



Geschäftsplan für ein DIN SPEC-Projekt nach
dem PAS-Verfahren zum Thema
**„Künstliche Intelligenz - Life Cycle
Prozesse und Qualitätsanforderungen –
Teil 3: Erklärbarkeit“**

Status:
Erarbeitung der DIN SPEC nach Annahme am
10.08.2022

Die Empfänger dieses Geschäftsplans werden gebeten, mit ihren
Kommentaren **jegliche relevanten Patentrechte**, die sie kennen, mitzuteilen
und unterstützende Dokumentationen zur Verfügung zu stellen.

Berlin, 17.03.2023 (Version 2)

Inhaltsverzeichnis

1. Status/Version des Geschäftsplans.....	3
2. Initiator und weitere Konsortialmitglieder.....	3
3. Ziele des Projekts.....	5
4. Arbeitsprogramm.....	7
5. Ressourcenplanung	8
6. Regeln der Zusammenarbeit im DIN SPEC-Konsortium	8
7. Kontaktpersonen	10
Anhang: Zeitplan (vorläufig).....	12

1. Status/Version des Geschäftsplans

- Zur Erarbeitung der DIN SPEC nach Annahme am 10.08.2022

Änderungsvermerk zur Vorgängerversion 01:

- Abschnitt 2: Tabelle der teilnehmenden Organisationen ergänzt
- Abschnitt 3.2: Anwendungsbereich präzisiert
- Abschnitt 4: Dokumentensprache wurde angepasst

2. Initiator¹ und weitere Konsortialmitglieder

- Initiatoren:

Person/Organisation	Kurzbeschreibung
Dr. rer. nat. Maximilian Poretschkin, Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS	Als Teil der größten Organisation für anwendungsorientierte Forschung in Europa ist das Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS mit Sitz in Sankt Augustin bei Bonn eines der führenden Wissenschaftsinstitute auf den Gebieten Künstliche Intelligenz, Maschinelles Lernen und Big Data in Deutschland und Europa. https://www.iais.fraunhofer.de/de/institut/ueber-uns.html
Felix Assion, Neurocat GmbH	Neurocat GmbH ist ein IT-Startup mit Sitz im Innovations- und Gründerzentrum Berlin-Adlershof. Ein Team, bestehend aus Mathematikern, Informatikern und Wirtschaftsingenieuren, beschäftigt sich mit Forschungsfragen im Bereich der Künstlichen Intelligenz und mit Anwendungsmöglichkeiten moderner KI-Technologien. Dabei liegt der Fokus auf den Forschungsfeldern Qualität und Sicherheit von Künstliche Intelligenz Systemen, mit dem Schwerpunkt <i>Explainable Artificial Intelligence</i> .

¹ Die in diesem Dokument gewählte männliche Form der geschlechtsbezogenen Begriffe wie z. B. „der Initiator“ gelten selbstverständlich auch für alle weiblichen Personen. Lediglich aufgrund der besseren Verständlichkeit des Textes wurde einheitlich die männliche Form gewählt.

- Organisationen³, die den Geschäftsplan angenommen haben:

