

ICS 35.080; 35.240.01

**Übertragung von sprachbasierten Daten zwischen
Künstlichen Intelligenzen –
Festlegung von Parametern und Formaten**

Transmission of language-based data between artificial intelligences –
Specification of parameters and formats

Gesamtumfang 22 Seiten

Inhalt

	Seite
Vorwort	3
Einleitung	4
1 Anwendungsbereich	5
2 Normative Verweisungen	5
3 Begriffe	5
4 Festlegung von Parametern und Formaten für den Austausch von Sprachbefehlen	7
4.1 Allgemeines	7
4.2 Klassifikation von Parametern	8
4.3 Zwingende Parameter	8
4.4 Empfohlene Parameter	8
4.5 Optionale Parameter	9
4.6 Serialisierung der Parameter und Formate	10
Anhang A (informativ) Analyse möglicher zukünftiger Anforderungen an NLP-Interoperabilität	11
A.1 Zukünftige Entwicklungen	11
A.2 Definition von Schnittstellen und Datenaustauschprozessen	11
A.2.1 Allgemeines	11
A.2.2 Rezipient	11
A.2.3 Protokoll der Kommunikation	12
A.2.4 Eigener Beitrag mit Begründung	12
A.2.5 Weitergabe der Datengrundlage (z. B. als Beleg)	12
A.3 Datenformate	13
A.4 Kompatibilitätserfordernisse für Internationalisierung	13
Anhang B (informativ) Allgemeiner Anwendungsfall	15
B.1 Anforderungsbeschreibung	15
B.2 Beschreibung der NLP-Schnittstelle	15
Anhang C (informativ) Anwendungsfall im fachgebietsspezifischen Umfeld	21
Literaturhinweise	22

Tabellen

Tabelle 1 — Zwingende Parameter	8
Tabelle 2 — Empfohlene Parameter	8
Tabelle 3 — Optionale Parameter	9
Tabelle B.1 — Protokollsyntax einer Ticketbuchung mithilfe eines Sprachassistenten	16

Vorwort

Diese DIN SPEC wurde nach dem PAS-Verfahren erarbeitet. Die Erarbeitung von DIN SPEC nach dem PAS-Verfahren erfolgt in DIN SPEC (PAS)-Konsortien und nicht zwingend unter Einbeziehung aller interessierten Kreise.

Die Erarbeitung und Verabschiedung des Dokuments erfolgten durch die nachfolgend genannten Initiator(en) und Verfasser:

- Alexander Klug IT Beratung
Alexander Klug
- ARS Computer und Consulting GmbH
Dr. Thorsten Gressling
- Cleopa GmbH
Dipl.-Ing. Detlef Olschewski
- Deutscher Terminologie-Tag e.V. (DTT)
Dr. Detlef Reineke
- Onsei GmbH
Jochen Emig
- PricewaterhouseCoopers GmbH Wirtschaftsprüfungsgesellschaft
Dr. Sebastian Dungs
- Whoelse UG
Sven Krüger
Tobias Martens
- Zuse Institute Berlin (ZIB)
Dr. habil. Marcus Weber

Für dieses Thema bestehen derzeit keine Normen im Deutschen Normenwerk.

DIN SPEC (PAS) sind nicht Teil des Deutschen Normenwerks.

Für diese DIN SPEC (PAS) wurde kein Entwurf veröffentlicht.

Trotz großer Anstrengungen zur Sicherstellung der Korrektheit, Verlässlichkeit und Präzision technischer und nicht-technischer Beschreibungen kann das DIN SPEC (PAS)-Konsortium weder eine explizite noch eine implizite Gewährleistung für die Korrektheit des Dokuments übernehmen. Die Anwendung dieses Dokuments geschieht in dem Bewusstsein, dass das DIN SPEC (PAS)-Konsortium für Schäden oder Verluste jeglicher Art nicht haftbar gemacht werden kann. Die Anwendung der vorliegenden DIN SPEC (PAS) entbindet den Benutzer nicht von der Verantwortung für eigenes Handeln und geschieht damit auf eigene Gefahr.

