

# Linked Data im Zeitalter der generativen Künstlichen Intelligenz – Neue Perspektiven oder überholte Vision? Symposium

12. – 13. Dezember 2024, Berlin



## Grußwort

Sehr geehrte Damen und Herren,

wir begrüßen Sie herzlich hier am Bundesinstitut für Risikobewertung (BfR) auch im Namen der Akteure der drei Forschungsprojekte KIDA, nfdi4health und FAIRagro, die dieses Symposium initiiert und organisiert haben.

Forschungseinrichtungen wie das BfR und viele der Partner in den Projektverbänden sind bereits seit vielen Jahren dazu verpflichtet, ihre Forschungsdaten entsprechend den FAIR Data Prinzipien zur Verfügung zu stellen, z.B. im Rahmen von Forschungsprojekten unter dem Dach der EU-Programme Horizon 2020 oder Horizon Europe. Diese FAIR Data Prinzipien — Findable, Accessible, Interoperable und Reusable — sind von zentraler Bedeutung, um wissenschaftliche Erkenntnisse effizient und nachnutzbar zu teilen. Die FAIR Prinzipien wurden 2016 entwickelt und stehen auch in engem Zusammenhang mit den Grundsätzen der guten wissenschaftlichen Praxis, denen Wissenschaftsorganisationen wie das BfR verpflichtet sind.

Die Frage der Interoperabilität von Daten ist somit ein wichtiger Aspekt unserer wissenschaftlichen Arbeit und bedeutet, dass Daten nicht nur zugänglich, sondern auch semantisch interoperabel bereitgestellt werden sollten. Dies geht auf das Semantic Web und Linked Data Konzept zurück, das erstmals im Jahr 2001 von Tim Berners-Lee formuliert wurde, und heutzutage den Aufbau sogenannter Wissensgraphen (Knowledge Graphs) ermöglicht. Damit können Daten über verschiedenste Fachbereiche hinweg verknüpft, in Beziehung gesetzt und intersektoriell genutzt werden.

Aus Sicht vieler Forschungseinrichtungen ist jedoch festzustellen, dass trotz jahrelanger Bemühungen, gerade die Bereitstellung von Daten in semantisch interoperabler Weise eine enorme Herausforderung darstellt. Ein Hindernis ist dabei sicherlich auch, dass die Forschenden selbst häufig nur wenig unmittelbaren Mehrwert aus dieser Arbeit für ihre eigene Forschungsarbeiten ziehen können.

Genau an dieser Stelle soll unsere gemeinsame Veranstaltung „Linked Data im Zeitalter der generativen Künstlichen Intelligenz – Neue Perspektiven oder überholte Vision?“ ansetzen. Sie bietet den beteiligten Projekten, Vertretern von Behörden, Forschungseinrichtungen und Universitäten eine Plattform, um über Lösungsansätze zu reden und dabei auch neue Möglichkeiten zu diskutieren, die sich durch den Einsatz von Künstlicher Intelligenz (KI) ergeben. Dazu werden im Rahmen unseres Symposiums „Linked Data Leuchttürme“ aus verschiedenen Sektoren präsentiert. Ziel der Veranstaltung ist es zudem, die Linked Data Community zu stärken sowie fach- und projektübergreifende Netzwerke aufzubauen. Ein besonderes Highlight ist auch die Möglichkeit, im Rahmen der Poster- und Softwaresession verschiedenste Softwaretools auszuprobieren bzw. mit den Entwicklerinnen und Entwicklern ins Gespräch zu kommen. Dafür bietet sich diese Veranstaltung in Präsenz in idealer Weise an.

Zum Abschluss wird es die Gelegenheit geben, im Rahmen einer Podiumsdiskussion konkrete „Action Points“ zu definieren, die dann ab 2025 von den beteiligten Partnerprojekten gemeinsam vorangetrieben werden können. Den Teilnehmenden wünschen wir eine interessante, interaktive Veranstaltung mit vielen guten Impulsen für diese wichtige zukunftsgerichtete Thematik.

Prof. Dr. Karsten Nöckler (Leiter der Abteilung „Biologische Sicherheit“ des BfR)  
Dr. Sören Preibusch (zukünftiger Leiterr der Abteilung Informationstechnik)

# Inhaltsverzeichnis

<b>Grußwort</b>	<b>2</b>
<b>Inhaltsverzeichnis</b>	<b>3</b>
<b>1 Programm</b>	<b>5</b>
<b>2 Abstracts – Vorträge</b>	<b>8</b>
2.1 KIDA: KI- und Datenkompetenzzentrum für die forschenden und beratenden Einrichtungen des BMEL	8
2.2 Services von NFDI4Health und der Stand der semantischen Standardisierung von Gesundheitsdaten	9
2.3 NFDI4DataScience: Eine Nationale Forschungsdateninfrastruktur für Datenwissenschaften und Künstliche Intelligenz	10
2.4 FAIRAgro	11
2.5 Dialog mit Daten: Von Sprachmodellen zu Wissensgraphen	12
2.6 Linked Data, Ontologien und Large Language Models – Gegensätze oder Partner?	13
2.7 AI-assisted approach to efficiently extract data and information from information sources into Knowledge Graphs	14
2.8 Linking data automatically from their origin – a demonstration of the LARAsuite open data platform	15
2.9 Pflanzenschutzmitteldaten: Flexible, anwendungsfallunabhängige Bereitstellung über Linked Data Technologien, Integration mit anderen Datenbeständen und Aufnahme in NLP-Systeme und Large Language Models	16
2.10 Innovative Anwendungsbeispiele von Linked Data und Knowledge Graphen	17
2.11 Wikidata: Wikimedia’s knowledge graph and your Rosetta Stone for the web	18
<b>3 Abstracts – Softwaresession</b>	<b>19</b>
3.1 eccenca Corporate Memory: an enabler for knowledge graph based generative AI solutions	19
3.2 Harnessing neuro-symbolic AI to help customers build knowledge-centric organizations	20
3.3 Automating the creation of linked data repositories with Fluree software	21
3.4 The Wikidata Open-Source Knowledge Graph, Semantic Search and AI	22
3.5 KGI4NFID – knowledge graph infrastructure for the research community	23
<b>4 Abstracts – Poster</b>	<b>24</b>
4.1 An ontology classifying residues from the bioeconomy	24
4.2 In the beginning there was complex data... ... or how local deep data annotation enables new perspectives on scientific data use, linked data publication, and large language model integration into workflows	25
4.3 Rapid Alert Supply Network EXtractor (RASNEX): Mining and structuring supply chain information from incident reports	26

---

4.4 Enhancing Proteomics Meta-Analysis through Linked Data and LLM-Driven Metadata Extraction	27
4.5 NFDI4Chem's Semantic Data Hub: Central Resource for FAIR Chemistry Data and Metadata	28
4.6 Enhancing the discoverability of Predictive Models with the FAIR Scientific Knowledge Exchange (FSKX) Format	29
4.7 AGROVOC und Large Language Models: Henne oder Ei	30
4.8 Linked Data Toolbox: Enabling Interoperability across Risk Assessment Research	31
4.9 Integrating Various Supply Network Abstractions Through a Uniform Model and Knowledge Graphs	32
4.10 Integration of knowledge graphs into LLMs: Are Graph Language Models better for reasoning in intelligent systems for biomedical research?	33
4.11 Operationalizing Linked Data in Public Health Agencies Through the One Health Model	34
4.12 Faktenerkennung für Knowledge-Graphs aus unstrukturierten Textkorpora	35
4.13 Agro-Intelligence Hub (AI-Hub): ein Pilotprojekt zur Verknüpfung von Datenbeständen des JKI und BfR mittels „Linked Data“ Technologien	37
4.14 SIMLEARN – Betriebliche Entscheidungsunterstützung durch ontologiegestützte Integration von Simulationsmodellen, Systemen für maschinelles Lernen und Planungsdaten	38
<b>5 Verzeichnis der Autorinnen und Autoren</b>	<b>39</b>

# 1 Programm

---

## Donnerstag, 12. Dezember 2024

---

13:00 – 13:30 Uhr Anmeldung und Willkommenskaffee

---

13:30 – 13:40 Uhr **Begrüßung**  
Prof. Dr. Karsten Nöckler, Bundesinstitut für Risikobewertung (BfR)

---

### Anknüpfungspunkte von Linked Data und Semantic Web im Kontext der Projekte

Moderation: Yvonne Mensching, Bundesinstitut für Risikobewertung (BfR)

---

13:40 – 13:50 Uhr **KIDA: KI- und Datenkompetenzzentrum für die forschenden und beratenden Einrichtungen des BMEL**  
Lea Löhn, Thünen-Institut

---

13:50 – 14:00 Uhr **Services von NFDI4Health und der Stand der semantischen Standardisierung von Gesundheitsdaten**  
Dr. Carina Vorisek, Berliner Institut für Gesundheitsforschung in der Charité

---

14:00 – 14:10 Uhr **NFDI4DataScience: Eine Nationale Forschungsdateninfrastruktur für Datenwissenschaften und Künstliche Intelligenz**  
Prof. Dr. Sonja Schimmler, TU Berlin und Fraunhofer FOKUS

---

14:10 – 14:20 Uhr **FAIRagro - FAIRes Datenmanagement in der Agrosystemforschung**  
Dr. Ulrike Stahl, Julius Kühn-Institut (JKI)

---

14:20 – 14:30 Uhr **Q&A**

---

### Aktuelle Entwicklungen im Bereich der generativen KI – Implikationen für das Linked Data Konzept

Moderation: Taras Günther, Bundesinstitut für Risikobewertung (BfR)

---

14:30 – 14:50 Uhr **Dialog mit Daten: Von Sprachmodellen zu Wissensgraphen**  
Prof. Dr. Andreas Harth, Fraunhofer-Institut für Integrierte Schaltungen IIS

---

14:50 – 15:10 Uhr **Linked Data, Ontologien und Large Language Models – Gegensätze oder Partner?**  
Dr. Ansgar Bernardi, Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI)