Person	Organisation
Felix Assion	Neurocat GmbH
Stephen Bäuerle	metafinanz Informationssysteme GmbH
Nikolas Becker	Gesellschaft für Informatik (gi)
Dr. Tarek R. Besold	DEKRA DIGITAL
Lukas Bieringer	QuantPi GmbH
André Bluhm	ai.dopt GmbH
Prof. Armin B. Cremers	Universität Bonn
Dr. Peter Deussen	Microsoft Deutschland GmbH
Marton Eifert	metafinanz Informationssysteme GmbH
Jörg Firnkorn	DEKRA DIGITAL
Antoine Gautier	QuantPi GmbH
Rebekka Görge	Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS
Prof. Dr. Stefan Haufe	PTB - Physikalisch-Technische Bundesanstalt
Andreas Hauschke	VDE Verband der Elektrotechnik Elektronik Informationstechnik e. V.
Dr. Janis Kesten-Kühne	PricewaterhouseCoopers GmbH
Prof. Dr. Kai-Uwe Kühnberger	Universität Osnabrück
Dr. Christian Kruschel	IAV GmbH
Erik Martori López	metafinanz Informationssysteme GmbH
Adriano Lucieri	Deutsches Forschungszentrum für Künstliche Intelligenz
Matthias Neumann-Borsig	IAV GmbH
Detlef Olschewski	Cleopa GmbH
Sebastian Palacio	Deutsches Forschungszentrum für Künstliche Intelligenz
Dr. Henrich Poehls	Universität Passau
Dr. rer. nat. Maximilian Poretschkin	Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS
Dr. Georg Rehm	Deutsches Forschungszentrum für Künstliche Intelligenz
Prof. Dr. Ute Schmid	Otto-Friedrich-Universität Bamberg
Annegrit Seyerlein-Klug	Technische Hochschule Brandenburg
Pierre Springer	Zentrale Stelle für Informationstechnik im Sicherheitsbereich
Yeji Streppel	Technical University Eindhoven

Person	Organisation
Dr. Rustam Tagiew	Deutsches Zentrum für Schienenverkehrsforschung beim Eisenbahn-Bundesamt
Hauke Timmermann	eco – Verband der Internetwirtschaft e.V.
Prof. Dr. Carlos Zednik	Technische Universität Eindhoven
Prof. Dr. Thomas Zielke	Hochschule Düsseldorf

3. Ziele des Projekts

3.1. Allgemeines

Die „Deutsche Normungsroadmap (NRM) Künstliche Intelligenz“, Stand November 2020, formuliert mehrere Forschungsbedarfe als Grundlage für die Standardisierung. Als ein großer Forschungsbedarf wird das Thema „Erklärbare künstliche Intelligenz - Explainable AI (xAI)“ identifiziert (Bedarf 6 – im Bereich Sicherheit).

Erklärbarkeit und Nachvollziehbarkeit dienen dem Ziel sicherzustellen, dass Teilnehmer, die mit einer künstlichen Intelligenz (KI) interagieren, verstehen, wie die KI zu ihren Entscheidungen kommt, wie sie entwickelt und getestet wird und wie sicher sie ist. Nur wenn diese Möglichkeiten gegeben sind, kann Sicherheit und Vertrauen der Wirtschaft und der Anwender in Künstliche Intelligenz geschaffen und eine Nachweisbarkeit von erforderlichen Eigenschaften von KI-Systemen sichergestellt werden.

Mit dem zunehmenden Einsatz von Modellen des maschinellen Lernens in scheinbar allen Bereichen der Industrietechnik und in vielen Produkten des täglichen Lebens stehen die KI-Nutzer vor der Herausforderung, zu verstehen, wie bestimmte Vorhersagen, die durch Black-Box- oder Grey-Box-Ansätze gewonnen wurden, zustande gekommen sind. Darüber hinaus ist es unmöglich, die Vertrauenswürdigkeit eines maschinellen Lernmodells vollständig zu beurteilen, ohne zu verstehen, wie und warum es seine Entscheidungen trifft und ob diese Entscheidungen gerechtfertigt und angemessen sind.

Vor dem Einsatz eines maschinellen Lernmodells in der Praxis ist es insbesondere bei hochrisiko-KI-Systemen von größter Bedeutung, seine Entscheidungsprozesse zu verstehen, da schlecht verstandene Modelle negative Auswirkungen auf das Geschäft oder die Gesellschaft haben und mit den gesetzlichen Vorschriften in Konflikt geraten können.

Die „Erklärbarkeit“ des Modells ist daher ein wichtiges Rückgrat des gesamten Qualitätskontrollprozesses des maschinellen Lernens.

Die Erklärbarkeit ist eine der Dimensionen bei der Bewertung der Korrektheit und der kontext- und aufgabenspezifischen Angemessenheit eines maschinellen Lernmodells durch eine Reihe von Methoden ("erklärbare KI"),

die es Menschen (Experten und/oder Nutzern) ermöglichen, die vom Modell erzeugten Ergebnisse zu verstehen und ihnen schließlich zu vertrauen.

