WOHIN GEHT DIE REISE BEI KÜNSTLICHER INTELLIGENZ?
DEEP LEARNING AND BEYOND

Prof. Dr. Stefan Wrobel | Fraunhofer IAIS | 3.7.2019

Fraunhofer IAIS

Deep Learning
Standard KI-Technologien werden Commodity

Google-Ankündigung am 26. Juni 2019:

»Introducing Deep Learning Containers: Consistent and portable environments«

Quelle: https://cloud.google.com/blog/products/ai-machine-learning/introducing-deep-learning-containers-consistent-and-portable-environments (Abrufdatum: 9.7.2019)

- Open Source Plattformen und ML-as-a-Service Angebote machen Standard KI-Technologien einfach zugänglich
- Personal und Qualifikation werden eine Schlüsselressource, Basis-Nutzung wird Wettbewerbs-Standard
- Vorsprung im Wettbewerb erfordert Adoption neuer Entwicklungen in der Künstlichen Intelligenz

Wohin geht die Reise bei Künstlicher Intelligenz?

- Hybride KI – Daten, Wissen, Modelle, Lernen
- Vertrauenswürdige KI – Zertifizierung als Wettbewerbsvorteil
- Verteilte KI – Effizientes und sicheres Lernen on the edge
- Quanten-KI – Neue Rechnerarchitekturen am Horizont

Hybride KI

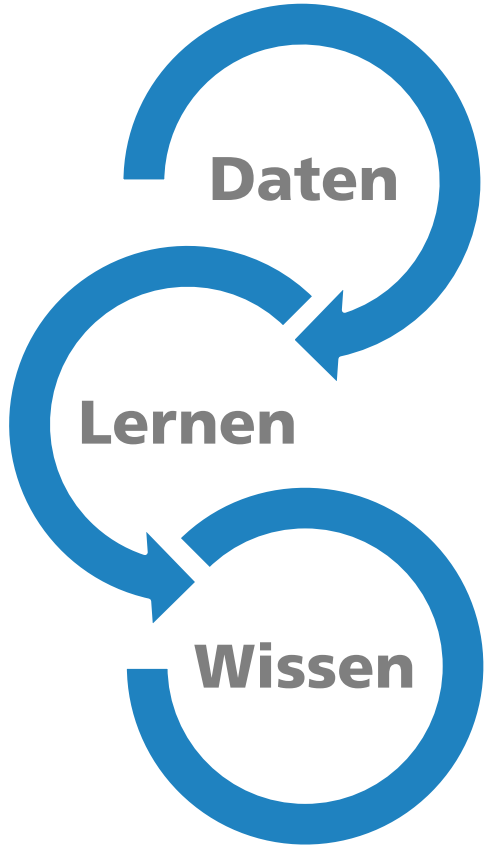
Hybride KI – Hybrides Maschinelles Lernen

Klassisches datengetriebenes Maschinelles Lernen gerät an Grenzen

- Wenn nicht genügend oder nicht die richtigen Daten für (zu) komplexe Probleme verfügbar sind
- Wenn out-of-area Generalisierung, Verlässlichkeit, Kausalität, Erklärbarkeit, Verantwortbarkeit erforderlich sind

Hybride Ansätze

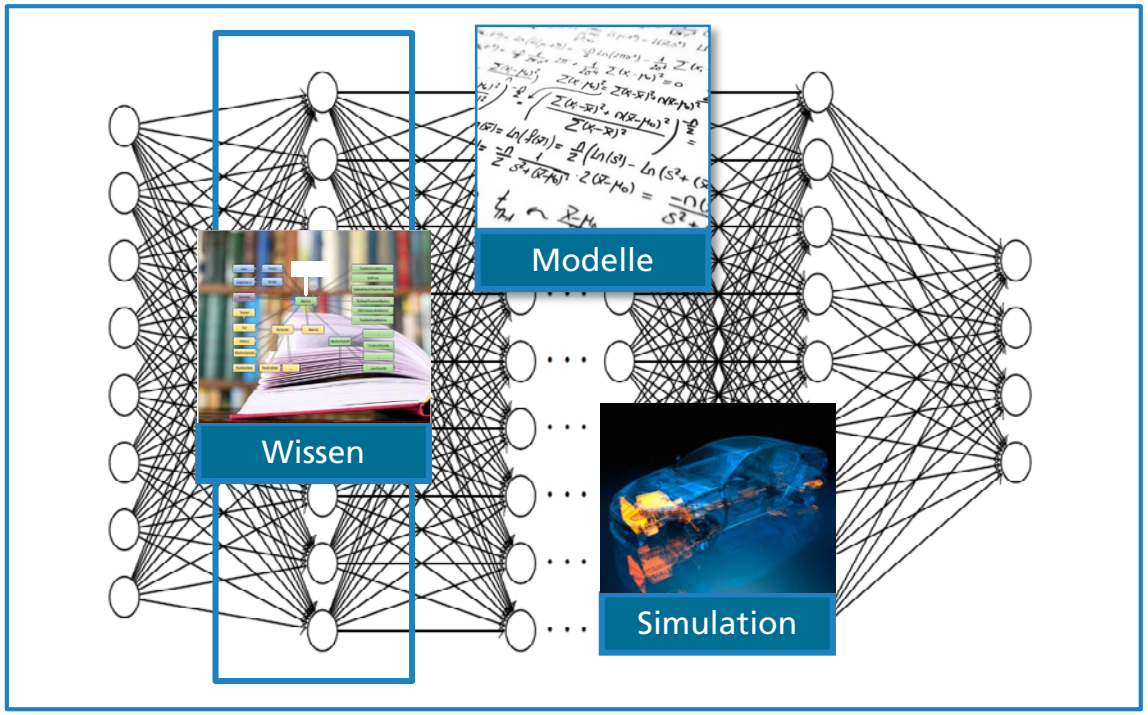
- Verbindung maschineller Lernverfahren mit modellhaftem Wissen
- Nutzung von menschlichem und technischem Wissen – Modelle, Simulationen, Erwartungen, Normen und Vorgaben
- Gemeinsames Vokabular erleichtert Kommunikation und Erklärbarkeit



