PADERBORN UNIVERSITY
The University for the Information Society

Exploring the data-driven world - Teaching AI and ML from a data-centric perspective

Raspberry Pi Foundation research seminar

Carsten Schulte, Yannik Fleischer, Lukas Höper

(Rolf Biehler, Daniel Frischemeyer, Sven Hüsing, Susanne Podworny)

Paderborn University

05.10.2021

Context

Deutsche Telekom Stiftung

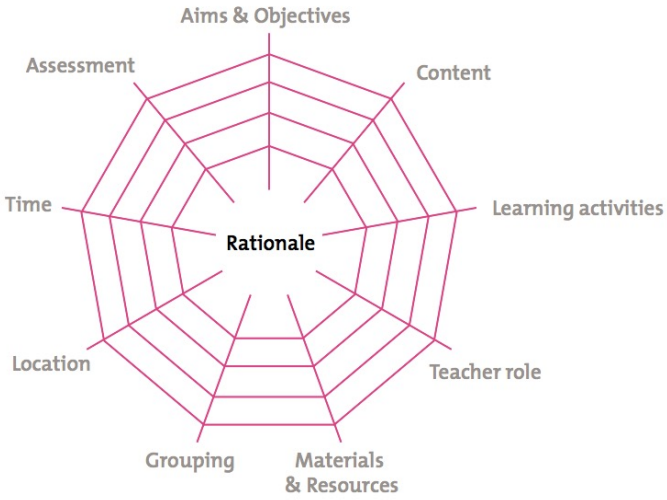
ProDaBi
Projekt Data Science und Big Data in der Schule

Prodabi

- Symposium on Data Science @ School 2017
- Curriculum development
- Project Course
- Stand alone modules
  - 15/16 year olds (grade 9710)
  - 11/12 year olds (grade 576)

An Initiative of the Federal Ministry of Education and Research
Science Year 2019
ARTIFICIAL INTELLIGENCE

Curriculum Design



Annette Thijs, and Jan van den Akker, eds. *Curriculum in Development.* Enschede: SLO - Netherlands Institute for curriculum development, 2009. , p. 11

PADERBORN UNIVERSITY
The University for the Information Society

Context

Deutsche Telekom Stiftung

ProDaBi
Projekt Data Science und Big Data in der Schule

Project Course

- Data Project Topic
- JIM-Data (Codap)
- Data Exploration (JN)
- Decision Trees
- Neural Nets
- Project

Modules 8/9

- JIM-Data (Codap)
- Data Exploration (JN, Arduino)
- Decision Trees & Neural Nets
- Data Awareness

Grade 7

An Initiative of the Federal Ministry of Education and Research

Science Year 2019

ARTIFICIAL INTELLIGENCE

Modules 5 / 6

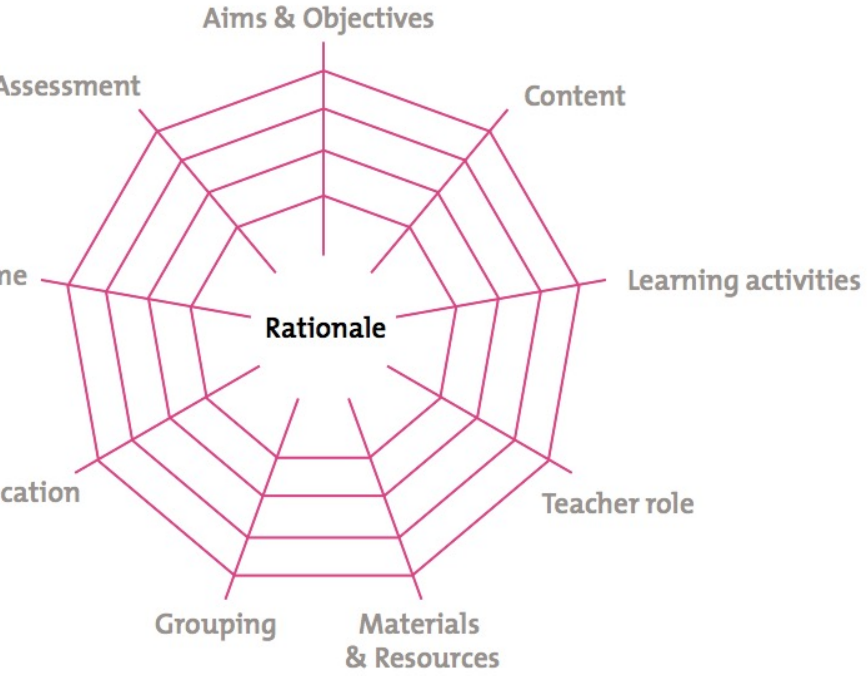
- Decision Trees & Data Cards
- Data Awareness

Context

Deutsche Telekom Stiftung

ProDaBi
Projekt Data Science und Big Data in der Schule

Project Course



Modules 8/9

Grade 7

An Initiative of the Federal Ministry of Education and Research
Science Year 2019
ARTIFICIAL INTELLIGENCE

Modules 5 / 6

Decision Trees & Data Cards

Data Awareness

Rationale

Machine Behaviour, 2019

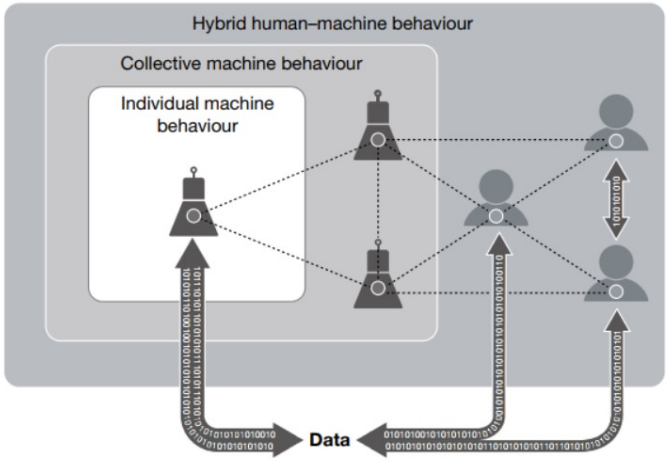


Fig. 4 | Scale of inquiry in the machine behaviour ecosystem. AI systems represent the amalgamation of humans, data and algorithms. Each of these domains influences the other in both well-understood and unknown ways. Data—filtered through algorithms created by humans—influences

Rahwan, Iyad ; Cebrian, Manuel ; Obradovich, Nick ; Bongard, Josh ; Bonnefon, Jean-François ; Breazeal, Cynthia ; Crandall, Jacob W. ; Christakis, Nicholas A. ; u. a.: Machine behaviour. In: Nature Bd. 568 (2019), Nr. 7753, S. 477–486. — tex.ids= Rahwan.2019, rahwanMachineBehaviour2019a

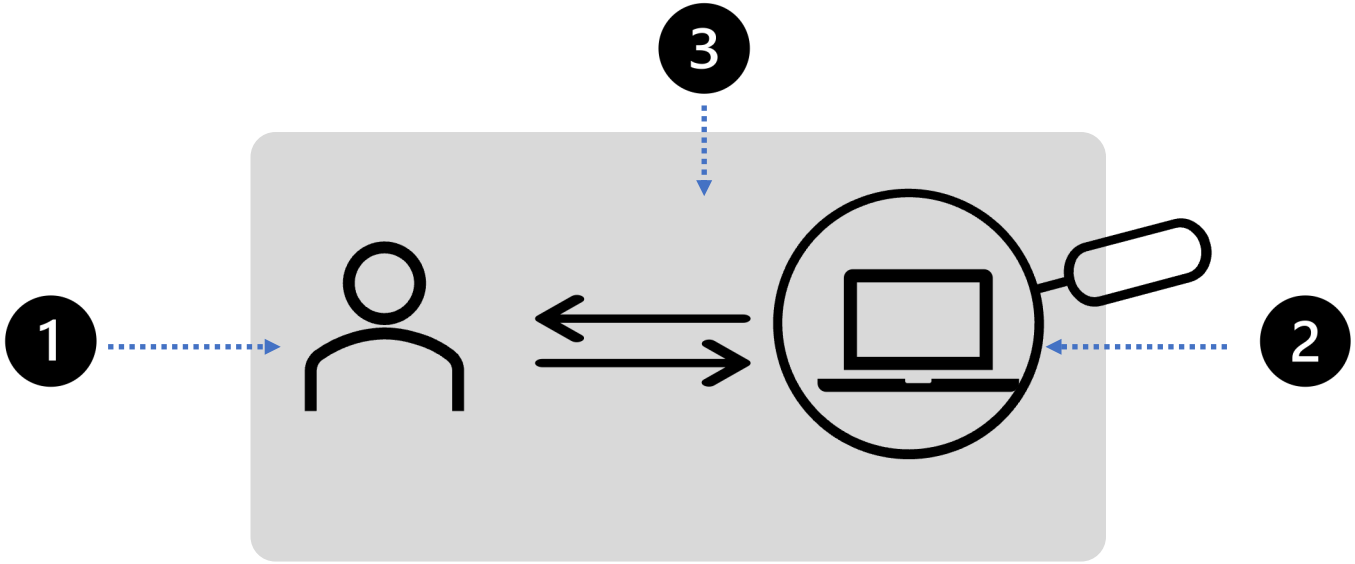
Computational Thinking, 2006

“Ideas, not artifacts”

1970 1980 1990 2000 2010 2020

WING, Jeanette M, 2006. Computational Thinking. *Communications of ACM*. March 2006. Vol. 49, p. 33–35. DOI 10.1145/1118178.1118215.
p. 35