---

15:10 – 15:30 Uhr **AI-assisted approach to efficiently extract data and information from information sources into Knowledge Graphs**  
Dr. Iurii Savvateev, Bundesinstitut für Risikobewertung (BfR)

---

15:30 – 16:00 Uhr Kaffeepause

---

## Linked Data Success Stories – Leuchttürme

Moderation: Taras Günther und Yvonne Mensching, Bundesinstitut für Risikobewertung (BfR)

16:00 – 16:30 Uhr	<b>Linking data automatically from their origin – a demonstration of the LARAsuite open data platform</b> Dr. Mark Dörr, Universität Greifswald
16:30 – 17:00 Uhr	<b>Pflanzenschutzmitteldaten: Flexible, anwendungsfallunabhängige Bereitstellung über Linked Data Technologien, Integration mit anderen Datenbeständen und Aufnahme in NLP-Systeme und Large Language Models</b> Daniel Martini, Kuratorium für Technik und Bauwesen in der Landwirtschaft e. V. (KTBL)
17:00 – 17:30 Uhr	<b>Innovative Anwendungsbeispiele von Linked Data und Wissensgraphen in den Bereichen Life Science und Industrie</b> Taras Günther, Bundesinstitut für Risikobewertung (BfR)
17:30 – 18:00 Uhr	<b>Wikidata: Wikimedia’s knowledge graph and your Rosetta Stone for the web</b> Lydia Pintscher, Wikimedia Deutschland e. V.
18:00 – 20:30 Uhr	Networking-Empfang mit Fingerfood-Buffer und Getränken

## Freitag, 13. Dezember 2024

### Softwaresession

Moderation: Daniel Martini, Kuratorium für Technik und Bauwesen in der Landwirtschaft e. V. (KTBL)

09:00 – 09:15 Uhr	<b>eccenca Corporate Memory: an enabler for knowledge graph based generative AI solutions</b> Magnus Knuth, eccenca GmbH
09:15 – 09:30 Uhr	<b>Harnessing neuro-symbolic AI to help customers build knowledge-centric organizations</b> Dr. Peter Haase, metaphacts
09:30 – 09:45 Uhr	<b>Automating the creation of linked data repositories with Fluree software</b> Eliud Polanco, Fluree und Tim Biedenkapp, adorsys
09:45 – 10:00 Uhr	<b>The Wikidata open-source knowledge graph, semantic search and AI</b> Lydia Pintscher und Philippe Saadé, Wikimedia Deutschland e. V.
10:00 – 10:15 Uhr	<b>KGI4NFID – knowledge graph infrastructure for the research community</b> Prof. Dr. Konrad Förstner, Deutsche Zentralbibliothek für Medizin (ZB MED) Informationszentrum Lebenswissenschaften
10:15 – 10:30 Uhr	<b>Q&amp;A</b>
10:30 – 11:30 Uhr	Kaffeepause, Postersession und Softwaredemos

---

### **Paneldiskussion: Linked Data–Aktionsplan**

Moderation: Matthias Filter, Bundesinstitut für Risikobewertung (BfR)

---

11:30 – 12:30 Uhr	Hilke Thordsen, Bundesministerium für Ernährung und Landwirtschaft (BMEL) Prof. Dr. Andreas Harth, Fraunhofer-Institut für Integrierte Schaltungen IIS Dr. Carsten Hoffmann, Leibniz-Zentrum für Agrarlandschaftsforschung (ZALF) e. V. Dr. Soeren Preibusch, Bundesinstitut für Risikobewertung (BfR) Prof. Dr. Konrad Förstner, Deutsche Zentralbibliothek für Medizin (ZB MED) Informationszentrum Lebenswissenschaften
12:30 – 12:45 Uhr	<b>Verabschiedung</b> Matthias Filter, Bundesinstitut für Risikobewertung (BfR)

---

## 2 Abstracts – Vorträge

### 2.1 KIDA: KI- und Datenkompetenzzentrum für die forschenden und beratenden Einrichtungen des BMEL

Dr. Nathalie Gottschalk, Lea Löhn

Thünen-Institut

**Keywords:** Künstliche Intelligenz, Beratung, BMEL

Mit dem KI- und Daten-Akzelerator (KIDA) fördert das Bundesministerium für Ernährung und Landwirtschaft (BMEL) seit 2022 den Aufbau eines einrichtungsübergreifenden KI- und Daten-Kompetenzzentrums für seine nachgeordneten Behörden.

Die Einsatzmöglichkeiten für KI in der wissenschaftlichen Datenverarbeitung sind vielfältig und ermöglichen neben der effizienteren Auswertung auch neue komplexere Analysen und damit Erkenntnisgewinne. Aufgrund diverser Hürden, u. a. technischer, methodischer aber auch rechtlicher Art, benötigt ein Großteil der Forschenden im Agrifood-Bereich Unterstützung beim individuellen Einsatz von KI. Mit einem breiten Angebot setzt KIDA hier zur Stärkung der KI- und Daten-Expertise an, schafft die notwendigen Voraussetzungen, unterstützt individuell und fördert die Mitarbeitenden der beteiligten Einrichtungen.

Hierzu begleitet KIDA die gesamte KI-Wertschöpfung in einem Ende-zu-Ende Ansatz - von technischer Infrastruktur und Daten über Beratung bis zu rechtlicher Klärung und Produktivsetzung. Die für den KI-Einsatz notwendige IT-Infrastruktur zur Verarbeitung großer Datenmengen wird beispielweise durch den gemeinsamen Betrieb von Hochleistungsrechnern ausgebaut. Um die Zugänglichkeit zu Daten für die Entwicklung neuer KI-Modelle zu erleichtern, befasst sich eine Expertengruppe mit der Gewährleistung der Interoperabilität von Daten und bietet den Forschenden individuelle Beratungs- und Unterstützungsleistungen im Themenfeld Linked Data an. Auch bei der Veröffentlichung entwickelter KI-Modelle nach den FAIR-Prinzipien können die Beschäftigten von KIDA-Expert\*innen unterstützt werden. Weiterhin profitieren die beteiligten Einrichtungen vom einrichtungsübergreifenden KIDA-Wissenstransfer durch unterschiedliche Angebote insbesondere zu Kompetenzaufbau und Vernetzung. Eine Rechtsberatungsstelle rundet durch Beratung zum rechtssicheren Umgang mit Daten und Modellen das Angebot von KIDA ab.



## 2.2 Services von NFDI4Health und der Stand der semantischen Standardisierung von Gesundheitsdaten

Prof. Dr. Juliane Fluck<sup>1,2</sup>, Dr. Carina Vorisek<sup>3</sup> im Auftrag des NFDI4Health Konsortiums<sup>4</sup>

<sup>1</sup> ZB MED Informationszentrum Lebenswissenschaften

<sup>2</sup> Rheinische Friedrich-Wilhelms-Universität Bonn

<sup>3</sup> Berliner Institut für Gesundheitsforschung in der Charité

<sup>4</sup> NFDI4Health

**Keywords:** Health data, FAIR, Clinical Studies, Epidemiologic studies, Public Health

Die Nutzung von Daten gewinnt in der Gesundheitsforschung zunehmend an Bedeutung und die Gesundheitsforschungsdateninfrastrukturen in Deutschland bereiten mit ihren Netzwerken und Förderstrukturen seit einigen Jahren den Weg, um die Nachnutzung dieser Daten zu erleichtern.

NFDI4Health bietet Konzepte und Dienstleistungen an, um in Deutschland erhobene Gesundheitsdaten zentral auffindbar und zugänglich zu machen. Für epidemiologische, klinische und Public Health Studien schafft der German Central Health Study Hub von NFDI4Health eine zentrale Auffindbarkeit über standardisierte Metadaten und forciert die Veröffentlichung von Datenkatalogen, die ohne direkten Zugriff auf die sensiblen Daten Auskunft über die erhobenen Datenelemente geben. Standardisierte Metadaten, semantisch annotierte Kerndatensätze sowie Harmonisierungsansätze sind die Basis für Interoperabilität und damit Voraussetzung für die Realisierung von Linked-Data-Anwendungen. Die Standardisierungsbemühungen im Gesundheitsbereich finden auf (inter)nationaler Ebene in enger Zusammenarbeit verschiedener Initiativen und Projekte statt. Aktuelle Themen sind die Einbettung der deutschen Gesundheitsforschungsdateninfrastrukturen in den European Health Data Space (EHDS) und die Entwicklung des nationalen/europäischen Metadatenkatalogs Health-DCAT-AP.

## 2.3 NFDI4DataScience: Eine Nationale Forschungsdateninfrastruktur für Datenwissenschaften und Künstliche Intelligenz

Prof. Dr. Sonja Schimmler<sup>1,2</sup>

<sup>1</sup> TU Berlin

<sup>2</sup> Fraunhofer FOKUS

**Keywords:** Forschungsdateninfrastruktur, Datenwissenschaften, Künstliche Intelligenz

Moderne Forschung auf international kompetitivem Niveau ist ohne digitale Unterstützung nicht mehr möglich. Es werden immer mehr Daten generiert und immer mehr technische Systeme eingesetzt. Dies führt zu einem Paradigmenwechsel, der aktuell von zahlreichen Aktivitäten im Bereich der Forschungsdateninfrastrukturen begleitet wird. Prominente Beispiele sind die Nationale Forschungsdateninfrastruktur (NFDI) und die European Open Science Cloud (EOSC).

Das NFDI-Konsortium NFDI4DataScience verfolgt das Ziel, eine nationale Forschungsdateninfrastruktur für Datenwissenschaften und Künstliche Intelligenz aufzubauen und zu etablieren. Im Fokus stehen dabei Transparenz, Reproduzierbarkeit und Fairness von Projekten. Dies soll durch die Verfügbarmachung und Verknüpfung aller digitaler Artefakte (auch Code und Modelle) sowie die Bereitstellung innovativer Werkzeuge und Dienste befördert werden.

