

Proteinhaltige Lebensmittel im Vergleich – eine Datenbank hilft weiter

Dr. Claudia Müller (ZHAW)

15. September 2023

Zürcher Hochschule
für Angewandte Wissenschaften

zhaw Life Sciences und
Facility Management



food data

Swiss
Food Data Mediator

Übersicht

- Ausgangslage
- Datenmodell und Datenbasis (PROTOTYP)
- Technische Aspekte
- Anwendungsbeispiele
- Ausblick

Ausgangslage allgemein

- Der Proteinbedarf weltweit steigt.
 - Der globale Fleischkonsum nimmt zu.
 - Negative Auswirkungen sind:
 - Hohe Treibhausgasemissionen
 - Hoher Landverbrauch
 - Mögliche gesundheitliche Folgen (v.a. durch Überkonsum von verarbeiteten Fleischprodukten)
 - Ethische Aspekte (Tierwohl)
- Mit einem Fleischkonsum, wie er momentan v.a. in Nordamerika und Europa typisch ist, können die vereinbarten Nachhaltigkeitsziele (SDGs) nicht erreicht werden.
- Wie kann die Weltbevölkerung nachhaltig mit ausreichend Proteinen versorgt werden?

Quellen:

Forum, W.E. (2019); Meat: the Future series Alternative Proteins, White Paper

Henchion M. et. al. (2017); Future Protein Supply and Demand: Strategies and Factors Influencing a Sustainable Equilibrium; Food, 6,53; doi:10.3390/foods6070053

Tilman D. et. al. (2011); Global food demand and the sustainable intensification of agriculture; Proc. Natl. Acad. Sci. U. S. A. vol. 108, pp. 20260–20264; doi: 10.1073/pnas.1116437108

Ausgangslage Lehre



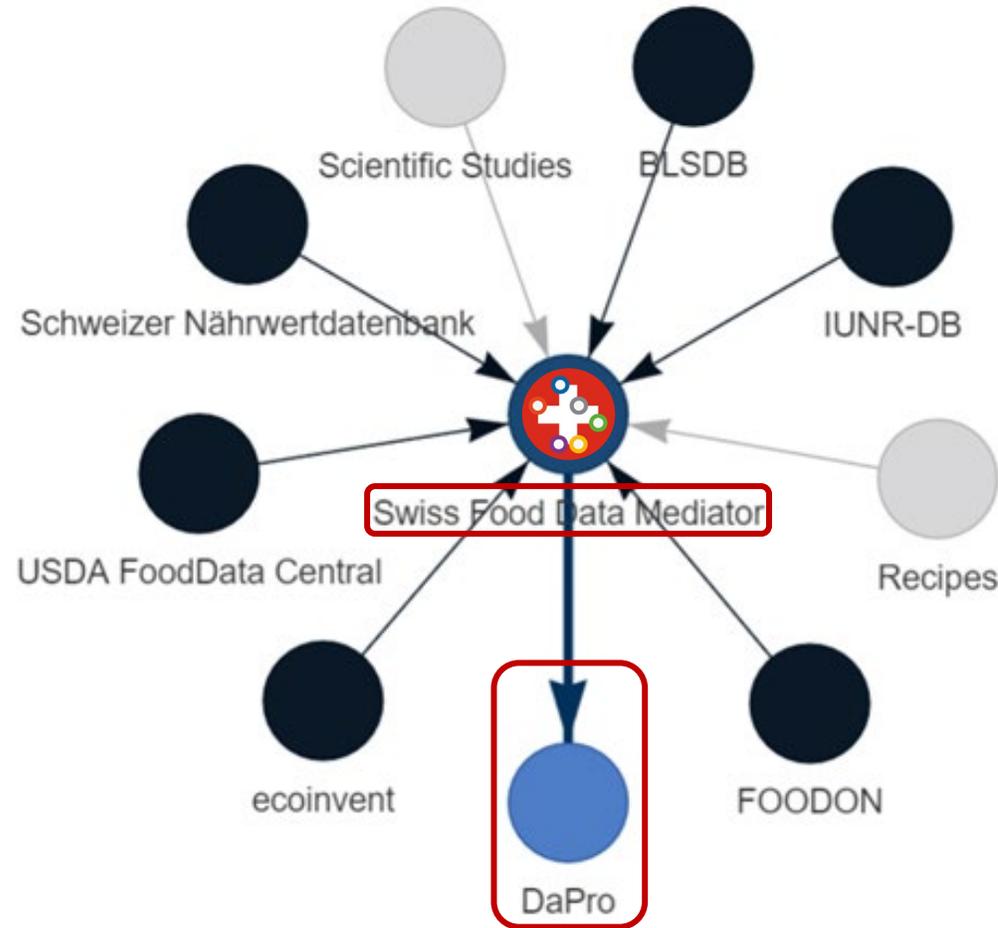
- Mastermodul “Sustainable Food Supply Chains”
- Studierende haben die Aufgabe die Wertschöpfungskette eines proteinreichen Lebensmittels bezüglich ökologischer, sozialer, gesundheitlicher und ökonomischer Werte quantitativ (KPIs*) zu bewerten und im Anschluss zu optimieren.
- Grosser zeitlicher Aufwand für Recherche
- Datenverfügbarkeit nicht immer gewährleistet
- Zuverlässigkeit der Datenquellen bzw. Datenqualität teilweise problematisch

→ Bedarf nach «zentralem Nachschlagewerk»

→ Idee «Datenbank Proteinreiche Lebensmittel» (DaPro)

*Key Performance Indicators

Datenmodell



Grafik: MA Christina Köck, FH Kufstein

- Existierende, validierte Informationen aus bestehenden Datenbanken
- Rezepte und Daten aus Einzelstudien, die ergänzt werden können
- Datenbank proteinreiche Lebensmittel

Swiss Food Data Mediator (CH-FDM)



- «Vermittelt» zwischen den verschiedenen Datenbanken.
- Gibt ein Schema vor, welches sich für alle involvierten Datentypen eignet.
- Bindet bekannte, etablierte und vor allem grosse Datenbanken ein.
- Bildet die Basis für beliebige Applikationen.



“Datenbank Proteinreiche Lebensmittel” (DaPro)