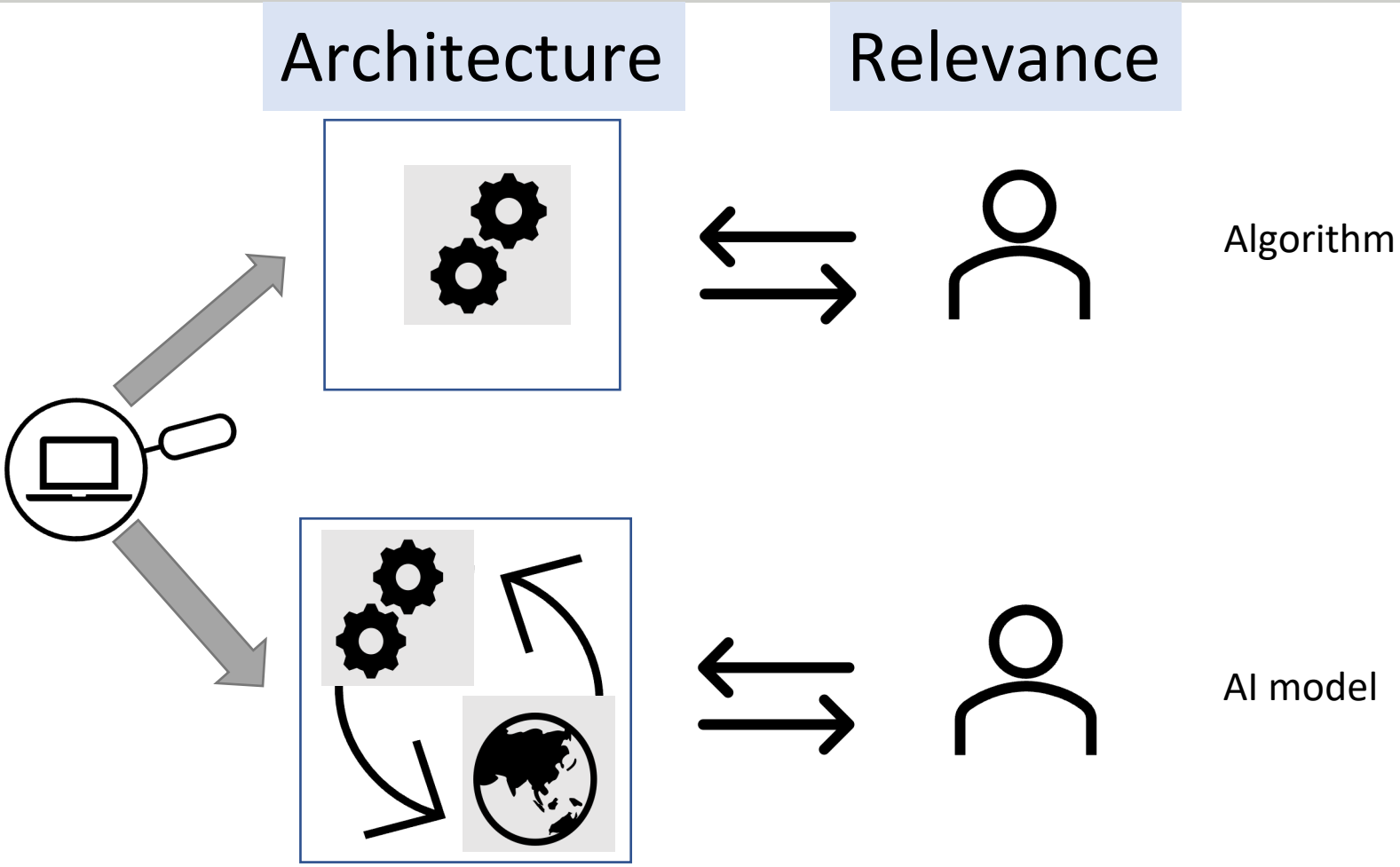
Hybrid Interaction System



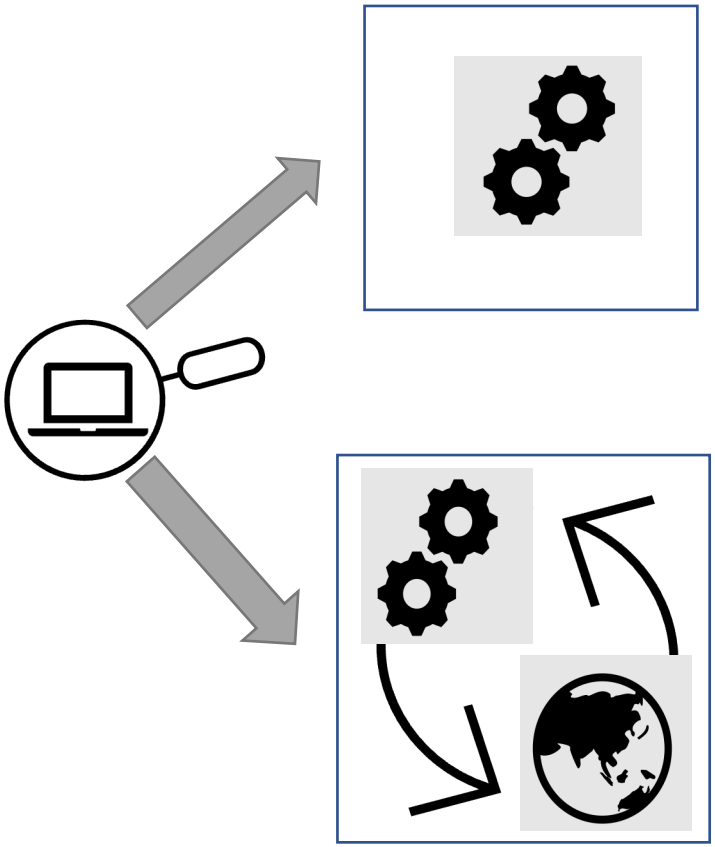
1. The role of the human: “program or be programmed”
2. The role of the artifact: replacement, augmentation, symbiosis
3. The role of the hybrid system: shaping and being shaped

SCHULTE, Carsten and BUDDE, Lea, 2018. A Framework for Computing Education: Hybrid Interaction System: The need for a bigger picture in computing education. In: *18th Koli Calling International Conference on Computing Education Research (Koli Calling ’18)*. Koli, Finland: ACM. 22 November 2018

(Dual nature of) digital Artifacts



Impact of change to ML: problem solving



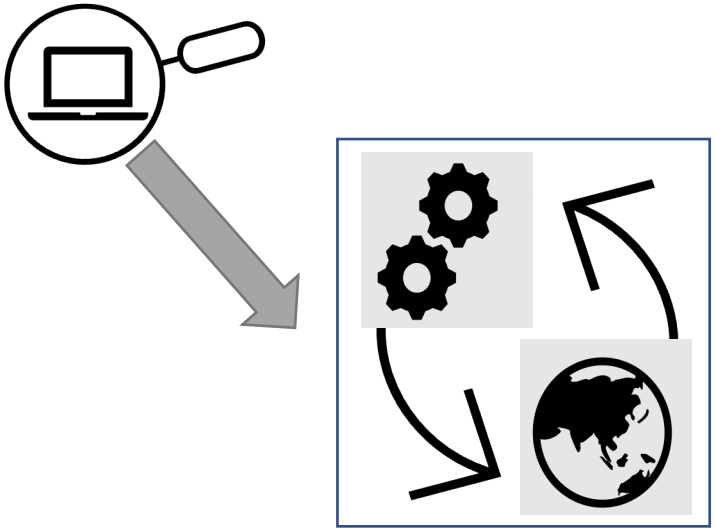
$$x \rightarrow ALG \rightarrow y$$

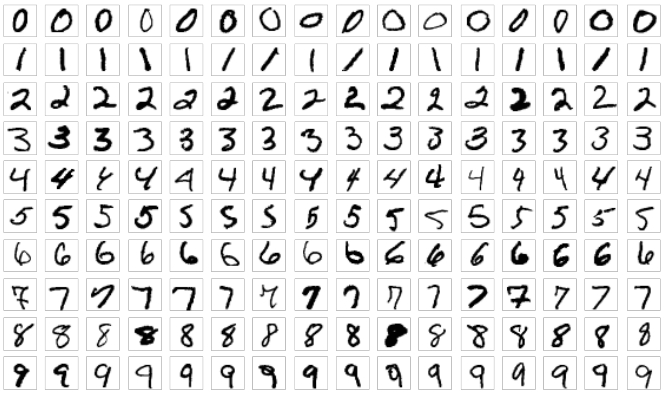
Understanding the problem, deriving a solution, understanding the solution

$$Data \rightarrow Learner \rightarrow (x \rightarrow Model \rightarrow y)$$

~~Understanding the problem, deriving a solution, understanding the solution~~

Impact of change to ML: accuracy





Josef Steppan, CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0>, via Wikimedia Commons

Error rate (%)
7.6[10]
0.52[23]
0.87[24]
3.3[10]
2.8[26]
0.56[27]
1.6[28]
0.7[28]
0.35[29]
0.31[30]
0.27[31]
0.25[17]
0.23[12]
0.21[19][20]
0.18[22]
0.17[33]

https://en.wikipedia.org/wiki/MNIST_database

Impact of change to ML: role of code

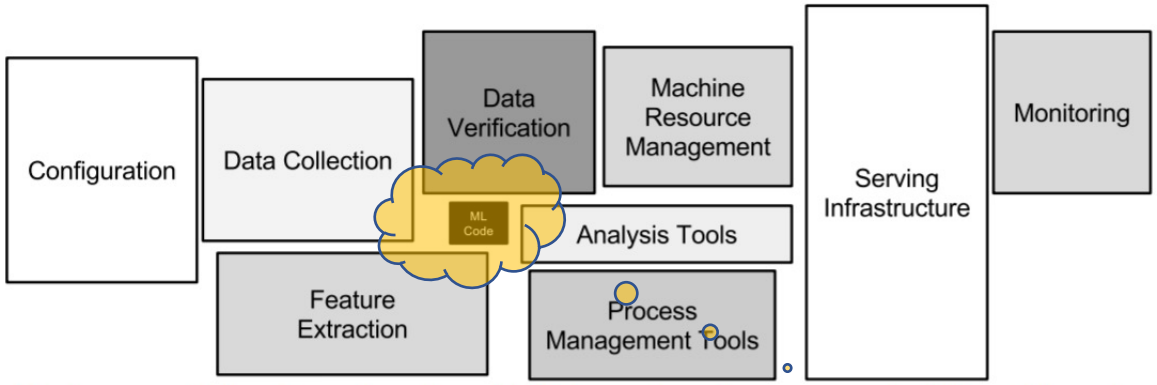
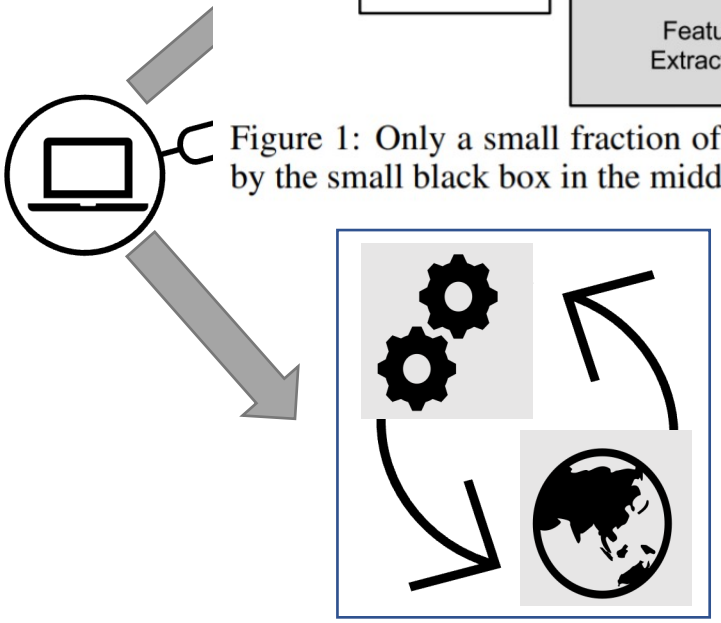


Figure 1: Only a small fraction of real-world ML systems is composed of the ML code, as shown by the small black box in the middle. The required surrounding infrastructure is vast and complex.

