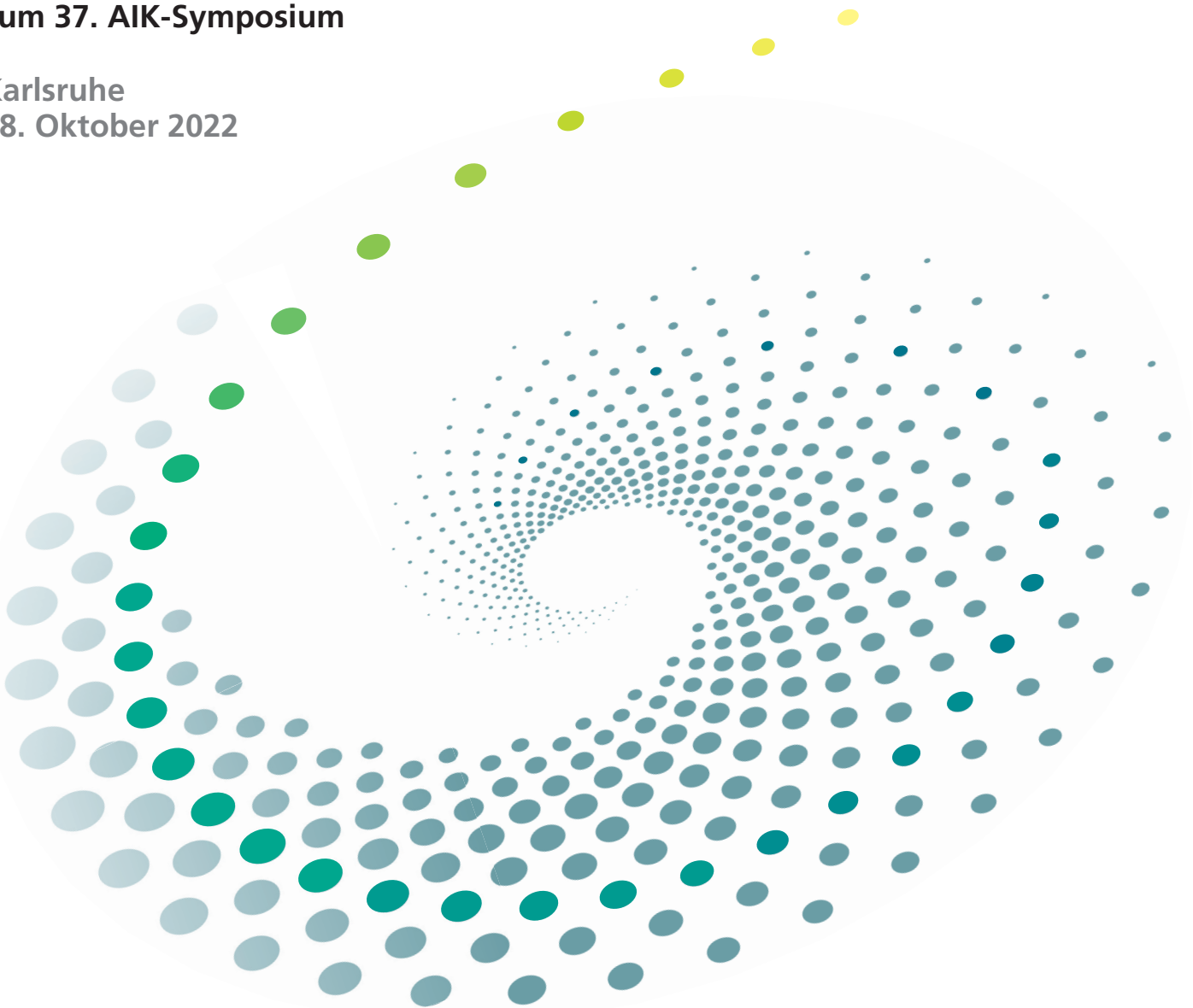


## Vertrauenswürdige Künstliche Intelligenz AIFB-Themenheft 2022

Einladung des Vereins  
Angewandte Informatik Karlsruhe e.V.  
zum 37. AIK-Symposium

Karlsruhe  
28. Oktober 2022





# Einladung des Vereins Angewandte Informatik Karlsruhe e.V. zum 37. AIK-Symposium „Vertrauenswürdige Künstliche Intelligenz“

Karlsruhe, 28. Oktober 2022  
ACHAT Hotel Karlsruhe City,  
Mendelssohnplatz – und im Internet

«Künstliche Intelligenz (KI), auch artifizielle Intelligenz (AI bzw. A. I.), englisch artificial intelligence, ist ein Teilgebiet der Informatik, das sich mit der Automatisierung intelligenten Verhaltens und dem maschinellen Lernen befasst. Der Begriff ist schwierig zu definieren, da es bereits an einer genauen Definition von „Intelligenz“ mangelt. (...)» Quelle: Wikipedia

KI breitet sich dennoch unaufhaltsam aus, denn ihr Potenzial für den Fortschritt ist immens. Das 37. AIK-Symposium thematisiert und diskutiert die Frage, wie künstlich intelligente Systeme gestaltet werden können und müssen, damit Menschen den Entscheidungen der Maschinen vertrauen können.

## Programm

- 13:00 Mitgliederversammlung AIK e.V.
- 14:15 Eröffnung und Begrüßung**  
Dr. Roland Schätzle, [adviiion GmbH](#),  
Vorsitzender Verein AIK e.V.  
Dr.-Ing. Michael Färber, [Institut AIFB](#),  
[Karlsruher Institut für Technologie \(KIT\)](#)
- 14:30 Wissensgraphen für die KI-gestützte Forschung der nächsten Generation**  
Dr.-Ing. Michael Färber, [Institut AIFB](#), [KIT](#)
- 15:00 “I'm sorry Dave, I'm afraid I can't do that” – Von Vertrauen, Verlässlichkeit und babylonischer Sprachverwirrung**  
Reinhard Heil, M.A., [Institut für Technikfolgenabschätzung und Systemanalyse \(ITAS\)](#), [KIT](#)
- 15:30 Kaffeepause**
- 16:15 Vertrauenswürdige KI-Frameworks: Von implizit bis explizit und überprüfbar**  
Dr. Clara Neppel, [Senior Director European Business Operations](#), [IEEE Technology Centre GmbH](#), [Wien](#)
- 16:45 Wollen wir einer KI mehr vertrauen als einem Thermostat? Modelle für die Zukunft in Richtung transformativer KI**  
Dipl.-Wi.-Ing. Jonas Andrulis, [Gründer und CEO](#), [Aleph Alpha GmbH](#), [Heidelberg](#)
- 17:15 Artificial Intelligence Evolution & Regulation**  
Marko Grobelnik, [Jožef Stefan-Institut](#), [Ljubljana](#),  
[Slowenien](#), [Gründer und CEO Quintelligence](#), [Ljubljana](#)
- 17:45 Abschlusspanel**  
Abschlusspanel mit allen Vortragenden
- 18:15 Preisverleihungen**
- anschließend **Gemeinsames Abendessen**

Wir planen unser 37. AIK-Symposium als Präsenzveranstaltung im ACHAT Hotel Karlsruhe City. Auf unserer Webseite halten wir Sie über Corona-Schutzmaßnahmen, die zum Veranstaltungszeitpunkt möglicherweise gelten, auf dem Laufenden. Alternativ können Sie an den Vorträgen auch gerne per Livestream teilnehmen.