## 2.4 FAIRAgro

Dr. Ulrike Stahl<sup>1</sup>, Daniel Martini<sup>2</sup>

<sup>1</sup> Julius Kühn-Institut (JKI)

<sup>2</sup> Kuratorium für Technik und Bauwesen in der Landwirtschaft e. V. (KTBL)

**Keywords:** Nationale Forschungsdateninfrastruktur (NFDI), interoperable Infrastrukturen, Helpdesk, Datastewards, FAIR-Prinzipien

Das FAIRAgro-Konsortium mit mehr als 25 Partnern baut ein FAIRes Forschungsdatenmanagement (Findable, Accessible, Interoperable, Reusable) für die Community der Agrosystemforschung auf. Es entwickelt eine interoperable und skalierbare Forschungsdateninfrastruktur, indem es verfügbaren Repositorien verbindet und gewährleistet somit die Auffindbarkeit und Wiederverwendbarkeit. Dazu wird FAIRAgro die FAIR-Datenkonzepte der beteiligten Forschungsdateninfrastrukturen konsolidieren, schema.org erweitern, die Interoperabilität der beteiligten Vokabulare und Ontologien verbessern, Metadatenmodelle erstellen und Publikationsleitlinien für die verschiedenen Arten von Daten erarbeiten, FAIRAgro wird kombinierte Datenanalysen erleichtern und ein mehrstufiges Supportsystem einrichten, indem es ein Data Steward Service Center mit Helpdesk aufbaut. Ebenso stellt es Informationsmaterial bereit und befasst sich eingehend mit Herausforderungen der Qualität und der Rechtssicherheit.

## 2.5 Dialog mit Daten: Von Sprachmodellen zu Wissensgraphen

Prof. Dr. Andreas Harth

Fraunhofer-Institut für Integrierte Schaltungen IIS

**Keywords:** Sprachmodelle, Wissensgraphen, Linked Data, Dialogsysteme, Faktentreue

Der Beitrag stellt eine Methode vor, welche die Stärken von Sprachmodellen und Wissensgraphen miteinander verbindet, um leistungsfähige und verlässliche Dialogsysteme zu schaffen.

Sprachmodelle bieten zwar einen intuitiven Zugang über natürliche Sprache, haben aber oft Probleme mit der Genauigkeit faktischer Informationen. Wissensgraphen hingegen liefern strukturierte, verlässliche Daten, setzen aber voraus, dass Nutzer komplexe, formale Abfragen stellen können.

Der zentrale Mechanismus der vorgestellten Methode ist es, Anfragen in natürlicher Sprache in strukturierte Abfragen an den Wissensgraphen zu übersetzen. Sprachmodelle werden dabei in einem iterativen, agentenbasierten Ansatz genutzt, um die natürlichsprachige Anfrage schrittweise in eine formale Abfrage zu überführen. So können Nutzer ihre Fragen unkompliziert formulieren und gleichzeitig sicherstellen, dass gezielt Informationen aus dem Wissensgraph abgerufen werden.

Durch den Einsatz von Linked Data lassen sich die zugrunde liegenden Wissensgraphen verteilt erstellen und kontinuierlich pflegen. Dies erhöht die Aktualität und Verlässlichkeit der Antworten.

Insgesamt entsteht so ein Dialogsystem, das die Bedienungsfreundlichkeit natürlicher Sprache mit der Genauigkeit und Verlässlichkeit strukturierter Wissensgraphen kombiniert.

## 2.6 Linked Data, Ontologien und Large Language Models – Gegensätze oder Partner?

Dr. Ansgar Bernardi

Deutsches Forschungszentrum für Künstliche Intelligenz GmbH – DFKI, Kaiserslautern, Germany

Lernende Systeme, Neuronale Netze, Generative KI und Large Language Models – in jüngster Zeit haben wir atemberaubende Fortschritte erlebt. Erstmals können wir in natürlicher Sprache effektive Dialoge mit riesigen Informationssammlungen führen und Antworten auf unsere Fragen erhalten.

Linked Data und formale Ontologien versprechen ebenfalls Zugang zu großen, weltweit verteilten Datensammlungen, bringen aber scheinbar vergleichsweise hohen Modellierungsaufwand mit sich. Lohnt es sich, diese Ansätze weiter zu verfolgen?

Vor dem Hintergrund verteilter, komplexer Prozesse mit unterschiedlichen teilnehmenden Parteien – wie etwa den offenen Wertschöpfungsketten in der Land- und Ernährungswirtschaft – beleuchten wir die unterschiedlichen Beiträge und Leistungsversprechen der verschiedenen Ansätze. Die formale Korrektheit der expliziten logischen Modellierung, die in Ontologien und Linked Data zum Tragen kommt, ermöglicht logische Schlüsse mit verlässlicher Korrektheit und zukunftsorientierten Wissenstransfer. Die Flexibilität und Breite wohltrainierter Large Language Models erschließt ungeahnte Informationsräume und bietet intuitiven Zugang auch zu komplexen Fragestellungen. Eine optimale Aufgabenverteilung zwischen den verschiedenen Ansätzen muss neben technischer Effizienz auch die Möglichkeit der Verantwortung für die Korrektheit der Lösungen und die Verteilung von Aufwand und Ertrag zwischen allen Beteiligten im System betrachten.

Es zeigt sich, dass jeder Ansatz eigene Stärken und Schwächen bietet, so dass die sorgfältige Auswahl einer zum jeweiligen Einsatzszenario passenden Kombination Erfolg verspricht und weiter erforscht werden sollte.

## 2.7 AI-assisted approach to efficiently extract data and information from information sources into Knowledge Graphs

Iurii Savvateev, Taras Günther, Julia Golz, Stefan Dehm, Matthias Filter

German Federal Institute for Risk Assessment (BfR)

Regulators and governmental agencies in Food Safety and Agriculture often deal with reports that include excessive amounts of multimodal data: plain text, tables, and figures. At the same time, the rapid and standardized processing of these data is a key differentiator ensuring the reliable communication across agencies, scientists and consumers. The presented work describes a pipeline to extract the information from data into a Knowledge Graph. This graph can then subsequently be linked to domain-specific ontologies, ultimately allowing standardized, interoperable and in-depth data analysis and exchange.

Recent developments in large language models (LLM) have made it possible to process data automatically to reliably extract the key entities and their relationships. The inherent capabilities of LLMs in natural language understanding enable the accurate extraction of complex information. Once the entities and relationships are identified, they are structured into a Knowledge Graph, which represents the interrelations between the data points, transforming content into an organized, interoperable, and accessible format. The constructed Knowledge Graph is further augmented by linking it with domain-specific ontologies, which provide a framework that categorizes and contextualizes the information, increasing its semantic value. By doing so, the interoperability of information across multiple sources can be ensured and enriched with background knowledge enabling new insights and connections.

The graphs resulting from this approach can subsequently be used as a basis for improving human or automated decision-making or as a basis for deductive reasoning. Importantly, the given solution is open source, can be potentially operated offline, and thus may be integrated into processes that deal with confidential data.

Funding acknowledgments:

This research has been funded by the German Federal Ministry of Food and Agriculture (BMEL) in the research project “KI- & Daten-Akzelerator (KIDA)” with project number 28KIDA004.

## 2.8 Linking data automatically from their origin – a demonstration of the LARAsuite open data platform

Dr. Mark Dörr<sup>1</sup>, Stefan Maak<sup>1</sup>, Uwe T. Bornscheuer<sup>1</sup>, Stefan Born<sup>2</sup>

<sup>1</sup> Universität Greifswald

<sup>2</sup> Technische Universität Berlin

**Keywords:** Lab automation, Data Standardisation, RDM, LIMS, ELN

How can we inter-link experimental data and additionally link the data to standardised terminologies / ontologies in a way that the scientist does not need to touch the data and the complexity of semantically annotated data with terms from ontologies?

The LARAsuite is our attempt to implement this vision. We built up an infrastructure from bottom up, starting at the measurement devices and sensor, enable them to send semantically annotated data. For this purpose we leverage the upcoming open source lab automation communication standard SiLA ([sila-standard.org](http://sila-standard.org)). Experiments and data aggregation are conducted by an orchestrator/ scheduler and collected by an data aggregator (time series data) or data collector (file based measurement results) into a central database (LARA-django) (source: [gitlab.com/opensourcelab](https://gitlab.com/opensourcelab) and [gitlab.com/larasuite](https://gitlab.com/larasuite)). During this automated process, data is semantically enriched and this semantic information transferred to a graph database / triple store). It is designed to reduce data inputs of scientist to the bare minimum and make data accessible and findable through deep query infrastructures (like SPARQL). This also enables advanced Machine Learning and AI applications to access data in a machine-understandable, "semantic" form. The full complexity of the data structure can be synchronised between different instances of LARA (or public data repositories).

All parts are realised as open source, mainly python based microservices, which are dockerised and deployed via docker-compose or kubernetes.