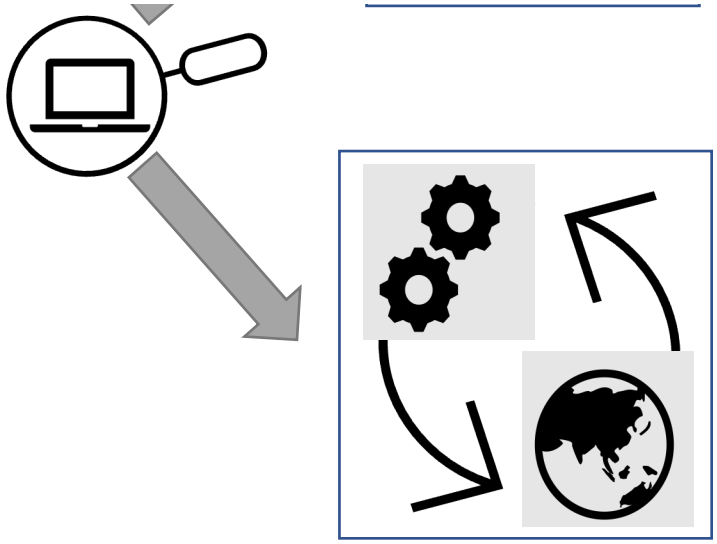


SCULLEY, D., et al., 2015. Hidden Technical Debt in Machine Learning Systems. In: *Advances in Neural Information Processing Systems* [online]. Curran Associates, Inc. 2015.

Impact of change to ML: role of data

Conceptual Framework	Data Collection		Data Management					
Introduction to Data	Data Discovery and Collection	Evaluating and Ensuring Quality of Data and Sources	Data Organization	Data Manipulation	Data Conversion (from format to format)	Metadata Creation and Use	Data Curation, Security, and Re-Use	Data Preservation

Data Evaluation							Data Application					
Data Tools	Basic Data Analysis	Data Interpretation (Understanding Data)	Identifying Problems Using Data	Data Visualization	Presenting Data (Verbally)	Data Driven Decisions Making (DDDM) (Making decisions based on data)	Critical Thinking	Data Culture	Data Ethics	Data Citation	Data Sharing	Evaluating Decisions Based on Data

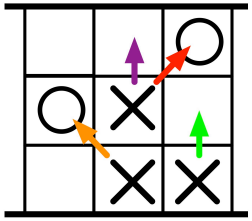


RIDSDALE, et al., 2015. Strategies and best practices for data literacy education: knowledge synthesis report [online]. Dalhousie University p.3

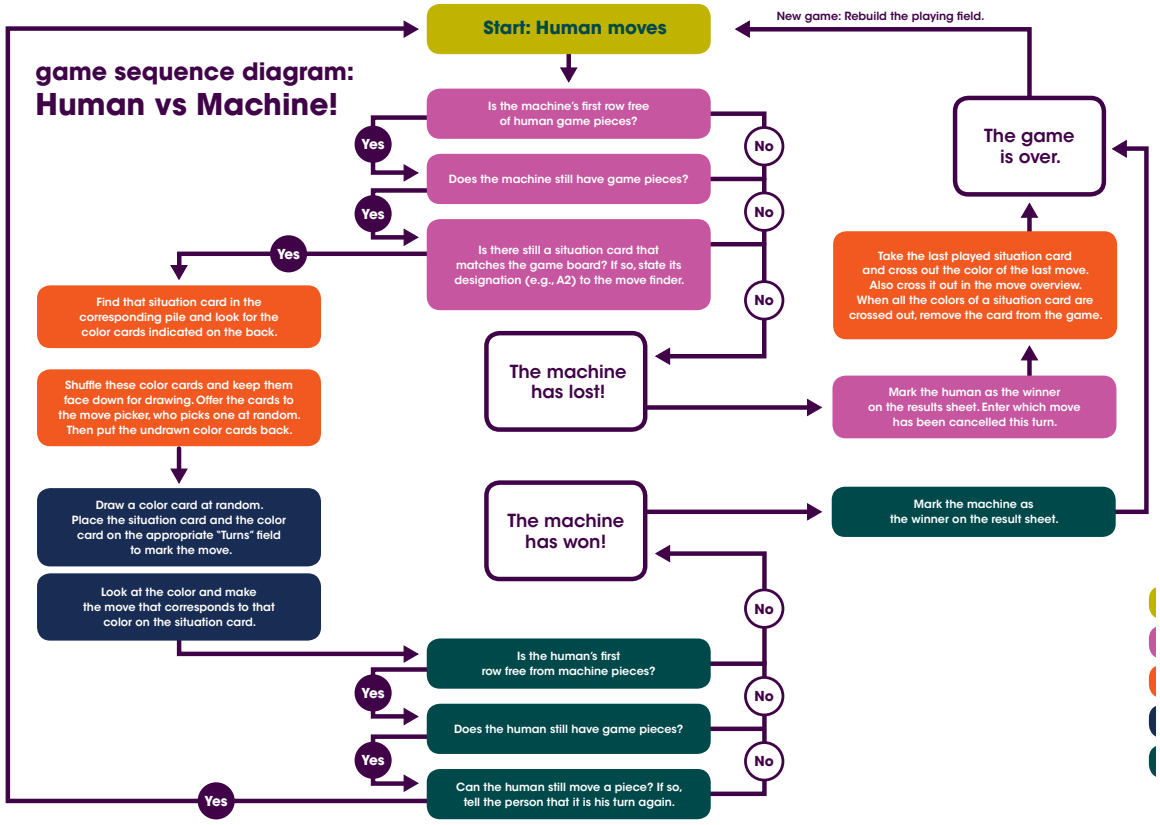
Example 1:Human vs Machine

https://www.prodabi.de/english-version-of-the-human-vs-machine-game/

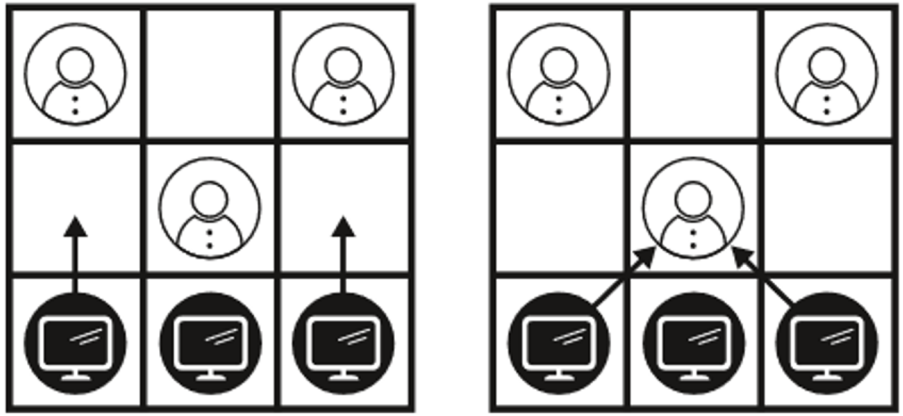
https://en.wikipedia.org/wiki/Matchbox_Educable_Noughts_and_Crosses_Engine



http://www.cs4fn.org/machinelearning/sweetlearningcomputer.php

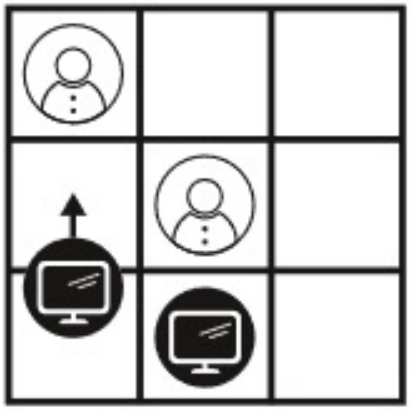

game sequence diagram:
Human vs Machine!
Start: Human moves
Is the machine's first row free of human game pieces?
Yes
No
Does the machine still have game pieces?
Yes
No
Is there still a situation card that matches the game board? If so, state its designation (e.g., A2) to the move finder.
Yes
No
Find that situation card in the corresponding pile and look for the color cards indicated on the back.
Shuffle these color cards and keep them face down for drawing. Offer the cards to the move picker, who picks one at random. Then put the undrawn color cards back.
Draw a color card at random. Place the situation card and the color card on the appropriate "Turns" field to mark the move.
Look at the color and make the move that corresponds to that color on the situation card.
Is the human's first row free from machine pieces?
Yes
No
Does the human still have game pieces?
Yes
No
Can the human still move a piece? If so, tell the person that it is his turn again.
Yes
The machine has lost!
The machine has won!
Mark the human as the winner on the results sheet. Enter which move has been cancelled this turn.
Mark the machine as the winner on the result sheet.
Take the last played situation card and cross out the color of the last move. Also cross it out in the move overview. When all the colors of a situation card are crossed out, remove the card from the game.
The game is over.
New game: Rebuild the playing field.

Hexapawn – poosible moves

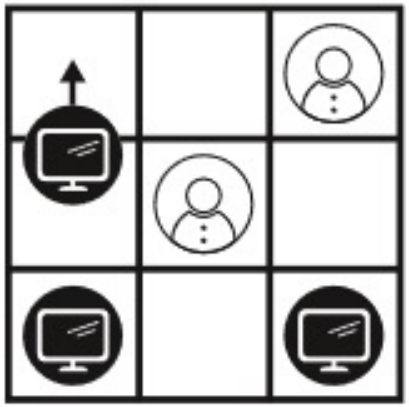


The pieces in Hexapawn can either move straight ahead, or capture an opponent's piece diagonally

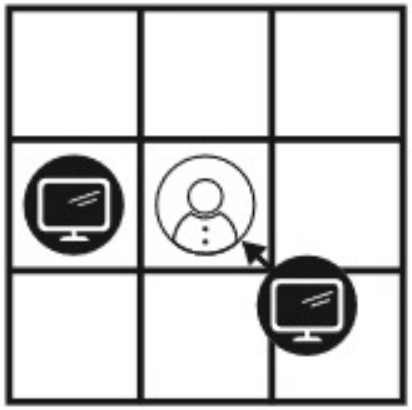
When is the game won?