Es wird auf die Möglichkeit hingewiesen, dass einige Elemente dieses Dokuments Patentrechte berühren können. DIN ist nicht dafür verantwortlich, einige oder alle diesbezüglichen Patentrechte zu identifizieren.

Die kostenfreie Bereitstellung dieses Dokuments als PDF-Version über den Beuth WebShop wurde im Vorfeld finanziert.

Aktuelle Informationen zu diesem Dokument können über die Internetseiten von DIN (www.din.de) durch eine Suche nach der Dokumentennummer aufgerufen werden.

Einleitung

Der Markt für Anwendungen im Bereich der maschinellen Verarbeitung natürlicher Sprachen (NLP, en: natural language processing) wächst rasant. Studien prognostizieren, dass bereits im Jahr 2025 mehr als zwei Drittel aller Benutzeranfragen in Internet-Suchmaschinen, industriellen IT- und E-Commerce-Anwendungen mithilfe von KI-basierten Sprachassistenten abgewickelt werden.

Die Veränderung des Benutzerverhaltens ist eine große Herausforderung für Anbieter von Hardware-Produkten, OEMs (en: original equipment manufacturer) oder Herstellern von Systemlösungen, die KI-Anwendungen einkaufen bzw. in Wertschöpfungsketten integrieren. Voraussetzung für den Vertrieb von Produkten in unterschiedlichen Märkten, ist die Kombination von NLP-Anwendungen für lokale Sprachen und Anwendungsfälle.

Sprachdialogsysteme funktionieren i. d. R. immer nur in einer Sprache. Außerdem beherrschen NLP-Anwendungen generell nur die Anwendungsfälle, für die ihr linguistisches Modell ausgelegt und trainiert ist. Der Austausch sprachbasierter Daten zwischen KI-Ökosystemen, wie etwa Informationen zur Absicht des Anliegens (Intents), erfolgt bislang vorwiegend über proprietäre Datenformate und Schnittstellen.

Konsumenten bemängeln, dass Sprachdialogsysteme z. T. eine schlechte Benutzbarkeit haben. Es ist oftmals nicht klar, welche Befehle eine NLP-Anwendung verstehen bzw. beantworten kann. Weiterhin ist Datenschutz ein Problem, da nicht nachvollziehbar ist, welche Benutzereingaben auf Basis von sprachbasierten Daten erhoben werden und welche Parteien Zugriff auf Nutzerdaten wie z. B. biometrische, Standort- oder anderweitige personenbezogene Daten haben.

Ziel dieses Dokuments ist es Industrieanwendern, Open Source-Gemeinschaften und Privatanwendern zu helfen, interoperable sprachbasierte Anwendungen zu entwickeln, Daten sprachbasierter Anwendungen zu überprüfen und nachzuvollziehen und Datenzugriff bzw. Datenschutz zu ermöglichen. Bei der Verbesserung der Interoperabilität KI-basierter NLP-Anwendungen geht es im Einzelnen um:

- Maßnahmen zur Herstellung von KI-Ökosystem-Interoperabilität, um einen schnelleren Datenaustausch und Vernetzung unterschiedlicher NLP-Anwendungen zu ermöglichen;
- Ansätze zur Komprimierung von Spracheingaben in KI-basierten Anwendungen, um eine einheitliche Verarbeitung in Maschinensprache sicherzustellen; und
- vereinheitlichte Verfahrensweisen für sprachbasierte Anwendungen, um übertragene Daten durch Dritte nachvollziehbar und überprüfbar zu machen und einen plattformübergreifenden Datenaustausch mit anderen KI-Anwendungen zu ermöglichen.