Hybride KI

Maschinelles Lernen mit Wissen, Modellen, Simulationen, Gedächtnis

Beispiel: Compositional Informed Learning

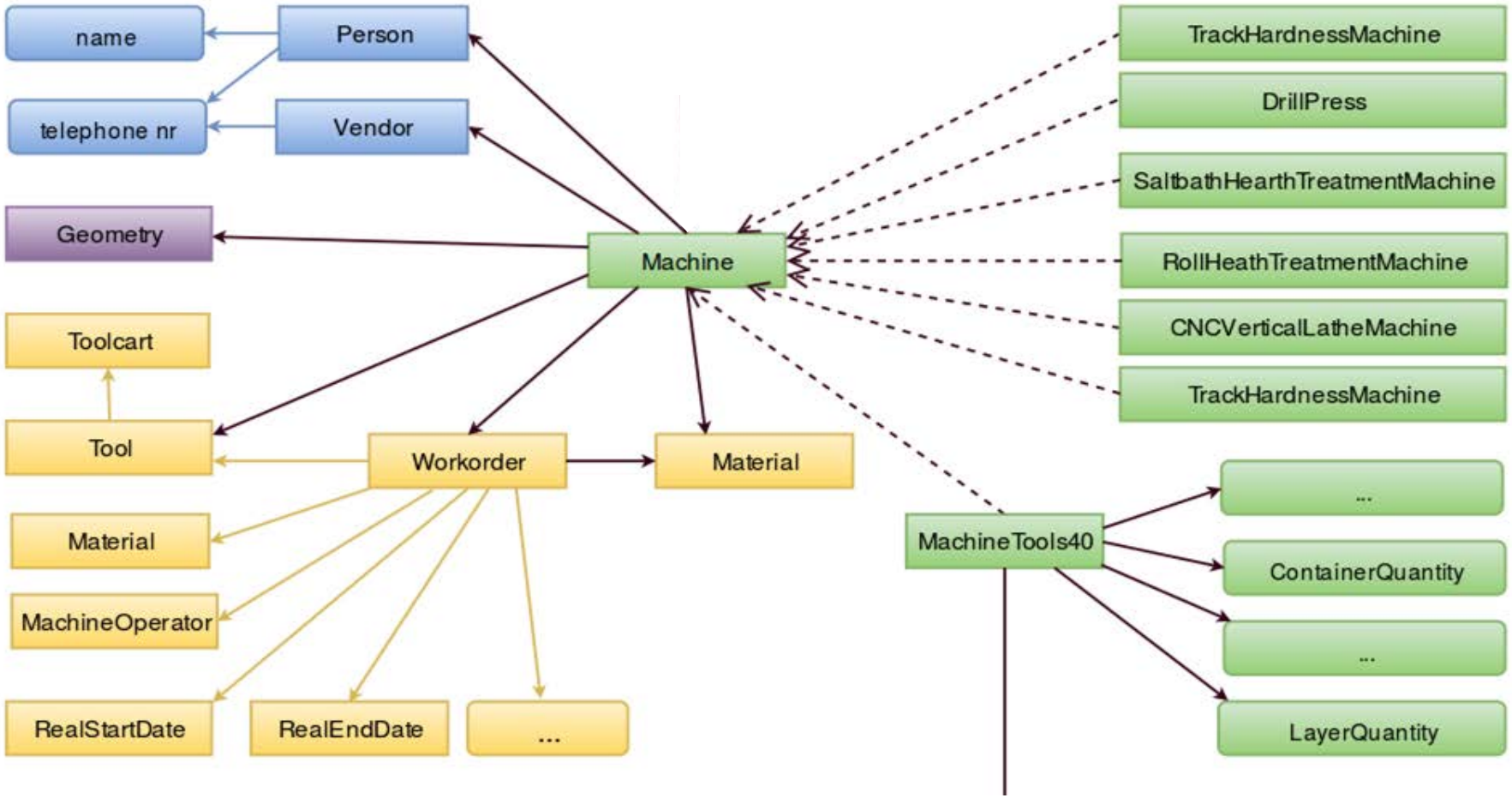


- Kann als eine Zusammensetzung aus elementaren »Modulen« - den individuellen Schichten – betrachtet werden
- »h(x) = f(g(k(x),m(x)),n(p(x)))« Funktionszusammensetzung
- Vorgegebene oder vortrainierte Modelle nehmen ähnliche Rolle ein und stellen Bausteine für das Lernen dar (cf. Reservoir Computing)

Bilder: Valkh / fotolia.de und Firstsignal / iStock

Weitere Ansätze: memory-augmented networks, neural Turing machines, differential neural computer, meta-learning

