Cambridge Large One

www.cambridge.org

Robotic Safe Adaptation In Unprecedented Situations: The RoboSAPIENS Project

Peter G Larsen,[*,1] Shaukat Ali,[2] Roland Behrens,[3] Ana Cavalcanti,[4] Claudio Gomes,[1] Guoyuan Li,[5] Paul De Meulenaere,[6] Mikkel L Olsen,[7] Nikolaos Passalis,[8] Thomas Peyrucain,[9] Jesús Tapia,[10] Anastasios Tefas,[8] and Houxiang Zhang[5]

[1]Aarhus University, Denmark
[2]Simula Research Lab, Norway
[3]Fraunhofer IFF, Germany
[4]University of York, UK
[5]Norwegian University of Science and Technology, Norway
[6]University of Antwerp, Belgium
[7]Danish Technological Institute, Denmark
[8]Aristotle University of Thessaloniki, Greece
[9]PAL Robotics, Spain
[10]ISDI Accelerator, Spain

Article-type

Keywords:
Robotics and automatic control, Artificial intelligence, intelligent systems, multi agent systems, Open-ended, uncertainty, adaptation, safety, trustworthiness

*Author for correspondence. Email: pgl@ece.au.dk

Abstract

The robots of tomorrow should be endowed with the ability to adapt to drastic and unpredicted changes in their environment and interactions with humans. Such adaptations, however, cannot be boundless: the robot must stay trustworthy. So, the adaptations should not be just a recovery into a degraded functionality. Instead, they must be true adaptations: the robot must change its behaviour while maintaining or even increasing its expected performance, and staying at least as safe and robust as before. The RoboSAPIENS project will focus on autonomous robotic software adaptations, and will lay the foundations for ensuring that they are carried out in an intrinsically trustworthy, safe, and efficient manner, thereby reconciling open-ended self-adaptation with safety by design. RoboSAPIENS will transform these foundations into 'first time right'-design tools and platforms, and will validate and demonstrate them.

Introduction

Whenever autonomy is introduced in physical systems that can potentially harm the environment, including humans, it is essential to provide the necessary evidence to assure the safety. Different standards are used in different domains to ensure the trustworthiness of such autonomous systems. The area of robotics is governed by what is called the machinery directive[1]. One requirement in the machinery directive prevents any robot that includes any learned element in its control system from being legally used. We believe that this requirement is too strict: our hypothesis is that in some cases it is possible to provide the necessary safety evidence. Our goal is to prove this hypothesis. To achieve this overall goal, the RoboSAPIENS project[2] will extend the state-of-the-art by pursuing four main objectives:

1. Enable robotic open-ended self-adaptation in response to unprecedented system structural and environmental changes;
2. Advance safety-engineering techniques to assure robotic safety not only before, but also during and after adaptation;
3. Advance deep learning techniques to actively reduce uncertainty in robotic self-adaptation;
4. Assure trustworthiness of systems that use both deep-learning and computational architectures for robotic self-adaptation.

To achieve these objectives, RoboSAPIENS will extend techniques such as MAPE-K (Monitor, Analyse, Plan, Execute, Knowledge) (Kephart and Chess 2003) (see Figure 1) and Deep Learning (DL) to set up generic adaptation procedures, including also for the Social Sciences and Humanities (SSH) dimension of a robotic system. RoboSAPIENS will

1. https://www.europarl.europa.eu/RegData/etudes/BRIE/2022/733576/EPRS_BRI(2022)733576_EN.pdf
2. The RoboSAPIENS project started January 2024 so, naturally, there are not many research results to report here. Instead this is an illustration for what to do enable the desired level of autonomy for robots while keeping the overall safety.

CAMBRIDGE
UNIVERSITY PRESS

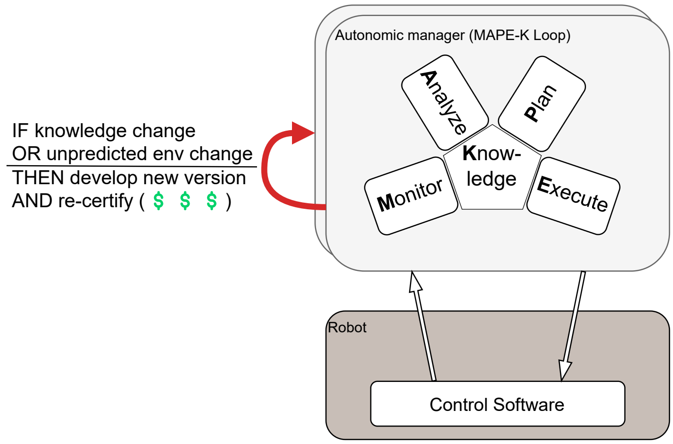

Figure 1. MAPE-K loop in an autonomic element.

demonstrate a novel approach to trustworthy robotic self-adaptation on four industry-scale use cases: an industrial disassembly robot, a warehouse robotic swarm, a prolonged hull of an autonomous vessel, and an application that requires interaction between humans and robots.