## Anmeldung

Den Link zur Anmeldung und weitere organisatorische Informationen finden Sie unter: [www.aik-ev.de](http://www.aik-ev.de)

	Präsenz	Livestream
Teilnahmebeitrag für AIK-Mitglieder	€ 35	kostenlos
Teilnahmebeitrag für Nichtmitglieder*	€ 75	kostenlos
Beitrag zum Abendessen	€ 35	

\*Bei gleichzeitigem Vereinsbeitritt sind nur der Mitgliedsbeitrag von €25 für das Jahr 2022 sowie der Teilnahmebeitrag für Mitglieder zu entrichten. Unternehmen, die eine Firmenmitgliedschaft im AIK e.V. haben, können bis zu drei Firmenangehörige zum Mitgliederpreis entsenden.

Für die Teilnahme vor Ort empfehlen wir wegen des Platzkontingents eine möglichst frühzeitige Anmeldung. Grundsätzlich bitten wir um Ihre Anmeldung und um Überweisung des Beitrags bis zum 14.10.2022. Bei Rücktritt bis zum 21.10.2022 werden die entrichteten Gebühren erstattet.

Übernachtungsmöglichkeiten bestehen im ACHAT Plaza Karlsruhe City, Mendelssohnplatz, 76131 Karlsruhe, Tel. + 49 721 3717-0  
E-Mail: [karlsruhe-plaza@achat-hotels.com](mailto:karlsruhe-plaza@achat-hotels.com)  
(Sonderrate unter Stichwort „AIK-Symposium“)

**Sehr geehrte Damen und Herren,  
liebe Freunde und Förderer des  
Instituts AIFB,**



ein vertrauenswürdiges KI-System zeichnet sich durch vielseitige Kriterien aus, die während des gesamten Lebenszyklus des KI-Systems erfüllt sein sollten. Gemäß den 2018 veröffentlichten „Ethik-Leitlinien für eine vertrauenswürdige KI“, erarbeitet von einer unabhängigen Expertengruppe im Auftrag der Europäischen Kommission, sind unter anderem der Respekt der menschlichen Selbstbestimmung, die technische Robustheit und Sicherheit, die Fairness, sowie die Erklärbarkeit zentrale Elemente einer solchen „erweiterten“ KI.

Bis zur vollständigen Realisierung einer solchen KI dürften noch viele Jahre ins Land gehen. Für die Forschung gilt es deshalb, das Potenzial, den bereits bekannten Nutzen und die neuen Chancen gegen die Gefahren abzuwägen und passende Lösungen zu entwickeln. „Verantwortungsvolle KI versteht sich als umfassendes Paradigma, die Nutzerinnen und Nutzer von KI-Systemen und deren Interaktionen grundlegend in die Entwicklung und Bereitstellung von KI-Systemen miteinzubeziehen. Im Falle von KI-gestützten Empfehlungssystemen gilt es beispielsweise, Systeme so zu konstruieren, dass die Nutzerinnen und Nutzer die Gründe nachvollziehen können, die zu den Empfehlungen geführt haben“, so Michael Färber, Leiter der Forschungsgruppe Web Science am Institut AIFB. Er organisiert mit seiner Gruppe die inhaltliche Ausrichtung des 37. AIK-Symposiums „Vertrauenswürdige Künstliche Intelligenz“, zu dem wir Sie herzlich einladen.

KI wird bereits in vielen Bereichen der Informatik eingesetzt. Erwähnenswert ist daher, dass auch andere Forschungsgruppen am Institut an Themen, Projekten und Herausforderungen einer vertrauenswürdigen KI arbeiten. Auf den Seiten 8/9 dieses Themenheftes geben sie einen kurzen Überblick über ihre Forschung und Entwicklung. Auf den Seiten 10/11 stellen sie jeweils ein spannendes Projekt aus ihrer Gruppe vor, und wir berichten über die Highlights des Jahres, das seit dem Festkolloquium zum 50-jährigen Bestehen des Instituts vergangen ist. Das Festkolloquium war gleich der erste Höhepunkt. Hoch interessante Vorträge ehemaliger AIFB-Institutsangehöriger reflektierten einen großartigen Effekt der Lehre und Forschung an unserem Institut.

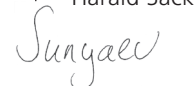
Der Erfolg geht weiter: Auch im vergangenen Jahr gab es für Beiträge, die Angehörige unseres Instituts zur Lehre und Forschung sowie zum Austausch in der wissenschaftlichen Fachgemeinschaft leisteten, wieder schöne Anerkennung und Auszeichnungen. Mehr dazu in den Zahlen und Fakten in diesem Heft.

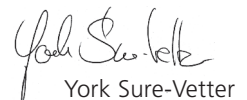
Und es gibt noch einen Grund zu großer Freude: Das AIFB wächst! Sanja Lazarova-Molnar tritt am ersten August in der Nachfolge von Hartmut Schmeck die Professur „Angewandte Informatik für Energiesysteme“ an. Sie kommt vom Maersk Mc-Kinney Moller-Institut der University of Southern Denmark, Odense. Alexey Vinel übernimmt zwei Monate später die neue Professur „Kooperative autonome Systeme im Reallabor“. Er war zuletzt an der Universität Passau. Beiden ein herzliches Willkommen in Karlsruhe!

**Wir danken Ihnen für Ihr Interesse an unserer  
Lehre und Forschung und würden uns freuen,  
Sie beim 37. AIK-Symposium begrüßen zu können.**

  
Andreas Oberweis

  
Harald Sack

  
Ali Sunyaev

  
York Sure-Vetter

  
Melanie Volkamer

  
J. Marius Zöllner

  
Michael Färber



# Vertrauenswürdige KI – eine würdige KI?

Dr.-Ing. Michael Färber  
Nicholas Popovic

**Insbesondere die technologischen Fortschritte der letzten 10 Jahre, welche zentral durch das Deep-Learning vorangetrieben wurden, erlauben den Einsatz von KI-Systemen in immer komplexeren Einsatzgebieten. Mittlerweile sind Anwendungsszenarien, welche noch vor kurzer Zeit nur in Science-Fiction Romanen zu finden waren, in greifbarer Nähe. Die Vertrauenswürdigkeit von künstlicher Intelligenz (KI) beschäftigt deshalb nicht nur die Forschung, sondern auch die Gesellschaft im Allgemeinen.**

Die Antwort auf die Frage, welche spezifischen Kriterien eine KI erfüllen muss, um als vertrauenswürdig eingestuft zu werden, ist alles andere als trivial. Auch die Frage, ob ein KI-System die gleichen oder sogar höhere Anforderungen als ein menschlicher Entscheidungsträger erfüllen muss, ist nicht einfach zu beantworten und wird oftmals stark diskutiert. Grundsätzlich werden von vertrauenswürdigen Entscheidungsträgern – ob Mensch oder Maschine – mannigfaltige Eigenschaften wie Fairness, Transparenz, Einhaltung von Datenschutz, Resilienz und Nachvollziehbarkeit erwartet. Daraus folgt für die Forschung und die Politik die gesellschaftliche Verantwortung, die Entwicklung von Systemen voranzutreiben, die diese Kriterien erfüllen. Neben technischen Herausforderungen rücken hierbei auch soziologische und psychologische Faktoren in den Vordergrund.

## Sprachmodelle und Weltmodelle

Ein wichtiger Grundstein für die Vertrauenswürdigkeit von KI-Systemen liegt in ihrer Erklärbarkeit, also der Nachvollziehbarkeit der von ihr getroffenen Entscheidungen. Voraussetzung hierfür ist, dass ein solches System ein Grundverständnis von der Realität aufweist, welches dem menschlichen Verständnis ähnelt oder sich aus diesem heraus nachvollziehen lässt. Hierfür sind die Verarbeitung von natürlicher Sprache und Wissensrepräsentation wichtige Bereiche der KI-Forschung. Besonders für die langfristige Vision einer generellen KI, oder Artificial General Intelligence (AGI), stehen diese Forschungsbereiche aktuell im Mittelpunkt.

