



Promoting harmonization of life cycle inventory and food composition databases through semi-automatic standardization

Authors: Cédric Furrer¹, Daniel Sieh², Anne-Marie Jank², Grégoire le Bras², Moritz Herrmann¹, Alba Reguant-Closa¹, Thomas Nemecek¹

¹ Research group LCA, Agroscope Reckenholz, CH-8046 Zurich, Switzerland

² themakers GmbH, Chausseestrasse 8A, DE-10015 Berlin, Germany

LCAfood Conference 2024, Barcelona, 10.09.2024





Introduction

- **Driving food systems more sustainable** is widely discussed due to the various dramatic implication of the intensification of food production on environment and health (Alemu, 2022; Sirdey *et al.* 2023; Zhu *et al.* 2023)
- Efforts are increasing to direct food production towards products with **low environmental impacts and high nutritional value** (Mazac *et al.* 2023)
- For that purpose, nutritional and environmental data of food products is needed **for multi-dimensional optimization**
- **Lack** of publicly available combined databases





Research questions

The paper aims to answer the following questions:

1

Do **LCI** and **FCDB databases** (Agribalyse and EuroFIR, respectively) provide **enough (meta) data to successfully describe and interlink** food entries between environmental and nutritional databases?

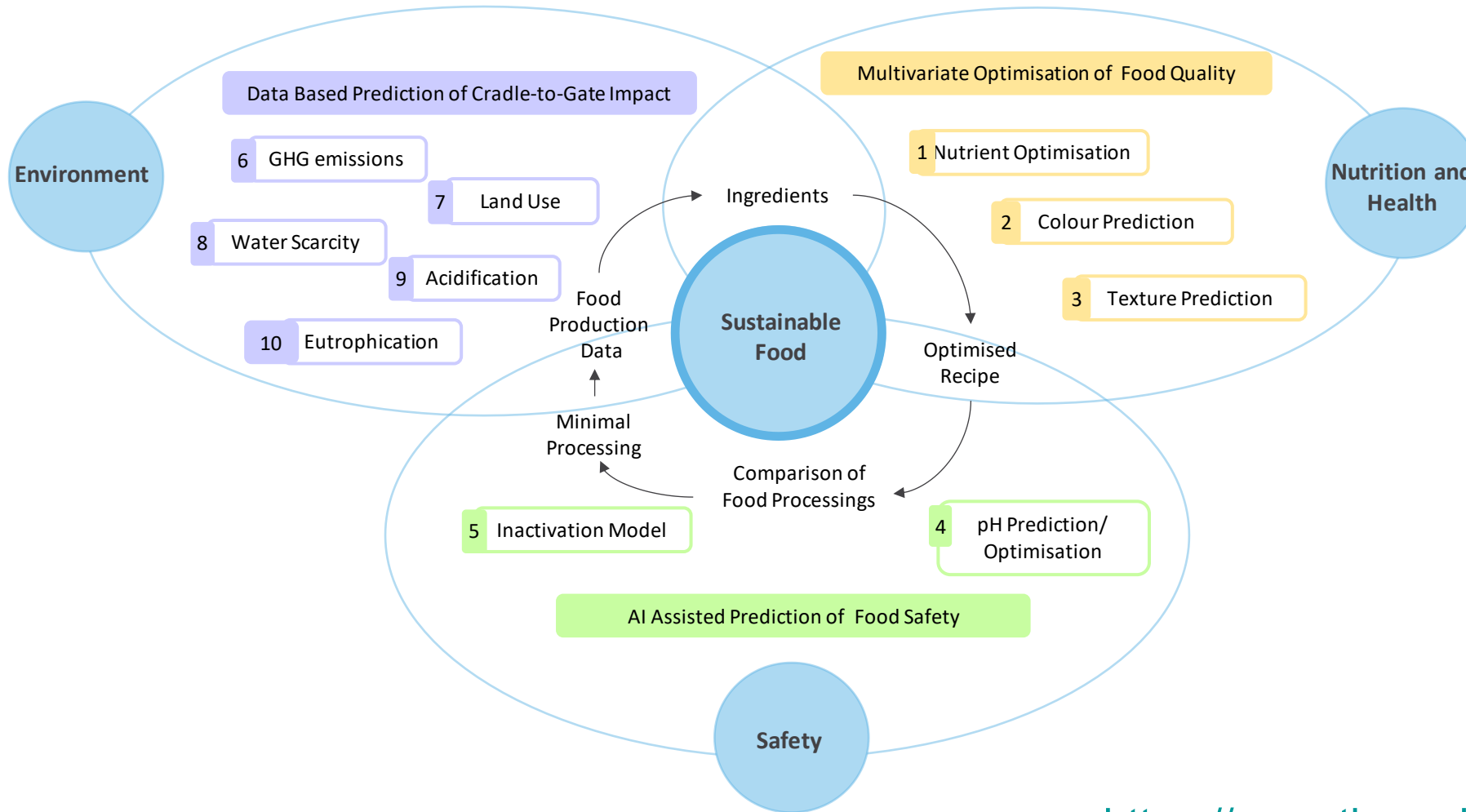
2

Can manual and automatized matching procedures be coupled into a **semi-automatized standardization approach** to facilitate data interlinkage?





EU project OptiSignFood



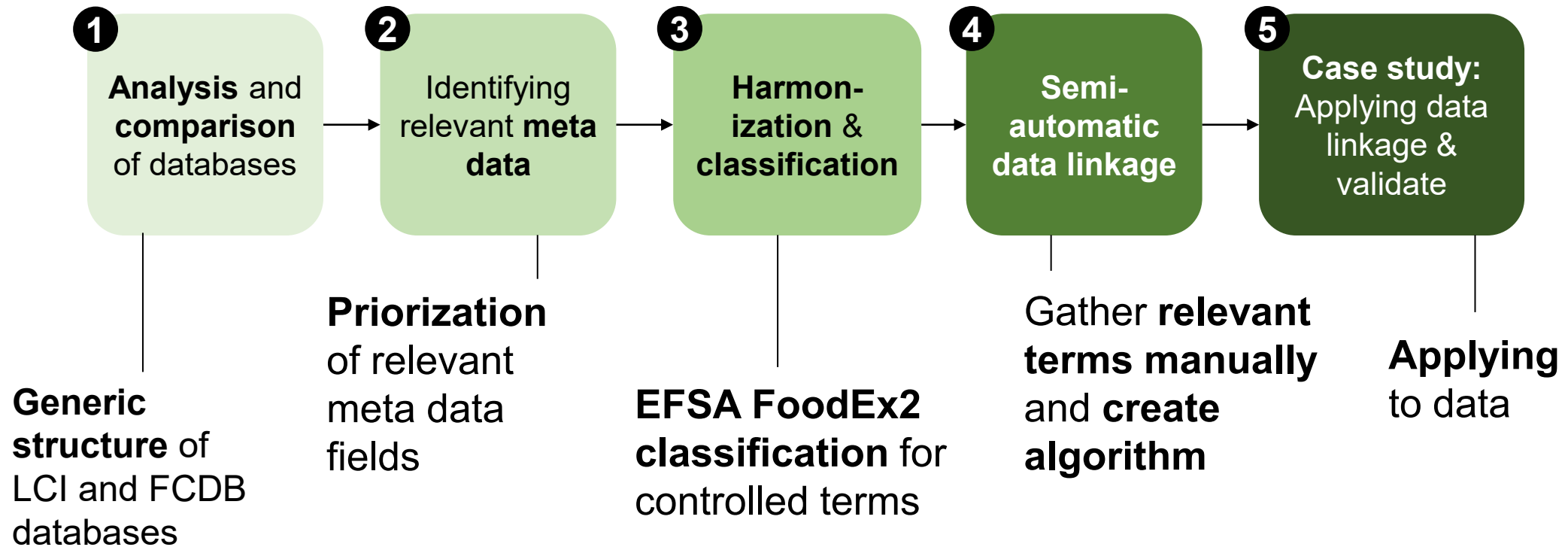
- Horizon 2020 call EIC-FTI-2018-2020
- Project OptiSignFood
- Title: Data Science and AI assisted holistic software to digitally design optimized high quality and safe food products with minor environmental impact
- 1.7.2021-30.6.2024