Datenbasis - aktueller Stand PROTOTYP

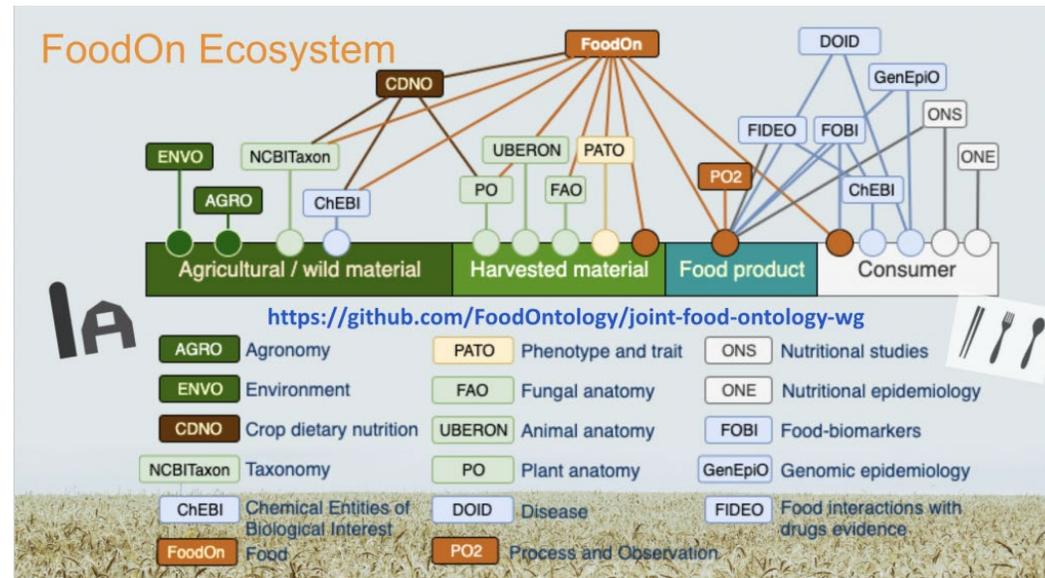


- FOODON
- SCHWEIZER NÄHRWERT-DATENBANK
- BUNDESLEBENSMITTELSCHLÜSSEL (BLS 3.02)
- NATIONAL NUTRIENT DATABASE (USDA)
- ECOINVENT
- CHEBI
- MENÜ-NACHHALTIGKEITS-INDEX (MNI)
- DATENBANK PROTEINREICHE LEBENSMITTEL (DaPro)

- 34'000 Klassen wie z.B. «cow milk (liquid)» (FOODON:03302116)
- Verlinkungen zu über 700 anderen Ontologien z.B. «Infectious Disease Ontology»



FOODON



Quelle: <https://foodon.org/>

SCHWEIZER NÄHRWERT DATENBANK

- Nährwerte von > 1'100 Lebensmitteln wie z.B. "Vollmilch, pasteurisiert"
- Bundesamt für Lebensmittelsicherheit und Veterinärwesen (BLV)

Vollmilch, pasteurisiert

ID: 62

Kategorie(n): Milch

Synonym(e):



**Schweizer
Nährwertdatenbank**



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Bundesamt für Lebensmittelsicherheit und Veterinärwesen BLV
Office fédéral de la sécurité alimentaire et des affaires vétérinaires OSAV
Ufficio federale della sicurezza alimentare e di veterinaria USAV

Quelle: <https://naehrwertdaten.ch/>

BUNDES- LEBENSMITTEL- SCHLÜSSEL (lizenzpflichtig)

- Nährwertdatenbank der Bundesrepublik Deutschland
- ~15'000 Lebensmittel z.B. "Kuhmilch" und Menükomponenten
- Pflege und Weiterentwicklung des BLS: Max Rubner-Institut



d	BlsCode	Name	Ste	Sz	Sl	Se	Gcal	Gj	Zw	z
144	M114300	Kuhmilch 3,5% Fett laktosefrei					65	272	87'399	
135	M110000	Kuhmilch					65	272	87'399	
134	M110232	Kuhmilch 1,5% Fett gekocht					49	205	89'072	
133	M110300	Kuhmilch 3,5% Fett					65	272	87'399	
132	M113200	Kuhmilch 1,5% Fett ultrahocherhitzt					48	201	89'267	
131	M113300	Kuhmilch 3,5% Fett ultrahocherhitzt					66	276	87'392	
IULL)	M114200	Kuhmilch 1,5% Fett laktosefrei					39	162	91'469	
IULL)	M114100	Kuhmilch < 1% Fett laktosefrei					27	111	92'650	
IULL)	M113100	Kuhmilch < 1% Fett ultrahocherhitzt					35	146	90'670	
IULL)	M112300	Kuhmilch Vorzugsmilch 3,5% Fett					67	279	87'240	
IULL)	M111300	Kuhmilch Trinkmilch 3,5% Fett					65	272	87'399	
IULL)	M111332	Kuhmilch Trinkmilch 3,5% Fett gekocht					68	284	87'014	
IULL)	M111200	Kuhmilch Trinkmilch 1,5% Fett					48	201	89'267	

Quelle: <https://www.blsdb.de/>

FOOD DATA CENTRAL (USDA)

- Nährwerte von ~8'000 Lebensmittel z.B. "Milk, whole"
- Amerikanische Landwirtschaftsbehörde USDA

Milk, whole

Data Type: Survey (FNDDS) **FDC ID:** 2340762 **Food Code:** 11111000
FDC Published: 10/28/2022

Nutrients **Portions** **Other Information** **Ingredients** **Ingredient**

Portion:

Name	Amount	Unit
Water	88.1	g
Energy	61	kcal
Protein	3.27	g
Total lipid (fat)	3.2	g
Carbohydrate, by difference	4.63	g
Fiber, total dietary	0	g
Sugars, total including NLEA	4.81	g
Calcium, Ca	123	mg
Iron, Fe	0	mg



U.S. DEPARTMENT OF AGRICULTURE
Agricultural Research Service

Quelle: <https://fdc.nal.usda.gov/>

ECOINVENT

(lizenzpflichtig)

Ergänzt durch Agrifood-
 Database der ZHAW
 (FG Ökobilanzierung)

- > 18'000 Life Cycle Datasets z.B. "cow milk (raw)"

Reference Product		
By-product classification: allocatable product\CPC: 02211: Raw milk of cattle		
cow milk		
+ allocation factor	0.85225	dimensionless
+ carbon allocation	0.066224	kg
+ <u>carbon content</u>	<u>0.53797</u>	<u>dimensionless</u>
+ carbon content, fossil	0	dimensionless
+ carbon content, non-fossil	– Ecological Scarcity 2021	
+ dry mass	+ climate change	
+ energy content	– <u>emissions to air</u>	
+ price	air pollutants and PM	
+ true value	carcinogenic substances into air	
+ water content	heavy metals into air	
+ water in wet mass	radioactive substances into air	
+ wet mass	– total	
UBP		



Quelle: <https://ecoinvent.org/>

CHEBI

- 60'000 Einträge von chemischen Entitäten
z.B. "Calcium" (CHEBI:22984)
- Elixir – European Life Sciences
Infrastructure



Search results for 'Ca (v0)':

ChEBI Name	calcium atom
ChEBI ID	CHEBI:22984
Stars	☆☆☆ This entity has been manually annotated by the ChEBI Team.
Supplier Information	No supplier information found for this compound.
Download	Molfile XML SDF

- Find compounds which contain this structure
- Find compounds which resemble this structure
- Take structure to the Advanced Search

Interconnected databases and consortia:

- ArrayExpress
- BioModels Database
- ChEBI
- ChEMBL
- Ensembl
- Ensembl Genomes
- International Nucleotide Sequence Databases
- Expression Atlas
- IMEx Consortium
- InterPro
- Protein Data Bank in Europe
- ProteomeXchange (PRIDE)
- RNAcentral
- PubMed Central International
- Reactome
- UniProt: The Unified Protein Resource

Quelle: <https://www.ebi.ac.uk/chebi/>

Anbindung weiterer Apps: z.B. MNI-TOOL

- ~1'400 Menüs, ~ 1'800 Komponenten z.B. "Milchreis" und über 1'000 Zutaten
- Fokus Nachhaltigkeit, Umweltaspekte (z.B. UBP; Grundlage ecoinvent und Agrifood-Database der ZHAW) aber auch Ernährungsphysiologie
- ZHAW-Entwicklung

Id	Name	Name	Anteil
884	Milchreis	Vanilleschote	
884	Milchreis	Salz	
884	Milchreis	Zucker	
884	Milchreis	Reis roh	
884	Milchreis	Kuhmilch 3,5% Fett	

Quelle: <https://www.zhaw.ch/de/lfsf/forschung/formen-der-zusammenarbeit/menue-nachhaltigkeits-index/>