Anwendungsfälle für dieses Dokument sind Lösungen mit hohem Transparenzbedarf in KI-Prozessen:

- im industriellen Sektor, wie z. B. Anwendungen in der industriellen Fertigung, Finanzdienstleistungen und Unternehmenssteuerung;
- im öffentlichen Sektor, wie z. B. Anwendungen in den Bereichen Gesundheit, Energie und Transport/Nahverkehr;
- im privaten Sektor, wie z. B. Anwendungen in Forschung, Gesundheit und andere Mandantendaten-Beziehungen.

1 Anwendungsbereich

Dieses Dokument legt Parameter und Formate für die Übertragung sprachbasierter Daten sowie Voraussetzungen für die Schnittstellen zwischen verschiedenen KI Ökosystemen fest.

Dieses Dokument legt nicht die Erkennung und Analyse von Daten innerhalb sprachbasierter Anwendungen fest.

2 Normative Verweisungen

Die folgenden Dokumente werden im Text in solcher Weise in Bezug genommen, dass einige Teile davon oder ihr gesamter Inhalt Anforderungen des vorliegenden Dokuments darstellen. Bei datierten Verweisungen gilt nur die in Bezug genommene Ausgabe. Bei undatierten Verweisungen gilt die letzte Ausgabe des in Bezug genommenen Dokuments (einschließlich aller Änderungen).

DIN EN ISO 3166-1, *Codes für die Namen von Ländern und deren Untereinheiten — Teil 1: Codes für Ländernamen*

DIN ISO 8601, *Datenelemente und Austauschformate — Informationsaustausch — Darstellung von Datum und Uhrzeit*

ISO 639-1, *Codes for the representation of names of languages — Part 1: Alpha-2 code*

3 Begriffe

Für die Anwendung dieses Dokuments gelten die folgenden Begriffe.

DIN und DKE stellen terminologische Datenbanken für die Verwendung in der Normung unter den folgenden Adressen bereit:

- DIN-TERMinologieportal: verfügbar unter <https://www.din.de/go/din-term>
- DKE-IEV: verfügbar unter <http://www.dke.de/DKE-IEV>

3.1

Sprache

systematische Verwendung von Lauten, Schriftzeichen, Symbolen oder Gebärden zur Kommunikation

[QUELLE: DIN ISO 20539:—¹, 3.1.1]

3.2

natürliche Sprache

historisch gewachsene, begrenzt rationale/standardisierte Sprache, deren Regeln überwiegend aus dem Gebrauch abgeleitet werden und deren Einhaltung nicht ausdrücklich vorgeschrieben ist

3.3

Künstliche Intelligenz

KI

Fähigkeit eines technischen Systems, Wissen und Fertigkeiten zu erwerben, zu verarbeiten und anzuwenden

[QUELLE: ISO/IEC 22989:—², 3.2.2, modifiziert — Anmerkung 1 zum Begriff wurde entfernt.]

¹ In Vorbereitung. Dokumentstufe zum Zeitpunkt der Veröffentlichung: E DIN ISO 20539:2020-07.

² In Vorbereitung.

3.4

NLU, en: natural language understanding
natürliches Sprachverstehen

Aufgabe, bei der für Menschen lesbarer Text ohne Bedeutungsverlust in strukturierten Text umgewandelt wird

[QUELLE: ISO/IEC 22989:—², 3.2.30, modifiziert — Anmerkung 1 zum Begriff wurde entfernt.]

3.5

NLP, en: natural language processing
Verarbeitung natürlicher Sprache

Informationsverarbeitung, die auf dem Verstehen natürlicher Sprache (3.2) beruht

[QUELLE: ISO/IEC 22989:—², 3.2.23, modifiziert — Anmerkung 1 bis Anmerkung 3 zum Begriff wurde entfernt.]

3.6

Sprachdialogsystem

Softwareanwendung, über die mithilfe gesprochener oder geschriebener Sprache interagiert werden kann

3.7

Sprachassistent

intelligenter persönlicher Assistent

Sprachdialogsystem, das Informationen in natürlicher gesprochener Sprache abfragt, Befehle ausführt und Assistenzdienste erbringt

3.8

Voice User Interfaces

VUI

stimmliche Benutzerschnittstelle

Benutzerschnittstelle von Sprachdialogsystemen, die über die menschliche Stimme angesprochen wird

Anmerkung 1 zum Begriff: Voice User Interfaces werden z. B. über Smart Speaker oder sonstige Hard- und Softwareprodukte zur Verfügung gestellt.