You win the game when the opponent's pieces are blocked

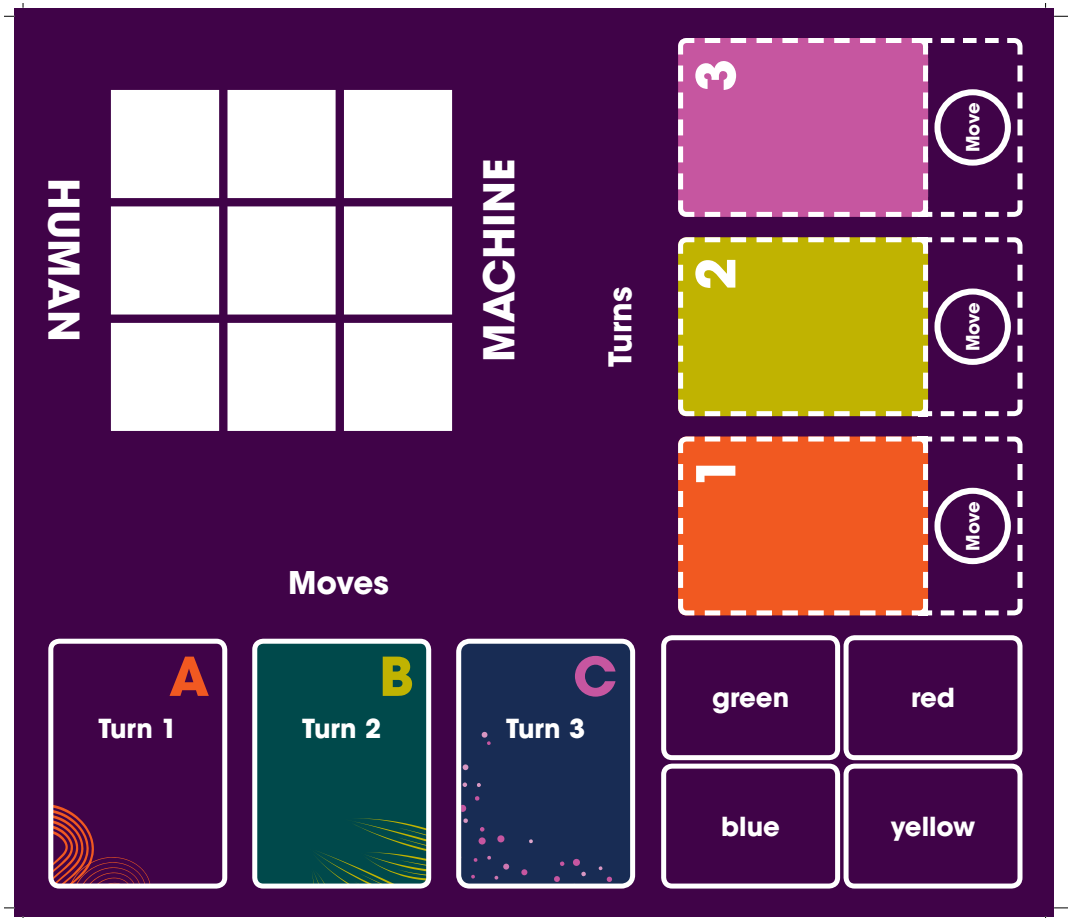


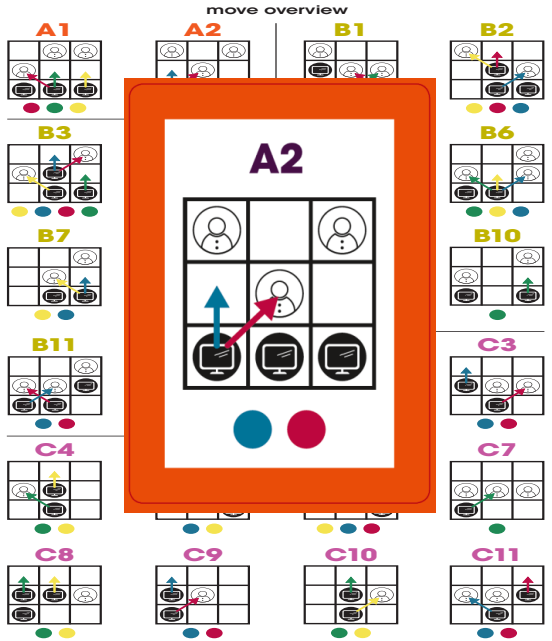
your own piece reaches the opponent's basic line



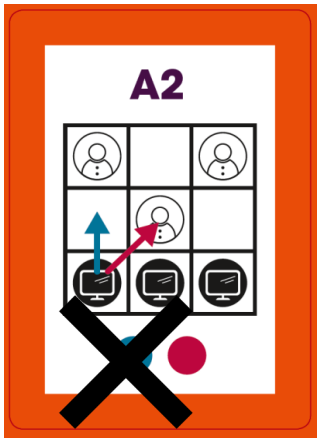
or when all the opponent's pieces have been captured.

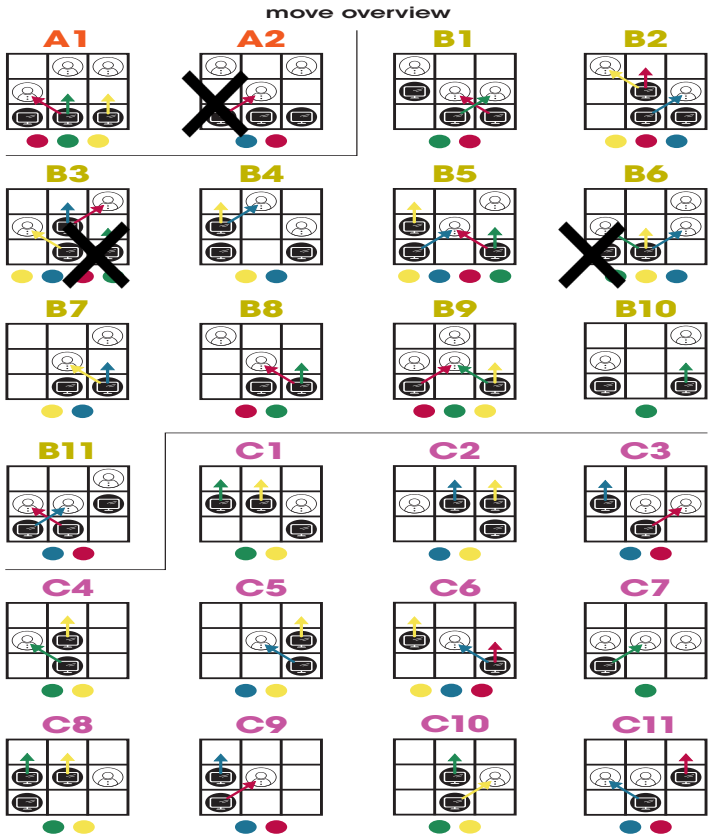
Game Board & overview



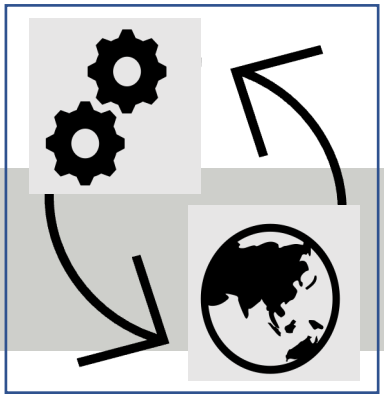


Game Board & overview





Game Board & overview



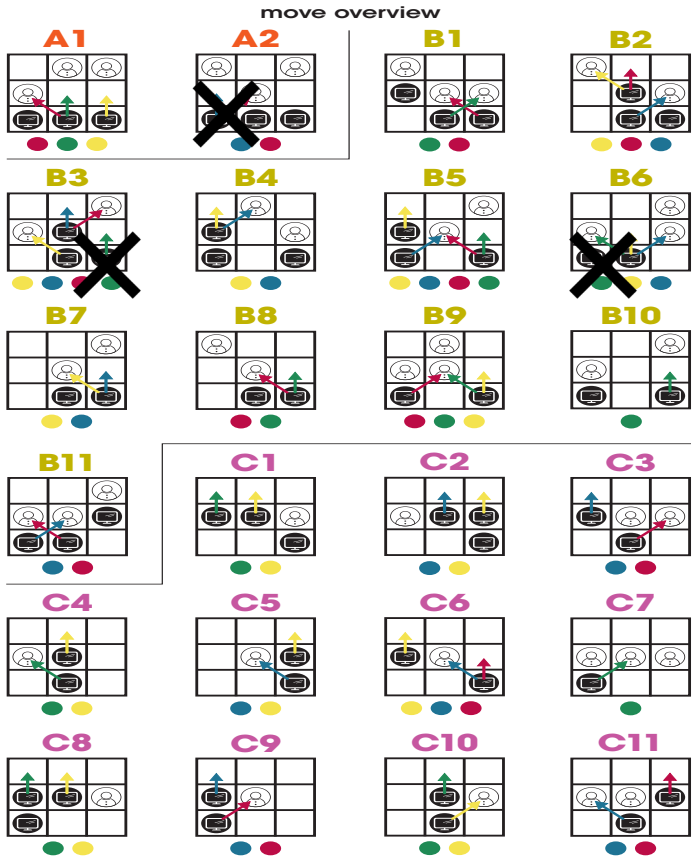
Result sheet

Play at least 10 games.

Record the outcomes of the games. If the machine loses, remove the color of the last turn from the situation card and the move overview. If all the colors on a situation card are crossed out, remove it from the game.

Also, if the machine loses, write down which colors were crossed out of which situation cards (e.g. "Red/C2" ...).

Game round	Winner	The following **color** was removed
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		



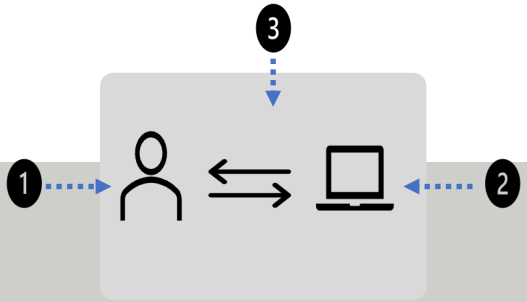
Reflection: What have we learned?

- Different groups= different games & different models
- Reflections on the process:
  - “machine” players just followed rules -> not ”intelligent“
  - Slow -> lot’s of data (games) needed
  - Machine “learns” only when it has lost
    - Why? Alternatives? ->human can shape the learning process
- …

Reflection
Hexapawn vs. autonomous cars

- Not only are there a lot more factors and sensor data to consider, you also have to react to sudden problems such as human fallibility:
  - A self-driving car must always be prepared for random events. For example, it cannot know that other road users make mistakes that cannot be calculated (heart attack, for example).
    So the AI must know how to react spontaneously.
- The other reason cited by many participants is that this learning method would be too dangerous or too expensive to use in the real world as both people and cars would be at risk if a self-driving car learned by punishment:
  - Cars have to learn without experience or mistakes, because mistakes could endanger human lives.

GROSSE-BÖLTING, Gregor and MÜHLING, Andreas, 2020. Students Perception of the Inner Workings of Learning Machines. In: *LATICE Seventh International Conference on Learning and Teaching in Computing and Engineering*. 2020.