Erklärbare KI wird zur Beschreibung und Analyse eines KI-Modells, seiner erwarteten Auswirkungen und potenziellen Verzerrungen verwendet. Sie hilft dabei, das Qualitätsprofil eines maschinellen Lernmodells zu charakterisieren.

Eine erklärbare künstliche Intelligenz ist somit eine der wichtigsten Grundlagen für geplante Prüfbarkeit auf Basis von Standardisierung, ggf. mit dem Ziel der Zertifizierung.

Die Internationale Normung arbeitet bereits intensiv im ISO/IEC JTC 1/SC 42 „Artificial Intelligence“ an Normen, u. a. auch am Thema Transparenz und Erklärbarkeit (siehe auch ISO/IEC DIS 22989).

3.2. Geplanter Anwendungsbereich

Dieses Dokument stellt einen branchenunabhängigen Leitfaden zu geeigneten Ansätzen und Methoden für die Förderung von Erklärbarkeit im gesamten Lebenszyklus eines KI-Modells dar. Es wird der Begriff "Opazität" definiert, die Ursachen und Auswirkungen von Opazität in der heutigen KI beschrieben und die Art und Weise betrachtet, wie Erklärungen eingesetzt werden können und sollten, um diese Auswirkungen für verschiedene Interessengruppen in verschiedenen Phasen des Lebenszyklus von KI-Systemen abzuschwächen.

Die in diesem Dokument beschriebenen Qualitätskriterien zielen darauf ab, für alle Arten von KI-Systemen und für alle Bereiche, in denen solche Systeme eingesetzt werden, anwendbar zu sein. Sie gelten insbesondere für KI-Systeme mit unterschiedlichem Grad an Opazität. Diese Qualitätskriterien erleichtern ein besseres Verständnis und eine effektivere Nutzung von erklärbarer KI (XAI) und ermöglichen es Unternehmen, Grundsätze des Vertrauens und der Transparenz in ihre KI-basierten Anwendungen einzubinden.

Sie liefern auch die konzeptionelle und prozedurale Grundlage für die Formulierung spezifischer Anforderungen an die Erklärbarkeit und den nachvollziehbaren Nachweis der technisch korrekten Funktionsweise.

Dieses Dokument ist keine Anforderungsspezifikation und enthält auch keine Umsetzungsvorgaben.

3.3. Verwandte Aktivitäten

Das Thema der geplanten DIN SPEC ist bisher nicht Gegenstand einer Norm. Es existieren jedoch die folgenden, themenverwandten Gremien, Normen und/oder Regelwerke, die im Zuge des Projekts berücksichtigt und ggf. einbezogen werden:

- ISO/IEC 22989 „Information technology — Artificial intelligence — Artificial intelligence concepts and terminology”
- ISO/IEC AWI TS 6254 „Information technology — Artificial intelligence — Objectives and approaches for explainability of ML models and AI systems”
- DIN SPEC 92001-1 „Künstliche Intelligenz - Life Cycle Prozesse und Qualitätsanforderungen - Teil 1: Qualitäts-Meta-Modell”
- DIN SPEC 92001-2 „Künstliche Intelligenz - Life Cycle Prozesse und Qualitätsanforderungen - Teil 2: Robustheit”
- DIN SPEC 92003 „Vorgehensmodell zur Planung und Einführung datengetriebener Softwareanwendungen“
- NA 043-01-42 GA DIN/DKE Gemeinschaftsarbeitsausschuss „Künstliche Intelligenz“
- NA 043-01-42-01 AK „KI-Qualität und Prüfverfahren“
- ISO/IEC JTC 1/SC 42 „Artificial intelligence”
- CEN/CLC/JTC 21 „Artificial Intelligence”

4. Arbeitsprogramm

Im Zuge des Projekts soll eine DIN SPEC nach dem PAS-Verfahren (vgl. www.din.de/go/spec) erarbeitet werden. Die DIN SPEC darf nicht in Widerspruch zum Deutschen Normenwerk stehen.