## **2.9 Pflanzenschutzmitteldaten: Flexible, anwendungsfallunabhängige Bereitstellung über Linked Data Technologien, Integration mit anderen Datenbeständen und Aufnahme in NLP-Systeme und Large Language Models**

Daniel Martini, Katharina Albrecht, Jascha Jung, Esther Mietzsch, Burkhard Golla, Aaron Roggenland

Kuratorium für Technik und Bauwesen in der Landwirtschaft e. V. (KTBL)

Die fachgerechte Durchführung von Pflanzenschutzmaßnahmen ist anspruchsvoll und erfordert in Planung und Vorbereitung eine Reihe von Informationen. Chemische Mittel sollten sowohl aufgrund von Risiken für Umwelt und anwendende Personen, möglicher Resistenzbildungen als auch aus Kostengründen so zielgerichtet und sparsam wie nur möglich eingesetzt werden. Daher ist es zunächst erforderlich, anhand von Beobachtungen im Feld und Warndienstinformationen festzustellen, ob überhaupt eine Indikation zur Anwendung vorliegt. Sofern dies der Fall ist, müssen für diese Indikation geeignete Mittel sowie die hierfür gültigen Anwendungsbestimmungen ermittelt werden. Von besonderer Bedeutung sind Auflagen, die vorgeben, welche Abstände zu schützenswerten Strukturen wie Hecken, Siedlungsflächen und Oberflächengewässern einzuhalten sind. Eine digitale Unterstützung bei diesen Vorbereitungsprozessen war in der Vergangenheit nur unzureichend möglich, da Informationen aus verschiedenen Quellen zusammengeführt werden mussten und diese teilweise in nur unstrukturierter Form vorlagen. Im Rahmen des PAM- und des HortiSem-Projektes wurden verschiedene Datenquellen – u.a. die Pflanzenschutzmittel-Registrierungsdaten des Bundesamtes für Verbraucherschutz und Lebensmittelsicherheit – als Linked Data aufbereitet und in einem Wissensgraphen als RDF zusammengeführt. Dadurch konnten die enthaltenen Daten integriert und für verschiedene Anwendungen und Zwecke gezielt und flexibel abfragbar gemacht werden. Warndienstmeldungen, die als Texte vorliegen, wurden mittels Named Entity Recognition (NER) bezüglich wichtiger Klassen wie Kulturen, Mittel usw. automatisch annotiert und die Annotationen ebenfalls als RDF abgespeichert. Aus den so vorbereiteten Datenbeständen konnten im Rahmen eines Hackathon die Informationen mit überschaubarem Aufwand auch in einer Form zurückgewonnen werden, die geeignet ist, um mit Hilfe von Large Language Models Beratungsassistenten umzusetzen. Potenziell wären auf Basis von Linked-Data-Technologien und Ontologien auch gewisse Konsistenzprüfungen möglich, deren Erprobung in dem Kontext jedoch bislang noch nicht erfolgt ist. Gerade in sensiblen Bereichen, in denen Maßnahmen gegebenenfalls auch schädliche Nebenwirkungen haben können, ist dies aber eine Anforderung, die LLMs alleine bislang nur unzureichend erfüllen können.



## 2.10 Innovative Anwendungsbeispiele von Linked Data und Knowledge Graphen

Taras Günther, Michael Zarske

Bundesinstitut für Risikobewertung (BfR)

**Keywords:** Knowledge Graphen, Linked Data, Angewandte Forschung, KI

Die rasanten Fortschritte im Bereich der künstlichen Intelligenz (KI) machen es immer wichtiger, große Datenmengen und das darin enthaltene Wissen strukturiert bereitzustellen. Schlüsseltechnologien wie Wissensgraphen (Knowledge Graphs), Ontologien und Linked Data spielen dabei eine zentrale Rolle, da sie eine Verknüpfung zwischen heterogenen und verteilten Datenbeständen über maschinenlesbare semantische Datenstandards ermöglichen können [1][2]. Damit stellen diese Technologien unter anderem eine ideale Ergänzung zu sogenannten großen Sprachmodellen (LLM) dar, da sie die Grundlage für eine Fakten-basierte Informationsbereitstellung mittels generativer KI schaffen können. Knowledge Graphen und Linked Data können zudem für die domänenübergreifende Verwaltung umfangreicher Datenbestände, Terminologien und Prozesse in Forschungseinrichtungen und Unternehmen genutzt werden [3][4].

Im Rahmen des Vortrages werden aktuelle Entwicklungen und konkrete Anwendungsbeispiele von semantischen Technologien aus den Lebenswissenschaften sowie der Industrie vorgestellt. Ziel ist es, einen Überblick über aktuelle Initiativen und Leuchtturmprojekte im internationalen Kontext zu vermitteln und das Potential von Linked Data und Knowledge Graphen an konkreten Beispielen zu veranschaulichen.

Funding acknowledgments:

Diese Forschung wurde vom Bundesministerium für Ernährung und Landwirtschaft (BMEL) im Rahmen des (BMEL) im Forschungsprojekt „KI- & Daten-Akzelerator (KIDA)“ mit der Projektnummer 28KIDA004 gefördert.

Referenzen:

- [1] Peng, C., Xia, F., Naseriparsa, M. et al. Knowledge Graphs: Opportunities and Challenges. *Artif Intell Rev* 56, 13071–13102 (2023). <https://doi.org/10.1007/s10462-023-10465-9>
- [2] Hofer, M.; Obraczka, D.; Saeedi, A.; Köpcke, H.; Rahm, E. Construction of Knowledge Graphs: Current State and Challenges. *Information* 2024, 15, 509. <https://doi.org/10.3390/info15080509>
- [3] Strömert, Philip, et al. "Towards a versatile terminology service for empowering FAIR Research Data: Enabling Ontology Discovery, Design, Curation, and Utilization Across Scientific Communities." *Knowledge Graphs: Semantics, Machine Learning, and Languages*. IOS Press, 2023. 53-69.
- [4] Meyers, B., Vangheluwe, H., Lietaert, P. et al. Towards a knowledge graph framework for ad hoc analysis in manufacturing. *J Intell Manuf* 35, 3731–3752 (2024). <https://doi.org/10.1007/s10845-023-02319-6>

## 2.11 Wikidata: Wikimedia's knowledge graph and your Rosetta Stone for the web

Lydia Pintscher

Wikimedia Deutschland e.V.

**Keywords:** Wikidata, Wikibase, Knowledge Graph

Wikidata is an open-source knowledge graph providing structured, standardized data for various applications. Launched in 2012 by Wikimedia Deutschland, it serves as a common repository of referenced statements on a vast array of topics anyone can access. With over 112 million data objects, a community of over 24,000 monthly volunteers, and available in over 300 languages, Wikidata is the world's largest collaboratively created open knowledge graph. Unlike Wikipedia's natural language articles, Wikidata's data can easily be processed by computers. This makes it invaluable for building applications, conducting data analysis, and enriching digital content across various domains.

Wikidata, however, does not just serve as a repository of linked open data about the world. It also acts as a hub in the Linked Open Data web, enabling you to access and switch between over 9,000 resources from libraries, social networks, archives, governmental databases, and much more. This enables access to a vast amount of data and supports the reconciliation between disparate data sources.

Wikibase, the software powering Wikidata, enhances these capabilities by enabling organizations to create and manage their own knowledge graphs. As a free and open-source software suite, Wikibase is maintained by Wikimedia Deutschland in collaboration with a global community and is widely adopted across various sectors, fostering connections within a broader ecosystem.

## 3 Abstracts – Softwaresession

### 3.1 eccenca Corporate Memory: an enabler for knowledge graph based generative AI solutions

Magnus Knuth

eccenca GmbH

**Keywords:** Knowledge Graphs, Generative AI, LLM

Linked Data and Knowledge Graphs enable a connected data landscape within an enterprise and over company borders. The eccenca enterprise knowledge graph platform Corporate Memory provides a flexible data integration solution for managing rules, constraints, capabilities, configurations and data in a single application. We consider generative AI and Linked Data to be mutually beneficial technologies. The factual knowledge from Linked Data can further increase the response quality of LLMs [Wang-KGP-2024]. On the other hand, generative AI can support knowledge representation tasks in data integration pipelines and knowledge graph enrichment.

As of now, we adopt artificial intelligence in eccenca products for various tasks: Ontology engineering is supported by extracting relevant concepts and relationships from unstructured data; sophisticated data integration mappings are suggested to build pipelines robust against unforeseen inputs; complex linking rules can be learned based on reference data; and queries are generated based on natural language user questions.

There are numerous options for how Generative AI and Knowledge Graph systems can work in tandem to eliminate their mutual weaknesses and play to their strengths. In the session, we will demonstrate how both technologies can be fruitfully combined to support knowledge engineers and end users with their everyday tasks.

References:

[Wang-KGP-2024] Wang, Y., Lipka, N., Rossi, R. A., Siu, A., Zhang, R., & Derr, T. (2024). Knowledge Graph Prompting for Multi-Document Question Answering. Proceedings of the AAAI Conference on Artificial Intelligence, 38(17), 19206-19214.  
<https://doi.org/10.1609/aaai.v38i17.29889>

### 3.2 Harnessing neuro-symbolic AI to help customers build knowledge-centric organizations

Dr. Peter Haase

metaphacts

**Keywords:** Knowledge Graphs, Symbolic AI, Neuro-symbolic AI, metaphactory, LLMs

In this session, we will provide an overview of how our platform metaphactory combines the power of symbolic AI (knowledge graphs) and sub-symbolic AI (LLMs) to help customers transition from a data-centric to a knowledge-centric culture and use trustworthy, explainable insights to drive innovation and scale business decisions. The session will dive into metaphactory's capabilities for capturing and organizing domain expertise in explicit semantic models, extracting insights from data, and sharing knowledge across the enterprise. All of this will be presented in the context of real-world use cases across diverse industries, like pharma and life sciences, engineering and manufacturing, and the academic space. Additionally, we will also look at how metaphactory was used to build the Dimensions Knowledge Graph, a knowledge graph 33 billion statements that brings together global research knowledge, public pharma ontologies and vocabularies, as well as public structured data sources.

### **3.3 Automating the creation of linked data repositories with Fluree software**

Mark Kram<sup>1</sup>, Eliud Polanco<sup>2</sup>, Tim Biedenkapp<sup>1</sup>

<sup>1</sup> adorsys GmbH & Co. KG

<sup>2</sup> Fluree

Today's modern data applications require being able to link both structured and unstructured sources of data into connected linked data repositories. In this presentation, Fluree will demonstrate how it uses machine learning, Generative AI and semantic inferencing to automatically convert data from database tables, spreadsheets and documents into a Knowledge Graph. Fluree will further demonstrate how new advances in semantic routing will enable more accurate natural language interaction and retrieval of factual information from linked data sources.

### 3.4 The Wikidata Open-Source Knowledge Graph, Semantic Search and AI

Lydia Pintscher, Philippe Saadé

Wikimedia Deutschland e.V.