Summary: Hexapawn

1 The role of the human: "program or be programmed"	"Trainer": provide suitable data Reflect on Training, maybe adapt
2 The role of the artifact: replacement, augmentation, symbiosis	It's a "machine", follows rules
3 The role of the hybrid system: shaping and being shaped	Training & use are shapeable
4 The role of data:	Data quality == model quality
5 Paradigm change in teaching?	Link ideas to artifacts; focus on reflections

Example 2: Decision Trees

Advantages
The issue concerns everyone (age, gender, ...)
Many possibilities for expansion, deepening, and connection

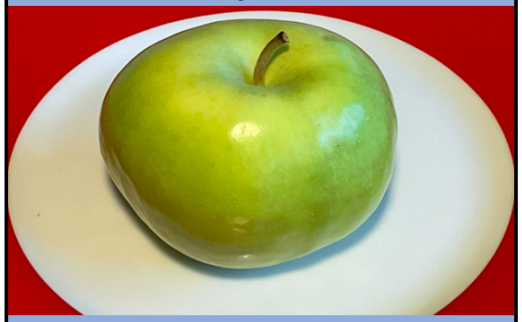
Difficulties
Labels are not clear for all foods, as the issue of nutrition is ambivalent and multifactorial (subjective labeling)

Teacher must convey sensitivity that a red bracket does not mean that the food should never be eaten again and vice versa

Chances
What role do humans and data have in this process?

Data Cards

Apfel



Nährwerte pro 100g	
Energie	52 kcal
Fett	0,2 g
davon gesättigte Fettsäuren	0,0 g
Kohlenhydrate	13,8 g
davon Zucker	11,0 g
Eiweiß	0,3 g
Salz	0,0 g

ProDaBi

Popcorn

Nährwerte pro 100g	
Energie	499 kcal
Fett	23,0 g
davon gesättigte Fettsäuren	13,8 g
Kohlenhydrate	57,0 g
davon Zucker	3,8 g
Eiweiß	10,7 g
Salz	1,8 g

ProDaBi

PADERBORN UNIVERSITY
The University for the Information Society

…more cards

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

Apfel

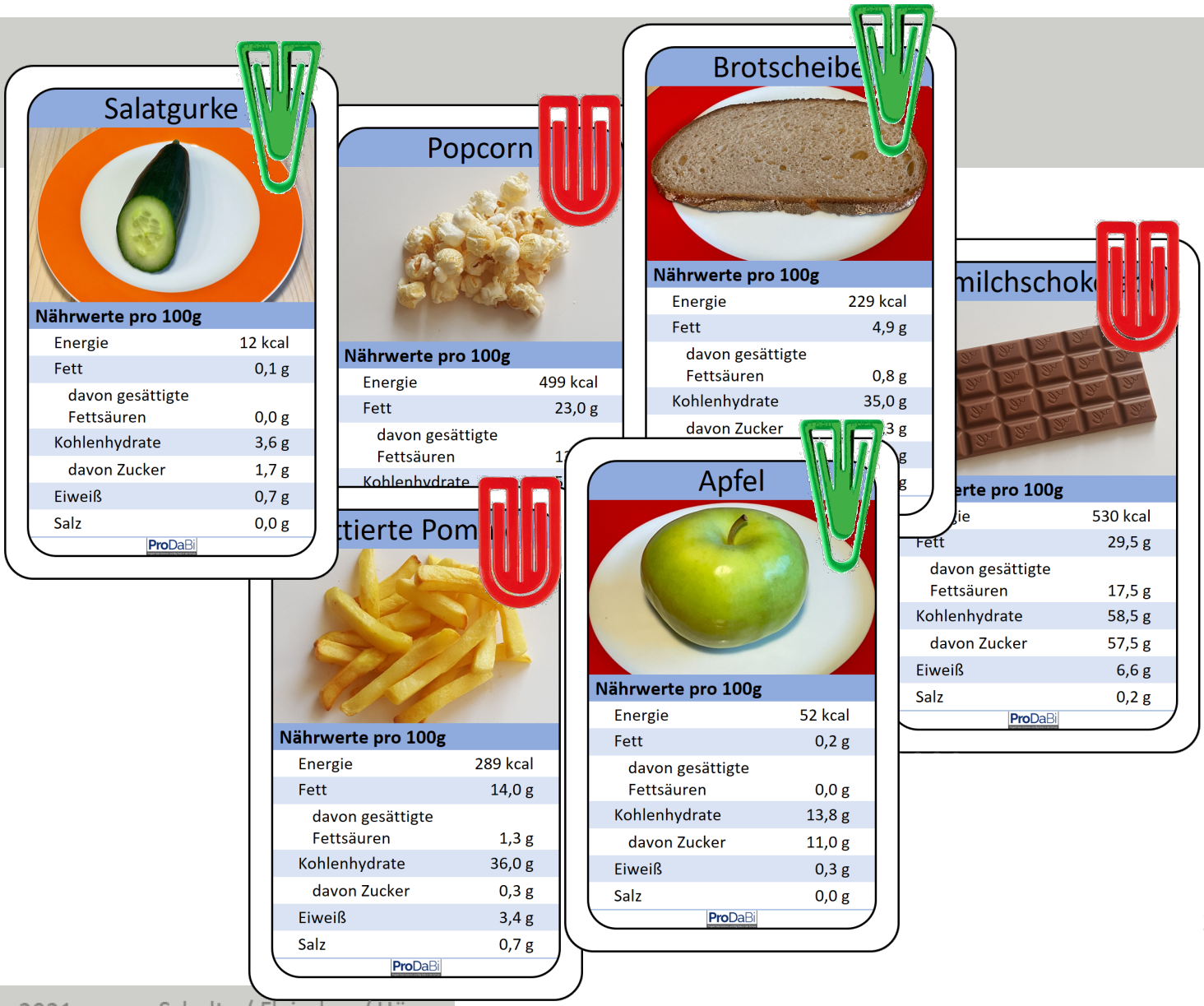
Nährwerte pro 100g	
Energie	52 kcal
Fett	0,2 g
davon gesättigte Fettsäuren	0,0 g
Kohlenhydrate	13,8 g
davon Zucker	11,0 g
Eiweiß	0,3 g
Salz	0,0 g

ProDaBi

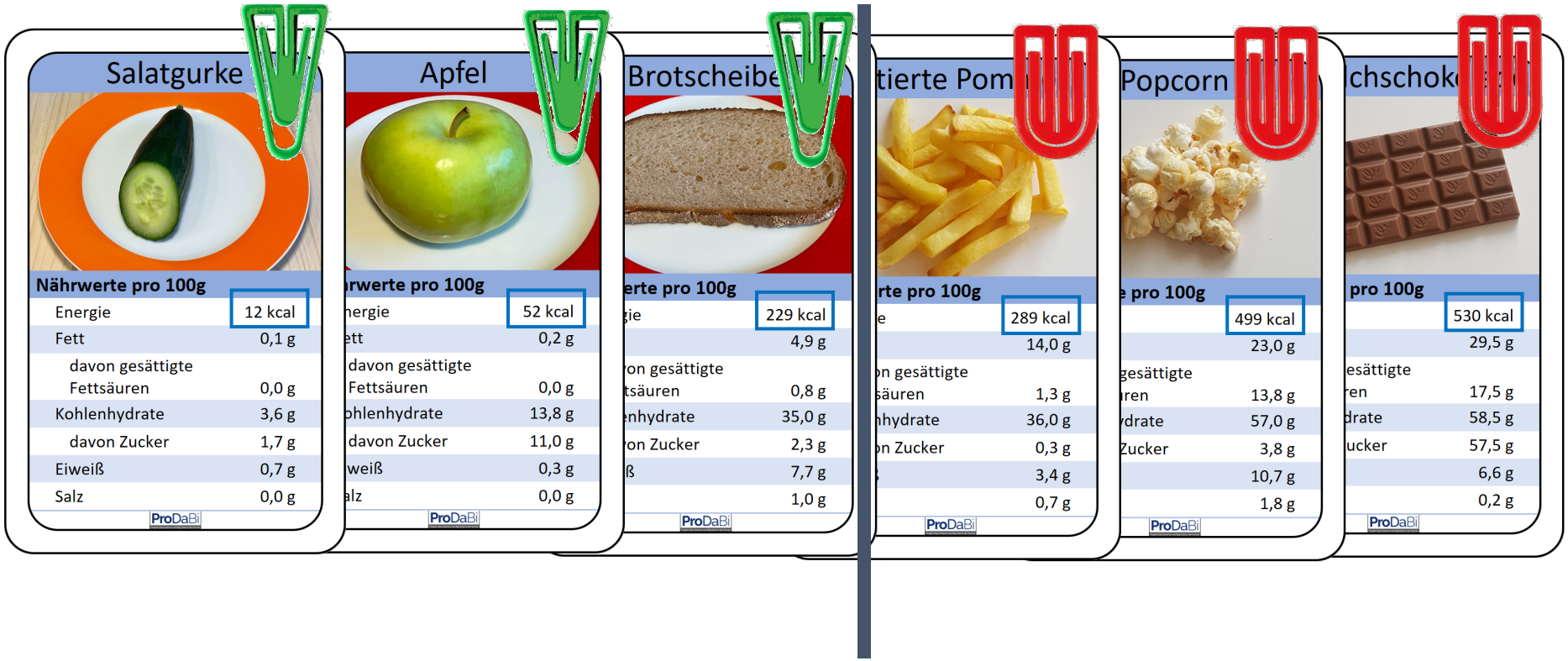
Popcorn

Nährwerte pro 100g	
Energie	499 kcal
Fett	23,0 g
davon gesättigte Fettsäuren	13,8 g
Kohlenhydrate	57,0 g
davon Zucker	3,8 g
Eiweiß	10,7 g
Salz	1,8 g