Die Idee, Sprache zu modellieren, um damit geschriebene und gesprochene Sprache automatisch verarbeiten zu können, ist nicht neu. Die Sprachmodelle haben sich aber dank neuer Architekturen, basierend auf Deep-Learning, und der stark gestiegenen Rechenmöglichkeiten sowie der Verfügbarkeit leistungsstarker Hardware erheblich weiterentwickelt. Durch statistische Methoden sind aktuelle Modelle in der Lage, Zusammenhänge zwischen Wörtern zu lernen, die weit über Grammatik und Syntax hinausgehen und zu einer internen Repräsentation von komplexen semantischen Zusammenhängen führen. Solche Modelle bauen durch das automatische „Lesen“ von Texten im großen Stil (z.B. Webseiten, Wikipedia, Publikationen, Patente, Nachrichtentexte) ein Modell der Beziehungen zwischen Begriffen und damit ein „Weltbild“ auf.

Schon jetzt werden diese Modelle in vielseitigen Anwendungen genutzt, etwa zur automatischen Übersetzung von Texten, zur Beantwortung von Fragen durch Chatbots, zur automatisierten Erhebung von themenbezogenen Stimmungsbildern in sozialen Medien oder zur Zusammenfassung von Texten aller Art. Diese kleine Auswahl von schon jetzt realisierten Anwendungen skizziert das Potenzial von Sprachmodellen. Der mögliche transformative Wert der Sprachmodelle für Gesellschaft, Ökonomie und Technik erscheint vor diesem Hintergrund enorm.

## Defizite und Herausforderungen von Sprachmodellen

Wenngleich die aktuellen Sprachmodelle ihren Vorgängern weit voraus sind, steht die Erforschung existierender Sprachmodelle und die Entwicklung neuer im Hinblick auf ihre Nutzbarkeit für vertrauenswürdige KI noch auf mehreren Ebenen vor großen Herausforderungen. Diese leiten sich aus den charakteristischen Eigenschaften der Deep-Learning Methoden ab. Dies betrifft zum einen das Training, also den Lernprozess, in dem die aus Hunderten von Millionen



Parametern bestehenden Modelle durch die Verarbeitung von Unmengen von Textdokumenten ihr „Weltverständnis“ formen. Dieser Prozess ist sehr langwierig, rechenintensiv und alles andere als trivial, was zu einer niedrigen Reproduzierbarkeit führt. Zum anderen sind die resultierenden Sprachmodelle in der Erzeugung ihrer Ausgaben nicht nachvollziehbar, was, wie erläutert, ein großes Hindernis bezüglich ihrer Vertrauenswürdigkeit darstellt. Des Weiteren stützt sich das „Weltbild“ eines Sprachmodells per Definition auf statistische Zusammenhänge, die aus den Trainingsdaten ermittelt werden. Dies hat zur Folge, dass etwaige Vorurteile oder gar Fehlinformationen aus dem Trainingsdatensatz übernommen werden. Klassische Lösungen, wie eine manuelle Vorverarbeitung zur Bereinigung der Datensätze, sind aufgrund der benötigten Masse von Textdaten jedoch schlichtweg nicht umsetzbar. Aus der mangelnden Vorverarbeitung und Filterung der Trainingsdaten entstehen zudem datenschutzrechtliche Bedenken: Persönliche Informationen wie etwa Adressen können im Trainingsdatensatz auftauchen und später potenziell abgerufen werden.

Ein Ansatz zur Bewältigung dieser Schwierigkeiten wird mit der Erforschung hybrider KI-Methoden, Neuro-Symbolic AI genannt, verfolgt. Solche Methoden haben zum Ziel, Modelle zu schaffen, die Informationen sowohl auf symbolischer Ebene als auch auf subsymbolischer Ebene verarbeiten können. Eine Informationsverarbeitung auf symbolischer Ebene ermöglicht nachvollziehbare, auf ihre Sinnhaftigkeit überprüfbare und gezielt anpassbare Systeme. In Kombination mit der aus subsymbolischen Systemen resultierenden Skalierbarkeit, Robustheit, und universellen Einsetzbarkeit sollen die oben genannten Schwächen ohne einen Verlust der Vorteile von Deep-Learning-Methoden behoben werden. Solche neuro-symbolischen KI-Methoden stehen noch in ihren Anfängen, werden jedoch von vielen als vielversprechender, von manchen gar als unausweichlicher nächster Schritt auf dem Weg zur vertrauenswürdigen KI gesehen.

### **Wissensgraphen**

Zentraler Bestandteil symbolischer Informationsverarbeitung ist eine geeignete symbolische Repräsentation der Informationen. Wissensgraphen (engl. Knowledge Graphs, kurz KGs) sind eine etablierte Form der Wissensmodellierung. Sie stellen Informationen als Netzwerke von Entitäten dar, die in Relationen zueinander stehen, um das Wissen symbolisch zu repräsentieren. Wissensgraphen werden manuell und auch vermehrt semi-automatisch oder vollautomatisch erstellt. Mit Wikidata, DBpedia, aber auch domänenspezifischen Wissensgraphen wie dem Microsoft Academic Knowledge

Graph (MAKG), der Informationen über wissenschaftliche Publikationen umfasst, bestehen bereits einige große Datenbanken, durch die Weltwissen in strukturierter Form abrufbar ist. Auch wenn bereits viele Ansätze erforscht wurden, um Wissensgraphen mit Deep-Learning-Methoden zu verknüpfen, ist der Sprung zur hybriden KI noch nicht gelungen. Anerkannt ist jedoch, dass eine richtige Verknüpfung das Training von Sprachmodellen transparenter und effizienter machen und ein gezieltes, manuelles Eingreifen in das interne Wissen einer KI ermöglichen könnte. Somit ist die Forschung an und mit Wissensgraphen ein zentraler Bestandteil der Entwicklung vertrauenswürdiger KI.

### **Mensch-Maschine-Interaktion**

Wissensgraphen sind eine wichtige und effektive Technologie zur Modellierung und Speicherung von Wissen. Sie alleine reichen jedoch nicht aus, um eine vertrauenswürdige KI zu erzeugen. Vertrauen wird oft als zweiseitiger Prozess gesehen, in dem Zweifel durch nachvollziehbare Argumente aus dem Weg geräumt werden können und sollen. Die KI-gestützte Entscheidungsfindung muss also nicht nur durch die statische Nutzung einer Wissensdatenbank, sondern als Ideenaustausch zu einem natürlichen und berechtigten Vertrauen führen. Daher werden oft Dialog- oder Frage-Antwort-Systeme (QA-Systeme) als Erweiterung von KI-gestützten Entscheidungsverfahren genutzt, sofern es sich nicht um zeitkritische Situationen, wie etwa beim autonomen Fahren, handelt. Strukturierte und in einem gemeinsamen „Weltmodell“ fundierte Erklärungen sind ein essentieller Bestandteil des Aufbaus von Vertrauen. So sollte etwa ein Arzt, der eine von einer KI erzeugte Empfehlung bezüglich einer Behandlung bekommt, diese nicht anwenden, ohne dass für ihn nachvollziehbar ist, wie es zu dieser kam. Es ist möglich, dass etwa unvollständige Informationen oder fehlendes Hintergrundwissen zu einer falschen Empfehlung führen. Die schlussendliche Entscheidung darf also nur in begründetem, nicht in blindem Vertrauen getroffen werden.



