complAI

Collaborative Model-Based Process Assessment for trustworthy AI in Robotic Platforms

Programm / Ausschreibung	Ideen Lab 4.0, Ideen Lab 4.0, Ideen Lab4.0 - Ausschreibung 2019	**Status**	abgeschlossen
Projektstart	01.02.2020	**Projektende**	31.01.2021
Zeitraum	2020 - 2021	**Projektlaufzeit**	12 Monate
Keywords	Geschäftsprozesse, Künstliche Intelligenz, Robotik		

Projektbeschreibung

Ausgangsituation, Problematik Motivation:

Um KI in Organisationen einzuführen, stehen Entscheidungsträger vor Herausforderungen die teilweisen unrealistischen Erwartungshaltungen zu bewerten, den zu erwartenden Nutzen im Vergleich zum erwarteten Aufwand abzuschätzen sowie ein Risiko im Verantwortungsbereich des Entscheidungsträgers abzuschätzen alle rechtlichen-, ethischen- und sicherheitsrelevanten Fragestellungen hinreichend zu berücksichtigen.

Aufgrund der Vielzahl verschiedener KI-Methoden mit unterschiedlichen Charakteristika – neben der präzisen aber aufwendigen symbolischen KI, und der abstrahierten aber flexiblen sub-symbolischen KI wird auch die Kombination mit der menschlichen Wissensrepräsentation im Sinne vom Wissensmanagement als intelligentes System verstanden – ist neben dem speziellen Domainwissen der Organisation auch ein tiefgreifendes Wissen über die KI notwendig um Entscheidungen verantwortungsvoll zu treffen.

Ziele und Innovationsgehalt:

Wir entwickeln deshalb ein Assistenzsystem um Organisationen bei der Verwendung von KI zu unterstützen. Relevante ethische, rechtliche und sicherheitsbezogenen Anforderungen werden durch ein modellbasiertes Risikomanagement mittels Softwaretool beurteilt. Als Sondierungsherausforderung wird die sichere, etische und strafrechtlich korrekt geprüfte Ausführung von Prozessen auf Roboterplattformen sichergestellt und die Anwendbarkeit eines solchen Systems getestet.

Die Roboterplattformen werden aufgrund der eindrucksvollen haptischen Demonstrationsmöglichkeiten gewählt um den Unterschied eines Prozesses mit und ohne KI zu verdeutlichen.

Der Innovationsgehalt ist:

• Modellbasiertes Assistenzsystem zur Entscheidungsunterstützung von KI Anwendungen
• Modellbasierte Ansteuerung von Robotern durch technische und fachliche Prozesse
• Signatur von Prozessen, die durch ein ethisches-, rechtliches- und sicherheitsspezifisches Assessment nachvollziehbar überprüft worden sind
• Die Konzeptualisierung und damit zur Verfügungstellung in einem Assistenzsystem von ethischen-, rechtlichen- und sicherheitsrelevanten Kriterienkatalogen
• Die Interpretation der Abhängigkeit von KI-Methoden und Kriterienkatalog.

Angestrebte Ergebnisse und Erkenntnisse:

• Im Sondierungsprojekt wird für den Anwendungsfall, sichere Robotik ein Modellbasiertes Assessmentsystem prototypisch entwickelt und auf ADOxx.org zur Verfügung gestellt.
• Der Mechanismus wie Prozess mittels Assessmentkriterien beurteilt und zertifiziert werden können wird erprobt.
• Roboterplattformen testen Mechanismen um ausschließlich vertrauensvolle Prozesse abzuarbeiten.
• Beispielmodelle von Assessmentkriterien und den Abhängigkeiten von Anwendungsfall spezifisch ausgewählten KI Methoden werden publiziert.

Erkenntnisse über die Machbarkeit, der Plausibilität sowie der Anwendbarkeit solcher Systeme wird im Rahmen der Demonstration zur Formulierung weiterführender Forschungsfragen verwendet.
Im Zuge dieser Sondierung wird auch das Risiko dieser Forschungsfragen speziell im Bezug auf die Möglichkeit die Assessmentkriterien zu Konzeptualisieren und somit mittels Modelle nutzbar zu machen – insbesondere die Überwindung der semantischen Lücke – bewertet.

Abstract

Starting Point, Challenges and Motivation

In order to introduce AI-based processes in organisations, the respective decision makers faces great challenges with regard to i) appropriately addressing unrealistic expectations, ii) assessing the expected added value compared to the expected effort in building and maintaining AI, and iii) the sufficient assessing the risk of legal, ethical, safety and security issues. Facing a variety of AI methods with different characteristics – the precise but expensive symbolic AI, the abstracted but flexible sub-symbolic AI as well as a combination including the human interpretation in the sense of knowledge management – requires a fundamental expertise about those AI mechanisms, chances and challenges on top of domain expert knowledge to make a traceable, transparent and informed decision.

Goal and Innovation

We develop an assistance system, which is to guide organisations in using AI. Relevant ethical, legal, safety and security issues will be assed in a model-based risk management software tool. For this investigation, we choose the safe, secure, ethical- and legal compliant execution of processes on industrial robotic platform, which perform both through certified and trustworthy processes.

The selection of the respective robotic platforms builds upon their impressive haptic demonstration potentials, which allow the visual demonstration of a certain process with and without AI.

Innovation is identified in:

• Model-based assistance system for decision support in using AI
• Model-based control of robotic platform with technical and domain specific processes
• Using electronic signature for processes to certify that they have been ethically, legally, safety- and security-wise approved.
• Conceptualisation of assessment criteria in order to enable computer interpretation of ethical-, legal-, security- and safety issues.
• Interpretation of assessment criteria and their relation to AI methods via the assessment system.