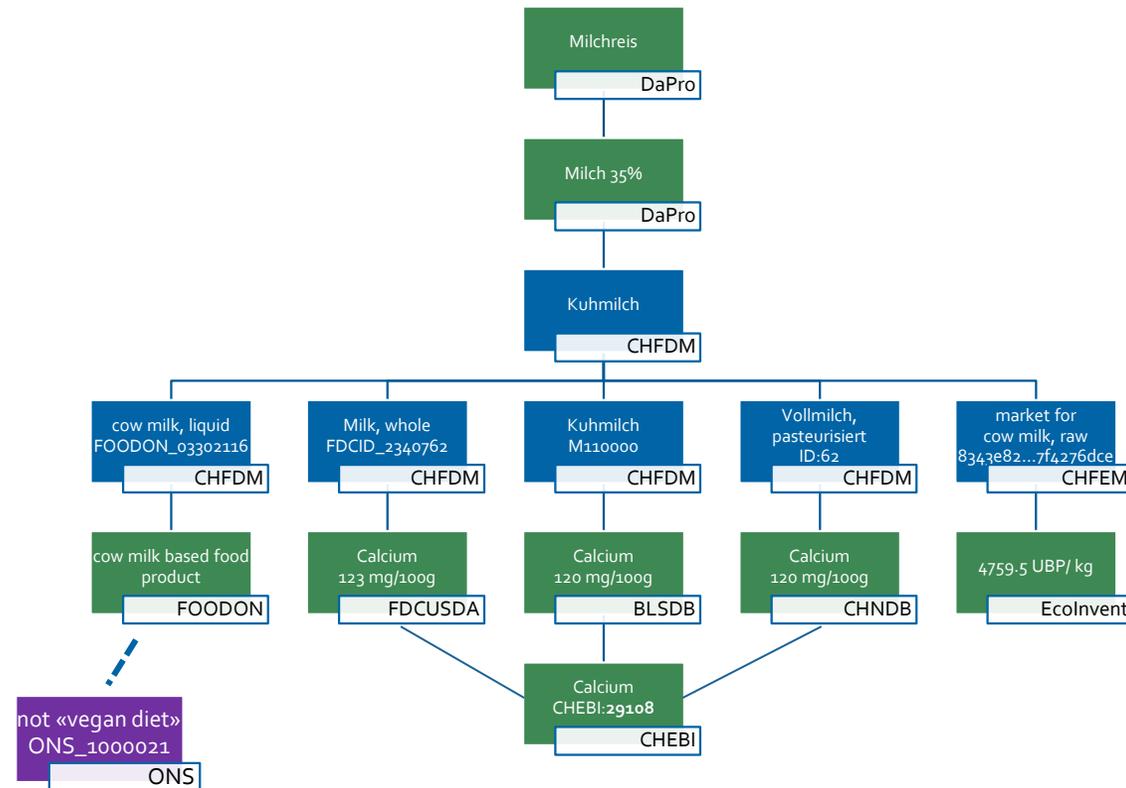

IN EINFACHEN 4 SCHRITTEN
**Mit dem MNI
mehr Nachhaltigkeit
auf dem Teller**

Mit dem Menü-Nachhaltigkeits-Index (MNI) kann ein Tellergericht in den Dimensionen Umweltfreundlichkeit und Ausgewogenheit bewertet werden. Diese Bewertung wird in ein einfaches Punktesystem übersetzt. So ist es möglich, den MNI auf verschiedenste Art und Weise zu nutzen und zu kommunizieren.



CH-FDM Swiss Food Data Mediator

- Momentan manuell kuratiert (proof of concept), Möglichkeit der automatisierten Datengenerierung und -überprüfung
- **Fachwissen** der ZHAW LSFM
- Verknüpfung der **vorhandenen** Datenquellen



DaPro - Single choice

>> Endpoint fuseki

Select Food Item

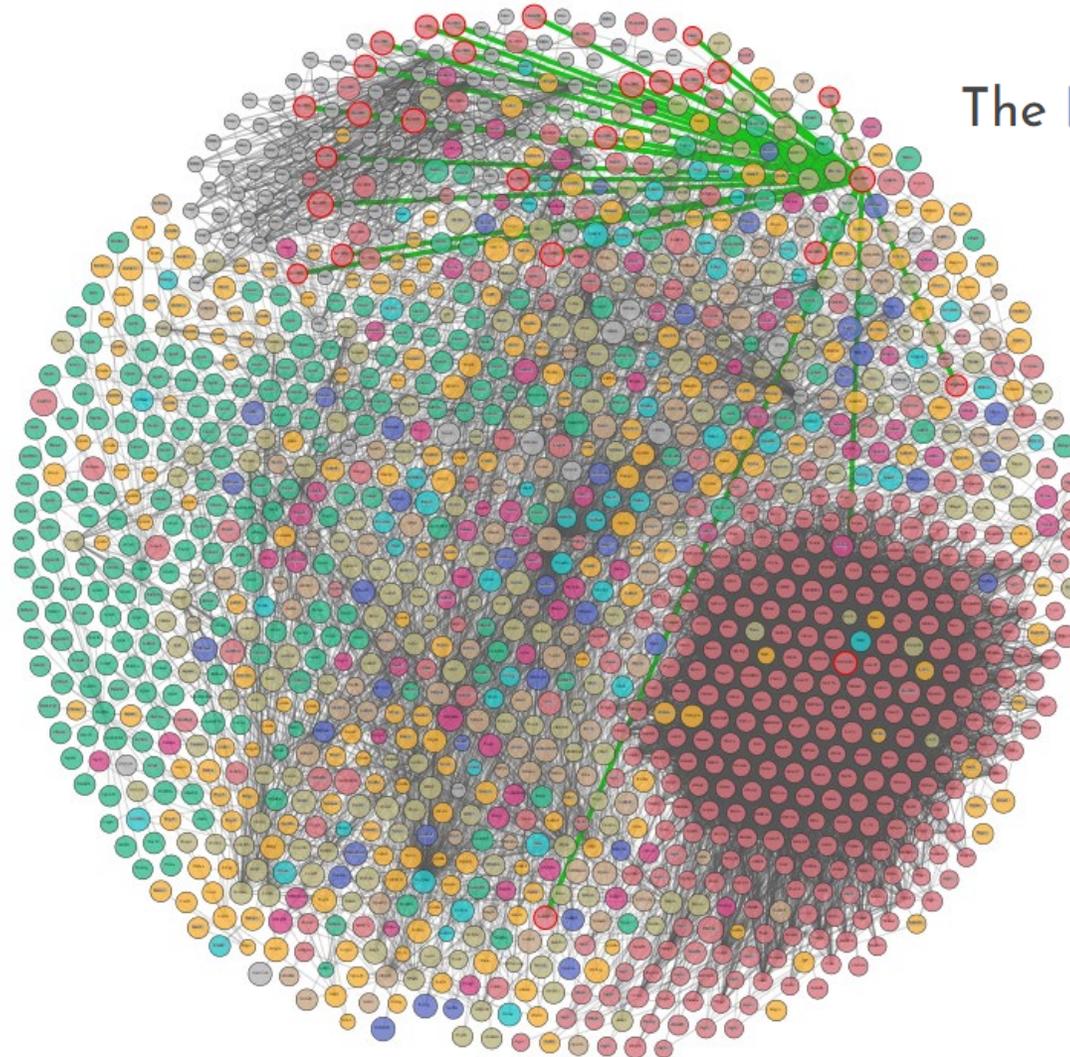
- Hartkäse
- Hartweizengriess
- Haselnuss
- Hirse
- Hühnerei
- Kabeljau
- Kalbfleisch
- Kaninchen
- Karpfen
- Kichererbse
- Kidneybohne
- Kuhbohne/Augenbohne
- Kuhmilch**
- Kürbiskern
- Lachs Wildfang
- Lachs Zucht
- Lammfleisch
- Leinsamen
- Limabohne
- Linsen