ProDaBi

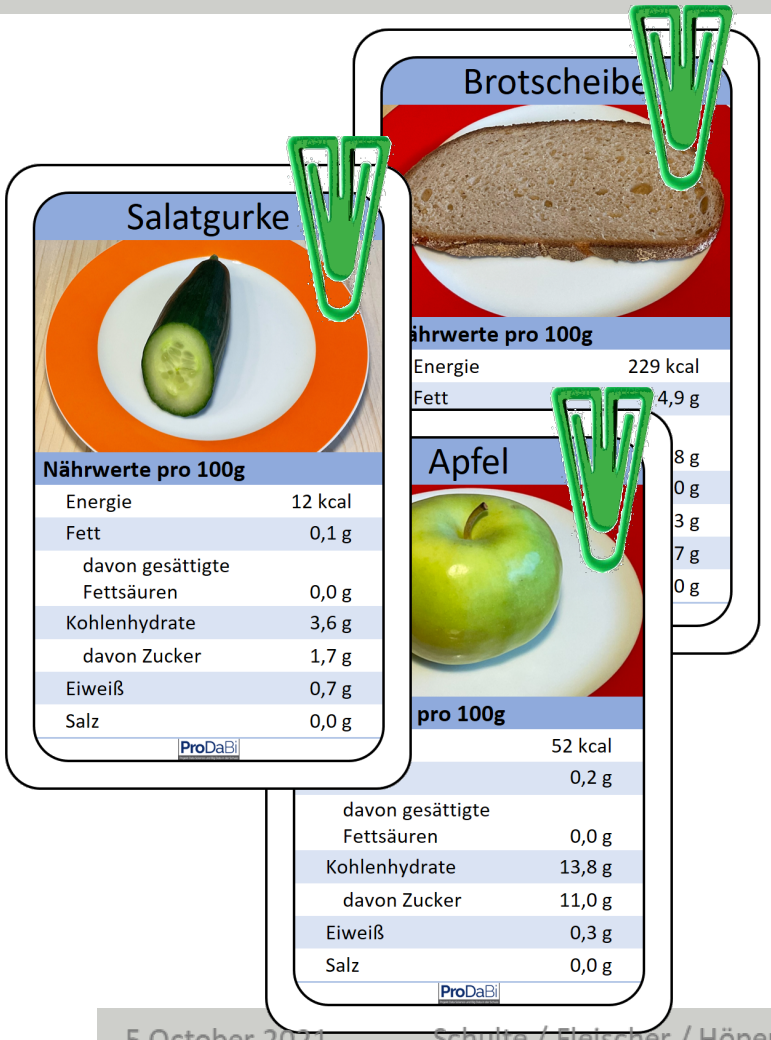

Salatgurke
Nährwerte pro 100g
Energie 12 kcal
Fett 0,1 g
davon gesättigte Fettsäuren 0,0 g
Kohlenhydrate 3,6 g
davon Zucker 1,7 g
Eiweiß 0,7 g
Salz 0,0 g
ProDaBi
Popcorn
Nährwerte pro 100g
Energie 499 kcal
Fett 23,0 g
davon gesättigte Fettsäuren
Brotscheibe
Nährwerte pro 100g
Energie 229 kcal
Fett 4,9 g
davon gesättigte Fettsäuren 0,8 g
Kohlenhydrate 35,0 g
Apfel
Nährwerte pro 100g
Energie 52 kcal
Fett 0,2 g
davon gesättigte Fettsäuren 0,0 g
Kohlenhydrate 13,8 g
davon Zucker 11,0 g
Eiweiß 0,3 g
Salz 0,0 g
ProDaBi
Nährwerte pro 100g
Energie 289 kcal
Fett 14,0 g
davon gesättigte Fettsäuren 1,3 g
Kohlenhydrate 36,0 g
davon Zucker 0,3 g
Eiweiß 3,4 g
Salz 0,7 g
ProDaBi
Nährwerte pro 100g
Energie 530 kcal
Fett 29,5 g
davon gesättigte Fettsäuren 17,5 g
Kohlenhydrate 58,5 g
davon Zucker 57,5 g
Eiweiß 6,6 g
Salz 0,2 g
ProDaBi

260 kcal

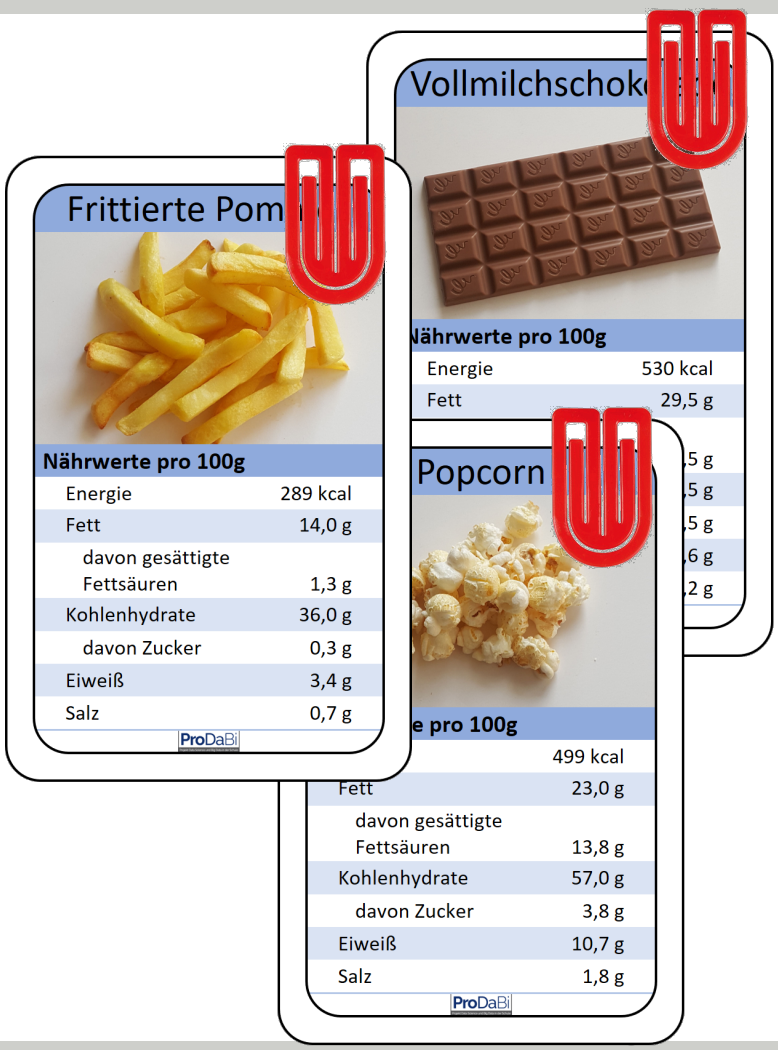

Salatgurke
Nährwerte pro 100g
Energie 12 kcal
Fett 0,1 g
davon gesättigte Fettsäuren 0,0 g
Kohlenhydrate 3,6 g
davon Zucker 1,7 g
Eiweiß 0,7 g
Salz 0,0 g
ProDaBi
Apfel
rwerte pro 100g
nergie 52 kcal
ett 0,2 g
davon gesättigte Fettsäuren 0,0 g
ohlenhydrate 13,8 g
davon Zucker 11,0 g
weiß 0,3 g
alz 0,0 g
ProDaBi
Brotscheibe
erte pro 100g
gie 229 kcal
4,9 g
on gesättigte tsäuren 0,8 g
enhydrate 35,0 g
on Zucker 2,3 g
ß 7,7 g
1,0 g
ProDaBi
tierte Pom
rte pro 100g
e 289 kcal
14,0 g
n gesättigte säuren 1,3 g
nhydrate 36,0 g
n Zucker 0,3 g
3,4 g
0,7 g
ProDaBi
Popcorn
e pro 100g
499 kcal
23,0 g
gesättigte ren 13,8 g
drate 57,0 g
Zucker 3,8 g
10,7 g
1,8 g
ProDaBi
chschok
pro 100g
530 kcal
29,5 g
esättigte ren 17,5 g
drate 58,5 g
ucker 57,5 g
6,6 g
0,2 g
ProDaBi