Das Kick-Off fand am 10.08.2022 per Webkonferenz (WebEx) statt. Die Projektlaufzeit beträgt ca. 8 Monate.

Das Kick-Off dient der Konstituierung des Konsortiums, der Abstimmung bzw. Klärung weiterer organisatorischer Punkte sowie ggf. der Aufnahme der inhaltlichen Arbeiten.

Die Veröffentlichung eines Entwurfs zur Kommentierung durch die Öffentlichkeit ist nicht vorgesehen.

Insgesamt werden voraussichtlich 2 Projektmeetings (Arbeitsmeetings) und 6 Webkonferenzen (Kick-off und Arbeitsmeetings) durchgeführt, um die jeweils bis dahin erarbeiteten Inhalte vorzustellen, abzustimmen und ggf. zu verabschieden. Die Erarbeitung der Inhalte kann durch einzelne Konsortialmitglieder oder Arbeitsgruppen erfolgen.

Die Terminierung der weiteren Projektmeetings und/oder Webkonferenzen erfolgt durch das Konsortium in Abstimmung mit DIN.

Die DIN SPEC wird in Deutsch erarbeitet (Sitzungssprache, Berichte, usw.). Die DIN SPEC wird in Englisch verfasst.

ANMERKUNG In der Kalkulation wurde nur eine Sprachfassung berücksichtigt. Die Erarbeitung weiterer Sprachfassungen verursacht zusätzliche Kosten und muss deswegen gesondert vereinbart werden. Wenn eine weitere Sprachfassung gewünscht wird, kann die Übersetzung auch durch Beuth/DIN erfolgen. Diese wäre nach Verabschiedung des Manuskripts zur Veröffentlichung der DIN SPEC zusätzlich zu beauftragen.

5. Ressourcenplanung

Jedes Konsortialmitglied trägt seine im Rahmen des Vorhabens anfallenden Aufwendungen selbst.

Genehmigt der Vorstand von DIN die Durchführung des Projekts, schließt der Initiator einen Vertrag mit DIN.

Die Mitgliedschaft im Konsortium und die Teilnahme an den Projektmeetings ist kostenfrei, da die Kosten, die DIN aufgrund der Durchführung des Projekts entstehen, durch Mittel aus dem Forschungsprojekt „Zertifizierung von Standard-KI-Anwendungen (Zertifizierte KI)“ – gefördert durch das Ministerium für Wirtschaft, Innovation, Digitalisierung und Energie des Landes Nordrhein-Westfalen (MWIDE NRW) im Rahmen der Richtlinie über die Gewährung von Zuwendungen zur Förderung von Vorhaben im Innovations- und Transferbereich nordrhein-westfälischer Spitzencluster und von für NRW strategisch bedeutsamen Kooperationsvorhaben und Netzwerkprojekten (Spitzencluster Förderrichtlinie) (Förderkennzeichen: 005-2011-0053) – finanziert werden.

6. Regeln der Zusammenarbeit im DIN SPEC-Konsortium

Das Projekt unterliegt den PAS-Verfahrensregeln. Alle Interessenten und Konsortialmitglieder sind dazu aufgefordert, sich unter <http://www.din.de/go/spec> über die Verfahrensregeln in Kenntnis zu setzen.

Die Konstituierung des Konsortiums erfolgt im Zuge des Kick-Offs. Der Kick-Off findet erst statt, nachdem der Geschäftsplan veröffentlicht und die Durchführung des Projekts durch die DIN-Geschäftsleitung genehmigt wurde. Das Konsortium muss sich aus mindestens drei Konsortialmitgliedern unterschiedlicher Organisationen² zusammensetzen. Es ist nicht notwendig, dass diese unterschiedliche interessierte Kreise repräsentieren. Durch Zustimmung zum Geschäftsplan erklären die Interessenten ihre Bereitschaft zur Mitarbeit im Konsortium und werden dadurch formell zu Konsortialmitgliedern mit den einhergehenden Rechten und Pflichten. Teilnehmer des Kick-Offs, die den Geschäftsplan nicht annehmen, erhalten nicht den Status eines Konsortialmitglieds und sind von weiteren Entscheidungen des Kick-Offs sowie vom weiteren Projekt ausgeschlossen.

