



netidee

PROJEKTE

Web of AI

Endbericht | Call 15 | Projekt ID 5158

Lizenz CC-BY

Inhalt

1	Einleitung.....	3
2	Projektbeschreibung.....	3
3	Verlauf der Arbeitspakete.....	3
3.1	Arbeitspaket 1 - <i>Detailplanung und Formales am Projektstart</i>	3
3.2	Arbeitspaket 2 - <i>Intelligence Task Ontology (ITO)</i>	4
3.3	Arbeitspaket 3 - <i>Interaktive Website</i>	5
3.4	Arbeitspaket 4 - <i>Studie</i>	7
3.5	Arbeitspaket 5 - <i>Dokumentation und Formales am Projektende</i>	9
4	Umsetzung Förderauflagen	9
	Die geforderte „Beschreibung & Mitberücksichtigung des Katalogbefüllungsprozesses“ wurde bereits in der Ergänzung vor Förderabschluss übermittelt; detailliertere Beschreibungen des Prozesses wurden in den beiden Publikation bereitgestellt.....	9
5	Liste Projektergebnisse	10
6	Verwertung der Projektergebnisse in der Praxis	11
7	Öffentlichkeitsarbeit/ Vernetzung.....	12
8	Eigene Projektwebsite.....	12
9	Geplante Aktivitäten nach netidee-Projektende.....	12
10	Anregungen für Weiterentwicklungen durch Dritte.....	12

1 Einleitung

Die Fortschritte in der AI verändern die Natur des World Wide Web, denn neben menschlicher Intelligenz wird es zusätzlich immer mehr zum Marktplatz und Interaktionsraum für AI — ein hybrides System aus Mensch und Maschine entsteht. Dabei fällt es immer schwerer, den Überblick über bestehende Services, deren Funktionalität und deren Performanz zu behalten, und verschiedene Services in komplexen Workflows optimal zu vereinen.

2 Projektbeschreibung

Web of AI hilft, AI im WWW besser zu konzeptualisieren und zu nutzen.

Wir entwickelten basierend auf der Web Ontology Language (OWL) zusammen mit internationalen Partnern ein umfassendes, formales Modell von Tasks und Systemen der künstlichen Intelligenz, genannt „Intelligence Task Ontology and Knowledge Graph“ (ITO). ITO deckt das gesamte Spektrum der AI von Bilderkennung über Textübersetzung bis zur EKG-Interpretation ab.

Mithilfe des entwickelten Modells können AI Services semantisch annotiert und dadurch besser gefunden und verwendet werden. Das entwickelte Modell ermöglicht fundamentale ökonomische und wissenschaftliche Analysen: Welche Problemstellungen/Geschäftsprozesse können mittlerweile gut durch AI Services adressiert werden? Welche Prozesse sind noch nicht gut adressiert?

3 Verlauf der Arbeitspakete

3.1 Arbeitspaket 1 - *Detailplanung und Formales am Projektstart*

Der Projektstart wurde erfolgreich abgewickelt (Vertragsprüfung, Unterzeichnung des Vertrages, Projektplanung basierend auf Excel-Vorlage, Start der Projektwebsite, Förderabruf für Förderrate 1).

3.2 Arbeitspaket 2 - Intelligence Task Ontology (ITO)

Skripte zum automatisierten Import externer Benchmark Daten wurden komplett neu implementiert und dadurch robuster und schneller gemacht. Die Modellierung der Daten wurde verbessert um Abfragen einfacher zu gestalten und das Modell etwas kompakter zu machen.

Eine umfassende Manuelle Kuratierung der Ontologie-Klassen und Performance-Metriken wurde durchgeführt. Version 1.0 der OWL Ontologie wurde publiziert. ITO und assoziierte Ressourcen wurden auf Github frei verfügbar gemacht (<https://github.com/OpenBioLink/ITO>).

Intelligence Task Ontology and Knowledge Graph (ITO). An AI ontology.