This article is a first response to the question: “How to ensure safety of learning-enabled cyber-physical systems?” (Paoletti and Woodcock 2023). This is accomplished by (see Figure 2): 1) adding an additional *Legitimate step* (validate and verify) of the safety of the suggested plan in a MAPE-K context (to become MAPLE-K); 2) adding a run-time trustworthiness checker to the actual robotics controller; and 3) establishing “continous” communication between the autonomic manager and the physical robot.

After this introduction

an overview of the envisaged RoboSAPIENS approach is presented. This is followed by a description of a small academic case study and four industrial scale case studies tested with the RoboSAPIENS technology. Finally, the paper is concluded with looking into what research will be conducted in the future in the framework of the RoboSAPIENS project.

The RoboSAPIENS Approach

To reconcile the opposite requirements for open-ended self-adaptation on the one hand, and safe and trustworthy behaviour of robotic systems even in circumstances not considered at design time on the other, RoboSAPIENS will provide the following extensions to the

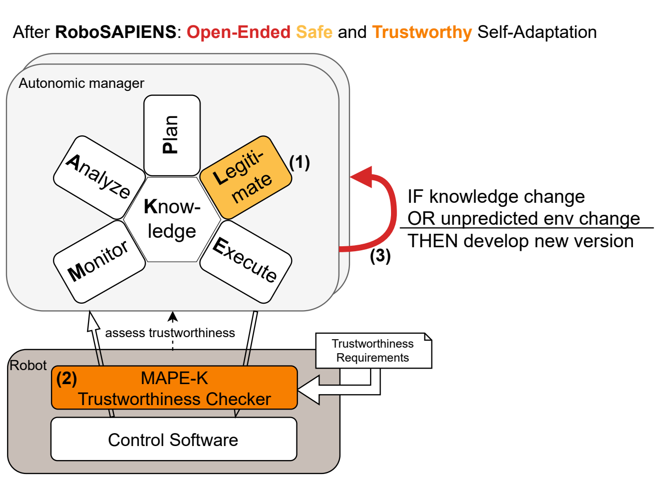

Figure 2. RoboSAPIENS impact in yielding robotic systems with advanced capabilities.

MAPE-K loop (see Figure 2):

(1) To “guarantee” the **safety and trustworthiness** of a self-adapting robot, RoboSAPIENS will add a Legitimate step (including validation and verification) to the MAPE-K loop (adjusting it to become a MAPLE-K loop)[3]. After the Monitor has detected a change in the robot or its environment, after having Analysed it, and after having Planned possible adaptations, the new **Legitimate** step will validate and verify whether all expected functionality can still be met safely (under the explicit assumptions mentions and taking the uncertainties into account). This includes not only a priori defined performance expectations (such as, correct execution of tasks, accuracy, velocity, etc), but also safety and other trustworthiness requirements. For these validation and verification tasks, experiments need to be conducted. Therefore, RoboSAPIENS will rely on a digital twin capability to conduct virtual experiments, and on real experiments (semi-)automatically defined by the Legitimate step and conducted on the robot itself.

(2) A second addition to achieve **open-ended, safe, and trustworthy** self-adaptation, will be the **MAPLE-K Trustworthiness Checker** also explicitly checking the assumptions. Any interaction between the MAPLE-K Loop and the managed robot must pass via this checker, at least for changes initiated to reduce knowledge uncertainty. For example, the Analyser may not have sufficient data to conclude with certainty the cause of an anomaly. So, the MAPLE-K loop may request the robot to execute sufficiently exploratory experiments to enable further analysis of the assumed change. The execution of such experiments may only be done under safe conditions and the results from such experiment should be trustworthy as well. For example, in the ship motion prediction case study (see below), the ship needs to be driven in a zig-zag path, to gather sufficient data to perform the self-adaptation. Such experiment can only be conducted with sufficient clearance of nearby objects and structures. At first one may think that when a plan already has been verified in the Legitimate step there is no need to have such an additional trustworthiness checker but the assumptions taken into account in the Legitimate step could still be wrong. Thus, we have opted for including this because the knowledge about the situation the robot is in may be different than the perception reached from the sensors.