### **Wie weit sind wir?**

Konzeptionell wurde vertrauenswürdige KI bereits vielfach durchdacht und wissenschaftlich und wissenschaftspolitisch diskutiert. Insbesondere auf EU-Ebene wurden Ethik-Leitlinien für eine vertrauenswürdige KI als grundlegende Imperative und Kernanforderungen entworfen und veröffentlicht. Auf Bundesebene werden diverse Forschungsprojekte gefördert, die zumindest gewisse Aspekte von vertrauenswürdiger KI in den Mittelpunkt rücken, wie etwa die Erforschung der Erklärbarkeit von KI-Modellen, inklusive soziologischer und psychologischer Aspekte wie dem kritischen Hinterfragen von Erklärungen und Gegebenheiten. Aber auch Recht und Philosophie sind relevante Disziplinen, die es einzubinden gilt. Erste formale Standards, die die Einhaltung bestimmter ethischer Anforderungen durch KI-Systeme messbar machen sollen, wie etwa die Spezifikation VDE Spec 90012, wurden bereits unter Mitarbeit des KIT entwickelt.

Zu den informationstechnischen Fragen kommen jedoch zahlreiche gesellschaftliche, regulatorische und technische Fragen bezüglich der vertrauenswürdigen KI hinzu, die ungeklärt sind. Auch die Frage nach der Haftbarkeit für eine fehlerhafte Implementierung von Aspekten einer vertrauenswürdigen KI und daraus möglicherweise folgende, finanzielle Schäden ist noch unbeantwortet.

Die Entwicklung vertrauenswürdiger KI wird daher von zahlreichen Forschungsbereichen innerhalb, aber auch außerhalb der Informatik geprägt sein. Genau wie das anvisierte begründete Vertrauen zwischen Gesellschaft und KI wird auch der Weg dorthin erst durch den offenen Austausch von Ideen möglich sein. Die technischen Fortschritte des vergangenen Jahrzehntes haben das gewaltige Potenzial aufgezeigt, welches das maschinelle Lernen in der großflächigen Verarbeitung von Informationen mit sich bringt. Für das nächste Jahrzehnt gilt es, dieses Potenzial verantwortungsvoll zu nutzen.

**KI-Systeme erkennen anhand antrainierter, beispielhafter Datenmodelle in riesigen Informationsmengen auftretende, vergleichbare Datenmuster und Strukturen, können ihre Funde interpretieren und das Ergebnis dem Menschen als Entscheidungshilfe präsentieren. Teilweise entscheiden sie gleich selbst, wenn es der Anwendungsfall erlaubt.**

**Beim 37. AIK-Symposium stellen fünf Vortragende KI-Anwendungen, Entwicklungsideen und Ansätze auf dem Weg zu einer vertrauenswürdigen KI vor. Sie erläutern den Stand der Technik, gehen auf technische, politische und gesellschaftliche Herausforderungen und Lösungsansätze ein und diskutieren, wie KI-Systeme so gebaut werden können, dass die Entscheidungsfindung der Maschinen für Menschen plausibel und somit vertrauenswürdig wird.**

### Wissensgraphen für die KI-gestützte Forschung der nächsten Generation

*Michael Färber*

In Anbetracht der ständig steigenden Zahl wissenschaftlicher Veröffentlichungen wird in diesem Vortrag ein Ansatz für ein neues Assistenzsystem vorgestellt, um Forscherinnen und Forscher bei wissenschaftlichen Recherchen zu unterstützen. Im Zentrum steht die Entwicklung neuer KI-Methoden, mit denen die in Publikationen enthaltenen Kernaussagen in großem Umfang automatisch extrahiert und in Form von Wissensgraphen konzentriert abgespeichert werden. Die Extraktionsquellen bleiben verknüpft, die ausgewerteten Einzelpublikationen auf diese Weise verfügbar. Solche Wissensgraphen können z.B. eingesetzt werden, um sich passende wissenschaftliche Methoden, Datensätze oder anderweitige Aspekte zu einem Forschungsvorhaben vorschlagen zu lassen. Die Anforderungen an ein solches Assistenzsystem sind hoch: Wo einfache, Schlüsselwort-gestützte Ansätze nur oberflächliche Zusammenhänge aufzeigen, haben neue KI-Methoden die Möglichkeit, tiefere Zusammenhänge zu finden. Das Fehlen einer präzisen und nachvollziehbaren Erklärung kann jedoch dazu führen, dass eine Empfehlung trotzdem nicht verwertbar ist. Insgesamt zeigt dieser Vortrag das Potenzial und die derzeitigen Grenzen der Forschung im Hinblick auf die KI-gestützte Forschung der nächsten Generation.



*Dr.-Ing. Michael Färber ist Vertretungsprofessor am Karlsruher Institut für Technologie (KIT), wo er die Forschungsgruppe Web Science am Institut AIFB leitet. Seine Forschungsschwerpunkte sind die Verarbeitung natürlicher Sprache, maschinelles Lernen und Wissensrepräsentation. Er ist Autor von mehr als 75 begutachteten wissenschaftlichen Publikationen und als Projektleiter in mehreren nationalen und internationalen Forschungsprojekten tätig.*

### “I'm sorry Dave, I'm afraid I can't do that” – Von Vertrauen, Verlässlichkeit und babylonischer Sprachverwirrung

*Reinhard Heil*

Vertrauen und Vertrauenswürdigkeit sind Wörter, denen wir unmittelbar eine, in den wohl allermeisten Fällen positive, Bedeutung beimessen. Vertrauen ist ein Urverhältnis: Dass wir vertrauen ist nicht die Ausnahme, sondern die Regel. Vertrauen ist der Kitt, der unsere Gesellschaft zusammenhält. Solange vertraut wird, bleibt das Vertrauen unthematized. Wir sprechen meist erst dann von Vertrauen und Vertrauenswürdigkeit, wenn das Vertrauen bereits in eine Krise geraten ist. Dass wir so viel über vertrauenswürdige KI reden, ja überhaupt darüber reden, weist darauf hin, dass wir ein Problem mit der sogenannten KI haben. KI scheint eine Sonderstellung einzunehmen. Warum reden wir nicht von vertrauenswürdiger Biotechnologie oder vertrauenswürdiger Gentechnik? Warum überhaupt „vertrauenswürdige KI“ und nicht „verlässliche KI“? Einfach ein Übersetzungsproblem mit Folgen? Oder ein gewichtiger semantischer Unterschied, der vielleicht sogar auf ein sich grundlegend veränderndes Mensch-Technik-Verhältnis hinweist? Oder verweist die Rede von vertrauenswürdiger KI doch nicht auf eine bestehende Vertrauenskrise, sondern auf eine „ai-phobia-phobia“?



*Reinhard Heil, M.A., studierte Philosophie, Soziologie und Literaturwissenschaft an der TU Darmstadt. Seit 2010 ist er wissenschaftlicher Mitarbeiter am Institut für Technikfolgenabschätzung und Systemanalyse (ITAS) am KIT, arbeitet dort in der Forschungsgruppe „Digitale Technologien und gesellschaftlicher Wandel“. Seine aktuellen Forschungsschwerpunkte sind die gesellschaftlichen Folgen Künstlicher Intelligenz, Vertrauenskonzepte und Transhumanismus.*




**Vertrauenswürdige KI-Frameworks:  
Von implizit bis explizit und überprüfbar**  
*Clara Neppel*

Der vorgeschlagene EU-Rechtsakt zur KI verfolgt einen risikobasierten Ansatz. Dabei wird das Risiko selbst neu definiert. Neben Dimensionen wie Safety und Security werden nun auch menschenrechtliche Risiken berücksichtigt. Um das Wohlbefinden von Mensch und Umwelt zu maximieren, müssen wir folglich auch die Erfolgsmaßstäbe neu definieren. Wie können diese Ziele erreicht werden und welche Möglichkeiten gibt es, sie während des gesamten KI-Lebenszyklus zu berücksichtigen, vom Design bis zum Einsatz? Darüber hinaus erfordert ethische Sicherheit ein globales Ökosystem, in dem sich Organisationen nicht nur dazu verpflichten, menschliche Werte, Würde und Wohlergehen zu wahren, sondern dies auch demonstrieren können, wenn es der spezifische Kontext erfordert, in dem sie tätig sind. Dieser Vortrag konzentriert sich auf mögliche Instrumente wie Standards und Zertifizierungssysteme sowie entsprechende Governance-Mechanismen eines solchen Human Centered AI-Ecosystems.