A comprehensive knowledge graph of artificial intelligence tasks and benchmarks

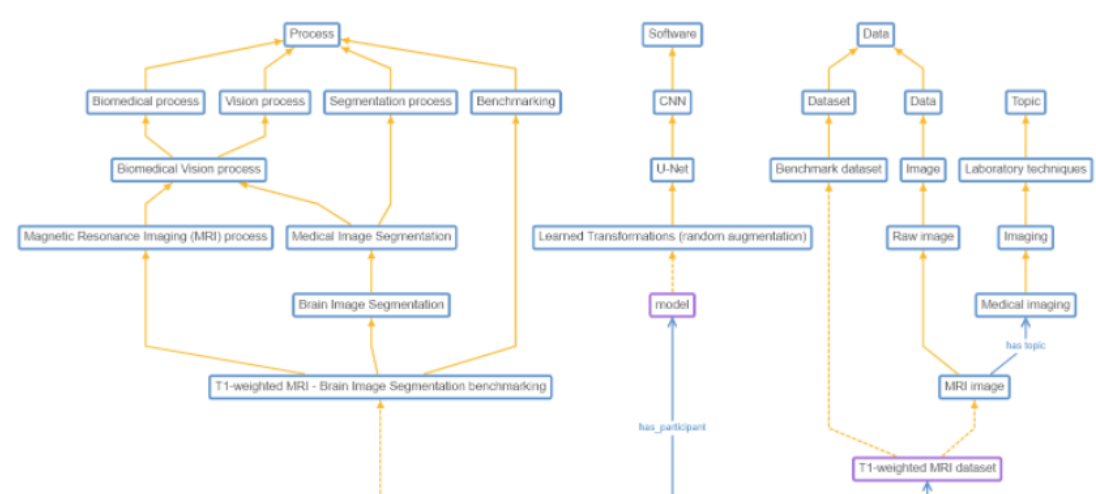
The Intelligence Task Ontology and Knowledge Graph (ITO) provides comprehensive, curated and interlinked data of artificial intelligence tasks, benchmarks, AI performance metrics, benchmark results and research papers.

You can explore ITO in our [ITO Explorer](#) dashboard or browse the ITO class hierarchy [online at BioPortal](#).

Examples

ITO aims to provide a richly structured hierarchy of processes, algorithms, data and performance metrics. Data on thousands of AI benchmark results have been imported from [Papers With Code](#) and are further curated.

An example of a benchmark result embedded in an ontological hierarchy:



The diagram illustrates an ontological hierarchy. At the top level, 'Process' branches into 'Biomedical process', 'Vision process', 'Segmentation process', and 'Benchmarking'. 'Biomedical process' further branches into 'Biomedical Vision process' and 'Magnetic Resonance Imaging (MRI) process'. 'Vision process' branches into 'Medical Image Segmentation' and 'Brain Image Segmentation'. 'Segmentation process' branches into 'Medical Image Segmentation' and 'Brain Image Segmentation'. 'Benchmarking' branches into 'T1-weighted MRI - Brain Image Segmentation benchmarking'. 'Data' branches into 'Dataset', 'Image', and 'Topic'. 'Dataset' branches into 'Benchmark dataset' and 'Raw image'. 'Image' branches into 'Raw image' and 'MRI image'. 'Topic' branches into 'Laboratory techniques' and 'Imaging'. 'Imaging' branches into 'Medical imaging'. 'MRI image' has a 'has topic' relationship with 'Medical imaging'. 'Software' branches into 'CNN' and 'U-Net'. 'U-Net' branches into 'Learned Transformations (random augmentation)'. 'model' has a 'has_participant' relationship with 'Learned Transformations (random augmentation)'. 'T1-weighted MRI dataset' has a 'has_participant' relationship with 'T1-weighted MRI - Brain Image Segmentation benchmarking'.

Ausschnitt der ITO Github Seite (<https://github.com/OpenBioLink/ITO>)

Google Colab bzw. Jupyter Notebooks wurden bereitgestellt, um EntwicklerInnen den Einstieg in die Arbeit mit den ITO Ressourcen möglichst einfach zu machen.