The MAPLE-K Trustworthiness Checker, therefore, contains a set of monitors to check whether elementary trustworthiness rules are respected under all circumstances, according to the domain’s trustworthiness requirements. One of the fundamental problems solved by the trustworthiness checker is: how can it be established that the relevant verification and validation activities have been carried out by the MAPLE-K Loop? For this purpose, the trustworthiness checker can rely on the partial observations of its interactions with the managed element, on historical data, on the use of models, as well as on the presentation of verification certificates. RoboSAPIENS will apply formal verification methods to accurately delineate the safe operation boundaries of the robot based on the readily available information. It is expected that this will be closely related to Run-Time Verification techniques (Falcone, Havelund, and Reger 2013).

(3) To achieve true **self-adaptation**, i.e., to deal with a broad range of unforeseen environments and structural changes, RoboSAPIENS will rely on two complementary solutions. The first is DL as a powerful self-adaptation technique. This takes place in the Planner, and it is expected that it will open up a plethora of robot adaptation possibilities. Nevertheless, it remains possible that some of the proposed changes are disapproved by the Legitimate step or deemed not trustworthy by the MAPLE-K Trustworthiness Checker. Therefore RoboSAPIENS will foresee the possibility of manual version updates of the autonomic manager. Besides validation failures, this manual update can also be applied in case of updates to the Knowledge base (such as the addition of new robot or human models).

3. The reason for writing “guarantee” in quotes is that there are various single point of failure situations that we cannot solve with the RoboSAPIENS solution, since if the sensors provide wrong information the perception will be incorrect.

Trustworthiness and Safety Assurance

A key aspect in an autonomic manager is its knowledge about the managed element and the world. Based on that knowledge, the MAPE-K (or MAPLE-K) loop monitors the managed element and its environment, including humans, and, when an anomaly is detected, constructs and executes plans based on the data gathered about the anomaly. In Figure 2 our suggested adjustment is sketched where an additional step is included between the plan and the execute elements. This step is indicated as Legitimate and consider this extension of the conventional MAPE-K architecture.

Trustworthiness in the context of RoboSAPIENS refers to the degree to which robots featuring the MAPLE-K architecture are perceived as robust, safe, and capable of performing tasks as expected during runtime. This includes their compliance to ethical or legal boundaries and their inability to cause harm to humans, living creatures, or the environment. The concept entails the following aspects that will be integrated into the RoboSAPIENS' MAPLE-K loop as internalised norms that are tightly linked to the ethics guidelines for trustworthy Artificial Intelligence (AI) of the European Union[4]. A second extension in our proposal is also visible in Figure 2 as a trustworthiness checker connected directly to the control software.

Levels of Adaptivity

Robot self-adaptation has been thoroughly studied, with different techniques and processes proposed to calculate control actuation following changes in a robot's environment, either predicted or monitored, to secure better customisation and performance. However, only a few attempts consider structural and functional changes, where functionality or hardware are upgraded or newly integrated (Alattas, Patel, and Sobh 2019; Silva et al. 2016). Evolutionary robotics has been introduced as a discipline to design and study autonomous adaptive modular robots (Alattas, Patel, and Sobh 2019; Tolley, Hiller, and Lipson 2011). Structural and functional changes to the robot add an extra dimension to the design complexity of self-adaptive robots, so that the adaptation space can be exponential with respect to the size of the newly added functionality. This makes safety verification and validation immensely challenging (White et al. 2005; Auerbach et al. 2014) and that is exactly what RoboSAPIENS targets to improve.

Correctness of Techniques

Across Europe, there are significant efforts to adapt and enhance modern Software Engineering techniques to robotics (Cavalcanti et al. 2021), including the application of formal, mathematically based approaches (Luckcuck et al. 2019).

A key part of a robust software development is the adoption of a robust architecture (Ahmad and Babar 2016). There are many more for robotic applications and many proposed architectures (Siciliano and Khatib 2016, Chap. 12). There are, however, no clear definitions of these architectures, and certainly no formalisation. In terms of formal approaches, the focus is on specific aspects of a system or even of just a component: reaction, time, neural network, uncertainty, or planning, for instance. This is particularly true for the verification of neural networks including DL: the techniques and tools are concerned with proofs of properties defined with respect to mathematical definitions of the input or output space, rather than system-level properties.