**Keywords:** Wikidata, Semantic Search, AI

Wikidata is an open-source knowledge graph providing structured, standardized data for a wide range of applications. Launched in 2012 by Wikimedia Deutschland, it serves as a common repository of referenced statements on a vast array of topics that anyone can access. With over 112 million data objects, a community of over 24,000 monthly volunteers and available in over 300 languages, Wikidata is the world's largest collaboratively created open knowledge graph. Unlike Wikipedia's natural language articles, Wikidata's data can easily be processed by computers. This makes it invaluable for building applications, conducting data analysis, and enriching digital content across various domains. However, with the rapid rise of Large Language Models and Generative AI, new techniques are needed to ensure that Wikidata remains accessible to advanced AI systems that rely on unstructured text and deep semantic understanding.

To address this, the Wikidata Embedding Project was initiated in partnership with Jina.AI and DataStax, aiming to enhance Wikidata's accessibility to advanced AI technologies by integrating vector-based semantic search and embedding techniques. Building vector representations of Wikidata entities improves search capabilities and facilitates integration in AI applications, bridging the gap between structured data and semantic understanding. The project supports the open-source machine learning community by making its tools and models freely accessible, encouraging developers to build AI projects on top of Wikidata. The initiative promotes global access and community collaboration, allowing AI projects to inherit Wikidata's commitment to inclusivity, transparency, and community involvement. By integrating Wikidata, the project encourages users to contribute actively, correct inaccuracies, and collaborate to continually improve data quality. With support for multiple languages, the project ensures accessibility, making these tools and data resources available to a global audience.

### **3.5 KGI4NFDI – knowledge graph infrastructure for the research community**

Prof. Dr. Konrad Förstner

Deutsche Zentralbibliothek für Medizin (ZB MED) Informationszentrum Lebenswissenschaften

This talk will explore the mission of the KGI4NFDI project to establish a centralized, reusable Knowledge Graph Infrastructure (KGI) designed to enhance research domain interoperability and further objectives of the NFDI (National Research Data Infrastructure). We will discuss key components of the initiative, including the development of a Knowledge Graph (KG) registry and services that enables access to KGs across NFDI projects. Additionally, we will cover how the service aims to empower research communities to create decentralized KG instances through standardized approaches, technologies, and expert guidance. By conducting surveys in the community, providing documentation as well as consulting services and ontology harmonization, KGI4NFDI seeks to facilitate KG creation and promote consistent ontological practices aligned with the "One NFDI" vision and FAIR data principles.

## 4 Abstracts – Poster

### 4.1 An ontology classifying residues from the bioeconomy

Kim Schmidt<sup>1,2</sup>, Kai Sven Radtke<sup>1</sup>, Marco Selig<sup>1,2</sup>

<sup>1</sup> DBFZ - Deutsches Biomasseforschungszentrum gGmbH

<sup>2</sup> KIDA, AI consultancy

Due to climate change and the Earth overshoot day, that is becoming earlier every year, it is getting more important to use renewable resources and, in particular, biogenic waste and residues instead of fossil resources. There is currently no general scheme to characterize different biogenic wastes and residues and their according legal regulations. We intend to create an ontology that contains all information necessary to (i) clearly identify biomasses and (ii) to enable dealing with it in a bioeconomic and industrial context, which is why the ontology shall be shared with researchers, industry and policy makers. Since this is an early stage work, we want to start with 77 residues from the DBFZ Resource Database developed at the DBFZ Leipzig, Germany.

The DBFZ Resource Database considers different types of national biomass potentials (e.g., theoretical or technical biomass potential) for Germany. It currently contains information on 77 biogenic wastes and residues for several years. The potentials, possible utilizations and the significance in target markets (e.g. transport sector) can be explored via an interactive web application (see <https://datalab.dbfz.de/resdb/potentials>).

We want to create a simple way to share information about residues and answer questions about their usage and characteristics. For a high generalization of all this information, we will comply with several national and European regulations (i.a. European Waste Catalogue, Renewable Energy Directives (RED II and RED III), DIN EN ISO 17225-1:2021-10 for solid biofuels) and work together with experts from the bioeconomy sector. Thus, we will develop a tool that can be used by farmers, industry but also researchers and policy makers.



## 4.2 In the beginning there was complex data...

### ... or how local deep data annotation enables new perspectives on scientific data use, linked data publication, and large language model integration into workflows

Dr. Fabian Billenkamp

Friedrich-Loeffler-Institut (FLI)

Farm animal research, particularly in animal nutrition, provides insights into animal health, food and feed safety, and the ecological impacts of agriculture. Heterogeneous and highly complex data are acquired throughout an animal nutrition trial. The ethical treatment of animals in experiments is guided by the 3R principle (**Reduce** number of animals in experiments, **Replace** experiments by alternative methods & **Refine** methods to guarantee a minimum required suffering of animals in an experiment).

A contemporary interpretation of the 3R principle emphasizes that data generated in animal experiments adheres to the **FAIR** guidelines (Findable, Accessible, Interoperable, Reusable) to make data usable beyond its original context. The heterogeneity and complexity of animal nutrition data render it unusable without a metadata context, which is larger than the original data. Beyond its primary research use, data from farm animal trials can act as a bridge between research and agricultural practice.

By linking public domain ontologies with local ontologies tailored for scenarios within an animal research institute, detailed annotation of data in linked data formats creates condensed and standardized information. Information can be stored in local databases and, with minimal effort, converted into public domain data via ETL interfaces (Extract, Transform, Load). This approach helps to simplify publishing data in public domain databases according to 3R principles.

Standardized data annotation in local databases enhances reproducibility of analyses when these are stored in a repository. Annotation standards in local ontologies and linked data schemas facilitate the creation of vectors for large language model-augmented querying of data and analyses.

### **4.3 Rapid Alert Supply Network EXtractor (RASNEX): Mining and structuring supply chain information from incident reports**

Marc Lorenzen, Matthew Salewski, Daria Savvateeva

Bundesinstitut für Risikobewertung (BfR)

Public authorities of member countries use the EU's Rapid Alert System for Food and Feed (RASFF) to report human-health, animal-health, and environmental risks arising from food or feed. During an incident it is of utmost importance to rapidly identify the involved supply chain actors (operators) and withdraw causative products, as impacts on human and/or animal health and economic and/or environmental damages grow with time.

Information collected in the RASFF system during an outbreak is made available to investigators in the form of PDF-documents along with supporting attachments, e.g. delivery lists, invoices, etc., which contain specific and detailed information in various formats: .pdf, .doc and .xls files, as well as images and scans. As Member States continuously provide and update relevant information, the timely extraction and analysis of critical information is a challenge for risk assessment and management as investigators must read all incoming information and identify the crucial events, locations, and actors.

To facilitate investigations, the German Federal Institute for Risk Assessment (BfR), in cooperation with the European Food Safety Agency (EFSA), is developing an open source tool called the Rapid Alert Supply Network Extractor (RASNEX) to (partially) automate the process extracting key information from the RASFF notifications and their attachments using AI. The overall aim is to extract information and construct a representation of the incident and its affected supply chain using the JSON-based Universal Traceability data eXchange (UTX) format.

As LLMs, document-parsing and graph-databases are growing in use for information processing, we investigate how these can be applied in our project. However, because the texts contain highly sensitive information, there are restrictions to how and where they can be processed, limiting the use of online platforms or APIs. Here we present our tested solutions of using locally-running LLMs and their integration within the RASNEX dashboard.

#### 4.4 Enhancing Proteomics Meta-Analysis through Linked Data and LLM-Driven Metadata Extraction

Aileen Bahl<sup>1</sup>, Verónica Dumit<sup>2</sup>, Taras Günther<sup>1</sup>, Janis Stiegeler<sup>3</sup>

<sup>1</sup> German Federal Institute for Risk Assessment, Department for Biological Safety, Study Centre Supply Chain Modelling and Artificial Intelligence

<sup>2</sup> German Federal Institute for Risk Assessment, Department for Chemical and Product Safety, Fibre and Nanotoxicology

<sup>3</sup> Federal Research Centre of Cultivated Plants, Institute for Grapevine Breeding, Digitalisation and Precision Viticulture

**Keywords:** Proteomics, Metadata Extraction, LLM, Meta-Analysis

Proteomics meta-analyses offer substantial advantages by combining data from multiple studies, thereby enhancing statistical power and reliability of findings. By aggregating results, these analyses can uncover broader trends and biomarker patterns that may not be evident in individual studies, thus increasing confidence in protein markers linked to specific diseases, biological processes or toxicological outcomes. Meta-analyses also aid identification and control for variability across studies, enabling a more accurate interpretation of proteomic profiles across diverse populations and experimental conditions.

Meta-analyses often include datasets from existing databases. While the advantage of such databases is that they contain large amounts of datasets, the reuse of the contained data is frequently hampered by missing metadata. Extraction of experimental conditions, sample types, data processing methods, and analytical parameters allows researchers to accurately combine datasets and address potential sources of variability. In meta-analyses, such metadata also enable stratification based on relevant criteria, ensuring that comparisons are meaningful and statistically valid. Furthermore, comprehensive metadata facilitate data linkage by aligning proteomics data with complementary datasets, such as other omics layers, clinical records or toxicological measurements through common descriptors, identifiers, and controlled vocabularies. By employing Large Language Models (LLMs), metadata extraction can be carried out in an automated way greatly supporting tedious manual extraction efforts.

In this work, an innovative approach to meta-analysis by leveraging linked data methodologies alongside advanced LLMs for metadata extraction is presented. We focus on the systematic extraction of experimental setup information from a variety of proteomics datasets, enabling a comprehensive comparison and synthesis of findings across studies. By employing LLMs, we automate the extraction of critical metadata elements, such as sample types, experimental conditions, and analytical methods, which are often inconsistently reported. Our approach not only streamlines the metadata curation process but also enhances the reproducibility and transparency of proteomics research.

Funding acknowledgments:

This research has been funded by the German Federal Ministry of Food and Agriculture (BMEL) in the research project “KI- & Daten-Akzelerator (KIDA)” with project number 28KIDA004.