<https://www.themakersfood.com/optisignfood>



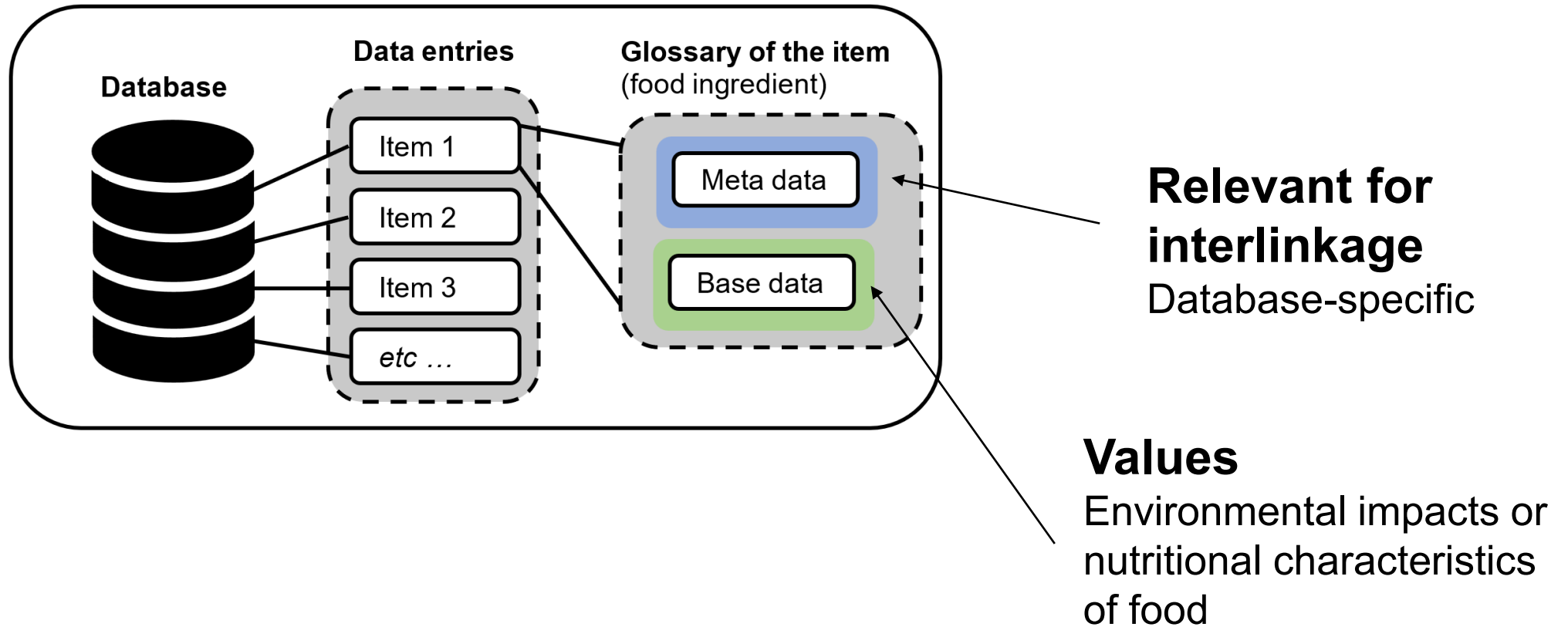
Methods

Relevant research areas for database interlinkage considered in this study





Data structure



Relevant for interlinkage
Database-specific

Values
Environmental impacts or nutritional characteristics of food



Data structure and availability

Comparing meta data

Description	Agribalyse	EuroFIR
ID		0211051
Name	Apple, conventional, electric platform, at orchard/kg/FR	Apple, fresh
Unit	Per kilogram	Per 100g edible portion
Category	AGRIBALYSE/Plant production/Fruits	Fruit or fruit product
Country	France	Switzerland
LanguaL™ code(s)		A0833, B1245, G0003, H0003, J0003, K0003 ...
Data quality	Technological representativeness = 2, Geographical representativeness = 2, Technological representativeness = 2, Completeness = 2, Precision/uncertainty = 2, Methodological appropriateness and consistency = 2	
Included processes	(1) the processes of soil preparation and cultivation, sowing, weed control, fertilisation, pest and pathogen control, harvest; (2) the machines and shed or surface used to park them; (3) all inputs as seed, fertilizers (mineral and organic), active substances, water for irrigation, fuels as well as the transport to the farm; (4) the direct emissions of the fuel combustion, the abrasion of tyres and the direct emissions on the field.	



Data structure and availability

Comparing meta data

Description	Agribalyse	EuroFIR	
ID		0211051	
Name	Apple, conventional, electric platform, at orchard/kg/FR	Apple, fresh	Useful meta data
Unit	Per kilogram	Per 100g edible portion	
Category	AGRIBALYSE/Plant production/Fruits	Fruit or fruit product	
Country	France	Switzerland	
LanguaL™ code(s)		A0833, B1245, G0003, H0003, J0003, K0003 ...	