Knowledge Graphs
Linked Machine Data Example



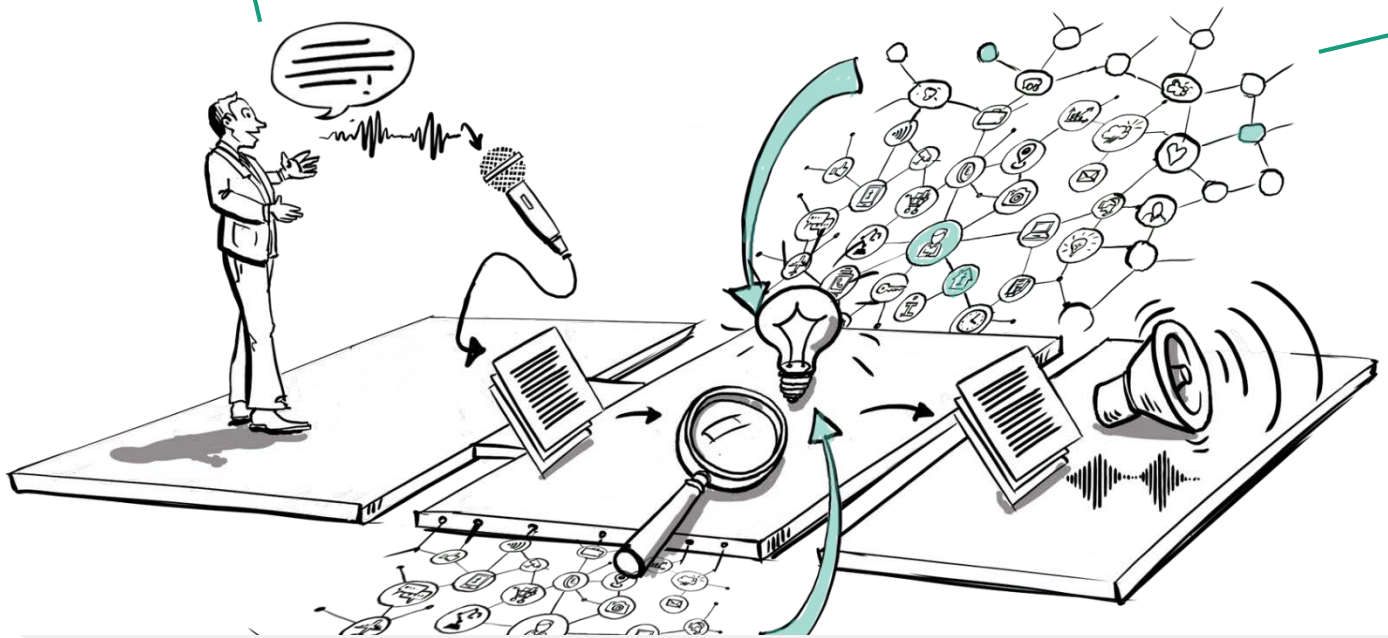
Fraunhofer-Sprachplattform: Komplettangebot für digitale Assistenten

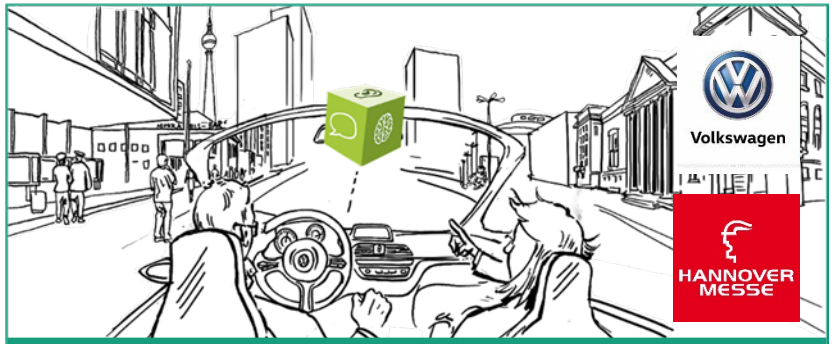
Hören, Verstehen, Sprechen, Antworten mit Hybrider KI

Komplette **sprachgesteuerte Dialogplattform** mit Fokus auf **domänenspezifisches Wissen**.

Kombination von Komponenten aus **Spracherkennung** und **Question/Answering** auf Basis von **Wissensgraphen** und **Sprachsynthese**.

Technologie und Algorithmen 100% **»made in Germany«** – Speicherung und Verarbeitung von Daten erfolgt in **sicheren Datenräumen**.





Aktuelle Anwendung: Prototyp (Sprachassistent im Auto) beantwortet Fragen zu bestimmten »Points of Interest« (Gebäuden)

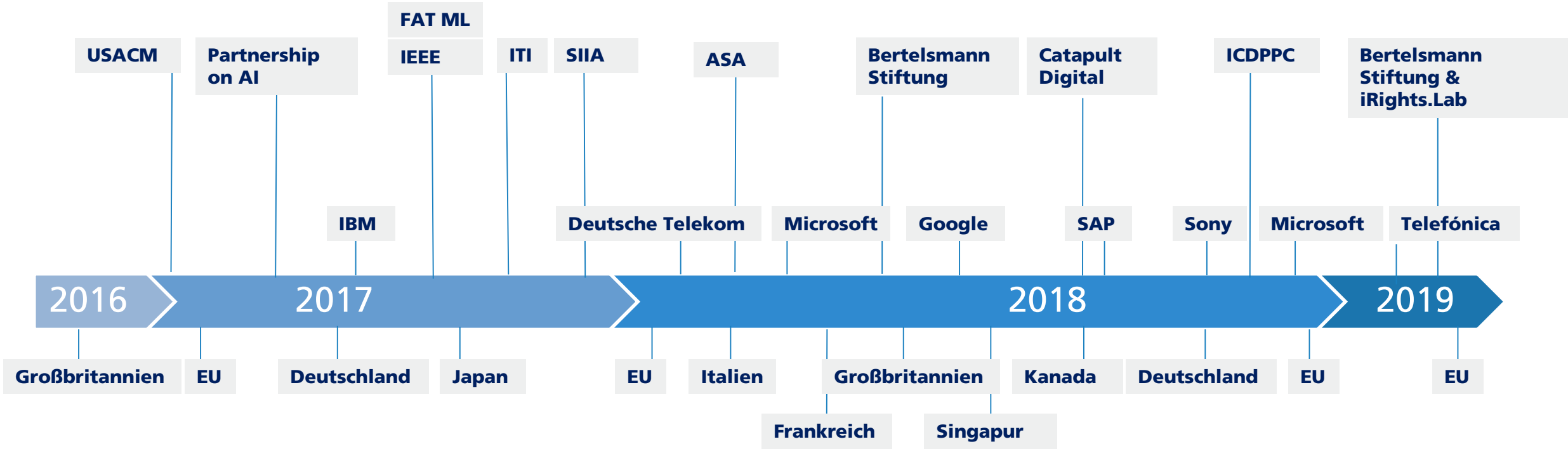
Einsatz in verschiedenen Bereichen unter Gewährleistung technologischer Souveränität

Fraunhofer IAIS

Vertrauenswürdige KI

Übersicht: KI-Leitlinien von staatl. Organisationen, Unternehmen und NGO‘s

In den letzten 2 Jahren wurden Leitlinien für eine ethische Gestaltung von KI veröffentlicht



WIE BAUT MAN KONKRET VERANTWORTUNGSVOLL KI-ANWENDUNGEN…?
…UND WORAN KANN MAN SIE ERKENNEN?