For the MAPE-K architecture, probabilistic model checking based on Markov chains to capture knowledge has been extensively used to improve the Analysis and Knowledge components (Fang et al. 2022). For runtime verification, where a software monitor is deployed that checks the system behaviour against a specification (Bartocci et al. 2018), a system approach is naturally adopted and can handle collections of adaptive systems (Calinescu, Gerasimou, and Banks 2015); existing work relies on the definition of mathematical models by hand and does not support for DL (Calisnescu et al. 2012). Formal techniques are popular in handling uncertainty (Hezavehi et al. 2021). The approach presented in this paper makes use of formal techniques in order to ensure trustworthiness and safety concerns in the new L element of the suggested MAPLE-K approach.

4. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

Deep Learning

Attempts to bridge the gap between perception and action have been made recently; active perception is a prominent example (Bajcsy, Aloimonos, and Tsotsos 2018; Tosidis, Passalis, and Tefas 2022). DL is also gradually shifting away from the traditional static training paradigm and delving into continual learning (De Lange et al. 2021), wherein DL models are designed to be capable of adapting as they receive more training data.

Several difficulties arise in continual learning and adaptation setups, such as catastrophic forgetting (Kemker et al. 2018), which can significantly deteriorate the performance of models if countermeasures are not taken. Anomaly detection methods (Pang et al. 2021), which are capable of identifying situations that have not been encountered in the past, have also seen significant advances. However, despite the progress in the aforementioned areas, little work has been done on developing complete self-adaptive pipelines on top of DL models, as also seen for traditional Machine Learning approaches (Saputri and Lee 2020). The approach presented in this article builds on the existing attempts of using DL in an autonomous robot setting without the need to re-certify the robot.

Active Uncertainty Reduction

Uncertainty quantification for DL models helps ensure their decisions' trustworthiness. To this end, there are two mainstream approaches: Bayesian and ensemble-based (Abdar et al. 2021), which have been applied to various tasks, e.g., medical imaging and natural language processing. Related to self-adaptive systems, recent works (Catak, Yue, and Ali 2021, 2022) propose a novel uncertainty quantification metric for DL models specifically trained for object detection in the context of self-driving cars. This metric was used to quantify the uncertainty in a DL model to evaluate the prediction's reliability, which was then improved by retraining. These works focus on classification tasks, and have not been used to quantify the uncertainty of embedded DL models in self-adaptive systems. Instead, the data produced was used to train DL models for uncertainty quantification. In the RoboSAPIENS approach it is targeted to provide "guarantees" in the presence of uncertainties, and propose methodologies for actively trying to reduce uncertainty and increase trustworthiness. This is to be used both inside the L part of the MAPLE-K loop, as well as inside the Trustworthiness checker.

The RoboSAPIENS Case Studies

This section starts with introducing an academic case study to demonstrate the proof of concept of the RoboSAPIENS approach. Afterwards, four industrial-scale case studies from RoboSAPIENS are described.

An Academic Case Study

A small academic case study based on a TurtleBot 4 has been defined. This will be used to illustrate the different RoboSAPIENS technologies as it is being developed and to be used in subsequent publications.

TurtleBot 4 is an open-source robotics platform designed for education and research. It comes equipped with an iRobot® Create3 mobile base, a Raspberry Pi 4 running ROS 2, an OAK-D spatial AI stereo camera, and a 2D LiDAR.

The robot, without any support from the MAPLE-K, should be able to autonomously navigate an unknown map using SLAM and a planner (referred to as the local planner to distinguish from the MAPLE-K Loop planner). Additionally, it must estimate the remaining useful life of the battery, assuming that the map floor is uniform.

With RoboSAPIENS technology, the aim is to demonstrate how this navigation can be improved for example to handle the following anomalies:

- Non-uniform floors, which cause the robot to consume more energy in certain areas.
- Partial obstruction of the LIDAR sensor.
- High vibration zones that should be avoided when the robot is carrying a load (to be implemented later as a demonstration of MAPLE-K continuous delivery).

To achieve these improvements, RoboSAPIENS will implement a MAPLE-K loop that complements the robot's local planner through multiple extension points. For example, rewards and punishments can be provided to influence the local planner's decision-making. Additionally, the sensor data accessible to the local planner can be modified by the MAPLE-K loop to enhance map information.