Entsendet eine Organisation (z. B. ein Verband) einen nicht-hauptamtlichen Mitarbeiter in das Konsortium, muss dieser von der Organisation autorisiert und DIN der Nachweis vorgelegt werden.

Jedes Konsortialmitglied erhält ein Stimmrecht und verfügt über jeweils eine Stimme. Entsendet eine Organisation mehrere Experten in das Konsortium,

² Organisationen sind teilnehmende juristische Personen, die die Experten in das DIN SPEC-Konsortium entsenden und einer Unternehmensstruktur i.S.v. § 15 Aktiengesetz oder § 271 Absatz 2 Handelsgesetzbuch zuzurechnen sind.

besitzt die Organisation, ungeachtet der Anzahl der entsendeten Teilnehmer, eine Stimme. Eine Übertragung von Stimmen auf andere Konsortialmitglieder ist nicht möglich. Bei Abstimmungen gilt einfache Mehrheit der abgegebenen Stimmen, wobei Stimmenthaltungen grundsätzlich nicht mitgezählt werden.

Das konstituierte Konsortium ist in der Regel geschlossen. Über die Aufnahme zusätzlicher Mitglieder entscheiden die bisherigen Konsortialmitglieder.

Im Zuge des Kick-Offs wählen die Konsortialmitglieder einen Konsortialleiter. Dieser leitet das Konsortium inhaltlich und führt die Entscheidungsfindung (Abstimmungen, Beschlüsse) herbei. Der Konsortialleiter wird hierbei durch den DIN-Projektmanager unterstützt, wobei DIN stets eine inhaltlich neutrale Position einnimmt. Darüber hinaus trägt der DIN-Projektmanager dafür Sorge, dass die Verfahrens- und Gestaltungsregeln von DIN bei der Erstellung der DIN SPEC eingehalten werden. Sollte der Konsortialleiter seine Funktion nicht mehr wahrnehmen können, werden vom DIN-Projektmanager Neuwahlen initiiert.

Die Organisation und Leitung des Kick-Offs erfolgt durch den DIN-Projektmanager in Abstimmung mit dem Initiator. Die übrigen Projektmeetings und/oder Webkonferenzen werden vom DIN-Projektmanager in Abstimmung mit dem Konsortialleiter organisiert.

Wenn Konsortialmitglieder bei der Verabschiedung der DIN SPEC bzw. des Entwurfs nicht anwesend sein können, sind diese über alternative Wege (z. B. schriftlich, elektronisch) in die Abstimmung einzubeziehen.

Alle Konsortialmitglieder, die für die Veröffentlichung der DIN SPEC bzw. des Entwurfs gestimmt haben, werden als Verfasser namentlich und mit der zugehörigen Organisation im Vorwort aufgeführt. Alle Konsortialmitglieder, die gegen die Veröffentlichung der DIN SPEC bzw. des Entwurfs gestimmt oder sich enthalten haben, dürfen nicht im Vorwort genannt werden.

Über eine nachträgliche Erweiterung des Konsortiums entscheiden die bisherigen Konsortialmitglieder. Dabei ist insbesondere zu berücksichtigen, dass

- a) die Erweiterung förderlich ist, die Projektdauer zu verkürzen bzw. ein drohender Verzug der geplanten Projektdauer vermieden bzw. abgewendet werden kann;
- b) die Erweiterung nicht zu einer drohenden Verlängerung der Projektdauer führt;
- c) das neue Konsortialmitglied keine neuen oder ergänzenden Sachverhalte abseits des im Geschäftsplans festgelegten und bewilligten Anwendungsbereiches thematisiert;
- d) das neue Konsortialmitglied ergänzendes Fachwissen mitbringt, damit die neuesten Erkenntnisse der Wissenschaft und der jeweilige Stand der Technik eingebracht werden;
- e) das neue Konsortialmitglied sich aktiv an der Manuskriptarbeit beteiligt durch Einbringen konkreter, aber nicht abstrakter Vorschläge und Beiträge.