Data quality	Technological representativeness = 2, Geographical representativeness = 2, Technological representativeness = 2, Completeness = 2, Precision/uncertainty = 2, Methodological appropriateness and consistency = 2		Partially useful
Included processes	(1) the processes of soil preparation and cultivation, sowing, weed control, fertilisation, pest and pathogen control, harvest; (2) the machines and shed or surface used to park them; (3) all inputs as seed, fertilizers (mineral and organic), active substances, water for irrigation, fuels as well as the transport to the farm; (4) the direct emissions of the fuel combustion, the abrasion of tyres and the direct emissions on the field.		Missing and/or incomplete





Relevant meta data

Parameter	Example	FCDB databases (e.g., EuroFIR)	LCI databases (e.g., Agribalyse)	Additional info
Food name	"Apple", "Mango", etc.	*** (III)	*** (III)	Information needs to be extracted from title of a database entry
Food specification	"Juice", "Oil", etc.	*** (II)	*** (II)	Information needs to be extracted from title of a database entry. Often inconsistently accessible information (e.g., "sunflower oil" vs. "oil, sunflower")
Food recipe	Percentage of water added to apple juice	*** (I)	*** (I)	Information, if provided, only in base data. Difficult to extract.
Food processing	"pasteurized"	*** (III)	*** (II)	Information needs to be extracted from title of a database entry
System boundaries	"at farm" or "at processing"	* (I)	*** (II)	Not always provided in the database entry in Agribalyse
Yield	Yield of apple from agricultural production	* (I)	*** (II)	Information only provided in base data. Difficult to extract.
Country of origin of food	"Germany", "France", etc.	** (I)	*** (III)	
Production system	"conventional", "organic", etc.	* (I)	*** (II)	Information needs to be extracted from title of a database entry