- ☰ Inhalt
- 🔍 Creating ontology statistics
- [x] Are there entities/properties with more than one label? (data quality check)
- ☐ Property stats on individual level (how many measurements with each property?)
- Property stats on benchmark level (how many benchmarks with each property?)
- Total number of benchmarks
- Number of different metrics per benchmark (for those where any metrics are listed)
- Number of different metrics per benchmark (for those where any metrics are listed – with topmost properties)
- Total number of papers
- Number of papers that have at least one annotated metric
- Time span of publications covered by PWC
- Number of metrics aggregated by high-level process / task
- Number of benchmarks for each high-level process
- Code for generating JSON for hierarchy visualization in d3js
- ☑ Bereich

+ Code + Text Verbinden ▾

	benchmark_count
	0 2298

▾ Number of different metrics per benchmark (for those where any metrics are listed)

```
[ ] result = graph.query("""
SELECT ?benchmark_label ?benchmark
(COUNT(DISTINCT ?performance_measure) AS ?performance_measure_count)
(GROUP_CONCAT(DISTINCT ?performance_measure_label;SEPARATOR=";") AS ?performance_measure_labels)
WHERE {
?performance_measure rdfs:subPropertyOf* ito:performance_measure ;
rdfs:label ?performance_measure_label .
?benchmark rdfs:subClassOf ito:Benchmarking ;
rdfs:label ?benchmark_label ;
rdfs:subClassOf ?property_restriction .
?property_restriction owl:Property ?performance_measure .
}
GROUP BY ?benchmark_label ?benchmark
ORDER BY DESC(?performance_measure_count)
""", inits = namespaces)
df = sparql_result_to_df(result)
df.to_csv('number-of-different-metrics-per-benchmark.csv', index=False)
df
```

	benchmark_label	benchmark	performance_measure_count	performance_measure_label
0	MSR-VTT - Video Retrieval benchmarking	https://ai-strategies.org/ontology/ITO_07273	10	text-to-video Median Rank;text-to-video R-at-1...
1	COCO test-dev - Keypoint Detection benchmarking	https://ai-strategies.org/ontology/ITO_11334	10	AR,AR50,AR75,ARM,AP,AP50,AP75,APM,APL,ARI...
2	MIT-States, generalized split - Compositional ...	https://ai-strategies.org/ontology/ITO_07763	9	H-Mean,Test AUC top 1,Test AUC top 2,Test AUC ...
3	COCO test-challenge - Keypoint Detection bench...	https://ai-strategies.org/ontology/ITO_11367	9	AR,AR50,AR75,ARM,AP,AP50,AP75,APL,ARI...
4	RCV1 - Text Classification benchmarking	https://ai-strategies.org/ontology/ITO_14703	9	nDCG-at-5;nDCG-at-3;nDCG-at-1;Accuracy,P-at-1...
...
2293	Tashkeela - Arabic Text Diacritization benchma...	https://ai-strategies.org/ontology/ITO_17254	1	Diacritic Error Rate
2294	CIFAR10-DVS - Event data classification benchm...	https://ai-strategies.org/ontology/ITO_17261	1	Accuracy
2295	Rotated MNIST - Rotated MNIST benchmarking	https://ai-strategies.org/ontology/ITO_17264	1	Test error
2296	StereoSet - Bias Detection benchmarking	https://ai-strategies.org/ontology/ITO_17269	1	ICAT Score
2297	Covers80 - Cover song identification benchmarking	https://ai-strategies.org/ontology/ITO_17288	1	MAF