- f) das neue Konsortialmitglied für eine verstärkte Anwendung der DIN SPEC sorgt.

Um die sachgerechte Vervielfältigung und Verbreitung der Ergebnisse der Standardisierungsarbeit zu ermöglichen, räumen die Konsortialmitglieder DIN die Nutzungsrechte an den ihnen erwachsenden Urheberrechten an den Ergebnissen der Standardisierungsarbeit ein. Die Einräumung der Urhebernutzungsrechte hindert die Mitglieder des Konsortiums nicht daran, ihr eingebrachtes Wissen, ihre Erfahrungen und Erkenntnisse weiterhin zu nutzen, zu verwerten und weiterzuentwickeln.

Die Konsortialmitglieder sind angehalten, DIN über relevante Patentrechte, die in Zusammenhang mit diesem DIN SPEC Projekt stehen, zu informieren.

Nachträgliche Änderungen am Anwendungsbereich (Abschnitt 3.2) oder an der Ressourcenplanung (Abschnitt 5) erfordern neben einer 2/3-Mehrheit aller abgegebenen Stimmen zusätzlich die Zustimmung von DIN.

7. Kontaktpersonen

- Konsortialleiterin:
Annegrit Seyerlein-Klug
Technische Hochschule Brandenburg
Magdeburger Straße 50
Brandenburg 14770
Tel.: +49 3381 355-290
E-Mail: annegrit.seyerleinklug@th-brandenburg.de
- Stellvertretender Konsortialleiter:
Prof. Dr. Carlos Zednik
Technische Universität Eindhoven
Department of Industrial Engineering & Innovation Science
Philosophy and Ethics (P&E) Group
P.O. Box 513, 5600 MB Eindhoven, Niederlande
E-Mail: c.a.zednik@tue.nl
- Projektmanager:
Sobhi Mahmoud
DIN Deutsches Institut für Normung e. V.
Burggrafenstraße 6
10787 Berlin
Tel.: + 49 30 2601- 2061
E-Mail: sobhi.mahmoud@din.de

- Initiatoren:
Dr. rer. nat. Maximilian Poretschkin
Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme
IAIS
Schloss Birlinghoven 1
53757 Sankt Augustin
Tel.: +49 2241 14-2260
E-Mail: maximilian.poretschkin@iais.fraunhofer.de

Felix Assion
Neurocat GmbH
IGZ Innovations- und Gründungszentrum
Rudower Chaussee 29
12489 Berlin
E-Mail: fa@neurocat.ai

Anhang: Zeitplan (vorläufig)

DIN SPEC-Projekt	2022/2023																									
	Mai	Jun	Jul	Aug	Sept	Okt	Nov	Dez	Jan	Feb	Mär	Apr	Mai													
Initiierung	■	■	■	■																						
1. Antrag und Prüfung	■	■	■																							
2. Erstellung des Geschäftsplans		■																								
3. Veröffentlichung des Geschäftsplans			■	■																						
Erstellungsphase				■	■	■	■	■	■	■	■	■	■	■	■											
4. Kick-Off / Konstituierung des Konsortiums				■																						
5. Erstellung der DIN SPEC				■	■	■	■	■	■	■	■	■	■	■	■											
6. Verabschiedung DIN SPEC im Konsortium												■														
Veröffentlichung														■	■											
7. Prüfung und Freigabe durch DIN														■												
8. Veröffentlichung der DIN SPEC														■	■											
Meilensteine					K / W		W	W		W		W	W		W		W		W		W		M / V			

- K** Kick-Off
- M** Projektmeeting
- W** Webkonferenz
- V** Verabschiedung der DIN SPEC