*Dr. Clara Neppel ist als Senior Director European Business Operations bei der IEEE Technology Centre GmbH in Wien tätig für die IEEE-Aktivitäten und -Präsenz in Europa. Sie hat an der TU München in Informatik promoviert und an der Universität Straßburg einen Masterabschluss in Intellectual Property Law erworben. Sie ist Mitglied des Aufsichtsrats von EIT Digital, des wissenschaftlich-industriellen Beirats der Research Studios Austria FSG, des Zuse Schools-Beirats im DAAD Programm „Konrad Zuse Schools of Excellence in Artificial Intelligence“ sowie des Independent Advisory Board des UK RI Centre for Doctoral Training in Accountable, Responsible and Transparent AI.*



**Wollen wir einer KI mehr vertrauen  
als einem Thermostat?  
Modelle für die Zukunft in Richtung  
transformativer KI**  
*Jonas Andrusis*

Mit neuen überraschenden und beeindruckenden Ergebnissen aus dem Training riesiger Sprach- und Bildmodelle werden Risiken und Verantwortungen neu diskutiert. Diese Modelle scheinen dabei auf den ersten Blick fast menschlich in ihren Fähigkeiten. Welche Ansprüche können wir an Systeme formulieren, die vielleicht schon in wenigen Jahren einen Großteil der informationsbasierten menschlichen Arbeit radikal transformieren? Welche Verantwortung haben Wissenschaft, Industrie und Politik? Anhand von Beispielen aus der aktuellen Forschung und Entwicklung von Aleph Alpha blicken wir auf Risiken und die Bedeutung von Vertrauen in der Mensch-Maschine-Interaktion.



*Dipl.-Wi.-Ing. Jonas Andrusis war nach seinem Studium des Wirtschaftsingenieurwesens am KIT als Berater tätig. Er gründete erfolgreich zwei KI-Softwareunternehmen. Nach drei Jahren als Engineering Manager für KI-Forschung in Apples Special Projects Group und später Siri gründete er 2019 mit der Aleph Alpha GmbH in Heidelberg das dritte KI-Start-Up. Seitdem hat das junge Unternehmen insgesamt 28 Mio. Euro Risikokapital eingesammelt für sein Ziel, unabhängige KI-Spitzenforschung zu etablieren und Basistechnologie für eine neue Art von KI-Wertschöpfungsprozess zu betreiben.*



**Artificial Intelligence  
Evolution & Regulation**  
*Marko Grobelnik*

The talk will cover the current efforts in the regulation of the field of Artificial Intelligence across the key international initiatives including European Union, OECD, Council of Europe, UNESCO and others. We will highlight the key elements of regulation frameworks such as definition, principles, classification, risk based approaches leading towards technology certification. In the second part we will present the "OECD AI Policy Observatory" observing the evolution of the field of Artificial Intelligence across several dimensions including academia, projects, patents, job market, media, perception, incidents and policies. Finally, we will present the impact and propagation of AI innovations throughout the research and commercial ecosystem.



*Marko Grobelnik is a researcher in the field of Artificial Intelligence (AI). He co-leads Artificial Intelligence Lab at Jožef Stefan Institute, Ljubljana, Slovenia. He cofounded UNESCO International Research Center on AI (IRCAI), and is the CEO of Quintelligence.com. He represents Slovenia in OECD AI Committee (ONE AI), in Council of Europe Committee on AI (CAHA), and Global Partnership on AI (GPAI). In 2016 Marko Grobelnik became Digital Champion of Slovenia.*



- 6 Forschungsgruppen
- 8 Promotionen
- 118 Publikationen
- 6 Auszeichnungen
- 6 Projekte
- 1 apl. Professor
- 30 Vorlesungen
- 3317 Prüfungen
- 50 Jahre, 220 Seiten
- 62 Mitarbeiterinnen und Mitarbeiter

## 6 Forschungsgruppen

### Angewandte technisch-kognitive Systeme (ATKS)

Professor Dr.-Ing. J. Marius Zöllner

ATKS untersucht Technologien und Anwendungen maschineller Intelligenz. Die Forschung und Entwicklung der Gruppe adressiert KI-Verfahren wie maschinelle Wahrnehmung, Situationsverstehen und kooperative Verhaltensentscheidung. Methoden des maschinellen Lernens und der probabilistischen Inferenz werden für alle Komponenten erforscht und auf neue technische Systeme wie autonome Serviceroboter, autonome Fahrzeuge oder Assistenzsysteme mit kognitiven Fähigkeiten angewandt. Im DFG-Projekt *Situationsverstehen und semantische Manöverplanung unter Berücksichtigung von Unsicherheiten für kooperative Fahrzeuge* beispielsweise werden probabilistische und lernende Methoden zur Situationserfassung und -prädiktion, sowie der Verhaltensentscheidung im realen Verkehr unter ausdrücklicher Berücksichtigung von Kooperation erforscht.

Das Testfeld *Autonomes Fahren Baden-Württemberg* – ein Reallabor für die Entwicklung und Erprobung neuer Mobilitätskonzepte – bietet eine Infrastruktur, um zukunftsorientierte Mobilitätslösungen für Individualverkehr und Öffentlichen Personennahverkehr (ÖPNV) zu erarbeiten. Es wurde unter der Leitung von J. Marius Zöllner mit drei Kommunen und vier weiteren Forschungseinrichtungen aufgebaut. Firmen und Forschungseinrichtungen können hier in realer, mit intelligenter Technik ausgestatteter Umgebung ihre Technologien und Dienstleistungen rund um das vernetzte und automatisierte Fahren im Alltagsverkehr erproben.

<https://taf-bw.de/>

### Betriebliche Informationssysteme (BIS)

Prof. Dr. Andreas Oberweis

BIS befasst sich mit Fragestellungen an der Schnittstelle zwischen Software Engineering und Business Process Engineering. Querschnittsthemen wie Digitalisierung, Informationssicherheit und Nachhaltigkeit werden in verschiedenen Forschungsschwerpunkten mit aktuellen Technologien und Anwendungsfeldern verknüpft. Mit Sprachen, Methoden und Werkzeugen der Angewandten Informatik entwickelt die Forschungsgruppe Lösungen für aktuelle Herausforderungen in Wirtschaft, Verwaltung und Gesellschaft.

Die aktuellen Projekte und Kooperationen spiegeln die breiten Anwendungsmöglichkeiten der F&E-Arbeiten der Gruppe BIS wider. In dem vom Land Baden-Württemberg geförderten KIT-Projekt „Science Data Center für Molekulare Materialforschung“ (MoMaF) werden Werkzeuge zur Verbesserung der Metadatenqualität entwickelt. Im Projekt „Software-Defined Car“ (SofDCar) setzt BIS in Zusammenarbeit mit namhaften Unternehmen Process-Mining-Methoden zur fahrzeugübergreifenden Prozessgestaltung ein. Für mehr Transparenz über den CO<sub>2</sub>-Rucksack von IT-Hardware sorgt das vom Bundesministerium für Umwelt (BMU) geförderte Projekt „Scope3Transparent“. BIS setzt zudem KI-Methoden zur genderspezifischen Analyse von Studienverläufen ein und entwickelt digitale Workflows zur Verbesserung der Lehrorganisation.