DaPro

- Datenbank Proteinreiche Lebensmittel
- ~100 Lebensmittel
- Fokus auf proteinreiche LM
- Vergleich von Produkten
- **Prototyp** für CH-FDM

Technische Grundlagen

The **Linked Open Data** Cloud



Legend

Cross Domain
Geography
Government
Life Sciences
Linguistics
Media
Publications
Social Networking
User Generated

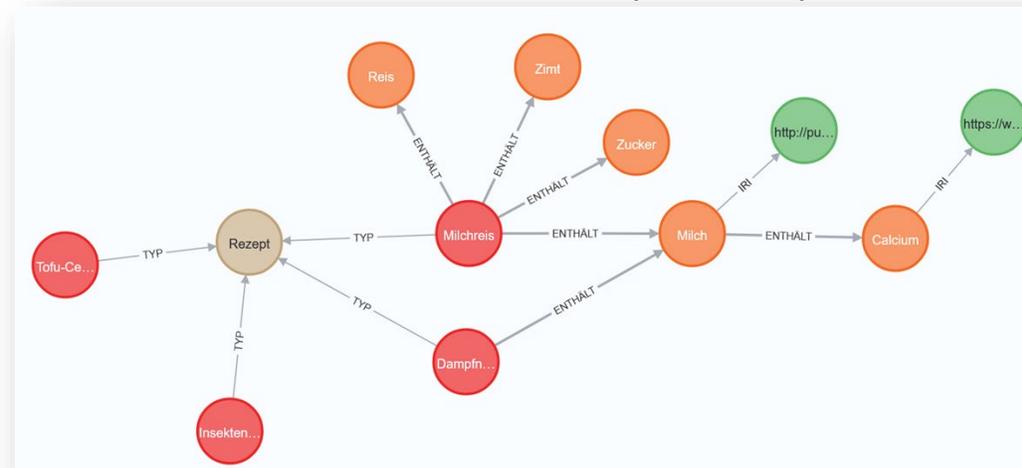
Quelle: <https://lod-cloud.net/clouds/lod-cloud.svg>

Technische Grundlagen

Graph-Datenbank

- Graphdatenbank: eine Art, Daten in **Tripeln** abzulegen z.B.

<i>Subjekt</i>	<i>Prädikat</i>	<i>Objekt</i>
MNITOOL:Milchreis	rdf:Typ	CHFDM:Rezept
MNITOOL:Milchreis	schema:Zutat	CHFDM:Milch
CHFDM:Milch	rdfs:definiertDurch	FOODON:03307455
CHFDM:Milch	rdfs:definiertDurch	BLSDB:M110000
BLSDB:M110000	rdfs:enthält	CHEBI:22984
BLSDB:M110000	rdfs:enthältMenge	CHEBI:130mg
CHEBI:22984	rdfs:Label	«Calcium»
- Die Datenbank lässt sich als verknüpften Graph darstellen

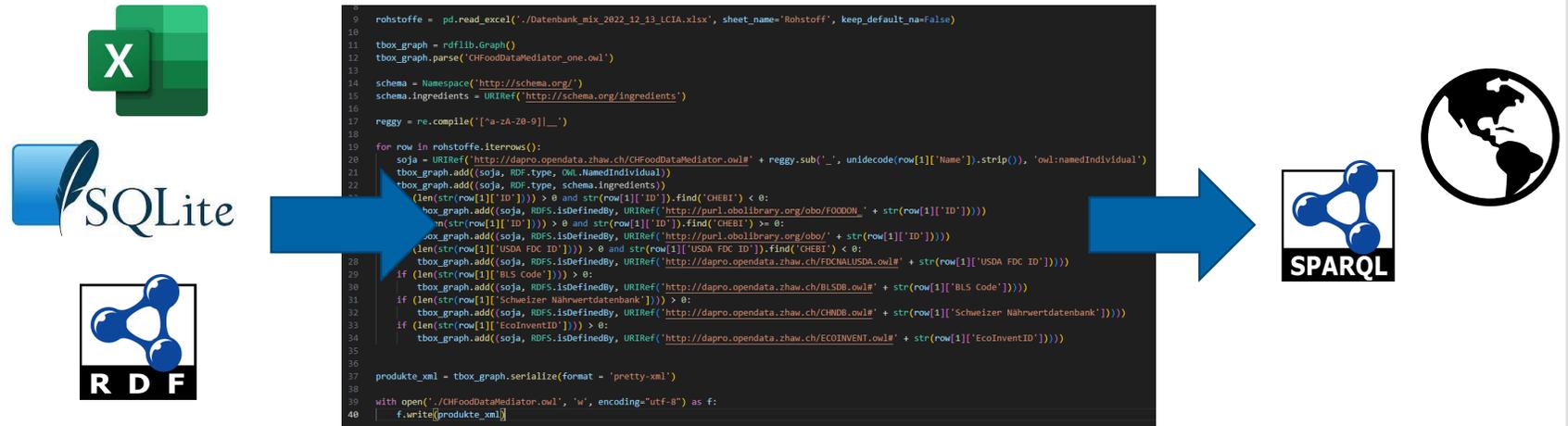


- Daten von externen Datenbanken werden über die Identifikationsnummer (Internationalized Resource Identifier IRI) referenziert und sind so über Updates hinweg konstant.

Grafik: semanticscholar.org

Datenbearbeitung

- Die momentan verwendeten Daten stammen aus verschiedenen **Quellformaten**: Excel, SQLite, OWL, API
- Für viele **Programmiersprachen** gibt es Bibliotheken, um Daten aus diesen Formaten zu lesen, sie zu bearbeiten (z.B. neu zu berechnen oder zu aggregieren) und sie dann an unser Modell anzuhängen.
- Wir verwenden momentan **Python**, das von vielen Forschenden in den Datendisziplinen eingesetzt wird. Hier ein kurzes Skript, welches aus den vorhandenen Excel-Daten unsere Grundlage für den CH-FDM erstellt.