Fraunhofer IAIS

Anforderungen an eine vertrauenswürdige KI

Ethik und Recht	Respektiert die KI-Anwendung gesellschaftliche Werte und Gesetze?
Fairness	Behandelt die KI alle Betroffenen fair?
Autonomie & Kontrolle	Ist eine selbstbestimmte, effektive Nutzung der KI möglich?
Transparenz	Sind Funktionsweise und Entscheidungen der KI nachvollziehbar?
Verlässlichkeit	Funktioniert die KI zuverlässig und ist sie robust?
Sicherheit	Ist die KI sicher gegenüber Angriffen, Unfällen und Fehlern?
Privacy	Schützt die KI die Privatsphäre und sonstige sensible Informationen?

Icons: www.flaticon.com

Framework zu KI-Zertifizierung

Drei Phasen des Lebenszyklus einer Anwendung

Design

- Die Konzeption und Architektur der KI-Anwendung, die sicherstellt, dass bestimmte Eigenschaften „per Konstruktion“ erfüllt sind, wie zum Beispiel Privacy-by-Design, Safety-by-Design, Prüfbarkeit-by-Design etc.

Entwicklung

Daten

- Die Auswahl, Augmentation, Vorverarbeitung der Trainings-, Test- und Inputdaten der KI-Anwendung als Grundlage für eine hohe Qualität der Anwendung.

KI-Komp.

- Die Auswahl einer Methode/ Algorithmus, das Training und Testen/Validieren des/der Modell(e), Aspekte zu Transparenz und Erklärbarkeit. Die Implementierung in Software.

Einbettung

- Die Einbettung der KI-Komponente in die Anwendung mit Fokus auf die Aspekte der Anwendung, deren Verhalten wesentlich auf der KI-Komponente beruht.

Betrieb

- Anwendungsbezogene Prüfung und Sicherstellung der Qualität der Modelle während des Betriebs. Überprüfbarkeit und Protokollierung des Verhaltens, das auf der KI-Komponente basiert.

Icons: www.flaticon.com

Fraunhofer IAIS

Framework zu KI-Zertifizierung

Struktur für den KI-Prüfkatalog

- Prüfkatalog bildet **typische Risiken, Kriterien, Maßnahmen** ab und bildet Framework zur Gesamtbeurteilung
- Entwickler **dokumentiert entlang der Vorgaben** des Prüfkatalogs
- Prüfer **beurteilt Plausibilität der Zielerreichung** (Basiszert.) bzw. führt **eigenständige Tests** (höhere Zert.) durch

Handlungsfeld

- **Prüfziel 1**
- **Prüfziel 2**
- . . .
- **Prüfziel N**

Für jedes Prüfziel

Risikoanalyse & Zielsetzung	Prüfziel-spezifische Risikoanalyse & Formulierung von Zielen, welche diese Risiken mitigieren
Kriterien zur Zielerreichung	Formulierung von Kriterien (KPIs oder qualitativ), welche die Zielerreichung erkennbar machen
Maßnahmen zur Zielerreichung	Dokumentation von Maßnahmen, welche zur Zielerreichung beitragen
Gesamtbewertung & Beurteilung	Würdigung aller Maßnahmen und Nachweis, dass die Zielsetzung erreicht wurde
Beherrschung der Dynamik	Formulierung von Prozessen, welche Risiken im Betrieb mitigieren und Anwendung beherrschbar machen

Fraunhofer IAIS

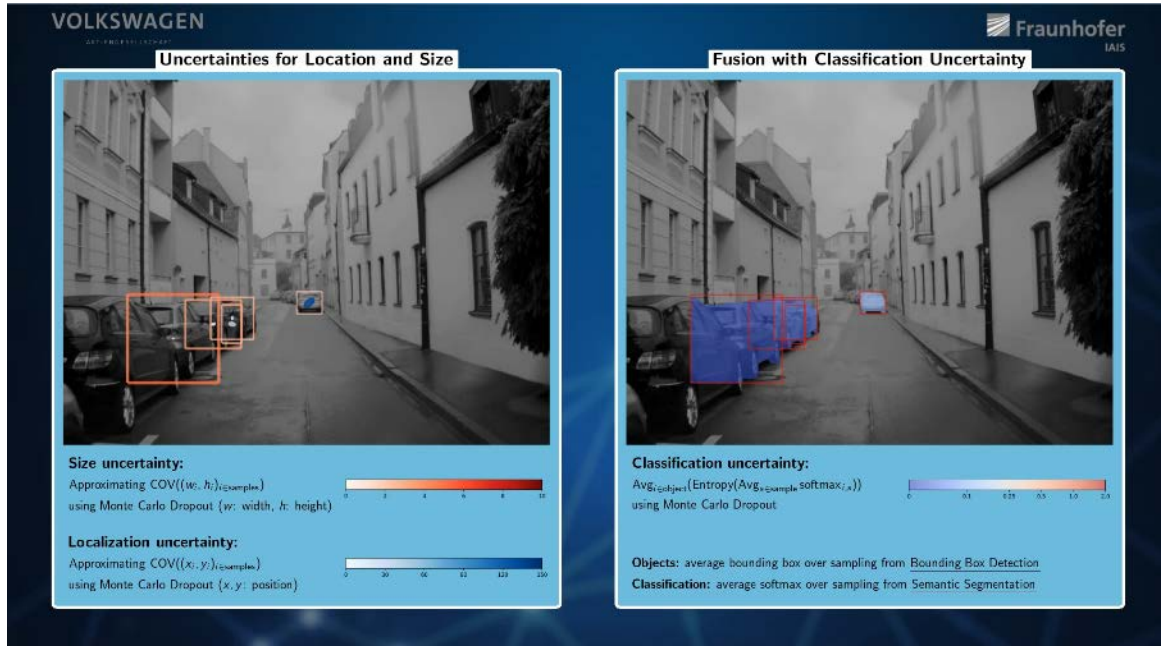
Vertrauenswürdige KI: Quantifizierung von Unsicherheit