## 4.5 NFDI4Chem's Semantic Data Hub: Central Resource for FAIR Chemistry Data and Metadata

Oliver Koepler<sup>1</sup>, Philip Strömert<sup>1</sup>, Johannes Hunold<sup>1</sup>, Bhavin Katabathuni<sup>1</sup>, Pooya Oladazimi<sup>1</sup>, Tillmann G. Fischer<sup>2</sup>, Steffen Neumann<sup>3</sup>

<sup>1</sup>TIB - Leibniz Information Centre for Science and Technology

<sup>2</sup>UFZ - Helmholtz Center for Environmental Research GmbH

<sup>3</sup>IPB - Leibniz Institute of Plant Biochemistry

**Keywords:** NFDI4Chem, Metadata, Terminologies, Knowledge Graph, Linked Data

Open linked chemistry data about molecules, their reactions, and their characterisation data are highly relevant for many related disciplines such as agricultural science, life sciences, material science or pharmacy. The NFDI4Chem Semantic Data Hub (SDH) is a comprehensive suite of services designed for the creation and management of linked FAIR chemistry research data. It facilitates the integration of electronic lab notebooks (ELNs) with data repositories, including Chemotion, RADAR4Chem, nmrXiv, and further downstream services. The Terminology Service (TS) is leveraged for the annotation of data with standardized concepts from various Terminologies. In conjunction with this, the Search Service collates semantically enriched metadata, thereby providing a unified gateway to data discovery. To further enhance the future SDH, plans include the development of a Metadata Schema Service (MSS) and a Knowledge Graph (KG). The MSS will assist experts in the formalization of Minimum Information about Chemical Investigation (MICHI) profiles into consistent metadata schemas for the annotation of FAIR semantic-rich data. Moreover, it will facilitate schema transformations. This service is indispensable for the harmonization of metadata, which is instrumental in the integration of disparate datasets. Furthermore, the MSS will support the construction of a Chemistry Knowledge Graph. The Knowledge Graph will facilitate data exploration and identification of data relation in the chemistry data space. It will enable complex analysis, thereby facilitating in-depth insights from the data for both human users and machines. In essence, the NFDI4Chem Semantic Data Hub, through collaborative and integrated services, addresses challenges of data integration and knowledge management. It provides semantically rich, well-structured chemistry data and enables future AI-application on FAIR chemistry data.

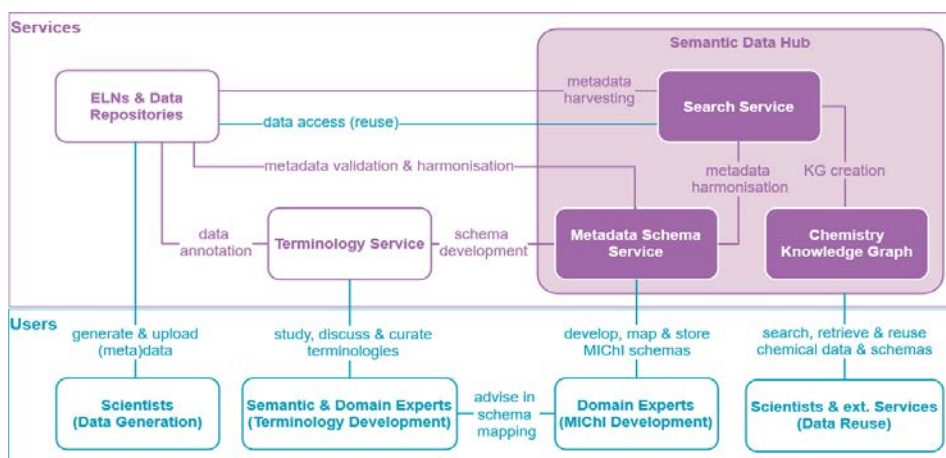


Image 1: Federation of Services, the Semantic Data Hub and stakeholder actions involved

## 4.6 Enhancing the discoverability of Predictive Models with the FAIR Scientific Knowledge Exchange (FSKX) Format

Uzair Jamil Khan, Taras Günther, Julia Golz, Thomas Schüler, Matthias Filter

German Federal Institute for Risk Assessment

**Keywords:** Linked Data, FSKX format, Knowledge Graph, metadata, ontology

Integrating metadata into Knowledge Graphs is important for enhancing the accessibility, interoperability and reusability of scientific data across disciplines. The FAIR Scientific Knowledge Exchange (FSKX) format, designed to standardize scientific models, includes information like metadata, model code, and simulation settings. This standardization facilitates efficient sharing of scientific models and improve their reproducibility.

In a recent project the FSKX metadata standard was converted into the FSKX Ontology (FSKXO) (<https://semanticlookup.zbmed.de/ols/ontologies/fskxo>) to enhance the standard towards the FAIR principles and make it compatible with existing interoperability standards (<https://www.w3.org/TR/json-ld>)

Here we follow up on this project and develop a Proof-of-Concept (PoC) FSKXO based Knowledge Graph. For this, we converted the metadata of existing FSKX models from the RAKIP-Web model repository (<https://knime.bfr.berlin/landingpage/RAKIP-Model-Repository>) to Resource Description Framework (RDF) and uploaded them together with FSKXO into a Virtuoso database (<https://virtuoso.openlinksw.com>). By doing this, querying the metadata of all available FSKX models becomes possible through the interconnected FSKX Knowledge Graph.

To test the FSKX Knowledge Graph, we created a simple front end with a faceted search to allow users complex queries over the FSKX Knowledge Graph. Thus, we enabled the possibility to query entities like parameters or as scholarly data in a flexible interface.

The results show that the FSKX metadata schema (<https://foodrisk-labs.bfr.bund.de/fskx-food-safety-knowledge-exchange-format>) can be used to create an FSKX Knowledge Graph that allows to run complex queries on model metadata across different models. Testing also showed that some parts of the ontology and conversion process should be adjusted to make the system easier to use. This project highlights the benefit of using standardized Internationalized Resource Identifiers (IRIs) to improve the accessibility and connection of scientific metadata. In the long run, these improvements will support better sharing, making it easier to reuse FSKX models, find information, and connect research across different areas.

Funding acknowledgments:

This research has been funded by the German Federal Ministry of Food and Agriculture (BMEL) in the research project “KI- & Daten-Akzelerator (KIDA)” with project number 28KIDA004.

## 4.7 AGROVOC und Large Language Models: Henne oder Ei

Esther Mietzsch, Daniel Martini, Jascha Jung

Kuratorium für Technik und Bauwesen in der Landwirtschaft e.V.

AGROVOC ist der multilinguale Thesaurus der FAO. Er umfasst Konzepte zu den Themen Landwirtschaft und Ernährung, aber auch aus Fischerei, Umwelt, Forstwirtschaft und verwandten Gebieten. Gegenwärtig enthält AGROVOC bereits mehr als 41.000 Konzepte mit mehr als 1,1 Millionen Bezeichnern in bis zu 42 Sprachen. Technisch gesehen ist AGROVOC ein SKOS-XL Konzeptschema und wurde als Linked Open Data (LOD)-Set veröffentlicht.

Die Konzepte sind in einer Hierarchie von Ober- und Unterkonzepten eingeordnet. Jedes Konzept hat einen bevorzugten Bezeichner (skos:prefLabel) mindestens in Englisch, meistens aber auch in den anderen FAO-Amtssprachen und häufig in weiteren Sprachen. Viele Konzepte haben auch weitere Bezeichner, also Synonyme (skos:altLabel). Neuere Konzepte sind durch Definitionen mit Quellenangabe näher erläutert. Außerdem sind die Konzepte durch nicht-hierarchische Relationen, wie z.B. agrontology:has-Product, miteinander vernetzt.

Um Lücken in AGROVOC zu füllen, existieren verschiedene Ansätze. Zum einen werden die Einträge in Vokabularen und Glossaren zu einem bestimmten Thema mit den existierenden Bezeichnern abgeglichen und bislang fehlende Konzepte hinzugefügt. Zum anderen werden Keywords aus Fachartikeln oder Indexeinträge aus wissenschaftlichen Literaturdatenbanken, z.B. AGRIS als neue Einträge vorgeschlagen. Schließlich wird diskutiert, ob Natural Language Processing sowie die Extraktion von Schlagworten mithilfe von Large Language Models die Aktualisierung von AGROVOC beschleunigen können (Panoutsopoulos und Brewster, 2022).

Auch wenn diese Methoden helfen können, mögliche Kandidaten für neue Konzepte zu identifizieren, spiegeln beispielsweise statistische Häufigkeiten nicht notwendigerweise die Relevanz eines Konzeptes wider. Um die Qualität von AGROVOC sicherzustellen, ist es notwendig, existierende und neue Konzepte, ihre Position in der Hierarchie und die Bezeichner in den verschiedenen Sprachen und die Definitionen sorgfältig zu prüfen. Dabei ist insbesondere auch die gute wissenschaftliche Praxis einer angemessenen Quellenrecherche und des Zitierens zu berücksichtigen. Dazu ist die Expertise der Editorinnen und Editoren und des Herausgeberteams unerlässlich.

## 4.8 Linked Data Toolbox: Enabling Interoperability across Risk Assessment Research

Julia Golz, Uzair Jamil Khan, Taras Günther

German Federal Institute for Risk Assessment

**Keywords:** Risk Assessment, Linked Data, Data Interoperability, Knowledge Graphs, Resource Description Framework (RDF)

Risk assessment research plays an important role in consumer protection. This requires the integration and analysis of data across diverse scientific domains, as has been highlighted by initiatives like One Health. However, the lack of interoperability between datasets and disciplines often leads to data silos, lowering efficient collaboration and knowledge acquisition. Improving data and knowledge interoperability across disciplines is essential to overcome these challenges. Technologies such as Knowledge Graphs and Linked Data offer promising solutions by operationalizing FAIR principles (e. g. as suggested by <https://faircookbook.elixir-europe.org/content/home.html>) through the use of ontologies and standards. Yet, their practical application in risk assessment research presents technical and methodological challenges.

Therefore, we developed the Linked Data Toolbox, consisting of a framework designed to generate linked risk assessment data and facilitate interoperability. The Toolbox includes the Linked Data Extractor, a tool using a large language model (LLM) to convert plain text into Resource Description Framework (RDF) representations. Since ontology matching is often labor-intensive, we also developed the Ontology Recommender App, which uses a LLM to suggest relevant ontologies based on input data concepts.