Speicherform

- Die Daten werden in «Mensch-lesbaren» Text-Dateien gespeichert. Sie ähneln von der Form her **XML**-Dateien.
- Hier ein Ausschnitt aus der BLSDB-Datei (welche 200'000 Zeilen aufweist). Wir sehen den Eintrag für Calcium in Kuhmilch aus dem BLSDB.



```

178873 .....</owl:NamedIndividual><CRLF
178874 .....</schema:ingredients><CRLF
178875 .....<schema:ingredients><CRLF
178876 .....<owl:NamedIndividual rdf:about="http://dapro.opendata.zhaw.ch/BLSDB.owl#M110000_110_MCA"><CRLF
178877 .....<rdf:type rdf:resource="http://dapro.opendata.zhaw.ch/DaPro.owl#ContainsIngredientPercentage"/><CRLF
178878 .....<rdfs:label>Calcium in Kuhmilch</rdfs:label><CRLF
178879 .....<schema:value rdf:datatype="http://www.w3.org/2001/XMLSchema#float">120.0</schema:value><CRLF
178880 .....<schema:ingredients rdf:resource="http://purl.obolibrary.org/obo/CHEBI_22984"/><CRLF
178881 .....<schema:unitCode>mg/100g</schema:unitCode><CRLF
178882 .....</owl:NamedIndividual><CRLF
178883 .....</schema:ingredients><CRLF
178884 .....<schema:ingredients><CRLF
178885 .....<owl:NamedIndividual rdf:about="http://dapro.opendata.zhaw.ch/BLSDB.owl#M110000_110_MMG"><CRLF
178886 .....<rdf:type rdf:resource="http://dapro.opendata.zhaw.ch/DaPro.owl#ContainsIngredientPercentage"/><CRLF
178887 .....<rdfs:label>Magnesium in Kuhmilch</rdfs:label><CRLF
178888 .....<schema:value rdf:datatype="http://www.w3.org/2001/XMLSchema#float">12.0</schema:value><CRLF

```

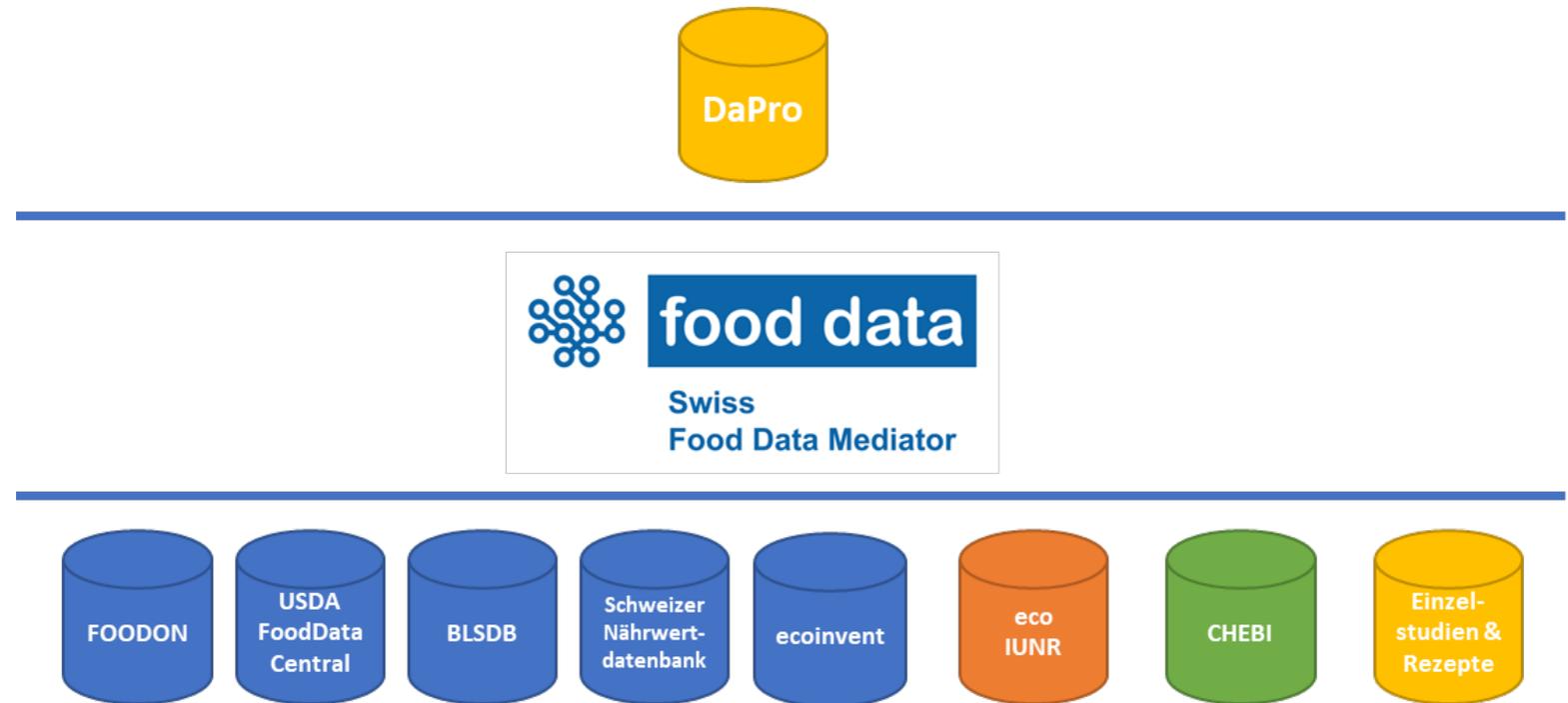
App- Entwicklung

- Die Weiterverwendung der Daten ist durch die verfügbaren Bibliotheken in verschiedensten Programmiersprachen gegeben.
- Unsere Demo läuft im Browser und verwendet JavaScript, um mit dem Datenbankserver zu kommunizieren, die Daten weiterzuverarbeiten und darzustellen.
- Andere Apps können auch mit gängigen proprietären Programmiersprachen auf mobilen Geräten entwickelt werden.



Stand des Prototyps CH-FDM + DaPro

- Als Graphdatenbank übersetzte Quellen
- Einbindung mit lokalem Datenbankserver (Apache Jena Fuseki)
- Lauffähiges Interface in HTML/JavaScript («Einzel» und «Vergleich»)



- Externe Daten
- Externe Daten – live eingebunden
- Interne Daten
- Projektdaten DaPro

Derzeit mögliche An- bzw. Auswertungen



Daten-Interface - Single Choice

Vergleichs-App - Choicer

Datensätze anlegen - Filler

DaPro - Single choice

>> Endpoint fuseki

Haferdrink

Haferdrink

Recipe

	[%]
water food product	87.00
whole oats (raw)	12.00

Haferdrink

Containings

Total alanine	69.96
Total arginine	82.32
Total aspartic acid	107.52
Total plant fiber (compressed)	1160.40
Total calcium atom	15.69
Total (1->4)-beta-D-glucan	139.20
Total chloride	16.89
Total cysteine	34.44
Total Disaccharide (2M)	129.12
Total Dodecansäure/Laurinsäure	0.84

Eicosansäure/Arachinsäure in Hafer

roh from whole oats	FAO: mg/g protein*	[mg/100g prod]	[mg/g prot]	covered
(1aut: http://dapro.opendata/BLSDB.owl#C130000 via http://BLSDB.owl#C130000_022_F200)				
L-isoleucine	30.00	49.92	38.97	1.30
Total Eicosansäure/leucine	59.00	185.52	144.82	2.45
Total lysine	45.00	52.80	41.22	0.92
Total (11Z)-icos-11-L-threonine	23.00	45.24	35.32	1.54
Total monounsaturatryptophan	6.00	40.56	31.66	5.28
Total iron atom valine	39.00	137.04	106.98	2.74
Total protein polyp histidine	15.00	48.00	37.47	2.50
Total Energie (Kilo: SAA (Met + Cys)	22.00	75.00	58.55	2.66
Total Energie (Kilo: *AAA (Phe/Trp/His/Tyr)	38.00	243.60	190.16	5.00
Total Essentielle				

* WHO HUMAN NUTRITION Report of a Joint WHO/FAO/UNU Expert Consultation WHO Technical Report Series 935

Interface / Steckbrief

Interface bietet Informationen zu allen vorhandenen Daten und wie diese berechnet werden.

Es dient somit auch der Kontrolle der Daten und dem Aufspüren von möglichen Fehlern.