### Critical Information Infrastructures (cii)

Prof. Dr. Ali Sunyaev

cii beschäftigt sich mit der Erforschung zuverlässiger, sicherer, zweckorientierter und dezentraler Informationssysteme mit Bezug zu kritischen Informationsinfrastrukturen, Health-IT-Anwendungen und Digital Health, Cloud- und Edge-Computing-Diensten, Distributed Ledger Technology und Blockchain, wirtschaftlichen Anwendungen von Künstlicher Intelligenz (KI) sowie der Auditierung und Zertifizierung von IT-Systemen. Im Berichtszeitraum durfte sich die Gruppe cii wieder über mehrere Auszeichnungen als Anerkennung ihrer Arbeit freuen (mehr dazu auf der nächsten Doppelseite) und es konnte eine Reihe neuer Forschungsprojekte eingeworben werden. Exemplarisch aufgeführt seien die Projekte „Accountable AI“ und „Digitale Transformation im Gesundheitswesen“, welche durch die DFG gefördert werden, sowie fünf neue, von den Bundesministerien BMBF, BMEL und BMWK geförderte Forschungsprojekte wie das Projekt „NephroCAGE“. Auch das Dissertationsprojekt „GaMeIT“ der Helmholtz Information and Data Science School for Health (HIDSS4Health) überzeugte.





Seit Mai 2021 ist Prof. Sunyaev Sprecher des Fachbereiches Wirtschaftsinformatik der Gesellschaft für Informatik (GI) e.V. Im Oktober 2021 veröffentlichte er erstmals den cii Student Papers Sammelband. Er soll ab sofort jährlich sehr gute studentische Seminararbeiten der Forschungsgruppe honorieren und publizieren.

## Information Service Engineering (ISE)

Prof. Dr. Harald Sack

ISE untersucht Modelle und Methoden zur Entwicklung und Bereitstellung von innovativen Informationssystemen. Im Fokus stehen effiziente semantische Erschließung, Aggregation und Retrieval umfangreicher heterogener und verteilter Datenquellen. Forschungsschwerpunkte liegen in der Optimierung und Anwendung von Deep Learning-basierten Verfahren zur Informations- und Wissensgewinnung aus heterogenen multimodalen Daten sowie, darauf aufbauend, in der Entwicklung semantischer und explorativer Suchtechnologien und Empfehlungssysteme. ISE ist an insgesamt vier Konsortien zum Aufbau der DFG-geförderten Nationalen Forschungsdateninfrastruktur (NFDI) beteiligt. Für NFDI4Culture (Research Data on Material and Immaterial Cultural Heritage), MaRDI (Mathematical Research Data Initiative), NFDI4Matwerk (NFDI for Materials Science and Engineering) und NFDI4DataScience (NFDI for Data Science and Artificial Intelligence) arbeitet ISE mit an der Konzeption und Implementierung von Ontologien und Wissensgraphen zur bereichsübergreifenden Vernetzung von Forschungsdaten. Generelles Ziel ist dabei die systematische Erschließung, nachhaltige Sicherung, Verfügbarmachung und internationale Vernetzung von Forschungsdaten.

## Security • Usability • Society (SECUSO)

Prof. Dr. Melanie Volkamer

SECUSO forscht zum Thema Sicherheit und Privatheit. Im Mittelpunkt der Forschung steht der Mensch. Untersucht werden Methoden zur Entwicklung und Evaluation von benutzerfreundlichen Maßnahmen zum Schutz der Privatsphäre und zur Erhöhung der Sicherheit, zur Bewusstseinsbildung bezüglich der Privatheit im Digitalen und für Sicherheitstrainings insbesondere für Unternehmen. Darüber hinaus forscht die Gruppe an Sicherheitsfragen zum Thema elektronische Wahlen (E-Voting).

SECUSO arbeitet aktuell an folgenden sieben Projekten: dem vom BMBF geförderten Forschungsprojekt INSPECTION, dem vom Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg geförderten Projekt digilog@bw, einem Google Faculty Research Award zum Thema Phishing, einem Facebook Research Award zum Thema

Authentifizierung in AR & VR (Augmented Reality & Virtual Reality), dem vom KIT-Strategiefonds geförderten Projekt „Effektive Security Awareness am KIT“ sowie an zwei im Rahmen der Exzellenzuniversität geförderten Projekten: Future Democracies und KD<sup>2</sup>Ex. SECUSO ist darüber hinaus im Subtopic „Engineering Secure Systems“ (ESS) des Forschungsfeldes Information (Key Technologies) in der Helmholtz-Gemeinschaft Teil der Forschungsgruppe *Human and Societal Factors*. SECUSO organisierte in 2021 die Konferenzen E-Vote ID sowie EuroUsec und veranstaltete Distinguished Lectures in Cybersecurity. Seit 2021 unterrichtet der SECUSO-Roboter STAR auf diversen Events wie z. B. der Hannover Messe das Erkennen von betrügerischen Nachrichten.

## Web Science

Prof. Dr. York Sure-Vetter (beurlaubt).  
Vertretung: Dr. Michael Färber

Die Gruppe Web Science beschäftigt sich mit der semantischen Wissensrepräsentation als Teilgebiet der Künstlichen Intelligenz (KI). Im besonderen Fokus stehen dabei Wissensgraphen, das maschinelle Lernen und die Verarbeitung natürlicher Sprache. Die Methoden werden vor allem entwickelt, um durch intelligente Systeme die Digitalisierung in Unternehmen voranzutreiben und durch Empfehlungssysteme die Informationsüberflutung einzudämmen.

Seit 2020 leitet York Sure-Vetter als Direktor den Aufbau der Nationalen Forschungsdateninfrastruktur (NFDI). Michael Färber vertritt ihn am KIT als Leiter der Forschungsgruppe Web Science und ist zudem als Helmholtz-AI-Associate aktiv. Die Gruppe Web Science ist aktuell an mehreren – teilweise sehr interdisziplinären – Projekten beteiligt. Hierzu gehören u.a. „digilog@bw“ mit Schnittpunkten zu den Sozial- und Geisteswissenschaften und „ChemKB“ mit der KIT-Fakultät für Chemie und Biowissenschaften sowie der KIT-Universitätsbibliothek. In Projekten wie „KIGLIS“, „IIDi“ und „MANDAT“ werden zudem KI-Technologien von der Forschung in die Unternehmen transferiert.

<https://www.aifb.kit.edu/web/Forschungsgruppen>

aifb

## 8 Promotionen

wurden seit der Publikation des letzten AIFB-Themenheftes erfolgreich abgeschlossen. Die neuen Doktorinnen und Doktoren, ihre Dissertationsthemen und ihre Betreuerinnen und Betreuer:

**Malte Greulich:** „Unveiling Adverse Effects of Organizational Information Security Practices“ (Ali Sunyaev)

**Lars Heling:** „Decentralized Query Processing over Heterogeneous Sources of Knowledge Graphs“ (York Sure-Vetter)

**Sebastian Lins:** „Unveiling and Enhancing the Effectiveness of Information System Certifications“ (Ali Sunyaev)

**Manuel Lösch:** „Utilization of Electric Prosumer Flexibility Incentivized by Spot and Balancing Markets“ (Hartmut Schmeck)

**Anna Nguyen:** „Human-Understandable Explanations of Neural Networks“ (York Sure-Vetter)

**Scott Thiebes:** „A Socio-Technical Analysis of Genetic Privacy and its Role in Genetic Data Sharing“ (Ali Sunyaev)

**Patrick Wiener:** „Managing Event-Driven Applications in Heterogenous Fog Infrastructures“ (York Sure-Vetter)

**Philipp Zehnder:** „Automating Industrial Event Stream Analytics: Methods, Models, and Tools“ (York Sure-Vetter)

## 118 Publikationen

wurden im Jahr 2021 aus dem Institut AIFB veröffentlicht. 1 Buch wurde herausgegeben. 4 Buchbeiträge und 25 Veröffentlichungen in Zeitschriften stammen von Angehörigen des Instituts AIFB. 9 Tagungsbände wurden gemeinsam mit Kolleginnen und Kollegen aus anderen Einrichtungen publiziert. 79 Beiträge in Tagungsbänden sowie weitere Publikationen und Vorträge zeugen von der aktiven Teilnahme am wissenschaftlichen Austausch.