Targeted Results and Insights

• Model-based assessment system for the investigation of using secure robotic developed on the open platform ADOxx.org
• Mechanisms how processes can be assessed using assessment criteria and, in case of traceable approval, are electronically signed.
• Robotic platform test mechanisms to ensure that exclusively trustworthy processes are executed.
• Sample models that show assessment criteria and their dependencies for application specific selection of AI methods are published.
Insights about feasibility, plausibility and applicability of such systems are used to express more detailed research questions during the demonstration of the prototypes.
The gained insights are further used to assess the risk of those research questions, especially with the possibility to conceptualise assessment criteria by considering the semantic gap that may arise in such transformation, and through this make these both computable and interpretable by software agents by considering the semantic gap that may arise in such transformation.

Endberichtkurzfassung

Das Sondierungsprojekt complAI entwickelt Modell-basierte Ansätze um die sichere Verwendung von KI und Roboter zu vereinfachen

Ergebnis 1: Modellbasiertes Assistenzsystem für sichere Anwendung von Künstlicher Intelligenz für Robotik.

Es wurde prototypisch ein Assistenzsystem entwickelt, das die Erstellung von sogenannten Workflow-Modelle zur Ansteuerung von Roboter Plattformen ermöglicht. Dabei wurde auf dem BPMN Modellierungsstandard aufgesetzt und der BPMN Prozess sowohl für die Beschreibung des Anwendungsfalles sowie zur technischen Ansteuerung der Roboter Plattform verwendet [D3.1]. Ein initiales Set an Workflows für eine „Pick and Place“ Aufgabe eines Roboterarm wurde mit unterschiedlichen Flexibilisierungsgraden – fix definierte Abläufe, Adaption der Abläufe vor der Ausführung, Adaption der Abläufe während der Ausführung –erarbeitet, erprobt und in beschrieben [D3.2] und kann auf ADOxx.org [1] heruntergeladen und weiterentwickelt werden.

Ergebnis 2: Mechanismus wie Prozesse mittels Assessmentkriterien beurteilt und zertifiziert werden können.

Weiters wurde ein Mechanismus erprobt, wie die vorhin modellierten Workflows mittels Kriterien Assessment beurteilt und zertifiziert werden können. Mittels Beispielkriterien wurde die Erstellung eines Fragebogens, die Modellierung vom Bewertungsmechanismus, die Verlinkung zu den vorhin beschrieben Workflows zu Roboter Ansteuerung, sowie die Beurteilung ob eine Freigabe des Workflows empfohlen oder verweigert werde soll. Bei einer Freigabe wird der Workflow digital signiert. Das erprobte Modellierungssystem verwendet BPMN Workflows als Zielobjekte, die es zu bewerten gilt. Fragebogenmodelle können mittels einer neuen Modellierungsmethode manuell aus dem Kriterienkatalog abgeleitet werden. Der Prototype ist auf ADOxx.org [1] zur Verfügung gestellt und wird im [D4.1] beschrieben.

Ergebnis 3: Mechanismen damit Roboterplattformen ausschließlich vertrauensvolle Prozesse abarbeiten.

Eine Simulationsumgebung stellt den identen Software Stack von industriellem Roboter zur Verifizierung von Workflows zur Verfügung. Somit können Workflow Modelle sowohl zur Ansteuerung in der Entwicklungsumgebung bspw. dem Roboter Arm „Dobot Magician“ im BOC §OMiLAB Innovation Corner“ als auch im industriellen Umfeld bspw. unter Verwendung des Roboters „Universal Robotics“ im JRR „Roboter Labor“ verwendet werden. Weiters kann diese Workflow Umgebung auf eine Verifikation zugreifen um zu überprüfen, ob der Workflow digital signiert und somit zertifiziert ist.

Ergebnis 4: Beispielmodelle und Assessmentkriterien sowie die Abhängigkeiten von Anwendungsfall spezifisch ausgewählten KI-Methoden.

Der Kriterienkatalog mit ethischen, rechtlichen und sicherheitsrelevanten Kriterien wurde ausgearbeitet und zur Verfügung gestellt. Dabei werden Kriterien konzeptuelle in Assessmentkriterien mittels eines Modellierungswerkzeuges umgewandelt. Die Zuordnung der Kriterienart – in unserem Fall (a) Ethisch, (b) Rechtlich und (c) Sicherheit – kann auf die unterschiedlichen Abstraktionsebenen – in unserem Fall (a) domänen-spezifische Anwendungsfall, (b) die AI Bearbeitungsmechanismen sowie (c) die technische Roboteransteuerung erfolgen. Das Lessons-Learned ist, dass technische Kriterien Anwendungsfall-unabhängig sein können, rechtliche oder ethische Fragestellungen jedoch zu konkreten Use Cases ausgearbeitet werden sollen.

Aufbereitung der Ergebnisse:

Alle geplanten Prototypen und Modelle können auf ADOxx.org [1] herunterladen werden. Weiters stehen alle Ergebnisse, Dokumentationen und Webinar auf der Projektwebseite [2] zur Verfügung.

[1] https://adoxx.org/live/web/complai/downloads

[2] Webseite: complai.innovation-laboratory.org/

Projektkoordinator

- BOC Asset Management GmbH

Projektpartner

- JOANNEUM RESEARCH Forschungsgesellschaft mbH
- Universität Linz
- Universität Wien