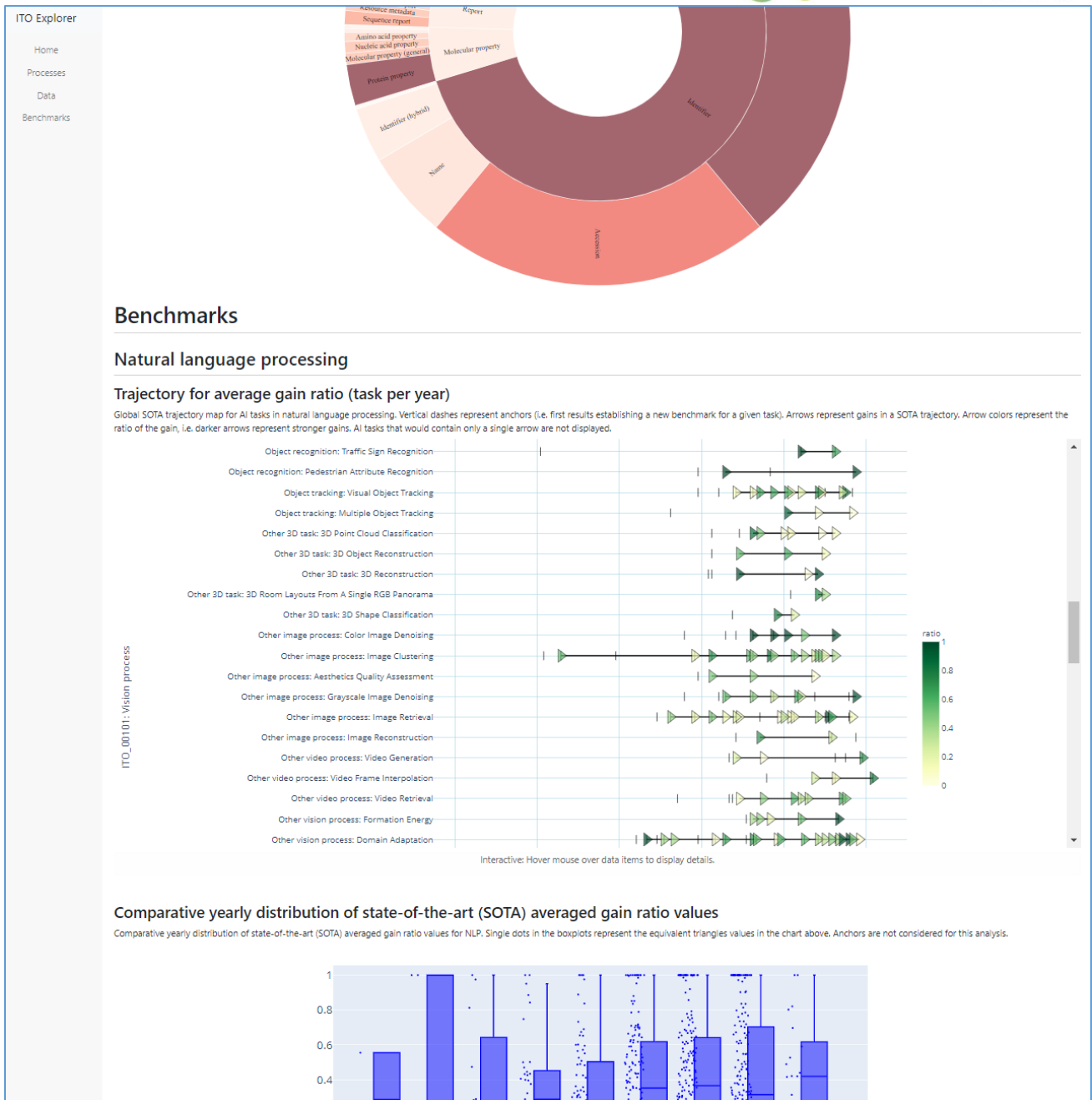
2298 rows × 4 columns

▾ Number of different metrics per benchmark (for those where any metrics are listed – with topmost properties)

Ausschnitt eines im Web zugänglichen Google Colab Notebooks mit Demonstrationen und Query-Templates für eigene Entwicklungen basierend auf ITO.

3.3 Arbeitspaket 3 - Interaktive Website

Eine interaktive Website zur Exploration der Daten in der ITO Wissensbasis wurde erstellt (<https://openbiolink.github.io/ITOExplorer/>).



Ausschnitt der interaktiven Website (<https://openbiolink.github.io/ITOExplorer/>)

Die Klassen- und Propertyhierarchie des Modells wurde über eine externe Website (<https://bioportal.bioontology.org/ontologies/ITO/>) einfach zugänglich gemacht.