Anwendungsfall autonomes Fahren

VOLKSWAGEN
AKTIENGESELLSCHAFT

Eine Quantifizierung von Unsicherheit erhöht die Qualität der Entscheidung autonomer Agenten.

Unsicherheit wird abgebildet durch eine vorgeschlagene Auswahl möglicher Antworten und korrespondierender Konfidenzintervalle.



- Höhere Geschwindigkeit des interaktiven Lernens
- Aktionen autonomer Agenten, die auf einem zu hohen Vertrauen beruhen, werden vermieden

Fraunhofer IAIS

Verteilte KI

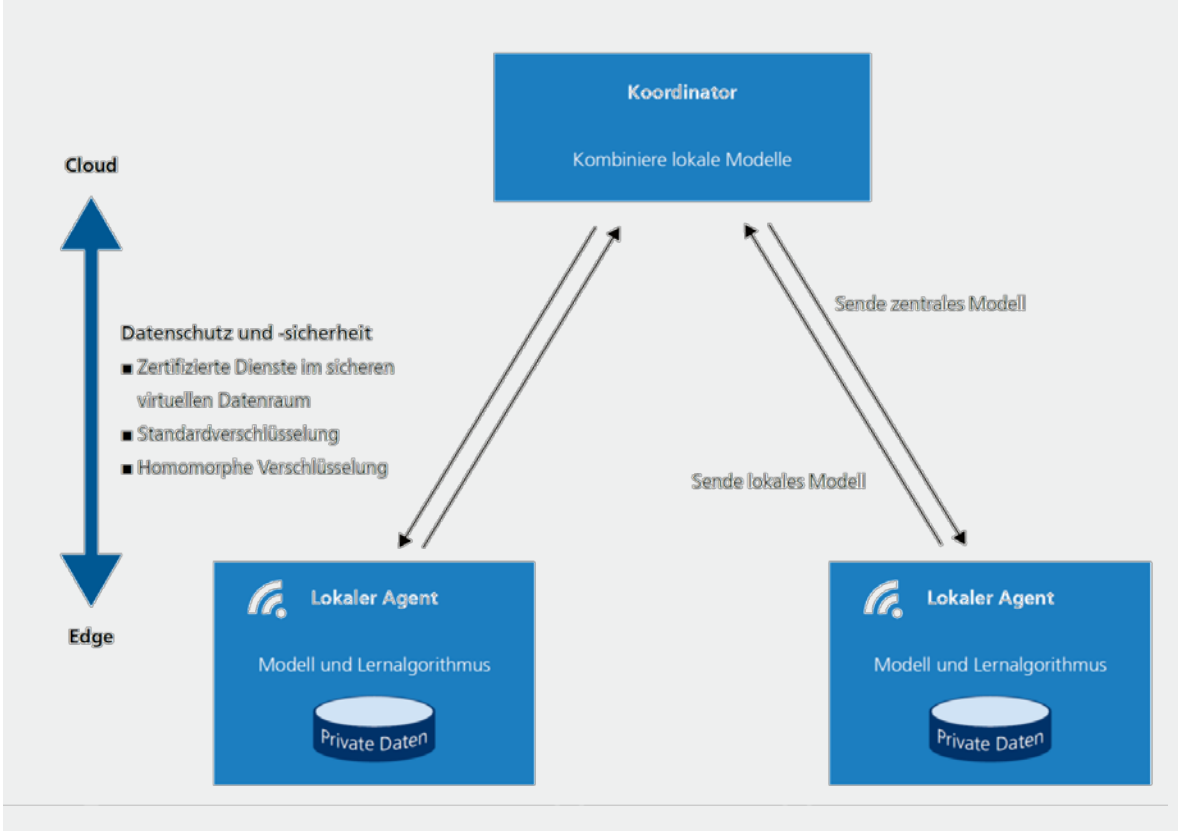
Maschinelles Lernen »on the edge«

Das Training neuronaler Netze findet bisher überwiegend in der Cloud statt

Herausforderungen:

- Datenschutz
- Hohe erforderliche Bandbreiten
- Hohe Kommunikationskosten
- Lange Reaktionszeiten