3.9

Speech to Text

STT

Umwandlung von Spracheingaben in Text

[QUELLE: ISO/IEC 2382:2015, 2120750, modifiziert — Die Benennung „speech-to-text conversion“ wurde als „Speech to Text“ übernommen.]

3.10

Intent

Anliegen

Klassifizierungsergebnis einer Benutzerabsicht

3.11

Entity

Entität

eindeutig zu bestimmendes Objekt

3.12**Konfidenzniveau**

Confidence Level

Gewissheit, mit der das Spracherkennungssystem ein Erkennungsergebnis für eine erhaltene Eingabe wiedergibt

[QUELLE: DIN EN ISO 9241-154:2013-05, 3.7, modifiziert — "Confidence Level" wurde als zugelassene Benennung ergänzt und Anmerkung 1 zum Begriff und Anmerkung 2 zum Begriff wurde entfernt.]

3.13**Internet der Dinge****IoT**

Infrastruktur, die aus miteinander verbundenen Entitäten, Personen, Systemen, Informationsressourcen und Diensten besteht, die Informationen aus der realen und der virtuellen Welt verarbeiten und auf diese Informationen reagieren

[QUELLE: ISO/IEC 20924:2018, 3.2.1]

3.14**OEM**, en: original equipment manufacturer

Erstausrüster

Ausrüstungshersteller, der ein Produkt herstellt oder entwickeln oder herstellen lässt und nicht selbst in den Einzelhandel bringt

[QUELLE: DIN 33867:2018-10, 2.36.3, modifiziert — Die zugelassene Benennung „Originalausrüstungshersteller“ wurde entfernt und in der Definition in „Ausrüstungshersteller“ geändert.]

3.15**Interoperabilität**

Fähigkeit zur Kommunikation, zur Ausführung von Programmen oder zur Übertragung von Daten zwischen verschiedenen Funktionseinheiten in einer solchen Weise, dass der Anwender nur geringe oder keine Kenntnisse über die einzelnen Merkmale dieser Einheiten besitzen muss

[QUELLE: DIN EN ISO 19118:2012-01, 4.25]

4 Festlegung von Parametern und Formaten für den Austausch von Sprachbefehlen**4.1 Allgemeines**

Eine wachsende Anzahl Internetdienste, Software- und IoT-Produkte kann über KI-basierte Sprachsteuerungen bedient werden. Hauptkomponenten von Sprachassistenten und Voice User Interfaces sind NLP-Technologien wie z. B. NLU-Software, die für die Erkennung und Klassifizierung der Absicht einer sprachbasierten Nutzereingabe zuständig sind. NLU-Software ist in der Regel das Bindeglied zu nachgelagerter Software- und Hardware-Funktionalität.

Die in 4.2 vorgenommene Klassifikation dient der Standardisierung der Parameter und Datenformate, die von NLP- bzw. NLU-Software typischerweise ausgetauscht und weiterverarbeitet werden. Hierfür wurden die Parameter und Formate unterschiedlicher Hersteller verglichen und bewertet, die aktuell in der industriellen Implementierung von Systemen für die Verarbeitung von Sprachbefehlen gängig sind. Aus Tabelle 1 bis Tabelle 3 können mögliche Parameter und Formate abgeleitet werden, mit welchen NLP-Technologien Informationen über Sprachbefehle austauschen.

Ergänzend werden in Anhang A Überlegungen zu möglichen zukünftigen Anforderungen an eine NLP-Interoperabilität mit Blick auf Aspekte wie Datensicherheit und Überprüfbarkeit erläutert. In Anhang B wird die praktische Umsetzung der festgelegten Parameter exemplarisch an einem allgemeinen