Regarding trustworthiness and safety, RoboSAPIENS envisions conducting formal verification on the local planner offline, covering a wide range of operational scenarios (though not necessarily the adaptations provided by the MAPLE-K loop). This verification will serve as the foundation for runtime verification during the validation of MAPLE-K loop activities. The trustworthiness checker will ensure that the MAPLE-K adheres to the best practices of mobile robots, and the legitimate block will employ simulation and model checking for validating new robot configurations.

Robotic remanufacturing

This case study, provided by the Danish Technological Institute (DTI), focuses on the remanufacturing process, where used products are repaired and restored to a like-new condition, maintaining the same quality, performance, and warranty. The remanufacturing process involves six steps: disassembly, cleaning, inspection, restoration, reassembly, and testing. This study emphasises the disassembly task, which is often the most time-consuming and labor-intensive phase. Traditionally, manual work is required for complex disassembly tasks involving high levels of uncertainty (Vongbunyong, Kara, and Pagnucco 2013). Tasks such as unscrewing, un-snap fitting, and destructive disassembly demand precision and adaptive control. While collaborative robots can be programmed by demonstration, their effectiveness highly depends on the task type and the expertise of the demonstrator. These robots are efficient for repetitive tasks but struggle with tasks requiring force-based control to compensate for inaccuracies.

The RoboSAPIENS project aims to bridge the gap between labor-intensive remanufacturing and adaptable robotic automation using the MAPLE-K framework. This technology enables robots to adapt to new and unforeseen situations while ensuring safety and trustworthiness.

In this context, the MAPLE-K framework is employed to enhance the adaptability and efficiency of robotic disassembly. The robot continuously monitors its environment and the state of the disassembly process using sensors and cameras. Upon detecting an anomaly, such as a difficult-to-remove screw, the system analyses the situation to determine the cause of the failure, leveraging historical data and real-time sensor inputs. Based on this analysis, the robot formulates a new plan to address the detected issue, such as switching tools or adjusting its force application strategy.

Before executing the new plan, the system validates and verifies it through simulations. This step ensures that the new plan will not compromise safety or performance. The validated plan is then executed by the robot, which adapts its behaviour in real-time to successfully complete the disassembly task, maintaining overall efficiency and safety. The outcomes of the executed plan are recorded and added to the system's knowledge base, enhancing future adaptations and sharing knowledge across different robots to improve their performance.

A demonstration of this use case will be set up at DTI's lab, involving a robot cell designed to disassemble electronic consumer waste, such as laptops. The demonstration will showcase the robot's ability to handle complex manipulations and adapt to unforeseen challenges using the MAPLE-K framework.

Autonomous Mobile Robots on Manufacturing Floor

Automated Guided Vehicles (AGVs) operating on shop floors are ad-hoc machines that require specific distribution and means of transport. Advancement towards Industry 4.0, however, calls for the use of Autonomous Mobile Robots (AMRs), a more versatile and affordable option than AGVs, consisting of robots equipped with a mobile base and even robotic arms, allowing them to autonomously navigate and perform dexterous tasks without the support of additional physical equipment. It is envisioned that these robots will be deployed as a fleet on the shop floor, able to navigate freely and safely, while taking into account changes in the fleet and the surroundings (e.g. change to the number of robots and blockages by humans) based on self and environmental awareness. RoboSAPIENS will provide a solution to dynamically adapt the work assigned to each member of the fleet and the navigation through paths when such changes occur. Such adaptation will take dynamic parameters into account, such as disconnected robots, battery status, proximity to goals, past human behaviour, etc.

The case study will use a fleet of robots from the TIAGo family developed by PAL Robotics, the TIAGo OMNI Base. This mobile base is equipped with omnidirectional mecanum wheels that allow the robot to move in any direction, two LIDAR sensors for an unobstructed 360° FOV and 2 depth camera to complement the other sensors and detect stairs, tables, etc. The scenario involve these robots set in a shop floor, controlled via a fleet management system. During their operation, one or more robots may come and go (e.g. due to low battery), communication between robots and the fleet management may drop, emergency exits may be blocked (due to stopped or malfunctioning robots cutting supply chains and endangering humans), or the floor plan itself may change.

Such anomalies will trigger the MAPLE-K at the fleet level, and the state of the fleet and the environment will be re-evaluated. The TIAGo robots are capable of Simultaneous Localisation and Mapping (SLAM), and their sensor readings are used to update the map and inform the fleet manager. The planning phase is carried out by adopting a genetic algorithm to reschedule tasks and paths of the robots. After the system is validated through simulation, and the self-adaptation process is deemed trustworthy, the model is deployed to the fleet manager.

