Workshop

Anforderungen an KI-Prüfungen – Standardisierungspotentiale identifizieren

Lena Krieger & Sobhi Mahmoud

11.09.2023

Agenda

10:00 Uhr	**Begrüßung und Motivation der Veranstaltung** (Lena Krieger, Sobhi Mahmoud, DIN)
10:15 Uhr	**Vorstellung des Projektes „Zertifizierte KI“** (Dr. Maximilian Poretschkin, Fraunhofer IAIS)
10:30 Uhr	**Ergebnisse der Normungsroadmap KI** (Jan Rösler, DIN)
10:50 Uhr	**Vorstellung ausgewählter Normungsbedarfe** (Taras Holoyad, Bundesnetzagentur)
11:25 Uhr	**Priorisierung der Normungsbedarfe**
11:30 Uhr	**Pause**
11:40 Uhr	**Diskussion der priorisierten Normungsbedarfe** (alle Teilnehmenden)
12:50 Uhr	**Zusammenfassung der Ergebnisse und weiteres Vorgehen** (Lena Krieger, Sobhi Mahmoud, DIN)
13:00 Uhr	**Ende der Veranstaltung**

Ziele des Workshops

→ Vorstellung und Priorisierung ausgewählter Handlungs- und Normungsbedarfe der Normungsroadmap KI (Version 2) - Arbeitsgruppe „Prüfung und Zertifizierung“

→ Diskussion der zwei am höchsten priorisierten Bedarfe zur Initiierung von Standards

→ Identifizierung interessierter Experten

Dr. Maximilian Poretschkin (Fraunhofer IAIS)

Vorstellung des Projektes „Zertifizierte KI“

Jan Rösler (DIN e. V.)

Ergebnisse der Normungsroadmap KI

Taras Holoyad (Bundesnetzagentur)

Vorstellung ausgewählter Normungsbedarfe

Priorisierung der Handlungsbedarfe

I Operationalisierung der „Erklärgüte“ von XAI-Methoden

a) Anforderungen an XAI-Methoden
b) Ground truth-Referenzdatensätze & Metriken
c) Verifizierung: XAI bei Qualitätssicherung

II Entwicklung von Funktionalitätsklassen für KI-Technologien

a) Definition und Klassifizierung von Funktionalitäten
b) Charakteristiken für Vertrauen in Funktionalitäten
c) Baukasten für Funktionalitätsklassen

III Entwicklung von Werkzeugkriterien für die Prüfung von KI-Systemen

a) Beschreibung von Werkzeugen für KI-Prüfungen
b) Prüfverfahren für Werkzeuge

Ergebnis zur Frage 1: Welcher Bedarf hat die höchste Priorität?

1a)	7/62 (11%)
1b)	9/62 (15%)
1c)	4/62 (6%)
2a)	4/62 (6%)
2b)	5/62 (8%)
2c)	1/62 (2%)
3a)	10/62 (16%)
3b)	4/62 (6%)

Keine Antwort:
18/62 (29%)

Priorisierung der Handlungsbedarfe

I Operationalisierung der „Erklärgüte“ von XAI-Methoden

a) Anforderungen an XAI-Methoden
b) Ground truth-Referenzdatensätze & Metriken
c) Verifizierung: XAI bei Qualitätssicherung

II Entwicklung von Funktionalitätsklassen für KI-Technologien

a) Definition und Klassifizierung von Funktionalitäten
b) Charakteristiken für Vertrauen in Funktionalitäten
c) Baukasten für Funktionalitätsklassen

III Entwicklung von Werkzeugkriterien für die Prüfung von KI-Systemen

a) Beschreibung von Werkzeugen für KI-Prüfungen
b) Prüfverfahren für Werkzeuge

Ergebnis zur Frage 2: Interessenslage zur Mitarbeit

1a)	24/62 (39%)
1b)	15/62 (24%)
1c)	17/62 (27%)
2a)	13/62 (21%)
2b)	15/62 (24%)
2c)	11/62 (18%)
3a)	25/62 (40%)
3b)	23/62 (37%)

Keine Antwort:
23/62 (37%)

Alle Teilnehmenden

Diskussion der priorisierten Normungsbedarfe

1. Ground truth-Referenzdatensätze & Metriken (slides 10 bis 12)
2. Beschreibung von Werkzeugen für KI-Prüfungen (slides 13 und 14)

Ground truth-Referenzdatensätze & Metriken

Welcher Typ DIN SPEC wird benötigt und ist realisierbar?

