research on domain-specific Edge AI usage, utilizing AI for mobile networks, especially for future 6G networks [47].

Deng et al. [44] present three „grand challenges” in AI on edge, namely data availability, model selection, and coordination mechanisms. Firstly, securing usable raw training data is crucial, requiring incentives for data provision. Model selection faces complexity in determining accuracy thresholds and resource allocation. Lastly, coordinating heterogeneous edge devices for uniform AI learning necessitates flexible mechanisms across hardware and middleware layers.

State-of-the-art Edge AI solutions pose challenges to the design of the Arrowhead AI Toolbox, namely, three following concerns can be formulated. First, Edge AI solutions extensively support real-time on-the-fly model deployment to provide the most up-to-date AI models at the edge. These models are mainly trained in the cloud but deployed onto edge devices automatically through numerous available platforms. However, Arrowhead lacks these features yet, and it is not decided how to implement on-the-fly model upgrade. Second, edge devices have limited resources and use special platforms to run AI models. Also, training and deployment of edge AI models split between the cloud and the edge device. While Arrowhead supports devices with limited capabilities, interoperability with edge AI platforms shall be investigated. Third, distributed learning and – in particular – federated learning are specific to Edge AI. Distributed and federated learning helps to employ the resources of numerous edge devices, converting limited resources into a serious computing capability. Especially federated learning also helps to realize the security objective – no other device gets information from data handled by edge devices. However, the implementation of distributed learning is also an area to be investigated in Arrowhead Framework.

IV. CASE STUDY VIA AUTHORIZATION DECISION IN A RESTRICTED AREA

In this section, we introduce a theoretical industrial use case and describe a possible solution with the help of Arrowhead AI Toolbox System on the different granularity scales. The problem is to alert the system if any unauthorized person or object (forklift, automated guided vehicle, etc.) enters a previously specified restricted area (see Figure 5). The high-level description of the proposed solution for this problem is described by Algorithm 1. The solution utilizes several AI and non-AI elements, including object detection, identification, authorization and geofencing methods. First, the camera stream is analyzed by an object detector (such as YOLOv5), to get the bounding boxes of the objects in the camera images. To project the bounding boxes onto the map of the area to get 2D coordinates, plane detection must first be performed. If the projected object is inside the restricted area, further examination is needed. In the case of a person, face detection and person identification is applied before the system checks if the person is authorized or not. In the case of non-human objects, a different authorization procedure is needed. The output of the process is a service call to alert the system if an unauthorized person or object is in the restricted area.

Algorithm 1

High level description of object detection, identification, and authorization based on visual geofencing in an industrial scenario.

Input:
Video stream
Configuration parameters (Restricted area, Authorization list, person ids, possible object types, etc.)

Output:
List of objects with authorization decision

Procedure

1: Object detection, segmentation, classification (Yolov5)
2: Plane detection
3: Location calculation
4: if Inside the restricted area then
5: if Human then
6: Face detection
7: Person identification (Vanilla CNN)
8: Authorization decision
9: else if Anything else then
10: Authorization decision
11: end if
12: else if Outside the restricted area then
13: Done
14: end if
identify, and authorization parts. The advantage of the solution is that the algorithms are very general, however, implementation of the application takes much more time. A further drawback is that the user has to comply with the interface specification for every used algorithm during the pipeline, which can be challenging and time-consuming in certain cases.

D. The role of the Methodology

As Figure 4 presents, the Methodology is fully orthogonal to the three implementation scales, and always present in any implementation since best practices and design considerations are discussed. However, pure methodology implementation can be imagined as a cookbook (or guide) on how to implement image-based authorization system (or geofencing) in the Arrowhead Framework.

V. CONCLUSION

The paper presented the concept of AI Toolbox for industrial automation applications, by suggesting a common methodology through different implementation scales of AI services compared to the diverse, vendor-custom designs of intelligent services in the Arrowhead Framework. While the emphasis is on the three implementation scales, the fourth theoretical methodology helps to bring together the distinct implementation scales – suggesting unified interfaces and design practices. Based on the concept as a starting point, a couple of future works and research challenges can be identified:

- Research regarding the unified interfaces of AI Toolbox services to provide composable services.
- Enumerating the required AI services in the AI Toolbox in the various Arrowhead Framework use-cases.
- Research on deployment and customization of AI services.
- Research on Edge AI issues, mainly deployment, and federated/distributed learning.
- Recommendations on Arrowhead Framework to implement AI Toolbox-specific services.

Carrying out the tasks in the future can lead to the architecture and implementation of the Arrowhead AI Toolbox which can decrease time-to-market while providing security, reliability, and scalability to industrial services.

ACKNOWLEDGMENT

The research leading to these results is funded by the EU KDT-JU organization under grant agreement 101112089, within the project AIMS5.0 and from the partners’ national programs and funding authorities.

REFERENCES