*: little relevant or irrelevant; **: moderately relevant; ***: highly relevant

I: not provided; II: sometimes provided; III: fully provided





Categories for connection list

Name

e.g., Apple

→ Describes **basic ingredient** without any further specification

Default

Not applicable

Specification

e.g., Juice

→ Describes a food in **more detail**

Default

None

Treatment

e.g., pasteurized

→ Any further **procedures** applied to the food

Default

Raw

Production System

e.g., Organic

→ Describes **how** the food is produced

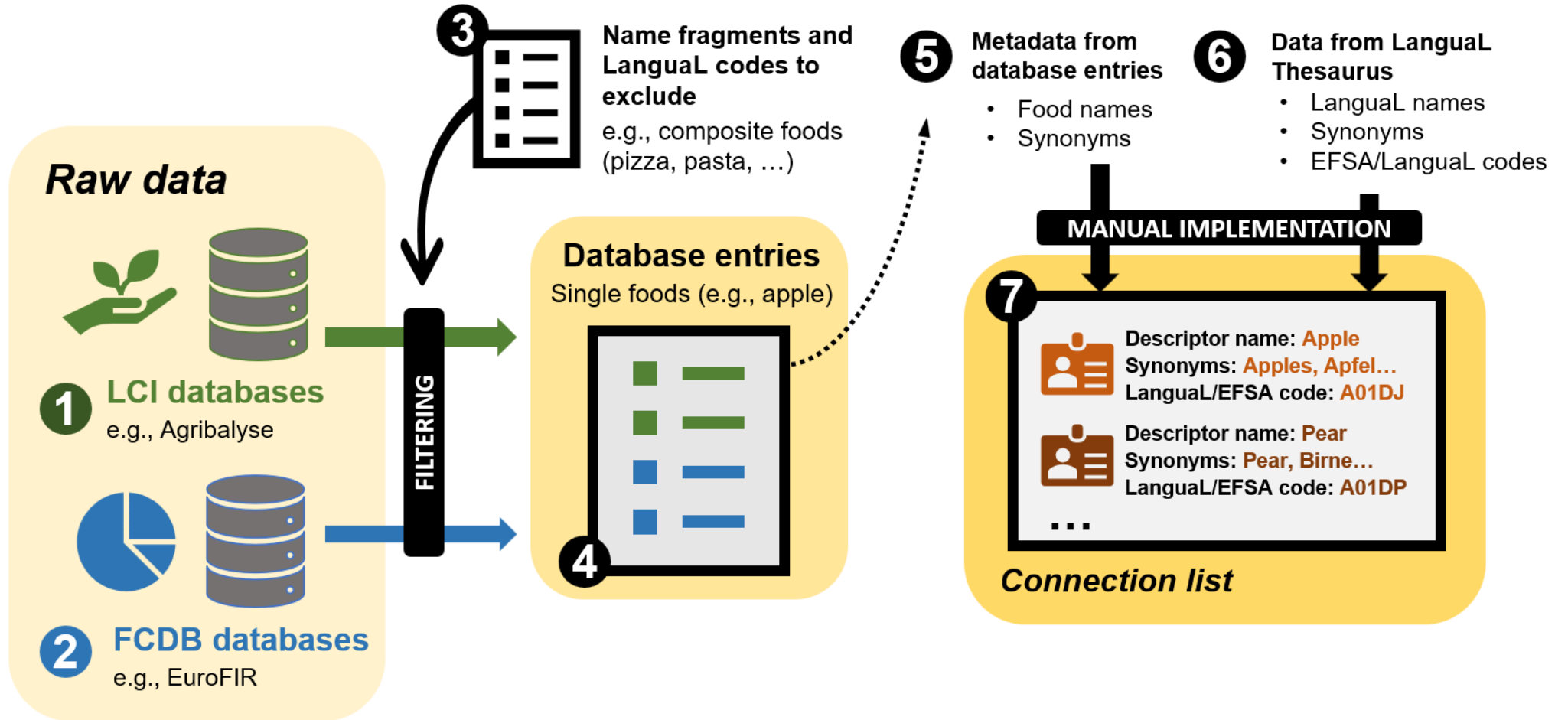
Default

Conventional



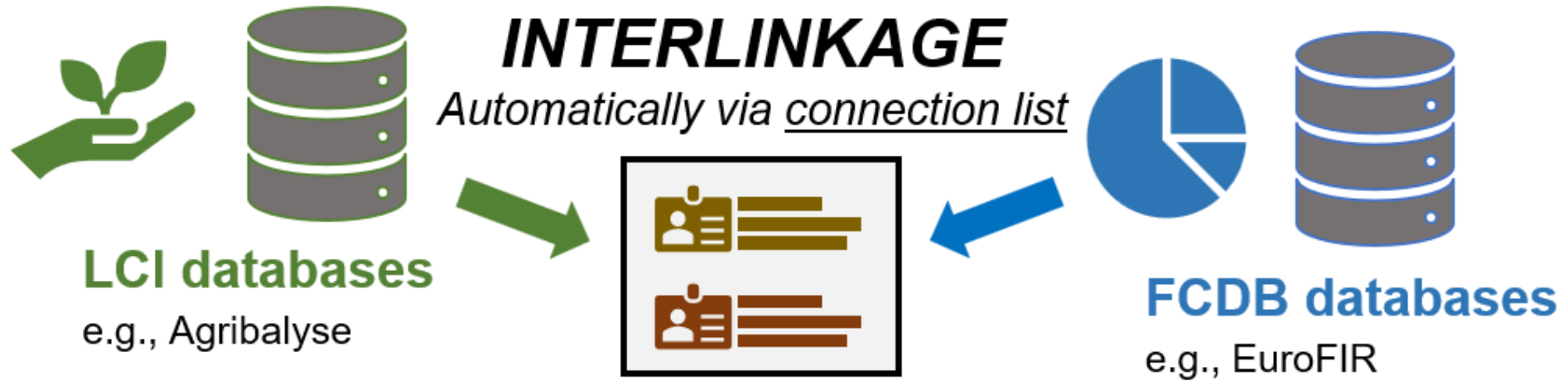


Connection list set-up





Semi-automatic data interlinkage



Name	Specification	Treatment	Production system	n (LCI)	n (EuroFIR)	Validation
Beef	Minced	Cooked	Conv.	4	6	
Cashew		Roasted	Conv.	2	6	
Cheese	Emmental	Raw	Conv.	5	8	
Rice	Flour	Raw	Conv.	2	6	
Sunflower	Oil	Raw	Conv.	4	6	1 wrong → Tuna canned in sunflower oil
Sweet potato		Cooked	Conv.	1	4	1 wrong → Puree with cream

Correct: **52** (17 LCI, 35 EuroFIR)

Incorrect: **2** (1 LCI, 1 EuroFIR)



Discussion

Classification systems

- Importance and relevance
- Implementation in Agribalyse and EuroFIR

Meta data

- Limited availability
- Often difficult to access
- Food water content
- Implementation in Agribalyse and EuroFIR

Complex technical infrastructure

- Data format
- Process automatization





Conclusion

The procedure developed **showed to work and connect data successfully**

Aspects that should be further elaborated

- Handling of **incomplete or missing data**
- **Technological complexity and data format**