PADERBORN UNIVERSITY
The University for the Information Society

Data Split with attribute Energy



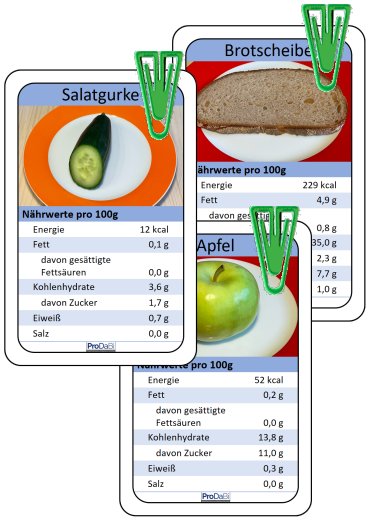
260 kcal

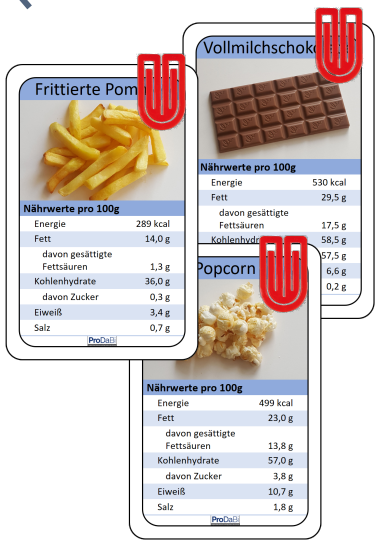


Energy

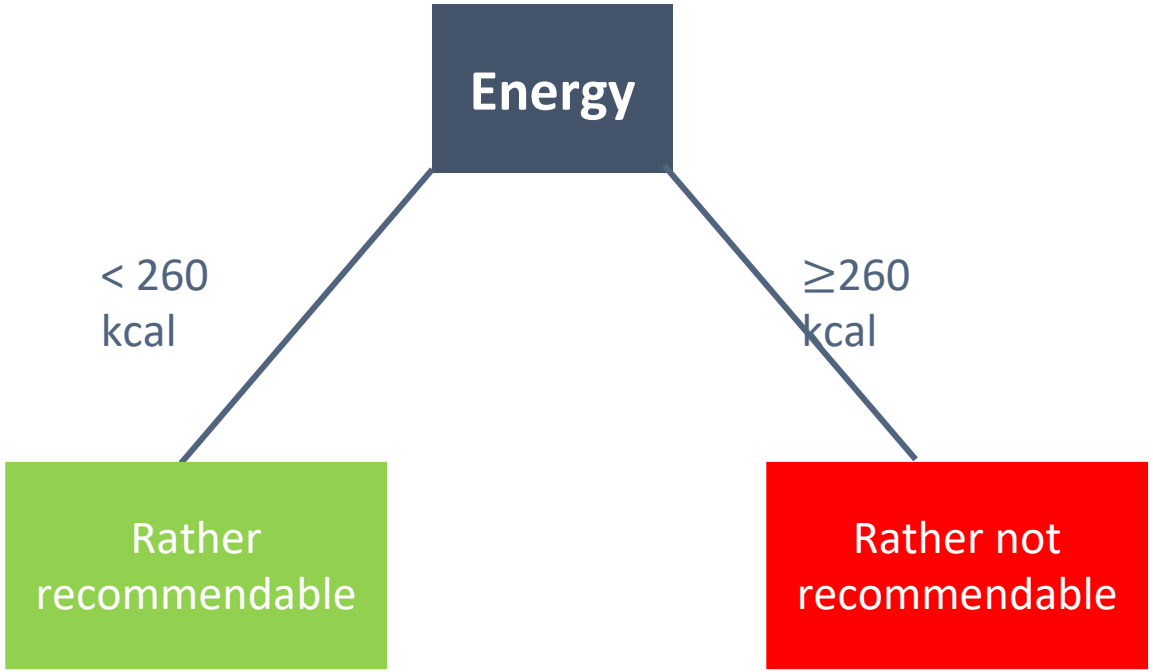
< 260 kcal

≥260 kcal



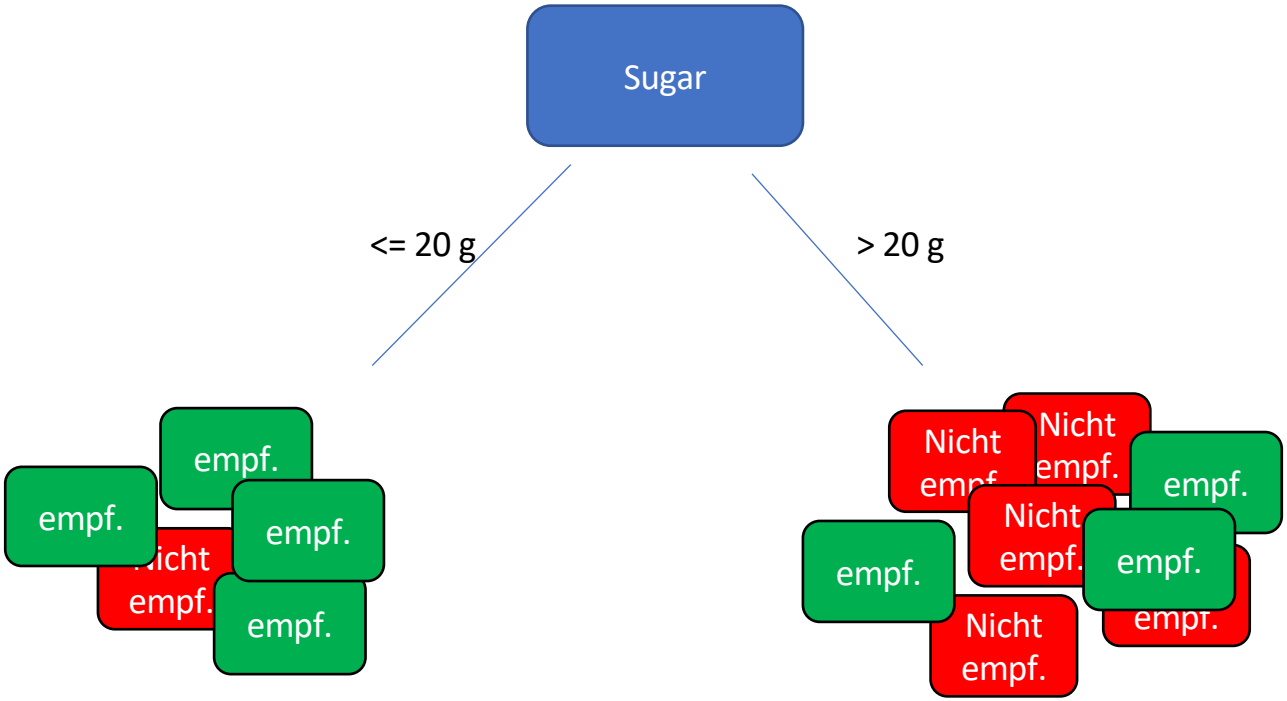


One split

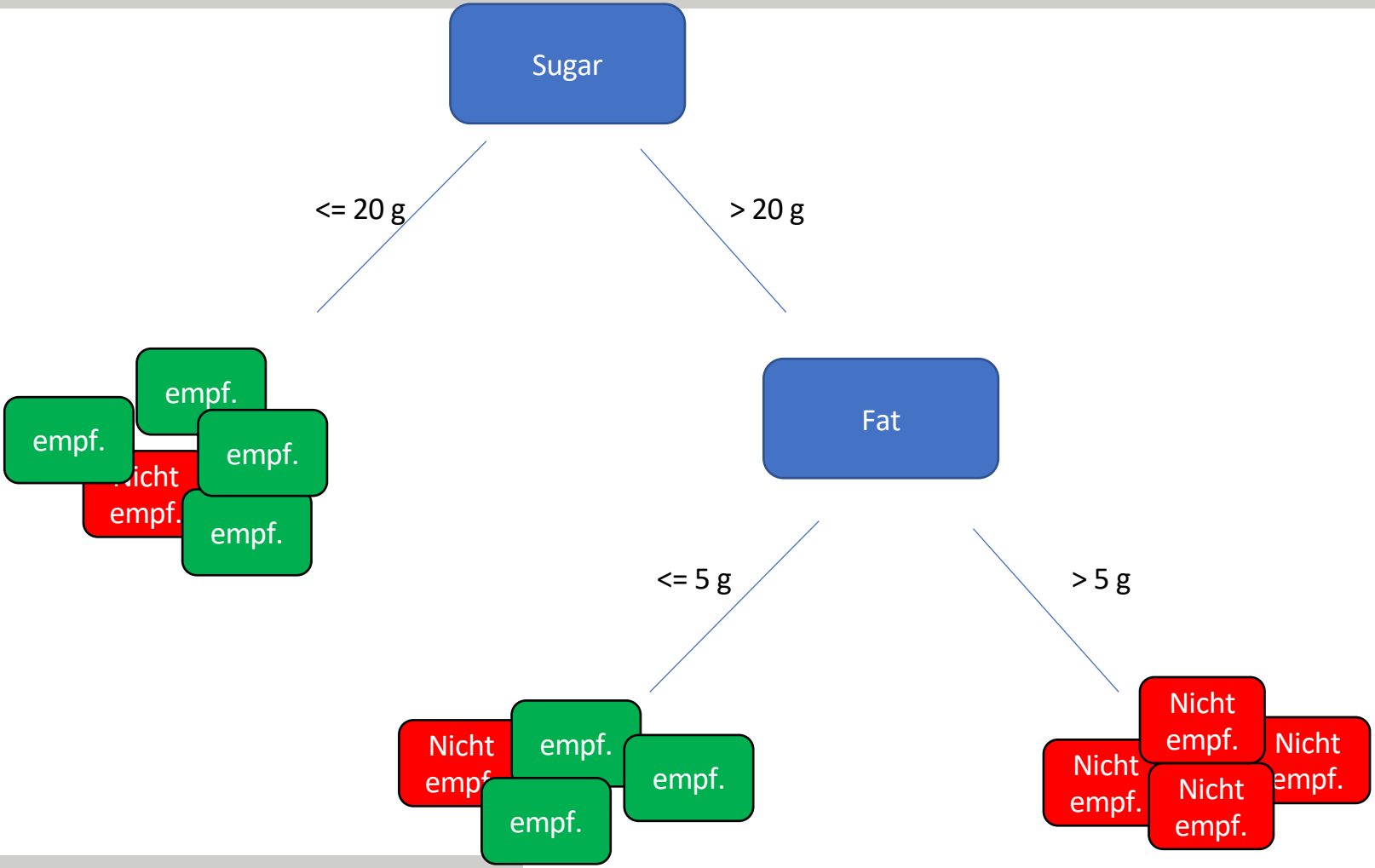


True for other food?

Better split possible?


Sugar
<= 20 g
> 20 g
empf.
empf.
empf.
empf.
empf.
Nicht empf.
Nicht empf.
empf.
Nicht empf.
empf.
empf.
Nicht empf.

A Tree


Sugar
<= 20 g
> 20 g
empf.
empf.
empf.
empf.
empf.
Fat
<= 5 g
> 5 g
empf.
empf.
empf.
Nicht empf.
Nicht empf.
Nicht empf.
Nicht empf.

1 Daten und Entscheidungsbäume - So macht das ein Computer

Erklärungstext - Von Datenkarte

Der Computer arbeitet mit Datentabel

Unten siehst Du den Apfel und das F

Die Daten der Tabelle muss ein Mens

Name	Energie	Fett	gesättigte F
Apfel	52	0.2	
Popcorn	499	23.0	

Aufgabe - Die Datentabelle ergänzer

Wähle für jedes Lebensmittel aus, ob der **"beiseite legen"**, dann wird es aus de

Am Ende siehst Du, dass die Datentabell

Apfel	eher empfeh
Banane	eher empfeh
Haselnussschnitte	eher empfeh
Chips	eher empfeh
Pommes	eher empfeh
Nudeln	eher empfeh
Erbsen	eher empfeh
Eisbergsalat	eher empfeh
Frikadelle	eher empfeh
Brotscheibe	eher empfeh
Popcorn	eher empfeh
Vanilleeis	eher empfeh
Marmorkuchen	eher empfeh
Chicken Nuggets	eher empfeh
Erdbeerjoghurt	eher empfeh
Salatgurke	eher empfeh
Brokkoli	eher empfeh
Schokomüsli	eher empfeh

6 Entscheidungsbaum automatisch erstellen lassen

Eklärungtext - Automatischer einen Entscheidungsbaum erstellen

Erstelle Baum

Super, Du hast den Computer einen Entscheidungsbaum erstellen lassen!

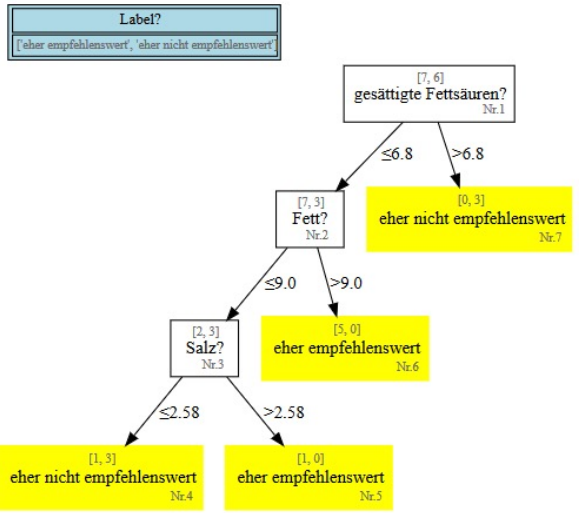
Aufgabe - Die Anzahl der Stufen verändern

Du kannst dem Computer noch sagen, wie viele Stufen der Entscheidungsbaum haben soll.

Probier mal aus, wie der Baum aussieht, wenn Du verschiedene Tiefen einstellst.

Max_Stufen 4

Erstelle Baum

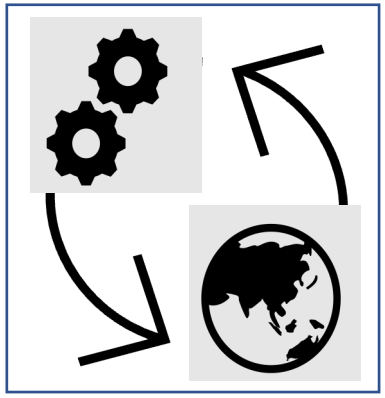



```
Korrekt Klassifikationen: 12

Fehlklassifikationen 1
```

Summary and Reflection

- Where are the digital artifacts?