At the robot level another MAPLE-K loop will be integrated. It will be a human tracker based on the sensing capabilities of the robot platform. The robot will be able to adapt its path and avoid humans at a socially acceptable distance while keeping track of the uncertainty of the human switching predicted path and crossing the robot planned path. Via RoboSAPIENS legitimate capability the new plan is then assessed and if it is decided as trustworthy, the updated path is then executed by the TIAGo OMNI base.

Autonomous Ship Motion Prediction

Estimating the motion of a ship in the immediate future, either from a dynamic model, or a data-driven one using adequate historical data, could support autopilots and thus improve the safety of autonomous ships. However, deploying the prediction system to new ships without sufficient prior knowledge of their dynamic behaviour deteriorates navigation capability, especially in the presence of environmental uncertainties such as wind, currents, and waves. Identifying model parameters via sea trials or collecting the needed data for ship motion modelling will take a relatively long time. In this case study, RoboSAPIENS leverages the dynamic model from a reference ship and the limited available data from the target ship to build up a transferrable model that can represent the target ship motion.

RoboSAPIENS will use the Norwegian University of Science and Technology (NTNU)'s Gunnerus research vessel as a case study. Gunnerus has gone through a thruster refit in 2015 and been extended by 5m in length in 2018. While there is a high-precision dynamic model of the original vessel, it cannot directly be used for the longer vessel, for which there are limited data available. In such a context, three objectives are considered in this case study, from dynamic system identification, to transfer learning of identified systems, to online model adaptation. RoboSAPIENS will first obtain a rough dynamic model of the longer vessel based on the dynamic model of the original vessel and then apply DL to the longer vessel, by combining the rough dynamic model with the limited real-motion data to generate a ship predictor.

In the MAPLE-K loop, a motion calibrator will be created based on the motion discrepancy from the hybrid predictor and real data, and further incorporated into that predictor for motion prediction. When the ship's motion predictor underperforms a monitor is triggered. Data is recorded from the trigger time to a predefined later time for generation of a new dataset in runtime, at the aims of analysing the main factor of prediction error in the analyse phase and updating the transferred prediction model trained using DL in the plan phase. If a better prediction performance is validated via RoboSAPIENS legitimate capability and it is deemed trustworthy, the updated model is deployed and executed, otherwise the system goes back to the plan phase. RoboSAPIENS will investigate what a suitable amount of data is needed for the transferable model, the impact on the prediction performance, and the generalisation of the transfer modelling.

Dynamic Risk Model for Cobots in Industry 4.0

Risk assessment is a mandatory procedure in human-robot interaction for cobots. It is an iterative process that systematically identifies hazards and specifies measures to reduce these hazards' probability. The procedure and requirements are specified in the Machinery Directive 2006/42/EC and harmonised safety standards.

The current manually operated and strongly heuristic practice contradicts the paradigm of Industry 4.0. Ignoring data during the risk assessment leads to a loss of efficiency in safety engineering and, most importantly, an unnecessarily decreased robot productivity. This is particularly evident in production systems featuring human-robot collaboration, where people and machines work closely together. In this case study, RoboSAPIENS will use system and sensor data in a dynamic human-robot safety model to automatically and continuously assess the risk, to improve the overall production system's efficiency, and to significantly reduce the costs associated with risk assessment.

An experimental production line will serve to test the benefits of the MAPLE-K loop technology in an industrial setting. The production line includes human workers, mobile platforms, a collaborative robot and various safety sensors that monitor positions, movements and states of human workers. The data from the safety sensors and the digital twins of the robots will be continuously analysed for incomplete data and changes, such as those that can occur when a human abruptly moves in another direction. In case of such abnormalities, the robots' motions and activities will be newly planned. The planning result will then be legitimated in simulations under worst-case conditions. Once this part of the MAPLE-K loop has evidently concluded that the newly planned robot motions will not lead to obvious or additional health risks, the new plan will be transmitted to the production line' management system and there executed if the Trustworthiness Checker confirms that all requirements from applicable standards, laws and other rules are fulfilled.

Thanks to RoboSAPIENS technology based on the MAPLE-K loop architecture, a dynamic risk management will be realised for a production system that includes multiple robots and, of course, freely-moving humans. Whenever a deviation from original assumptions or even abnormalities is detected, the production system will automatically adapt by itself to mitigate current risks. Only if any self-adaptation is deemed trustworthy, the production system will finally implement and execute the measures planned.