Lösung: Verteiltes Lernen – Training der neuronalen Netze auf den Endgeräten

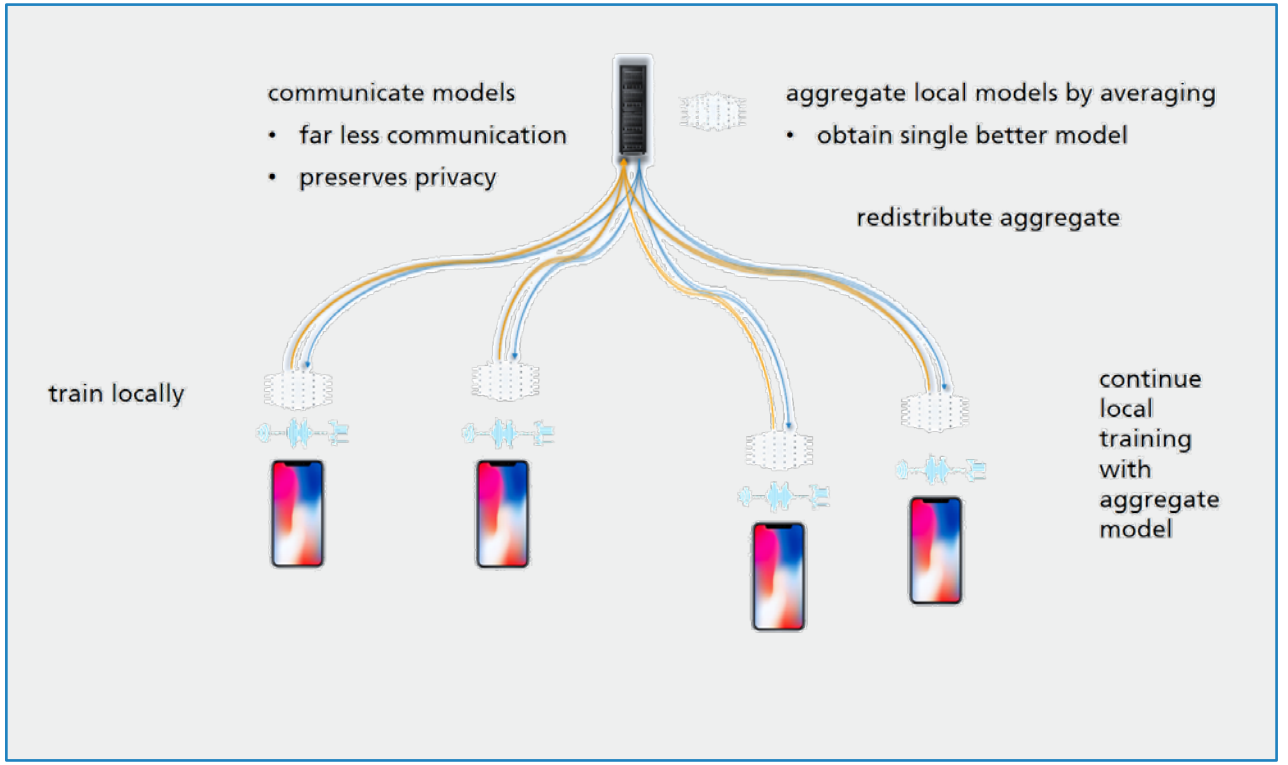


Quelle: Fraunhofer IAIS (2019), Maschinelles Lernen »on the edge« - Konzepte und Vorteile am Beispiel des autonomen Fahrens

Verteiltes Maschinelles Lernen

In-situ Processing – Von der Cloud zum Edge Computing & Fog Computing

- Tiefe neuronale Netze können mit minimierter Synchronisation erfolgreich verteilt trainiert werden
- Dies sichert Effizienz, Robustheit und Privatsphäre

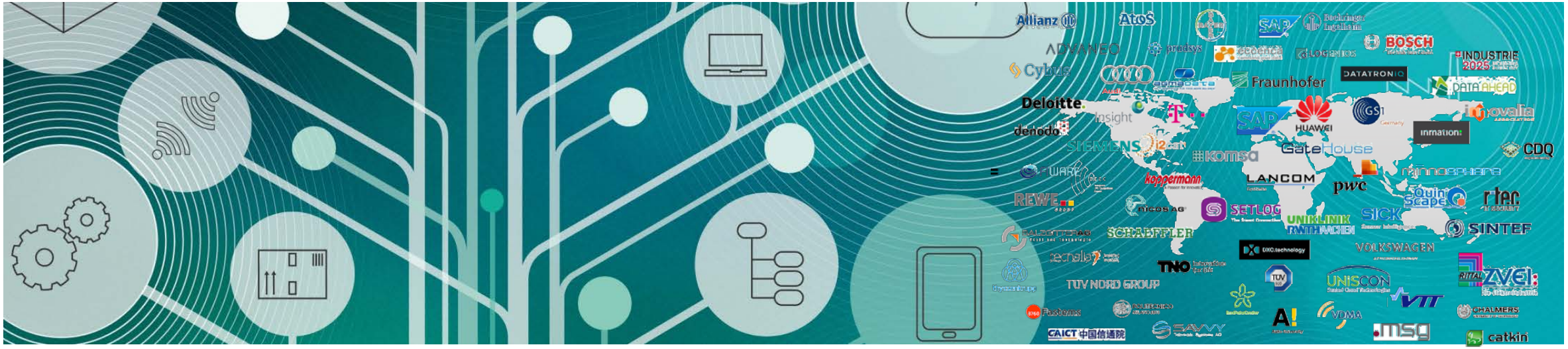


Quelle: Michael Kamp, Linara Adilova, Joachim Sicking, Fabian Hüger, Peter Schlicht, Tim Wirtz, Stefan Wrobel: Efficient Decentralized Deep Learning by Dynamic Model Averaging. ECML/PKDD (1) 2018: 393-409

International Data Space (IDS)

Auf dem Weg zum Marktstandard für digital souveränen Datenräume

- Grundlegende Fraunhofer Data Space Referenzmodellarchitektur für sichere und datensouveräne internetbasierte Wertschöpfung - Offenheit, Interoperabilität, Investitionsschutz
- Branchen-neutral: bereichsspezifische Data Spaces werden unterstützt und sind interoperabel
- International Data Space Association seit 2016 als Unternehmensplattform mit bereits über 95 Mitgliedsunternehmen