Intelligence Task Ontology

Last updated: January 13, 2022

↓
🔗
🏠

Summary

Classes

Properties

Notes

Mappings

Widgets

Jump to:

- ⊕ Data
- ⊕ Data format
- ⊕ Meta: Class requiring curation
- ⊕ Meta: Deprecated class
- ⊕ Process
 - ⊖ AI process
 - ⊕ Adversarial process
 - ⊕ Art-related process
 - ⊕ Audio process
 - ⊕ Benchmarking
 - ⊕ Biomedical AI process
 - ⊕ Computer code process
 - ⊕ Fundamental AI process
 - ⊕ Graph process
 - ⊕ Knowledge base process
 - ⊕ Miscellaneous process
 - ⊖ Natural language processing
 - ⊕ Biomedical natural language processing
 - ⊕ Chinese language processing
 - ⊕ Code generation
 - ⊕ Cross-lingual natural language processing
 - ⊕ Dialog process
 - ⊕ Image tagging
 - ⊕ Multimodal deep learning
 - ⊕ Natural language analysis
 - ⊖ Natural language generation
 - ⊕ Chatbot
 - ⊕ Dialogue rewriting
 - ⊕ Goal-oriented dialog
 - ⊕ Graph-to-sequence
 - ⊕ Image captioning
 - ⊕ Language modelling
 - ⊕ Machine translation
 - ⊕ Multimodal text prediction
 - ⊕ Question answering
 - ⊕ Question generation
 - ⊕ Speech-to-text translation
 - ⊕ Task-oriented dialog systems
 - ⊕ Text generation
 - ⊕ Text simplification
 - ⊕ Text summarization
 - ⊕ **Abstractive text summarization**
 - AMI - Text Summarization benchmarking
 - arXiv - Text Summarization benchmarking
 - BBC XSum - Text Summarization benchmarking
 - BigPatent - Text Summarization benchmarking

Details

Visualization

Notes (0)

Preferred Name	Abstractive text summarization
ID	https://identifiers.org/ito:ITO_01067x
hasDefinition	<p>**Abstractive Text Summarization** is the task of generating a short and concise summary that captures the salient ideas of the source text. The generated summaries potentially contain new phrases and sentences that may not appear in the source text. (Source: paperswithcode.com)</p> <p>Generating a summary of a text by condensing and paraphrasing the text.</p>
ito:papers_with_code_id	Abstractive Text Summarization
label	Abstractive text summarization
prefixIRI	ito:ITO_01067x
prefLabel	Abstractive text summarization
subClassOf	https://identifiers.org/ito:ITO_01066x

Interaktive Darstellung des ITO Modells im BioPortal Browser
(<https://bioportal.bioontology.org/ontologies/ITO/>)

3.4 Arbeitspaket 4 - Studie

Die Verfassung von zwei wissenschaftlichen Manuskripten wurde abgeschlossen, beide Manuskripte befinden sich noch in der Endphase von internationalen peer-review Verfahren. Das erste Manuskript beschreibt das Datenmodell und dessen Modellierung.

Preprint

A curated, ontology-based, large-scale knowledge graph of artificial intelligence tasks and benchmarks

Kathrin Blagec¹, Adriano Barbosa-Silva¹, Simon Ott¹, Matthias Samwald¹

Affiliations

1. Institute of Artificial Intelligence, Medical University of Vienna, Vienna, Austria.

Corresponding author: Matthias Samwald (matthias.samwald [at] meduniwien.ac.at)

Abstract

Research in artificial intelligence (AI) is addressing a growing number of tasks through a rapidly growing number of models and methodologies. This makes it difficult to keep track of where novel AI methods are successfully – or still unsuccessfully – applied, how progress is measured, how different advances might synergize with each other, and how future research should be prioritized.

To help address these issues, we created the Intelligence Task Ontology and Knowledge Graph (ITO), a comprehensive, richly structured and manually curated resource on artificial intelligence tasks, benchmark results and performance metrics. The current version of ITO contain 685,560 edges, 1,100 classes representing AI processes and 1,995 properties representing performance metrics.

The goal of ITO is to enable precise and network-based analyses of the global landscape of AI tasks and capabilities. ITO is based on technologies that allow for easy integration and enrichment with external data, automated inference and continuous, collaborative expert curation of underlying ontological models. We make the ITO dataset and a collection of Jupyter notebooks utilising ITO openly available.

Abstract des ersten Manuskriptes (<https://arxiv.org/abs/2110.01434>)

Ein zweites Manuskript analysiert und visualisiert den globalen Fortschritt in AI über verschiedene Anwendungsdomänen hinweg. Für dieses zweite Manuskript wurden umfangreiche Analyse- und Visualisierungsskripte entwickelt. Der Aufwand für die Erstellung dieser Analysen war höher als ursprünglich angenommen.