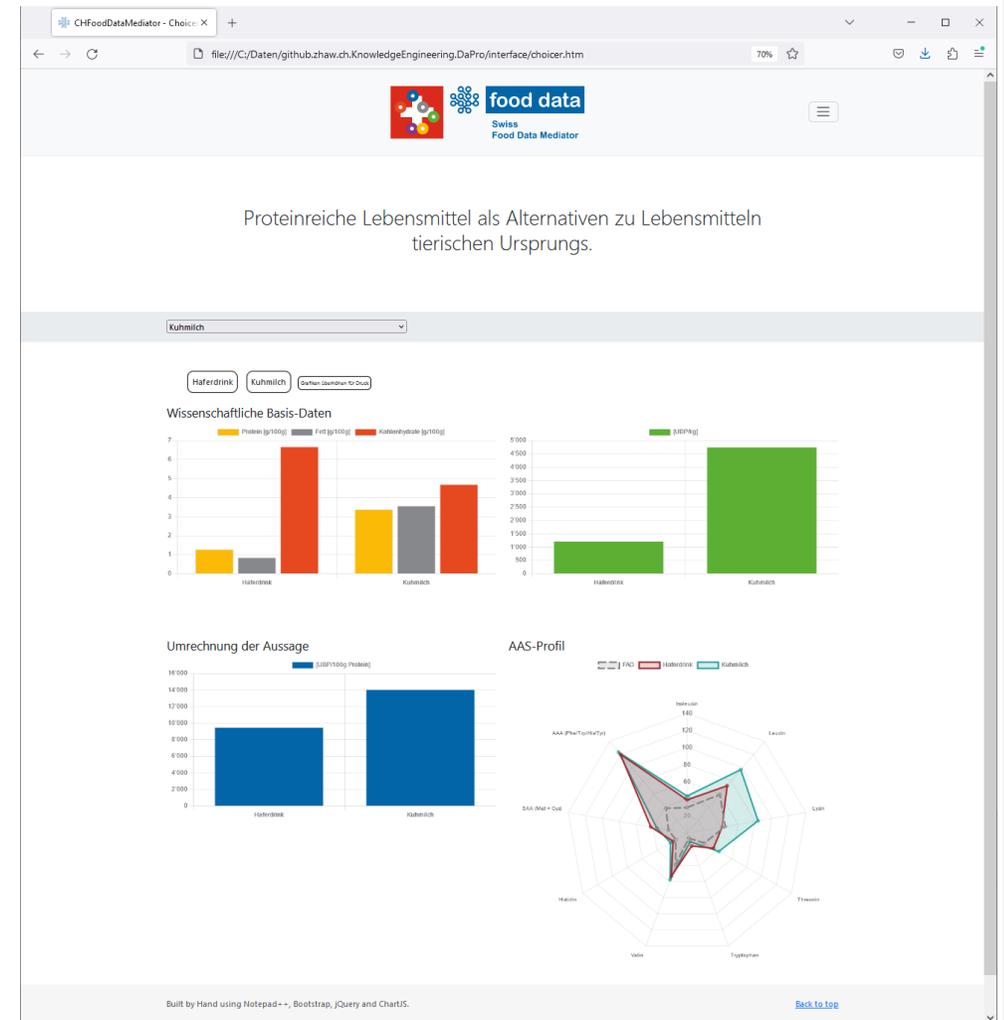
Daten-Interface - Single Choice

Vergleichs-App - Choicer

Datensätze anlegen - Filler

Choicer / Vergleich

- Mit dem «Choicer» können beliebig viele Lebensmittel miteinander verglichen werden.
- Die Vergleichsmöglichkeiten stammen aus der DaPro-Perspektive.



Daten-Interface - Single Choice

Vergleichs-App - Choicer

Datensätze anlegen - Filler

Filler / Daten erfassen

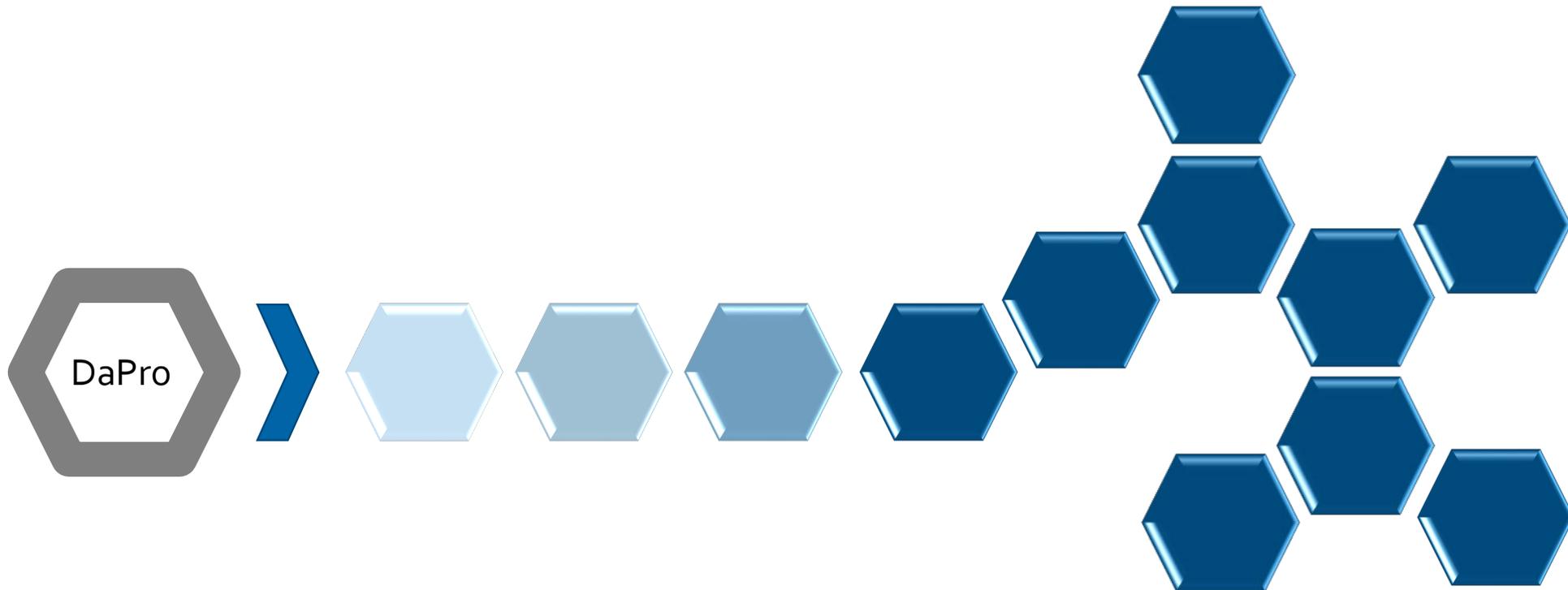
- Mit dem «Filler» können Rezepte manuell erfasst, angepasst und als OWL geteilt werden.
- Das OWL-Format lässt sich dann in die CH-FDM-Umgebung hochladen.

The screenshot shows a web browser window titled 'Filler' with the URL 'file:///C:/Daten/github.zhaw.ch.KnowledgeEngineering.DaPro/interface/filler.htm'. The page header includes the 'food data' logo and 'Swiss Food Data Mediator'. The main heading is 'Filler - Neue Rezepte anlegen und OWL herunterladen'. Below this, there is a section for uploading a recipe template, with a text input field containing '@' and a file selection button 'rezept_Hummus_DEMO.owl', and a 'Hochladen' button. The 'Neues Rezept' section contains a form with the following fields:

Zutat	Anteil	remove
Kichererbse	0.5882	remove
Tahini	0.2941	remove
Olivenöl	0.0735	remove
Knoblauch	0.0441	remove
Total:	3.0188	

At the bottom of the form, there are three buttons: 'Neue Zutat' (yellow), 'Umrechnung Rezept - Prozente' (orange), and 'Download' (blue).