INTERNATIONAL DATA SPACES ASSOCIATION

Fraunhofer IAIS

95
Unternehmen und Organisationen

5
Arbeitsgruppen

20
Use Cases

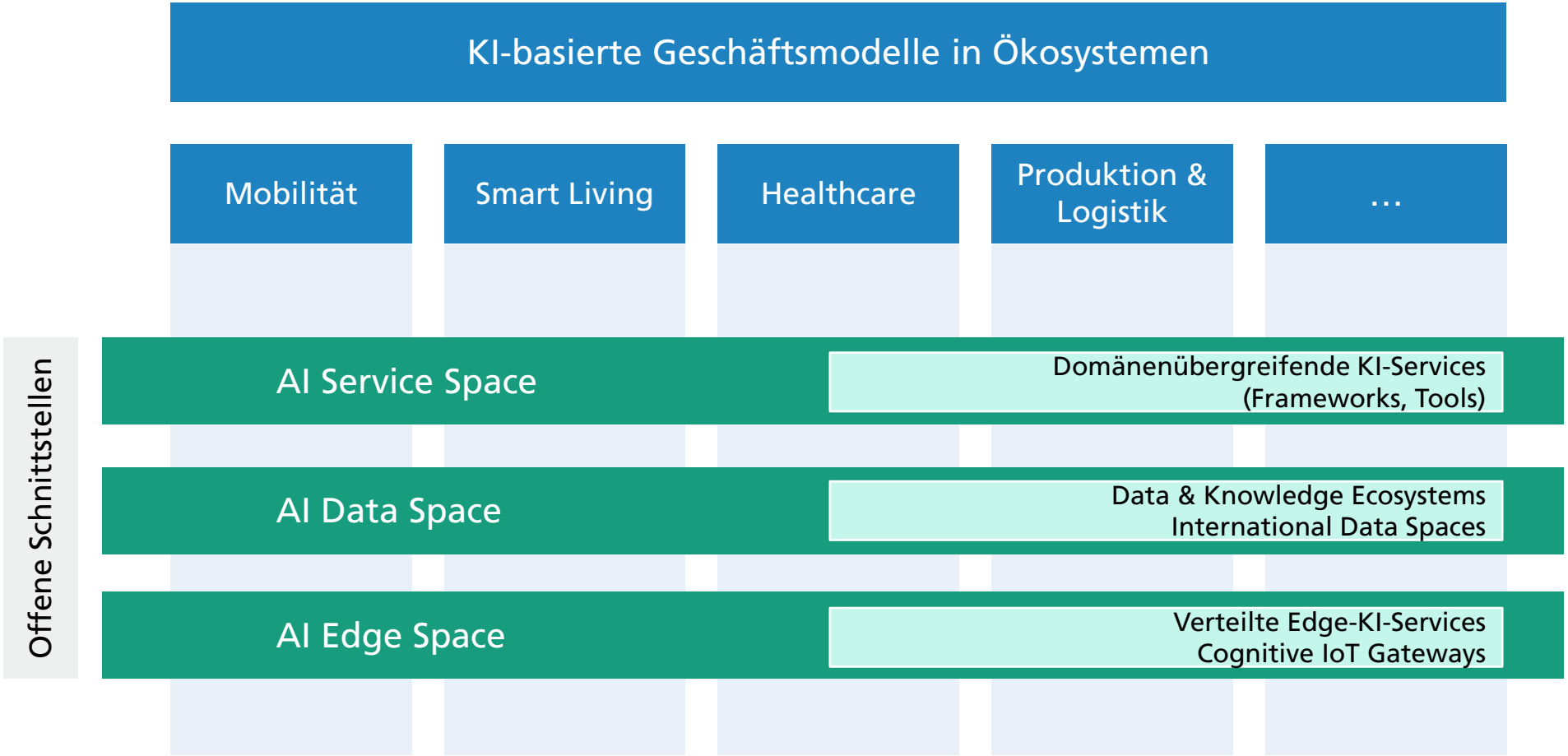
1
Ökosystem

INTERNATIONAL DATA SPACES ASSOCIATION

Fraunhofer IAIS

Ökosystem für Daten und KI-Anwendungen
KI-Infrastruktur »made in Germany«

Cognitive Internet Technologies
Fraunhofer Cluster of Excellence



Fraunhofer IAIS

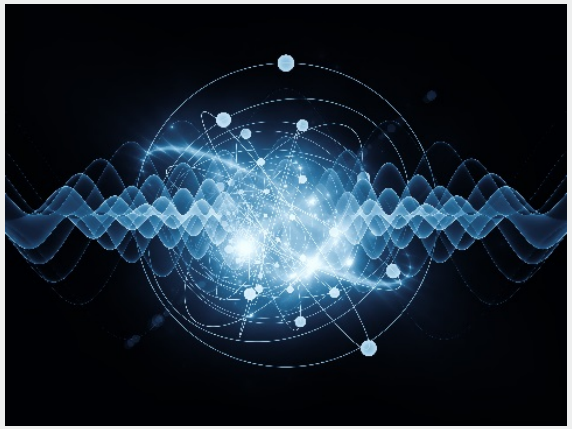
Quanten-KI

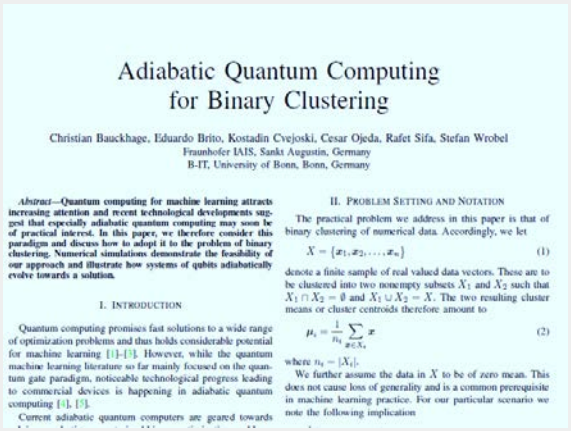
Quantum Machine Learning – Quantum AI

Neue Rechnerarchitekturen für exponentielle Probleme

Such- und Optimierungsprobleme könnten in subexponentieller Zeit gelöst werden

- Quantenrechner mit 30-72 vollen Qbits bzw. 2.048 simple Qbits verfügbar (IBM, Google, D-Wave)
- Industrielle Skalierung in Sicht
- Erste ML-Algorithmen weisen disruptives Potenzial nach
  (z.B. Quantum Clustering)