Additionally, the Linked Data Toolbox includes a workflow to enrich data with metadata using the Data Catalog Vocabulary (DCAT) standard, widely adopted by official agencies to enable seamless metadata exchange. Another key component facilitates RDF generation, ensuring effective data structuring and linking. To further enhance interoperability, the Toolbox incorporates a process for establishing an internal vocabulary aligned with the Simple Knowledge Organization System (SKOS) standard, integrating ontology matchings across domains.

By combining these components, the Linked Data Toolbox enables the semi-automatic conversion of data into Linked Data. This approach ensures cross-domain interoperability, and prepares data for advanced application in artificial intelligence, paving the way for more cohesive and advanced risk assessment research.

Funding acknowledgments:

This research has been funded by the German Federal Ministry of Food and Agriculture (BMEL) in the research project “KI- & Daten-Akzelerator (KIDA)” with project number 28KIDA004.

## 4.9 Integrating Various Supply Network Abstractions Through a Uniform Model and Knowledge Graphs

Daniel Henselmann<sup>1</sup>, Andreas Harth<sup>1,2</sup>

<sup>1</sup>Fraunhofer IIS

<sup>2</sup>Friedrich-Alexander University Erlangen-Nürnberg

Increases in supply network disruptions and sustainability efforts bring along diverse business questions like “Which industries highly depend on the oil industry?”, “What are critical products in the supply networks of an enterprise?”, or “What percentage of aluminum in an item is recycled?”. People’s understanding of a supply network, inherited in these questions, differs. Different abstractions of supply networks answer different business questions but are still interrelated.

We represent the structure of supply networks on various abstractions as ontologies, using a uniform model. Relations between associated entities of different abstraction interlink various representations. Inference of supply network structures from lower to higher abstraction and aggregation of data enable to integrate data from available sources with varying abstraction in a knowledge graph. The trade volume from BYD to BMW (enterprise abstraction) for example is the aggregate of the value of BYD’s Blade Batteries used to build BMW’s cars like the iX (product abstraction). The relation between iX and its associated entity BMW bridges the abstractions.

The model provides a simple yet semantically found structure of a supply network with the flexibility to append data for specific use cases. Thus, the model is easy to instantiate as Linked Data and furthermore allows for simple queries regarding business questions on integrated data of varying abstraction. We demonstrate our approach with an example of real-world data on the semiconductor industry.



#### 4.10 Integration of knowledge graphs into LLMs: Are Graph Language Models better for reasoning in intelligent systems for biomedical research?

Kerstin Neubert

Bundesinstitut für Risikobewertung

**Keywords:** Large Language Models, Knowledge Graphs, Reasoning, Information systems

Knowledge graphs (KGs), such as the Unified Medical Language System (UMLS) are used to store and retrieve explicit structured knowledge in critical fields like healthcare. Relation extraction (RE) is a significant part of an Information Extraction (IE) pipeline for the construction of KGs from unstructured data. It identifies all relationships between pre-defined entities in a text and assigns them to a category.

Large Language models (LLMs) showed remarkable performance in different Natural Language Processing (NLP) tasks due to their abilities in language understanding. They have limitations when complex reasoning is needed together with a high level of factual consistency, trustworthiness and explainability. To combine the strength of LLMs and KGs, current research investigates methods to augment LLMs with knowledge from KGs. Recently, the graph language model (GLM) was proposed as a new model architecture to fully integrate graphs and texts during training (Plenz and Frank, 2024).

In this work, an approach for relation extraction by using a GLM is presented. The heterogeneous Pharmacological Medical Biochemical Network (PharMeBInet) was selected as a source knowledge graph, because of its integration of several popular databases (e.g. CTD, DrugBank, ClinVar). The implemented method is applied to the ChemDisGene dataset, a collection of research abstracts that was annotated with Chemicals, Disease and Gene entities and their pairwise relations. To obtain aligned graphs for the texts, entities are linked with semantic triplets (subject-predicate-object expressions) using Cypher queries to the Neo4j graph database of PharMeBInet. After jointly training with graph and text data the accuracy of the GLM is evaluated for relation classification.

It was shown that GLMs significantly improve reasoning over text and graph-structured data from common sense knowledge (ConceptNet, Wikipedia) compared to graph models and LLMs. We assume and try to demonstrate, that this approach can be transferred to biomedical contexts as well.

Funding acknowledgments:

This research has been funded by the German Federal Ministry of Food and Agriculture (BMEL) in the research project “KI- & Daten-Akzelerator (KIDA)” with project number 28KIDA004.

## 4.11 Operationalizing Linked Data in Public Health Agencies Through the One Health Model

Thomas Art Burke, Antje Risius

University of Göttingen

**Keywords:** One Health, EPCIS, Ontologies, food safety, antibiotic resistance

Advancements in food traceability promise comprehensive supply chain digitalization and novel food safety informatics. However, collaborating successfully across administrative and disciplinary silos in public health regulatory programs requires sophisticated adoption of novel data management tools to handle multi-dimensionality, mixed identifier systems, latent semantics, and unstructured data that accompany new information streams. Despite the existence of knowledge graphs, ontologies, and linked data standards, their adoption into scientific processes present operational hurdles.

Simultaneously, One Health is envisioned to be a primary framework for public health subject matter expert collaboration but often struggles to move past qualitative applications. Herein, we propose how linked data tools and standards may be used as the basis for a novel One Health methodology, particularly through the introduction of software development techniques and strategies. The capture of batch/lot information at critical tracking events better enables the usage of temporally and geographically relevant environmental data in food supply chain public health initiatives. Iterative data extraction, transformation, and load processes can become part of One Health frameworks to allow for the operationalization, documentation, and dissemination of quantitative inquiry. Embedding the expectation of data reconciliation into the scientific process may be able to better link new data management strategies and tools into existing organizational structures and operating patterns. In our example framework, we use linked data standards (EPCIS 2.0 and GS1 Digital Link) as a basis for using an institutional knowledge graph, leveraging the FoodOn ontology.

Using antibiotic application and corresponding resistance in food supply chains as a hypothetical case, we identify data types, relevant subject matter expertise, methodologies, and data discovery hurdles that would fit within the digital One Health working process. We provide example user stories, linked data cross-linkages, and graph database deployments. We plan to test with an institution as a complementary process to novel graph database introduction.

## 4.12 Faktenerkennung für Knowledge-Graphs aus unstrukturierten Textkorpora

Stefan Reichert<sup>1</sup>, Kiara Marnitt Ascencion Arevalo<sup>2</sup>, Prof. Dr. Roland Zimmermann<sup>2</sup>, Prof. Dr. Ralph Blum<sup>1</sup>

<sup>1</sup>Fraunhofer IIS

<sup>2</sup>TH Nürnberg

**Keywords:** Ontology Engineering, Large Language Models, Natural Language Processing, Ontology-based Information Extraction

Der methodische Fokus der Forschungsgruppe liegt auf automatisierter Identifikation, Extraktion und Verknüpfung von Entitäten und Beziehungen in unstrukturierten Datenquellen (z. B. Patenten, Publikationen, Pressemitteilungen). Diese Methoden werden zur Markt- und Technologiebeobachtung sowie zur strategischen Vorausschau eingesetzt. Dafür werden Wissensgraphen (KG) für Wissensrepräsentation und große Sprachmodelle (LLMs) für Natural Language Processing (NLP) eingesetzt.

Bisheriger Schwerpunkt im Bereich NLP war die Faktenextraktion im Sinne einer automatisierten Extraktion von Einzelinformationen aus textbasierten Quellen und der anschließenden Aggregation von Wissen zu semantisch abgeschlossenen Themenclustern. Im Bereich der Wissensrepräsentation wurden zur Strukturierung von Informationen Ontologien entwickelt [1] und diese als Basis für die automatisierte, gezielte Extraktion von Entitäten und Relationen mittels LM-basierter Ansätze verwendet. Damit konnten komplexe Wissensstrukturen in Wissensgraphen mit Entitäten, Beziehungen und Attributen generiert werden, die komplexes Wissen übersichtlich und maschinenlesbar speichern.

Die noch weitergehende Kombination beider Forschungsgebiete soll die Erstellung von KG vereinfachen und steht in einem Anwendungsszenario des DFG-Verbundprojektes „DuraFuelCell“ im Fokus: Darin gilt es, Trends und Szenarios für langlebige, effiziente Wasserstoff-Brennstoffzellensysteme zu erforschen und eine KI-basierte Wissensdatenbank zu definieren.

Zwei relevante Meilensteine sind:

1. Eine domänenspezifische Ontologie für den Bereich Brennstoffzellen soll in Zusammenarbeit mit Domänenexperten erstellt werden. Diese Ontologie erlaubt es, Akteure sowie technologische Entwicklungen zu klassifizieren und ihre Beziehungen zueinander zu identifizieren. Hierfür wird im Jahr 2025 eine automatisierte Anforderungserhebung von Domänenexperten für das anschließende Ontologie-Engineering entwickelt und getestet (Rapid Ontology Prototyping).
2. Durch die Verknüpfung von NLP-basierten Algorithmen zur automatischen Extraktion und Aggregation strukturierter Daten aus Textquellen sowie durch automatisierte Erkennung von Themenclustern soll die Definition von Ontologien vereinfacht werden. Dazu wird der AutoOnto-Ansatz (Automated Domain Ontology Generation [2]) weiter ausgebaut.

References:

- [1] Ascencion Arevalo, K. M., Neunsinger, C., Zimmermann, R., Blum, R., Weakly, K. (2023): HOLY: An ontology covering the Hydrogen market. International Semantic Web Conference 2023 (ISWC), DOI: [10.1007/978-3-031-47243-5\\_1](https://doi.org/10.1007/978-3-031-47243-5_1)
- [2] Ascencion Arevalo, K. M., Ambre, S., Dorsch, R. (2024): AutOnto: Towards A Semi-Automated Ontology Engineering Methodology. Knowledge Graphs and Semantic Web. KGSWC 2024.