Mapping global dynamics of benchmark creation and saturation in artificial intelligence

Adriano Barbosa-Silva¹, Simon Ott¹, Kathrin Blagec¹, Jan Brauner^{2,3} and Matthias Samwald^{1,*}

¹Institute for Artificial Intelligence, Medical University of Vienna. Währingerstraße 25a, 1090, Vienna, Austria.

²Oxford Applied and Theoretical Machine Learning (OATML) Group, Department of Computer Science, University of Oxford, Oxford, UK.

³Future of Humanity Institute, University of Oxford, Oxford, UK.

*Corresponding author.

Abstract

Benchmarks are crucial to measuring and steering progress in artificial intelligence (AI). However, recent studies raised concerns over the state of AI benchmarking, reporting issues such as benchmark overfitting, benchmark saturation and increasing centralization of benchmark dataset creation. To facilitate monitoring of the health of the AI benchmarking ecosystem, we introduce methodologies for creating condensed maps of the global dynamics of benchmark creation and saturation. We curated data for 1688 benchmarks covering the entire domains of computer vision and natural language processing, and show that a large fraction of benchmarks quickly trended towards near-saturation, that many benchmarks fail to find widespread utilization, and that benchmark performance gains for different AI tasks were prone to unforeseen bursts. We conclude that future work should focus on large-scale community collaboration and on mapping benchmark performance gains to real-world utility and impact of AI.

Abstract des zweiten Manuskriptes (<https://arxiv.org/abs/2203.04592>)

3.5 Arbeitspaket 5 - Dokumentation und Formales am Projektende

Dokumentation und formales am Projektende wurden gemäß der Vorgaben abgewickelt. Gemäß der Vorgaben wurden Dokumentations-PDFs generiert; wir möchten aber darauf hinweisen, dass die Dokumentation über die diversen Projektwebseiten leichter zugänglich ist.

4 Umsetzung Förderauflagen

Die geforderte „Beschreibung & Mitberücksichtigung des Katalogbefüllungsprozesses“ wurde bereits in der Ergänzung vor Förderabschluss übermittelt; detailliertere Beschreibungen des Prozesses wurden in den beiden Publikation bereitgestellt.

5 Liste Projektergebnisse

Projektzwischenbericht	CC-BY-3.0 AT	https://netidee.at/web-ai
Projektendbericht	CC-BY-3.0 AT	https://netidee.at/web-ai
<p>Entwickler_innen-DOKUMENTATION des Projektergebnisses für andere Entwickler_innen ("Dritte"), die das Projektergebnis nach Projektende nutzen/weiterentwickeln wollen</p> <p><u>Für Entwickler_innen (Systemkonzept, ggf. Grobspezifikationen):</u></p> <p>a. WAS IST ES b. FÜR WEN IST ES /WEM HILFT ES WODURCH c. WIE FUNKTIONIERT ES (für Entwickler_innen: Übersicht und detailliertes Systemkonzept, SW-Struktur)</p>	CC-BY-3.0 AT	https://github.com/OpenBioLink/ITO https://netidee.at/web-ai
<p>Anwender_innen-DOKUMENTATION des Projektergebnisses für Anwender_innen, die das Projektergebnis nach Projektende nutzen wollen</p> <p><u>Für Anwender_innen ("Bedienungsanleitung"):</u></p> <p>a. WAS IST ES b. FÜR WEN IST ES /WEM HILFT ES WODURCH c. WIE FUNKTIONIERT ES</p>	CC-BY-3.0 AT	https://github.com/OpenBioLink/ITO https://netidee.at/web-ai
<p>Veröffentlichungsfähiger Einseiter</p> <ul style="list-style-type: none"> * Kurzfassung WAS FÜR WEN WIE * Liste Projektergebnisse - also diese Liste, ggf. kompromiert * mit Angabe Open Source Lizenz/Webadresse * wo finden Dritte die Projektergebnisse (inkl. Dokumentation Anwender_innen bzw. Entwickler_innen) * mögliche Weiterentwicklungen/ weitere Einsatz-/ Nutzungsmöglichkeiten 	CC-BY-3.0 AT	https://netidee.at/web-ai