<https://www.aifb.kit.edu/web/Veroeffentlichungen>

## 6 Auszeichnungen

wurden an Institutsangehörige verliehen. Allen Ausgezeichneten herzliche Gratulation!

Die amerikanische, international agierende Fachgesellschaft Association for Information Systems (AIS) hat **Ali Sunyaev** im Juli 2021 für sein Engagement für die Vereinigung zum „AIS Distinguished Member – Cum Laude“ ernannt.

Die Forschungsgruppe **Security • Usability • Society (SECUSO)** hat mit ihren Privacy Friendly Apps den zum ersten Mal verliehenen Digital Autonomy Award 2022 gewonnen. Die Apps verbessern auf Smartphones den Schutz der Privatsphäre. Der Beitrag, den die 33 Open Source Apps aus dem SECUSO-Projekt zur Steigerung der individuellen digitalen Souveränität leisten, überzeugte die zehnköpfige Fachjury ebenso wie die Teilnehmenden des offenen Online-Votings. Privacy Friendly Apps fordern bei der Anwendung lediglich für die Funktionalität notwendige Berechtigungen an und enthalten keine Tracking-Mechanismen. Die kostenlosen Apps gibt es für Fitness und Gesundheit, Spiele und Sicherheit, Wetterbericht, Passwortgenerator uvm. Der Digital Autonomy Award wird vom *Digital Autonomy Hub – Technik souverän nutzen* vergeben, einem jungen Kompetenzzentrum und Netzwerk aus 43 Instituten und Organisationen, das vom BMBF gefördert wird. Ziel der Zusammenarbeit ist die Steigerung der digitalen Selbstbestimmung von Nutzerinnen und Nutzern. Der Preis wurde gemeinsam mit dem iRights Lab, Wikimedia Deutschland, der Stiftung Digitale Chancen und der Stiftung Verantwortung ins Leben gerufen und mit ihnen ausgelobt.

**Jan Bartsch** und **Ali Sunyaev** haben gemeinsam mit Kollegen aus anderen Einrichtungen für ihren Konferenzbeitrag *Linking Data Sovereignty and Data Economy: Arising Areas of Tension* einen Best Paper Award der 17th International Conference on Wirtschaftsinformatik (WI 2022) erhalten. Die Veröffentlichung ist im Rahmen des Projekts „DaWID – Datenzentrierte Wertschöpfungskettenplattform“ entstanden. Mitgewirkt haben Florian Lauf, Simon Scheider, Jan Bartsch, Philipp Herrmann, Marija Radic, Marcel Rebbert, André T. Nemat, Christoph Schlueter Langdon, Ralf Konrad, Ali Sunyaev und Sven Meister.

### Auch im Bereich der Lehre gab es wieder ausgezeichnete Leistungen:

**Heiner Teigeler**, Doktorand in der Forschungsgruppe *cii*, hat auf Basis der Evaluation durch Studierende für die Übungen zu „Angewandte Informatik – Internet Computing“ im Sommersemester 2021 ein Zertifikat für gute Lehre erhalten.

Die Studierenden **Franziska Esser**, **Simon Stohrer**, **Frederic Puzicha** und **Lukas Bernhardt** haben beim Fakultätspartner-Preis „Teamprojekt Wirtschaft und Technologie“ im Sommersemester 2021 den 1. Platz belegt. Ausgezeichnet wurde ihr Vortrag zu ihrem Projekt *Innovative Designs für mHealth Privacy Assistenten*. Im Projekt wurden sie von *cii* betreut.

Die Forschungsgruppe ISE beteiligte sich mit einem studentischen Team am *Coding DaVinci Baden-Württemberg 2022 Kultur-Hackathon* und wurde als Preisträger für die „Most FAIR“ Application ausgezeichnet. Drei Masterstudierende und drei postgraduierte Mentorinnen und Mentoren verknüpfen Flugblätter aus der deutschen Revolutionszeit 1918-20 mit Inhalten des Deutschen Reichsanzeigers sowie des Preußischen Staatsanzeigers und machen diese in einer interaktiven Anwendung explorierbar. Die Datensätze wurden von der Universitätsbibliothek Mannheim und dem Haus der Geschichte Baden-Württemberg zur Verfügung gestellt.

<https://www.aifb.kit.edu/web/Neuigkeiten>

## 6 Projekte

an denen unsere Forschungsgruppen aktuell arbeiten, zeigen die Spannweite der Anwendungsgebiete der Forschung des Instituts AIFB.

Im BMBF-Projekt **SofDCar – Software-Defined Car** werden flexible Elektrisch-elektronische (E/E)- und Software (SW)-Architekturen für zukünftige Fahrzeuge erforscht, in denen Funktionen je nach Bedarf entweder im Fahrzeug oder außerhalb bereitgestellt werden können, ohne dass die Verteilung und der komplette Leistungsumfang bereits zum Entwicklungszeitpunkt der Funktionen festgelegt werden müssen. Das Fahrzeug wird dabei als Teil einer vernetzten Fahrzeug- und Systemumgebung verstanden. Die Gruppe **Angewandte technisch-kognitive Systeme (ATKS)** konzentriert sich im Rahmen ihrer Forschung und Entwicklung insbesondere auf Verfahren zur Robustheitssteigerung automatisierter und KI-basierter Fahrfunktionen, mit Fokus auf Perception und Aktionsplanung. Systemkomponenten mit der Fähigkeit zur Selbsteinschätzung ihrer Leistung sollen diese laufend bewerten und aktiv neue Daten für eine Steigerung ihrer Fähigkeiten sammeln (lernen). Im Zusammenspiel mit einem digitalen Zwilling werden die Funktionskomponenten auf diese Weise kontinuierlich verbessert, behalten die funktionale Robustheit bei und können als Software-Over-The-Air-Update bereitgestellt werden.

[https://www.aifb.kit.edu/web/Angewandte\\_Technisch-Kognitive\\_Systeme](https://www.aifb.kit.edu/web/Angewandte_Technisch-Kognitive_Systeme)

Im Verbundprojekt **KEA-Mod – Kompetenzorientiertes E-Assessment für die grafische Modellierung** im Rahmen des Förderprogramms „Digitale Hochschulbildung“ des BMBF hat ein Hochschulkonsortium unter der Leitung der Forschungsgruppe **Betriebliche Informationssysteme (BIS)** ein digitales Fachkonzept für die Lehre zur grafischen Modellierung erstellt und eine spezielle Lernplattform dazu entwickelt. Im Anschluss an Piloteinsatz und Evaluation wird diese für den Einsatz bei Prüfungen erweitert.

[https://www.aifb.kit.edu/web/Betriebliche\\_Informationssysteme](https://www.aifb.kit.edu/web/Betriebliche_Informationssysteme)

Das vom Bundesministerium für Bildung und Forschung (BMBF) geförderte Forschungsprojekt **BloG<sup>3</sup> – Blockchain-basiertes Gesundheitsdatenmanagement für gesamtgesellschaftliche Gesundheitsprofile** wird federführend von der Forschungsgruppe **Critical Information Infrastructures (cii)** geleitet. Eine Blockchain-basierte Smartphone-App, die im Projekt entwickelt wird, soll Onkologie-Patienten der Charité in Berlin die selbständige Verwaltung von Zugriffsrechten auf ihre Gesundheitsdaten ermöglichen.