### **4.13 Agro-Intelligence Hub (AI-Hub): ein Pilotprojekt zur Verknüpfung von Datenbeständen des JKI und BfR mittels „Linked Data“ Technologien**

Matthias Filter<sup>1</sup>, Carsten Kneuer<sup>1</sup>, Taras Günther<sup>1</sup>, Vikas Kumar<sup>1</sup>, Steffen Kecke<sup>2</sup>

<sup>1</sup>Bundesinstitut für Risikobewertung

<sup>2</sup>Julius Kühn-Institut

**Keywords:** Angewandte Forschung, Datenintegration, Wissensgraphen, Semantic Web

Das für 2025 geplante Pilotprojekt Agro-Intelligence Hub (AI-Hub) zielt darauf ab, bestehenden Datenbestände des BfR und JKI mit Hilfe von Linked Data Technologien zu verknüpfen. Das Pilotprojekt konzentriert sich dabei auf Datenbestände, die einen Beitrag zur Politikberatung im Themenfeld „Veränderte Gesundheitsrisiken durch Klimawandel an der Schnittstelle zwischen Pflanzenschutz und Verbrauchersicherheit“ leisten können.

Mit dem Pilotprojekt soll erprobt werden, wie bestehende Daten unter Einhaltung der rechtlichen Vorgaben verknüpft werden, sodass die Eigentümerschaft und Schutzbedarfe der Daten gewahrt bleiben. Die Projektpartner des BfR und JKI werden dazu bestehende Datenbanken aus den Bereichen der chemischen und mikrobiellen Pflanzenschutzmittel-Bewertung vernetzen, welche Informationen zu Feldversuchs- und Geodaten der Anwendung und Exposition von Pflanzenschutzmittel sowie toxikologische Informationen zu Gesundheitsgefahren durch die enthaltenen Stoffe und in der Umwelt neu entstehenden Abbauprodukte enthalten. Darüber hinaus werden Informationen aus öffentlich zugänglichen Wissensgraphen für weiterführende, innovative Abfragen und Analysen eingebunden, die praktischen Mehrwert für die Nutzer liefern. Aktuell werden diese Informationen in verschiedensten Datenformaten und an verschiedenen Standorten vorgehalten und gepflegt. Über das geplante KI-gestützte Datenportal wird demonstriert, dass eine derartige Datenintegration einen erheblichen Mehrwert für die Wissenschaft, die Allgemeinheit und Akteure der Land- und Ernährungswirtschaft sowie die Überwachungsorgane generiert. Damit steht diese Initiative im Einklang mit umfassenderen Nachhaltigkeitszielen und fördert Innovation und Widerstandsfähigkeit landwirtschaftlicher Praktiken angesichts der Herausforderungen des Klimawandels.

#### **4.14 SIMLEARN – Betriebliche Entscheidungsunterstützung durch ontologiegestützte Integration von Simulationsmodellen, Systemen für maschinelles Lernen und Planungsdaten**

Nils Reinosch<sup>1</sup>, Alexander Münzberg<sup>2</sup>, Daniel Martini<sup>1</sup>, Alexander Niehus<sup>1</sup>, Liv Seuring<sup>1</sup>, Christian Troost<sup>3</sup>, Rajiv Kumar Srivastava<sup>4</sup>, Thomas Berger<sup>3</sup>, Thilo Streck<sup>4</sup>, Ansgar Bernardi<sup>2</sup>

<sup>1</sup>Das Kuratorium für Technik und Bauwesen in der Landwirtschaft

<sup>2</sup>Deutsches Forschungszentrum für Künstliche Intelligenz

<sup>3</sup>Universität Hohenheim, Ökonomik der Landnutzung in den Tropen und Subtropen

<sup>4</sup>Universität Hohenheim, Biogeophysik

**Keywords:** Ontologie, Simulation, Maschinelles Lernen, Datenmanagement, Smart Data

Das für 2025 geplante Pilotprojekt Agro-Intelligence Hub (AI-Hub) zielt darauf ab, bestehenden Datenbestände des BfR und JKI mit Hilfe von Linked Data Technologien zu verknüpfen. Das Pilotprojekt konzentriert sich dabei auf Datenbestände, die einen Beitrag zur Politikberatung im Themenfeld „Veränderte Gesundheitsrisiken durch Klimawandel an der Schnittstelle zwischen Pflanzenschutz und Verbrauchersicherheit“ leisten können.

Mit dem Pilotprojekt soll erprobt werden, wie bestehende Daten unter Einhaltung der rechtlichen Vorgaben verknüpft werden, sodass die Eigentümerschaft und Schutzbedarfe der Daten gewahrt bleiben. Die Projektpartner des BfR und JKI werden dazu bestehende Datenbanken aus den Bereichen der chemischen und mikrobiellen Pflanzenschutzmittel-Bewertung vernetzen, welche Informationen zu Feldversuchs- und Geodaten der Anwendung und Exposition von Pflanzenschutzmittel sowie toxikologische Informationen zu Gesundheitsgefahren durch die enthaltenen Stoffe und in der Umwelt neu entstehenden Abbauprodukte enthalten. Darüber hinaus werden Informationen aus öffentlich zugänglichen Wissensgraphen für weiterführende, innovative Abfragen und Analysen eingebunden, die praktischen Mehrwert für die Nutzer liefern. Aktuell werden diese Informationen in verschiedensten Datenformaten und an verschiedenen Standorten vorgehalten und gepflegt. Über das geplante KI-gestützte Datenportal wird demonstriert, dass eine derartige Datenintegration einen erheblichen Mehrwert für die Wissenschaft, die Allgemeinheit und Akteure der Land- und Ernährungswirtschaft sowie die Überwachungsorgane generiert. Damit steht diese Initiative im Einklang mit umfassenderen Nachhaltigkeitszielen und fördert Innovation und Widerstandsfähigkeit landwirtschaftlicher Praktiken angesichts der Herausforderungen des Klimawandels.

Funding acknowledgements:

Das Projekt SIMLEARN wird mit Mitteln des Bundesministeriums für Bildung und Forschung (BMBF) unter Förderkennzeichen 01|S19073 gefördert

## 5 Verzeichnis der Autorinnen und Autoren

Albrecht, Katharina .....	16	Löhn, Lea.....	8
Ascencion Arevalo, Kiara Marnitt .....	35	Lorenzen, Mark.....	26
Bahl, Aileen .....	27	Maak, Stefan.....	15
Berger, Thomas .....	38	Martini, Daniel .....	11, 16, 30, 38
Bernardi, Ansgar.....	13, 38	Mietzsch, Esther .....	16, 30
Biedenkapp, Tim.....	21	Münzberg, Alexander.....	38
Billenkamp, Fabian .....	25	Neubert, Kerstin .....	33
Blum, Ralph .....	35	Neumann, Steffen.....	28
Born, Stefan .....	15	Niehus, Alexander.....	38
Bornscheuer, Uwe T. ....	15	Oladazimi, Pooya .....	28
Burke, Thomas Art.....	34	Pintscher, Lydia.....	18, 22
Dehm, Stefan .....	14	Polanco, Eliud .....	21
Dörr, Mark.....	15	Radtke, Kai Sven .....	24
Dumit, Verónica .....	27	Reichert, Stefan .....	35
Filter, Matthias.....	14, 29, 37	Reinosch, Nils.....	38
Fischer, Tillmann G. ....	28	Risius, Antje .....	34
Fluck, Juliane .....	9	Roggenland, Aaron .....	16
Förstner, Konrad .....	23	Saadé, Philippe .....	22
Golla, Burkhard .....	16	Salewski, Matthew.....	26
Golz, Julia .....	14, 29, 31	Savvateev, Iurii.....	14
Gottschalk, Nathalie.....	8	Savvateeva, Daria .....	26
Günther, Taras .....	14, 17, 27, 29, 31, 37	Schimmler, Sonja .....	10
Haase, Peter .....	20	Schmidt, Kim .....	24
Harth, Andreas .....	12, 32	Schüler, Thomas.....	29
Henselmann, Daniel .....	32	Selig, Marco .....	24
Hunold, Johannes.....	28	Seuring, Liv .....	38
Jung, Jascha.....	16, 30	Srivastava, Rajiv Kumar.....	38
Katabathuni, Bhavin .....	28	Stahl, Ulrike .....	11
Kecke, Steffen .....	37	Stiegeler, Janis .....	27
Khan, Uzair Jamil .....	29, 31	Streck, Thilo .....	38
Kneuer, Carsten.....	37	Strömert, Philip.....	28
Knuth, Magnus.....	19	Troost, Christian .....	38
Koepler, Oliver .....	28	Vorisek, Carina .....	9
Kram, Mark .....	21	Zarske, Michael.....	17
Kumar, Vikas .....	37	Zimmermann, Roland .....	35

### Über das BfR

Das Bundesinstitut für Risikobewertung (BfR) ist eine wissenschaftlich unabhängige Einrichtung im Geschäftsbereich des Bundesministeriums für Ernährung und Landwirtschaft (BMEL). Es berät die Bundesregierung und die Bundesländer zu Fragen der Lebensmittel-, Chemikalien- und Produktsicherheit. Das BfR betreibt eigene Forschung zu Themen, die in engem Zusammenhang mit seinen Bewertungsaufgaben stehen.

### Impressum

BfR Abstract

Linked Data im Zeitalter der generativen KI – Neue Perspektiven oder überholte Vision?

Für den Inhalt der Abstracts sind deren Autorinnen und Autoren verantwortlich.

Herausgeber:

#### **Bundesinstitut für Risikobewertung (BfR)**

Max-Dohrn-Straße 8-10

10589 Berlin

T +49 30 18412-0

F +49 30 18412-99099

bfr@bfr.bund.de

bfr.bund.de

Anstalt des öffentlichen Rechts

Vertreten durch den Präsidenten Professor Dr. Dr. Dr. h.c. Andreas Hensel

Aufsichtsbehörde: Bundesministerium für Ernährung und Landwirtschaft

USt-IdNr: DE 165893448

Berlin 2024

40 Seiten