Concluding Remarks and Future Work

We believe that RoboSAPIENS to a large extent is set up to answer the research question 'How to ensure safety of learning-enabled cyber-physical systems?" asked in the Cambridge University Press journal called "Research Directions: Cyber-Physical Systems". The RoboSAPIENS focus is naturally autonomous robots but it is expected that some of the research results that will be delivered will be of more general nature. The expectation is that more detailed publications will be published for the RoboSAPIENS technology, initially using the academic case study. Subsequently, it is expected that the usefulness of the conducted research will be demonstrated in the four industrial-scale case studies and separate publications will be made for each of these. We believe that each of these publications will be submitted as follow up papers to the same question.

Acknowledgement

We would like to thank all the researchers and collaborators of the RoboSAPIENS project, who are making the vision expressed here a reality.

Funding Statement The work presented here is partially supported by the RoboSAPIENS project funded by the European Commission's Horizon Europe programme under grant agreement number 101133807. The work of Ana Cavalcanti is funded by the UKRI (UK Research and Innovation Council) under Grants No EP/R025479/1 and EP/V026801/1, and by the Royal Academy of Engineering under Grant No CiET1718/45.

Ethics Statement Ethical approval and consent are not relevant to this article type.

Author Contribution Statement Peter Gorm Larsen is the overall coordinator of the RoboSAPIENS project and he has written the majority of the article here. All the other co-authors have taken an active role in the writing of the RoboSAPIENS proposal and have provided constructive comments and improvements of drafts of the article.

Data Availability Statement Data availability is not applicable to this article as no new data were created or analysed in this study. In the future we expect to use the open data principles in follow up publications.

Competing Interests None

Connections References

Paoletti, Nicola, and Jim Woodcock. 2023. How to ensure safety of learning-enabled cyber-physical systems? Research Directions: Cyber-Physical Systems 1:e2. https://doi.org/10.1017/cbp.2023.2.

Notes

References

Abdar, Moloud, Farhad Pourpanah, Sadiq Hussain, Dana Rezazadegan, Li Liu, Mohammad Ghavamzadeh, Paul Fieguth, et al. 2021. A review of uncertainty quantification in deep learning: Techniques, applications and challenges. *Information Fusion* 76 (December): 243–297. ISSN: 1566-2535, accessed November 15, 2022. https://doi.org/10.1016/j.inffus.2021.05.008.

Ahmad, Aakash, and Muhammad Ali Babar. 2016. Software architectures for robotic systems: a systematic mapping study. *Journal of Systems and Software* 122:16–39.

Alattas, Reem J, Sarosh Patel, and Tarek M Sobh. 2019. Evolutionary modular robotics: Survey and analysis. *Journal of Intelligent & Robotic Systems* 95 (3): 815–828. ISSN: 1573-0409.

Auerbach, Joshua, Deniz Aydin, Andrea Maesani, Przemyslaw Kornatowski, Titus Cieslewski, Grégoire Heitz, Pradeep Fernando, Ilya Loshchilov, Ludovic Daler, and Dario Floreano. 2014. Robogen: Robot generation through artificial evolution. In *ALIFE 14: The Fourteenth International Conference on the Synthesis and Simulation of Living Systems,* 136–137. MIT Press.

Bajcsy, Ruzena, Yiannis Aloimonos, and John K Tsotsos. 2018. Revisiting active perception. *Autonomous Robots* 42 (2): 177–196.

Bartocci, Ezio, Jyotirmoy V. Deshmukh, Alexandre Donzé, Georgios E. Fainekos, Oded Maler, Dejan Nickovic, and Sriram Sankaranarayanan. 2018. Specification-based monitoring of cyber-physical systems: A survey on theory, tools and applications. In *Lectures on runtime verification — introductory and advanced topics,* edited by Ezio Bartocci and Yliès Falcone, 10457:135–175. Lecture Notes in Computer Science. Springer.

Calinescu, Radu, Simos Gerasimou, and Alec Banks. 2015. Self-adaptive software with decentralised control loops. In *Fundamental approaches to software engineering,* edited by Alexander Egyed and Ina Schaefer, 235–251. Springer Berlin Heidelberg.

Calisnescu, Radu, Carlo Ghezzi, Marta Kwiatkowska, and Raffaela Mirandola. 2012. Self-adaptive software needs quantitative verification at runtime. *Communications of ACM* 55 (9): 69–77.

Catak, Ferhat Ozgur, Tao Yue, and Shaukat Ali. 2021. Prediction Surface Uncertainty Quantification in Object Detection Models for Autonomous Driving. In *2021 IEEE International Conference on Artificial Intelligence Testing (AITest),* 93–100. August. https://doi.org/10.1109/AITEST52744.2021.00027.