Konkrete Anwendungsbeispiele



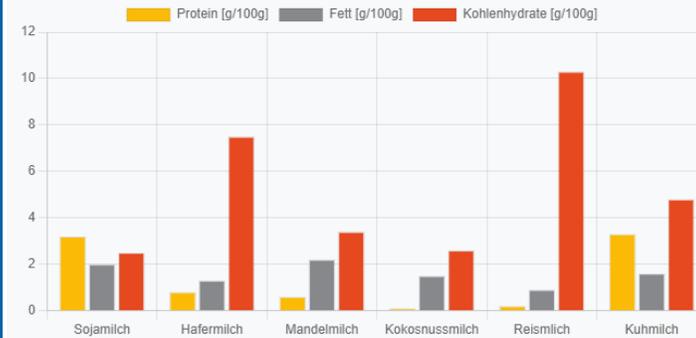
Demo-Version

Milch und Milchalternativen

Wie steht Milch und ihre Alternativen da in Bezug auf Nährwerte und Umwelteinflüsse?

Erforschen

Milch und ihre Alternativen Wissenschaftliche Basis-Daten

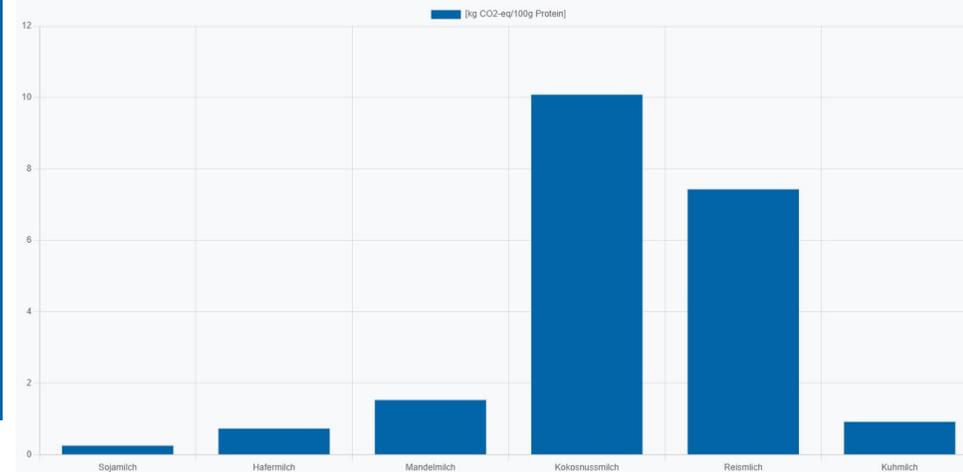


Quelle: Public Health Nutrition: 25(5), 1416–1426 (doi:10.1017/S1368980022000453) - aggregiert durch Swiss food Data Mediator



Quelle: ecoinvent, Agrofood-DB (IUNR, ZHAW) - aggregiert durch Swiss food Data Mediator

Berechnung



Anwendungs- Beispiel 1

Vergleich: Milch und Milchalternativen

Fleisch und Fleischalternativen

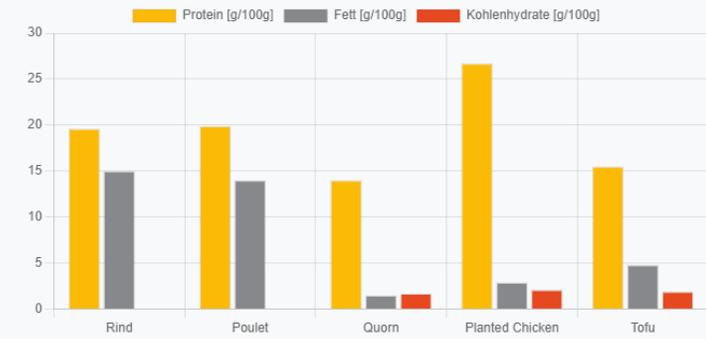
Wie steht Fleisch gegenüber seinen Alternativen da in Bezug auf Nährwerte und Umwelteinflüsse?

Erforschen

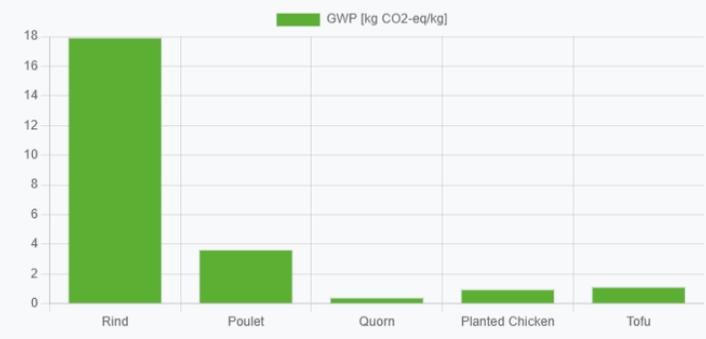
Demo-Version

Fleisch und seine Alternativen

Wissenschaftliche Basis-Daten

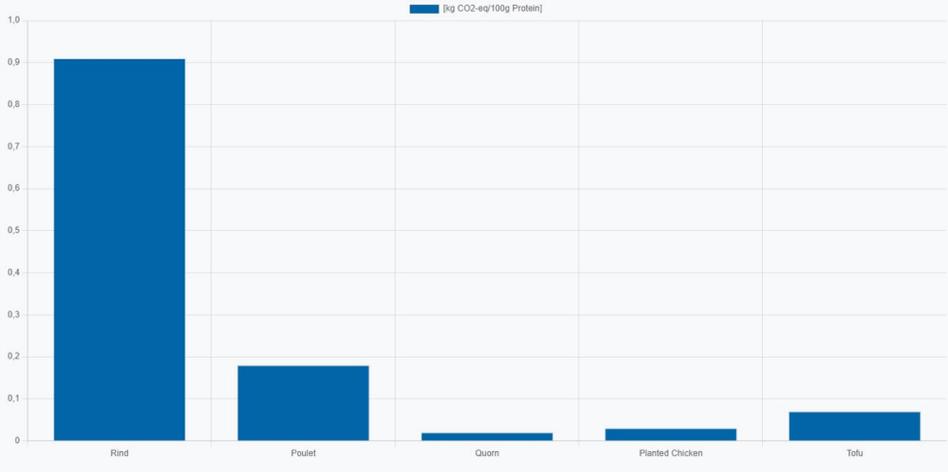


Quelle: Packungsangaben der Hersteller, Bundeslebensmittelschlüssel - aggregiert durch Swiss food Data Mediator



Quelle: ecoinvent, Agrofood-DB (IUNR ZHAW) - aggregiert durch Swiss food Data Mediator

Berechnung



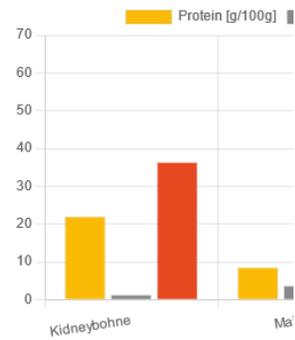
Quelle: ecoinvent, Agrofood-DB (IUNR ZHAW) - aggregiert durch Swiss food Data Mediator