- **Generisch/**~~sektorspezifisch/use case-spezifisch~~

Wer ist die Zielgruppe dieses Standards?

- Alle Anwender von Ground Truth Referenzdatensätzen

Ground truth-Referenzdatensätze & Metriken

Was soll Inhalt der DIN SPEC sein?

- Eigenschaften und Qualitätsmerkmale der Datensätze, die diese zu der Eigenschaft werden lässt, als Referenz zu dienen; Edge Cases, z.B. CV-HAZOP https://vitro-testing.com/cv-hazop/
- Wie werden die Daten erhoben; wie werden sie gespeichert, insbesondere um Manipulationen zu verhindern?
- Wie beschriebt man eine ODD (operational design domain)?
- Referenzen auf bestehende Normen, z.B. Datenschutz (Recht auf Vergessen, Anonymisierung etc); Kompatibilität mit bestehenden Normen, z.B. Managementsystemnormen
- Traceability
- Definition von Explainability, wie weit fassen wir hier explainability?
- Um welche Datentypen geht es? Bilddaten etc…
- Definieren von Metriken für Datenqualität (ISO/IEC DIS 5259)

Ground truth-Referenzdatensätze & Metriken

Was soll nicht enthalten sein?

-

Was gibt es bereits auf dem Markt und wie unterscheidet sich diese geplante DIN SPEC davon?

- ISO/IEC DIS 5259
- Referenzen auf bestehende Normen, z.B. Datenschutz (Normen aus JTC 1/SC 27)
- ISO 9001; ISO 13485
- ISO/IEC AWI TR 42103 Information technology — Artificial intelligence — Overview of synthetic data in the context of AI systems
- EASA Roadmap 2.0

Beschreibung von Werkzeugen für KI-Prüfungen

Welcher Typ DIN SPEC wird benötigt und ist realisierbar?

- **Generisch**/~~sektorspezifisch/use case-spezifisch~~

Wer ist die Zielgruppe dieses Standards?

Anwender/Entwickler von Werkzeugen für KI-Prüfungen

Was soll Inhalt der DIN SPEC sein?

- ~~Erkennung stehender Gegenstände im Straßenverkehr~~
- ~~Erkennung eines Warnsignals im Straßenverkehr (Störfaktoren vorhanden)~~
- ~~Personalplanung/Ressourcensteuerung/Routingsteuerung~~
- Qualitätskriterien generischer Natur für Werkzeuge;
  - was muss ein Werkzeug leisten können?
  - Wie leistungsfähig ist ein Mensch vs KI?
  - Kann das System, was es verspricht?

Beschreibung von Werkzeugen für KI-Prüfungen

Was soll nicht enthalten sein?

-

Was gibt es bereits auf dem Markt und wie unterscheidet sich diese geplante DIN SPEC davon?

- AI Act

Lena Krieger & Sobhi Mahmoud (DIN e. V.)

Weiteres Vorgehen

- Ausformulierung der geplanten Anwendungsbereiche für die beiden zuvor diskutierten Themen
  - Ground truth-Referenzdatensätze & Metriken
  - Beschreibung von Werkzeugen für KI-Prüfungen
- Erarbeitung der Standards (DIN SPECs) wird formell initiiert
  - Veröffentlichung des Geschäftsplans zur zweimonatigen öffentlichen Kommentierung und Bekundung von Interesse zur Mitarbeit am Projekt
  - Information zur GP-Veröffentlichung wird breit gestreut, und zusätzlich auf der DIN-Website bekannt gegeben

Lena Krieger
Senior-Projektmanagerin

Lena.Krieger@din.de
+49 (0) 30 2601-2810

Sobhi Mahmoud
Senior-Projektmanager

Sobhi.Mahmoud@din.de
+49 (0) 30 2601-2061

DIN
Deutsches Institut für Normung e. V.
Am DIN-Platz
Burggrafenstraße 6
10787 Berlin

www.din.de