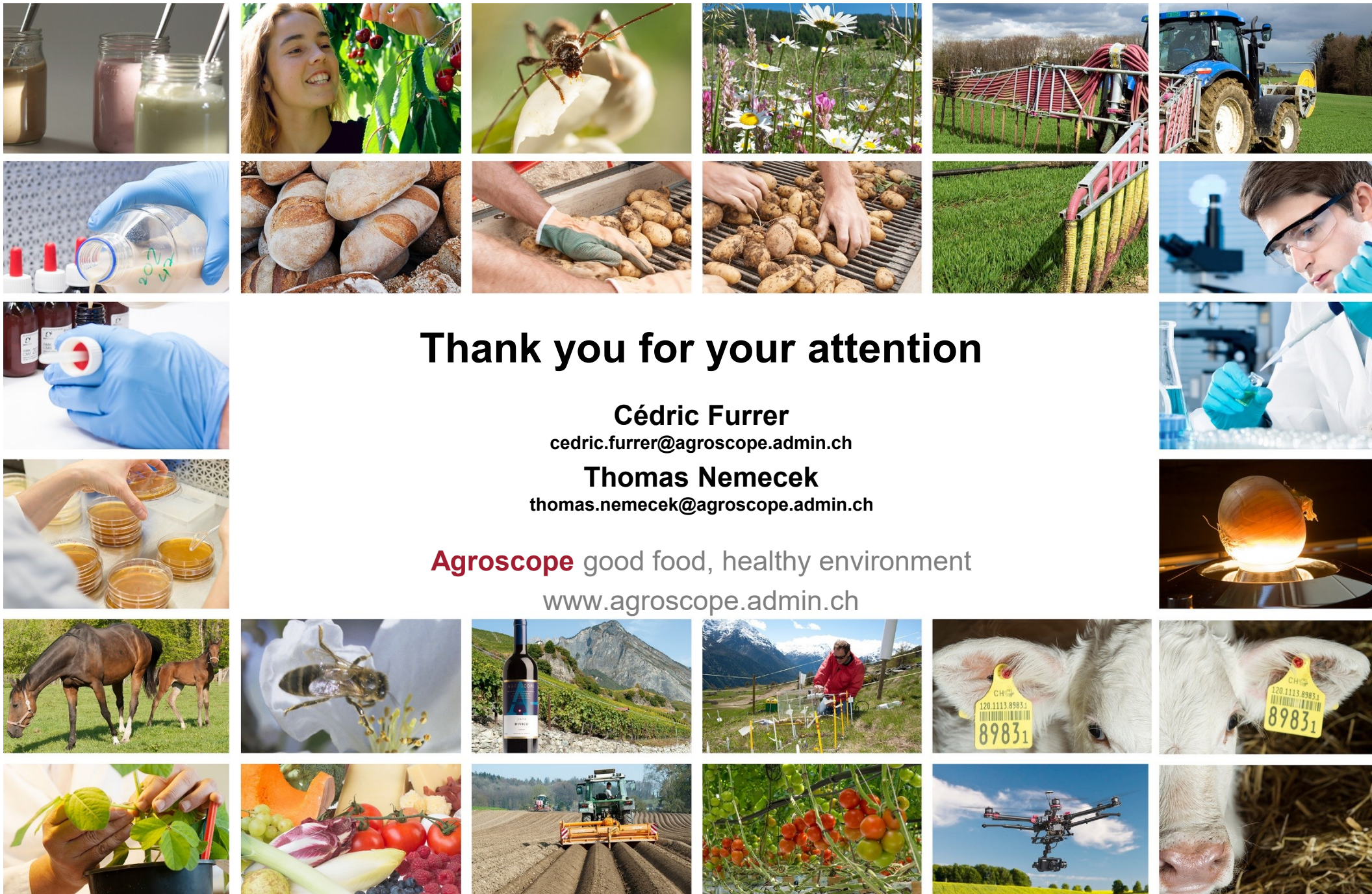
Standardization remains a **challenge**

- Several attempts, none of them has been shown to satisfy the needs → incomplete matching achieved, time-consuming process for manual validation of data

Use of **artificial intelligence**

Furrer C., Sieh D., Jank A. M., le Bras G., Herrmann M., Reguant-Closa A., & Nemecek T., 2024. Interlinking environmental and food composition databases: an approach, potential and limitations. J Cleaner Prod, 143198. <https://doi.org/10.1016/j.jclepro.2024.143198>





Thank you for your attention

Cédric Furrer
cedric.furrer@agroscope.admin.ch

Thomas Nemecek
thomas.nemecek@agroscope.admin.ch

Agroscope good food, healthy environment
www.agroscope.admin.ch