<https://cii.aifb.kit.edu/>

Das vom BMBF geförderte Projekt **Plattform MaterialDigital**, an dem unsere Forschungsgruppe **Information Service Engineering (ISE)** beteiligt ist, hat zum Ziel, die Vielfalt der Materialien und der mit ihrer Herstellung und Nutzung verbundenen Prozesse sowie die Komplexität der Materiallebenszyklen im Rahmen einer gemeinsamen Innovationsplattform zu bündeln, um eine gemeinsame Basis zur Digitalisierung der Materialien voranzutreiben. Kern der Aktivitäten von ISE in diesem Projekt sind das Design und die Implementierung einer semantisch ausdrucksstarken Ontologie der Material- und Werkstoffwissenschaften, die als verknüpfendes Element Forschungsdaten integrieren und nachnutzbar zur Verfügung stellen soll.

[https://www.aifb.kit.edu/web/Information\\_Service\\_Engineering](https://www.aifb.kit.edu/web/Information_Service_Engineering)

Das 2021 gestartete Projekt **Effektive Security Awareness am KIT** unserer Forschungsgruppe **Security • Usability • Society (SECUSO)** in Kooperation mit dem Steinbuch Centre for Computing (SCC) und dem KIT Digital Office hat zum Ziel, Security-Awareness-Maßnahmen für KIT-Mitarbeiter:innen in einem nutzerzentrierten Ansatz zu entwickeln und diese am KIT zu etablieren. Zu den Maßnahmen gehören zunächst die Sensibilisierung für Informationssicherheit sowie die Themen sicherer Arbeitsplatz und E-Mail-Sicherheit.

<https://secuso.aifb.kit.edu/>

Das Projekt **KD4RE – Knowledge Distillation for Relation Extraction** der Forschungsgruppe **Web Science** beschäftigt sich mit der Entwicklung von effizienten Informationsextraktionsmethoden. Aktuelle, sehr große Sprachmodelle sind aufgrund hoher technischer Anforderungen oft nur extern als Off-Premise-Lösung realisierbar. KD4RE strebt daher die Entwicklung von Methoden an, welche die Kapazitäten sehr großer Sprachmodelle in kleinere Modelle übertragen und so neben geringeren Kosten auch neue On-Premise-Lösungen (d.h. Software auf eigenen Servern) ermöglichen.

[https://www.aifb.kit.edu/web/Web\\_Science](https://www.aifb.kit.edu/web/Web_Science)



## 1 apl. Professor

gehört seit Juli 2022 zum Institut AIFB. Der Senat des KIT hat PD Dr. Pradyumn Kumar Shukla auf Antrag unserer Fakultät in Anerkennung seiner Leistungen die Bezeichnung „außerplanmäßiger Professor“ verliehen. Pradyumn Kumar Shukla begann 2009 als wissenschaftlicher Mitarbeiter in der Forschungsgruppe Effiziente Algorithmen von Hartmut Schmeck und habilitierte sich 2016 mit einer Arbeit zu „Theoretical and Algorithmic Advances in Preference-based Multi-criteria Optimization“. Danach arbeitete er als Professurvertreter, Privatdozent und Lehrbeauftragter weiter am Institut. Neben seiner Tätigkeit am KIT ist Pradyumn Kumar Shukla Lecturer im Bereich Decision Cognitive Sciences an der Alliance Manchester Business School der University of Manchester (UK).

## 50 Jahre, 220 Seiten

umfasst der Bericht „50 Jahre AIFB“, den Wolffried Stucky anlässlich unseres Institutsjubiläums herausgegeben hat. Das Buch liegt inzwischen in einer 2. Auflage vor, in der zahlreiche Rückmeldungen zur 1. Auflage berücksichtigt wurden. Der Bericht ist gedruckt verfügbar und steht auf der Webseite des AIK-Vereins zum Download zur Verfügung.

<https://www.aik-ev.de/index.php/aifb50/>

## 62 Mitarbeiterinnen und Mitarbeiter

arbeiten am Institut AIFB. Sie stehen den Professorinnen und Professoren sowie den Studierenden im Lehrbetrieb zur Seite, gestalten und bearbeiten die Forschungsprojekte, viele im Rahmen einer Dissertation. Die Kolleginnen und Kollegen in Verwaltung und Technik sorgen für eine funktionierende Infrastruktur, und leisten so einen wichtigen Beitrag zu den Forschungs- und Lehraktivitäten des Instituts. 1 Honorarprofessor, 1 apl. Professor und 5 Lehrbeauftragte bereichern das Lehrangebot zusätzlich mit einem attraktiven Themenspektrum. Dem Institut eng verbunden sind 4 emeritierte bzw. pensionierte Professoren, die ihr Fachwissen und ihre Arbeitskraft noch oft zur Verfügung stellen. 4 junge Menschen haben zurzeit einen Ausbildungsplatz am Institut. Je nach Semester sind zwischen 90 und 110 studentische und wissenschaftliche Hilfskräfte in den Projekten bzw. als Tutoren für Lehrveranstaltungen aktiv. Zu unseren Forschungsgruppen gehören zudem etwa 50 weitere Doktorandinnen und Doktoranden, die am FZI Forschungszentrum Informatik und im FIZ Karlsruhe arbeiten oder in kooperierenden Unternehmen beschäftigt sind. Insgesamt umfasst das Institut derzeit über 220 Personen.

<https://www.aifb.kit.edu/web/Personen>

## 30 Vorlesungen

mit jeweils bis zu 600 Zuhörerinnen und Zuhörern sowie 30 Seminare und Praktika mit insgesamt 218 Teilnehmenden wurden im Sommersemester 2021 und im Wintersemester 2021/22 vom Institut angeboten.

[https://www.aifb.kit.edu/web/Lehrangebot\\_des\\_Instituts\\_AIFB](https://www.aifb.kit.edu/web/Lehrangebot_des_Instituts_AIFB)

## 3317 Prüfungen

wurden im Sommersemester 2021 und im Wintersemester 2021/22 am Institut AIFB abgenommen. 103 Abschlussarbeiten legten Studierende im gleichen Zeitraum vor. 54 Masterarbeiten und 49 Bachelorarbeiten wurden geschrieben und betreut.

<https://www.aifb.kit.edu/web/Abschlussarbeiten>

Institut AIFB: Wir machen  
Angewandte Informatik am KIT.  
Unser Ziel: Innovative Lösungen  
für Wirtschaft und Gesellschaft.  
Gerne auch für Sie und mit Ihnen.



Sprechen Sie uns bitte an!  
<https://aifb.kit.edu/web/Personen>



#### Kontakt

Karlsruher Institut für Technologie (KIT)  
Institut AIFB  
Postfach 6980  
76049 Karlsruhe  
[www.aifb.kit.edu](http://www.aifb.kit.edu)

#### Herausgeber

Karlsruher Institut für Technologie (KIT)  
Präsident Professor Dr.-Ing. Holger Hanselka  
Kaiserstraße 12  
76131 Karlsruhe  
[www.kit.edu](http://www.kit.edu)

Karlsruhe © KIT 2022

**Redaktionelle Bearbeitung:**  
Dr. Daniel Sommer, Institut AIFB  
[daniel.sommer@kit.edu](mailto:daniel.sommer@kit.edu)  
Vera Münch, Alfeld  
[vera-muench@kabelmail.de](mailto:vera-muench@kabelmail.de)

**Gestaltung:**  
Studio Quitta, München  
[quitta@studio-quitta.de](mailto:quitta@studio-quitta.de)

**Druck:**  
Systemedia GmbH, Wurmberg  
[info@systemedia.de](mailto:info@systemedia.de)

August 2022  
ISBN 978-3-944361-08-6