———. 2022. Uncertainty-aware prediction validator in deep learning models for cyber-physical system data. *ACM Transactions on Software Engineering and Methodology (TOSEM)* 31 (4): 1–31. ISSN: 1049-331X.

Cavalcanti, Ana, Will Barnett, James Baxter, Gustavo Carvalho, Madiel Conserva Filho, Alvaro Miyazawa, Pedro Ribeiro, and Augusto Sampaio. 2021. Robostar technology: a roboticist's toolbox for combined proof, simulation, and testing. In *Software engineering for robotics,* edited by A. L. C. Cavalcanti, B. Dongol, R. Hierons, J. Timmis, and J. C. P. Woodcock, 249–293. Springer International Publishing. https://doi.org/10.1007/978-3-030-66494-7_9.

De Lange, Matthias, Rahaf Aljundi, Marc Masana, Sarah Parisot, Xu Jia, Aleš Leonardis, Gregory Slabaugh, and Tinne Tuytelaars. 2021. A continual learning survey: defying forgetting in classification tasks. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44 (7): 3366–3385.

Falcone, Yliès, Klaus Havelund, and Giles Reger. 2013. A Tutorial on Runtime Verification. *Engineering Dependable Software Systems* 34 (January 1, 2013): 141–175.

Fang, Xinwei, Radu Calinescu, Colin Paterson, and Julie Wilson. 2022. Presto: predicting system-level disruptions through parametric model checking. In *2022 international symposium on software engineering for adaptive and self-managing systems (seams),* 91–97. https://doi.org/10.1145/3524844.3528059.

Hezavehi, Sara M., Danny Weyns, Paris Avgeriou, Radu Calinescu, Raffaela Mirandola, and Diego Perez-Palacin. 2021. Uncertainty in self-adaptive systems: a research community perspective. *ACM Trans. Auton. Adapt. Syst.* 15 (4).

Kemker, Ronald, Marc McClure, Angelina Abitino, Tyler Hayes, and Christopher Kanan. 2018. Measuring catastrophic forgetting in neural networks. In *Proceedings of the aaai conference on artificial intelligence,* vol. 32. 1.

Kephart, Jeffrey O, and David M Chess. 2003. The vision of autonomic computing. *Computer* 36 (1): 41–50.

Luckcuck, Matt, Marie Farrell, Louise A. Dennis, Claire Dixon, and Michael Fisher. 2019. Formal Specification and Verification of Autonomous Robotic Systems: A Survey. *ACM Computing Surveys* 52 (5).

Pang, Guansong, Chunhua Shen, Longbing Cao, and Anton Van Den Hengel. 2021. Deep learning for anomaly detection: a review. *ACM Computing Surveys (CSUR)* 54 (2): 1–38.

Paoletti, Nicola, and Jim Woodcock. 2023. How to ensure safety of learning-enabled cyber-physical systems? *Research Directions: Cyber-Physical Systems* 1:e2. https://doi.org/10.1017/cbp.2023.2.

Saputri, Theresia Ratih Dewi, and Seok-Won Lee. 2020. The application of machine learning in self-adaptive systems: a systematic literature review. *IEEE Access* 8.

Siciliano, Bruno, and Oussama Khatib, eds. 2016. *Springer handbook of robotics.* Springer Handbooks. Springer.

Silva, Fernando, Miguel Duarte, Luís Correia, Sancho Moura Oliveira, and Anders Lyhne Christensen. 2016. Open issues in evolutionary robotics. *Evolutionary computation* 24 (2): 205–236. ISSN: 1063-6560.

Tolley, Michael T, Jonathan D Hiller, and Hod Lipson. 2011. Evolutionary design and assembly planning for stochastic modular robots. In *New Horizons in Evolutionary Robotics,* 211–225. Springer.

Tosidis, Pavlos, Nikolaos Passalis, and Anastasios Tefas. 2022. Active vision control policies for face recognition using deep reinforcement learning.

Vongbunyong, Supachai, Sami Kara, and Maurice Pagnucco. 2013. Application of cognitive robotics in disassembly of products. *CIRP Annals* 62 (1): 31–34.

White, Paul, Viktor Zykov, Josh C Bongard, and Hod Lipson. 2005. Three dimensional stochastic reconfiguration of modular robots. In *Robotics: Science and Systems,* 161–168. Citeseer.