Dokumentation Externkommunikation zur Erreichung Sichtbarkeit /Nachhaltigkeit (separates Dokument oder als Teil des Endberichtes) * Welche Maßnahmen wurden in welchem Umfang gesetzt * Jeweils Bewertung Aufwand / Nutzen * Lessons Learned / Empfehlungen für andere Projekte	CC-BY-3.0 AT	https://netidee.at/web-ai
Intelligence Task Ontology (ITO) Eine OWL Ontologie zu Prozessen, Datentypen, Services und Anwendungsgebieten künstlicher Intelligenz. Im netidee Projekt wird die Struktur von ITO grundlegend verbessert.	CC-BY-SA-3.0 AT	https://github.com/OpenBioLink/ITO/
Software für interaktive Website zum Durchsuchen, Visualisieren und Analysieren von ITO.	MIT	https://github.com/OpenBioLink/ITO/
Website zum Durchsuchen, Visualisieren und Analysieren von ITO.	CC-BY-SA-3.0 AT	https://openbiolink.github.io/ITOExplorer/
Studie, in der die derzeitige Struktur von AI-Modellen und den dadurch ermöglichten Prozessen analysiert werden.	CC-BY-SA-3.0 A	arxiv.org

6 Verwertung der Projektergebnisse in der Praxis

Neben den in den Arbeitspaketen genannten Analysen die als Teil des Projektes durchgeführt wurden, wurde das im Projekt geschaffene ITO Modell bereits in zwei weiteren wissenschaftlichen Arbeit in unserer Arbeitsgruppe verwendet:

Kathrin Blagec, Georg Dorffner, Milad Moradi, Simon Ott, Matthias Samwald. “A Global Analysis of Metrics Used for Measuring Performance in Natural Language Processing.” In Workshop on Efficient Benchmarking in NLP at ACL 2022, 2022. <http://arxiv.org/abs/2204.11574>

Kathrin Blagec, Jakob Kraiger, Wolfgang Frühwirth, Matthias Samwald. “Benchmark Datasets Driving Artificial Intelligence Development Fail to Capture the Needs of Medical Professionals”. arXiv:2201.07040, January 2022. <http://arxiv.org/abs/2201.07040>

Das im Projekt geschaffene ITO Modell wurde vom internationalen Papers With Code Projekt (der größten Datenquelle für Entwicklungen im AI Bereich) herangezogen um die Datenqualität zu erhöhen (siehe z.B. <https://twitter.com/paperswithcode/status/1351540645907857410>).

7 Öffentlichkeitsarbeit/ Vernetzung

Eine Vernetzung mit relevanten Stakeholdern des Europäischen Joint Research Center zum Thema „Measuring Progress in AI“ ist erfolgt.

Das Projekt wurde bei einem Keynote Vortrag von Matthias Samwald bei der ISMB 2021 Konferenz vorgestellt.

Zwei wissenschaftliche Manuskripte für Open Access Journale, welche die Entwicklungen des netidee Projektes dokumentieren, sind derzeit unter Review und bereits als Preprints frei zugänglich, zwei weitere auf Projektergebnissen aufbauende Manuskripte wurden ebenfalls publiziert.

8 Eigene Projektwebsite

Github page: <https://github.com/OpenBioLink/ITO/>

9 Geplante Aktivitäten nach netidee-Projektende

Die ITO Ressource wird weiter entwickelt werden; wir planen insbesondere detaillierte Erweiterungen für bestimmte, wichtige Domänen der AI wie z.B. Natural Language Processing.

10 Anregungen für Weiterentwicklungen durch Dritte

Die im Projekt geschaffenen Ressourcen können sowohl als Datenmodell für die Annotierung eigener Daten verwendet werden, sowie als reichhaltige, Netzwerk-zentrierte Wissensbasis zur Beantwortung eigener Problemstellungen. ITO kann von Dritten auch als Basis für die Entwicklung eigener, domänenspezifischer Wissensmodelle verwendet werden, was durch die verwendeten Datenstandards (z.B. OWL) einfach möglich ist.