Adiabatic Quantum Computing for Binary Clustering

Christian Bauckhage, Eduardo Brito, Kostadin Cvejoski, Cesar Ojeda, Rafet Sifa, Stefan Wrobel
Fraunhofer IAIS, Sankt Augustin, Germany
B-IT, University of Bonn, Bonn, Germany

Bild: Andrew Ostrovsky / Adobe Stock

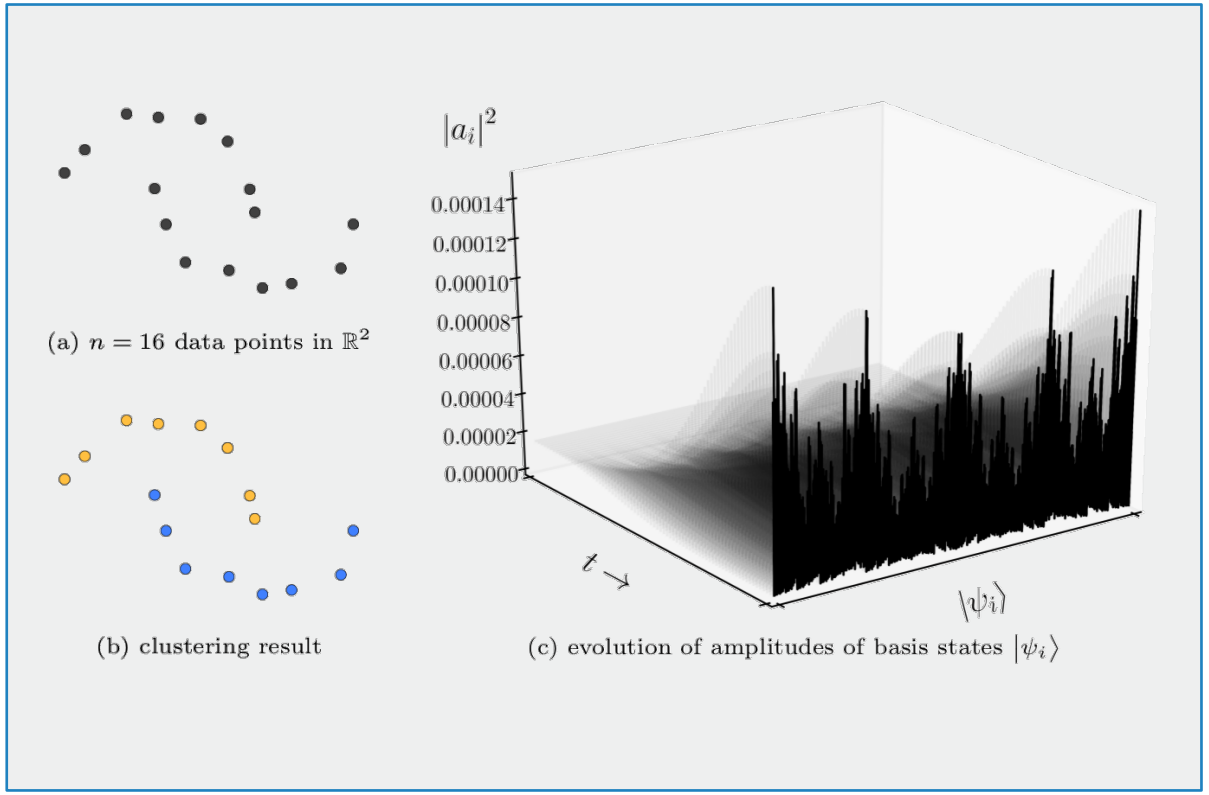
Quelle: Christian Bauckhage, Eduardo Brito, Kostadin Cvejoski, Cesar Ojeda, Rafet Sifa, Stefan Wrobel: Adiabatic Quantum Computing for Binary Clustering, 2017.

Fraunhofer IAIS

Quantum Computing für unüberwachtes Lernen

Adiabatic Quantum Computing for Kernel k = 2 Means Clustering*

- Eines der Kernthemen im Fraunhofer-Forschungszentrum Maschinelles Lernen (FZML)
- Das FZML konnte zeigen, wie unüberwachtes Lernen auf D-Wave Quantencomputern implementiert werden kann
- Dieser theoretische Ansatz wurde inzwischen von Physikern in Los Alamos praktisch auf einem Quantencomputer verifiziert



* Quelle: C. Bauckhage, C. Ojeda, R. Sifa, and S. Wrobel: Adiabatic Quantum Computing for Kernel k=2 Means Clustering. Proc. LWDA, 2018.

Wohin geht die Reise bei Künstlicher Intelligenz?

- Hybride KI – Daten, Wissen, Modelle, Lernen
- Vertrauenswürdige KI – Zertifizierung als Wettbewerbsvorteil
- Verteilte KI – Effizientes und sicheres Lernen on the edge
- Quanten-KI – Neue Rechnerarchitekturen am Horizont

WOHIN GEHT DIE REISE BEI KÜNSTLICHER INTELLIGENZ?
DEEP LEARNING AND BEYOND

Prof. Dr. Stefan Wrobel | Fraunhofer IAIS | 3.7.2019

Fraunhofer IAIS

Disclaimer

Zuständige Ansprechpartnerin: Gabriele Hühn
E-Mail: gabriele.huehn@iais.fraunhofer.de