Anwendungs- Beispiel 2

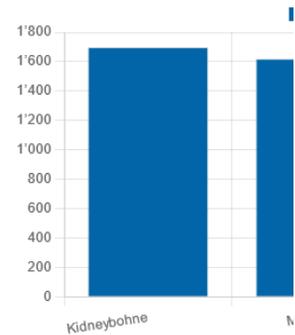
Vergleich:
Fleisch und
Fleischalternativen

Kidneybohne Mais

Wissenschaftliche Basis:

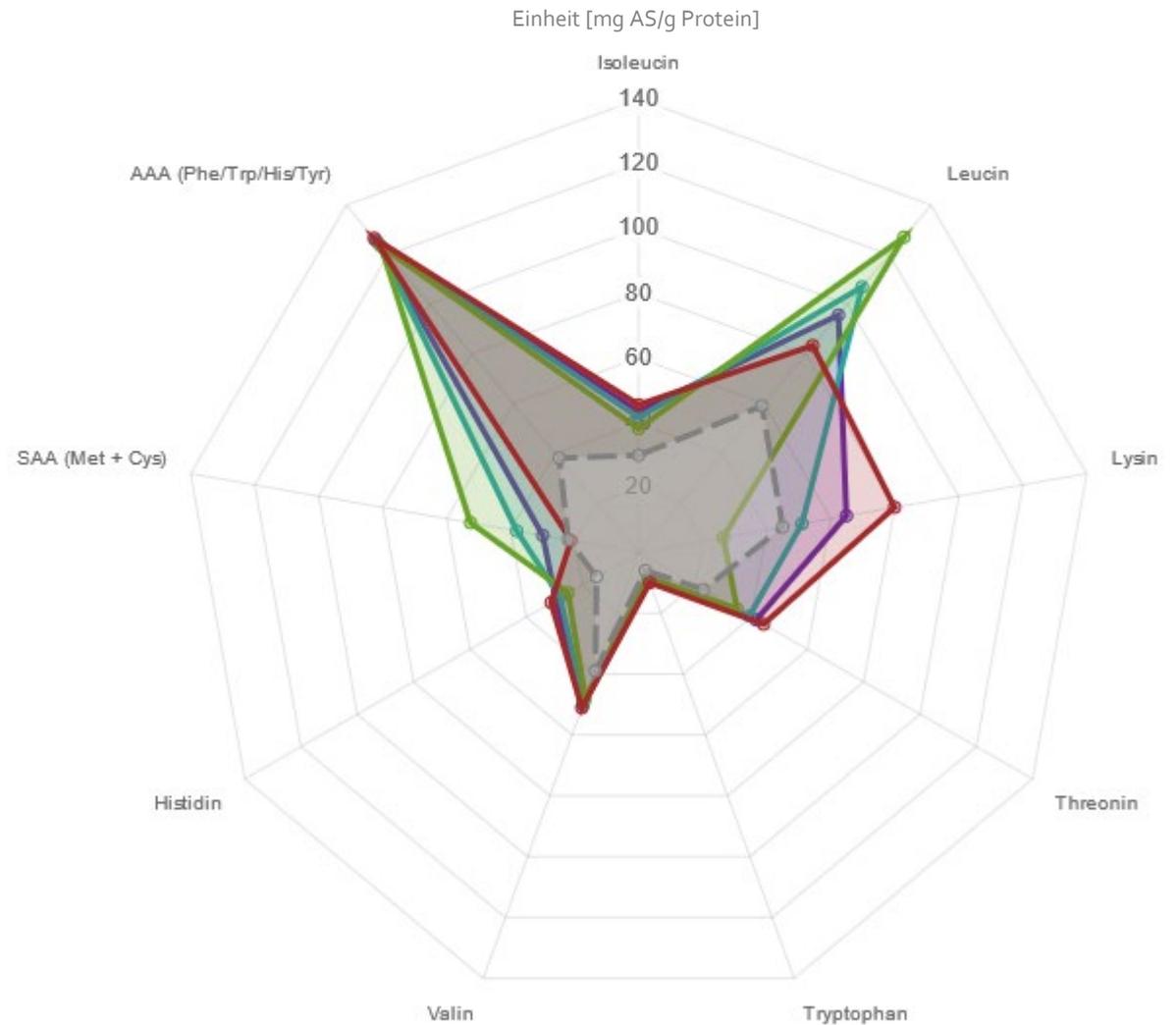


Umrechnung der Aussa:



AAS-Profil

FAO Kidneybohne Mais Protein-Mix Bohne/Mais 25/75 Protein-Mix Bohne/Mais 50/50



Anwendungs- Beispiel 3

*Kombination von
Lebensmitteln
(Proteinqualität):
Mais und Bohnen*

Ausblick

- **Datengewinnung**
 - Einzeleingabe von Daten (Studienergebnisse, Rezepte etc.)
 - Automatische Verlinkung von bestehenden Datenbanken durch automatisierte Sprachmodelle (evtl. mittels KI)
- Externe **Datenanbieter** ins Boot holen
 - z.B. Schweizer Nährwertdatenbank auf LINDAS portieren
 - Ziel: Daten direkt an der Quelle verfügbar machen, Redundanzen vermeiden
- **Infrastruktur** aufbauen
 - Eigene Daten auf SPARQL-Endpoint anbieten
- Tools zur **Datenverwendung**
 - Bestehende Projekte portieren ([MNI](#), [SFCM*](#), ...)
 - Forschung
 - Lehre
 - «Fancy Apps» für breite Masse (lizenrechtliche Fragen sind hier vorab zu klären!)

* *Sustainable Food Chain Model*

Projektteam:



Erich Zbinden, wissenschaftlicher Mitarbeiter,
FG Data Management and Visualisation, ZHAW
erich.zbinden@zhaw.ch



Dr. Claudio Beretta, wissenschaftlicher Mitarbeiter,
FG Lebensmitteltechnologie, ZHAW



Beatrice Baumer, Dozentin,
FG Lebensmitteltechnologie, ZHAW

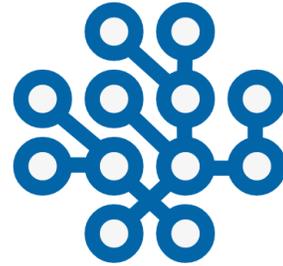
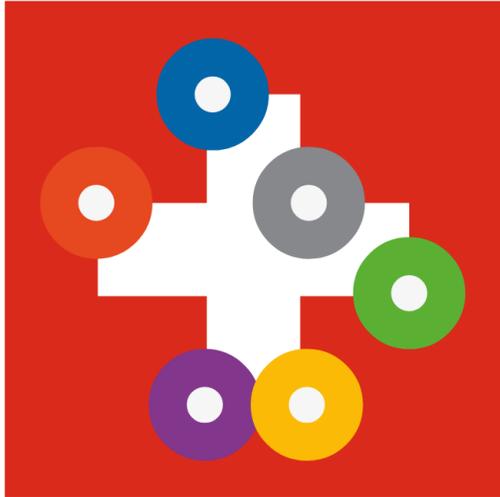
Unterstützt durch:



Matthias Stucki und **Silvan Wanner**
FG Ökobilanzierung, ZHAW

Vielen Dank





food data

**Swiss
Food Data Mediator**

**FRAGEN oder
ANREGUNGEN?**

Kontaktdaten

Dr. Claudia Müller

claudia.mueller@zhaw.ch

Tel.: 058 934 5453