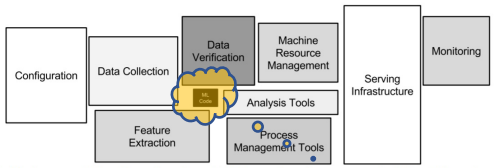
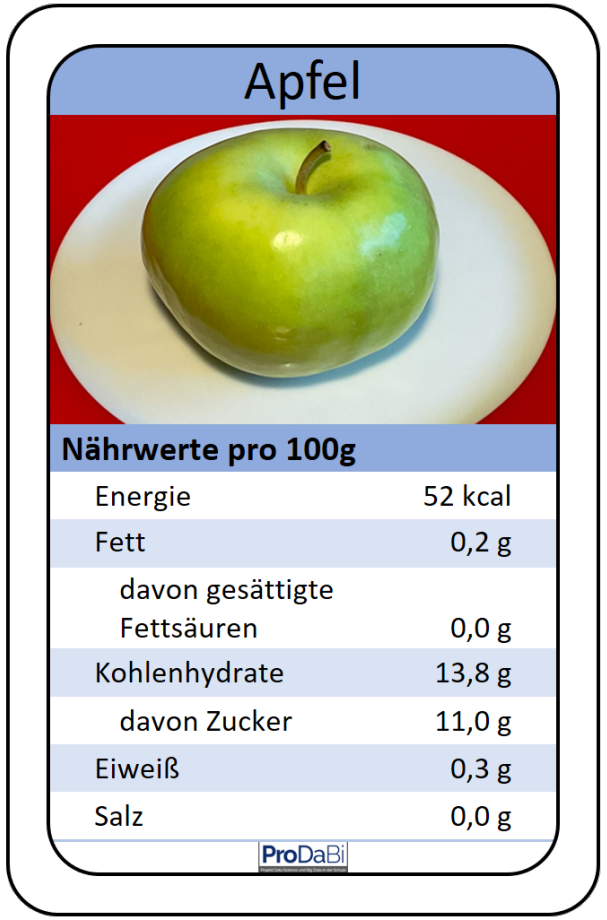


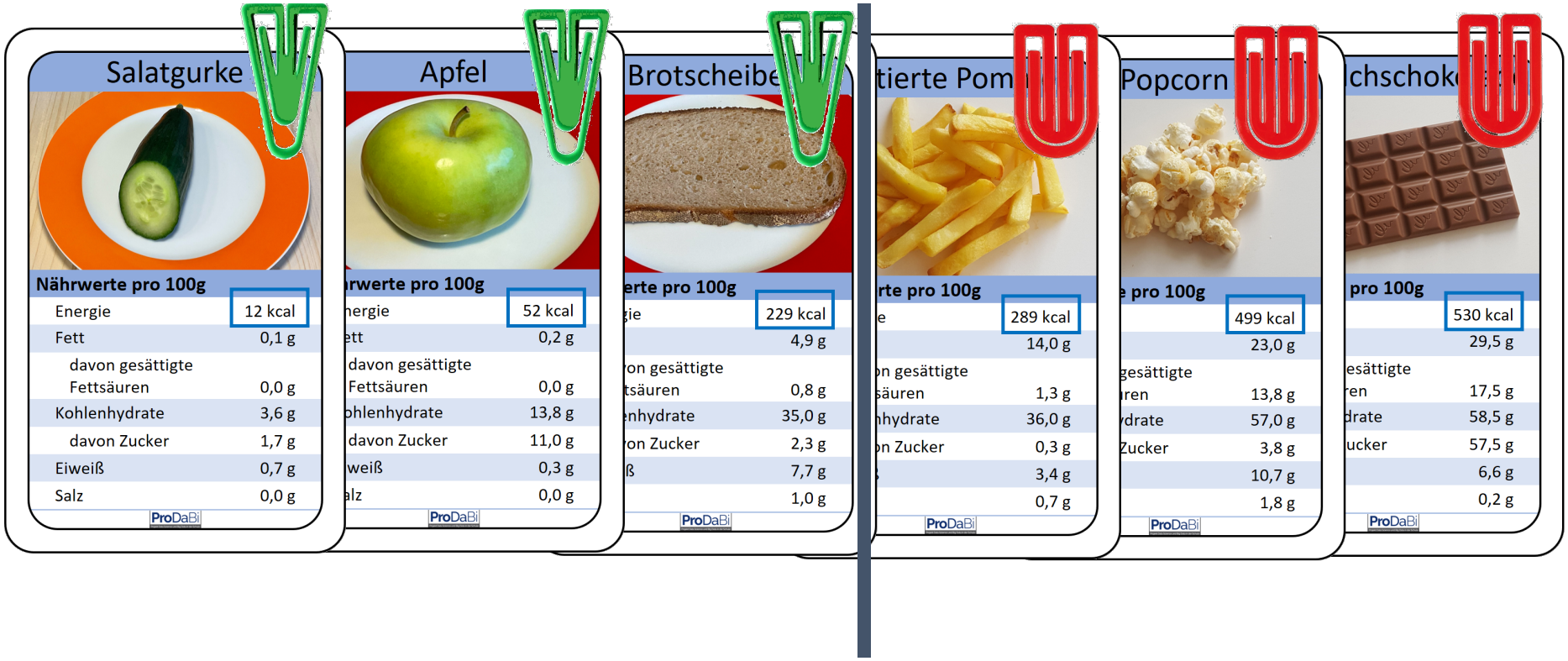
Figure 1: Only a small fraction of real-world ML systems is composed of the ML code, as shown by the small black box in the middle. The required surrounding infrastructure is vast and complex.



PADERBORN UNIVERSITY
The University for the Information Society

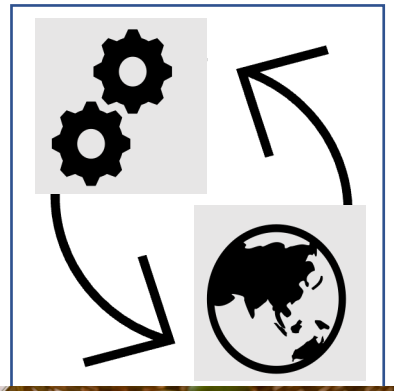
Sorting by Attribute and Attribute values is based on model decisions

260 kcal

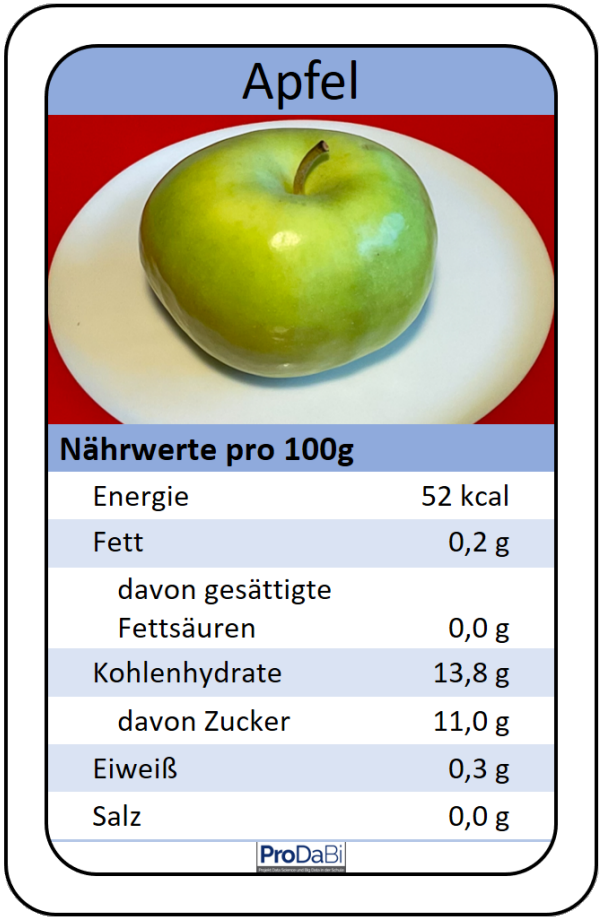


Summary and Reflection

- What is an apple?

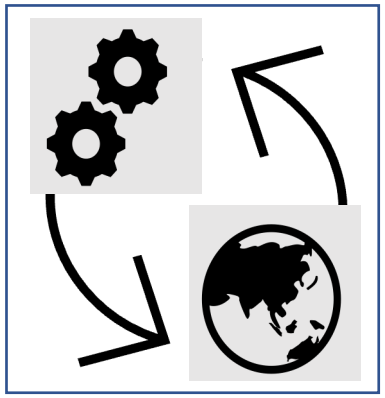


variability



Summary and Reflection

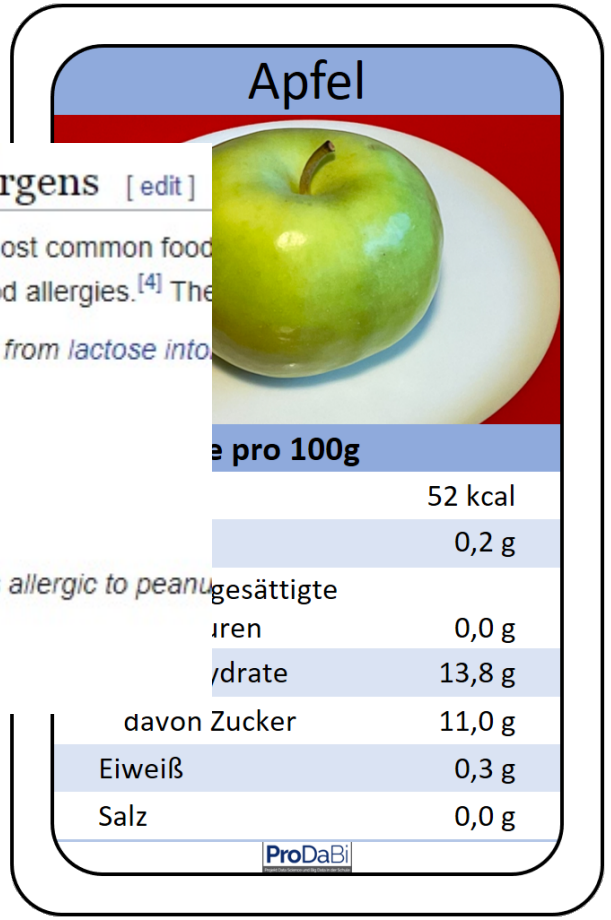
- Why these attributes?



Eight "major" food allergens [edit]

This law is in regard to the eight most common food
They account for about 90% of food allergies.[4] The

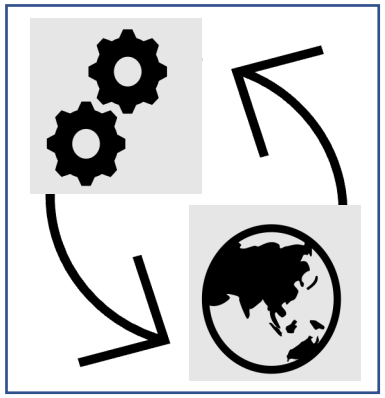
- Milk - *A milk allergy is different from lactose into*
- Eggs
- Fish
- Crustacean shellfish
- Tree nuts
- Peanuts - *Not everyone who is allergic to peanu*
- Wheat
- Soybeans



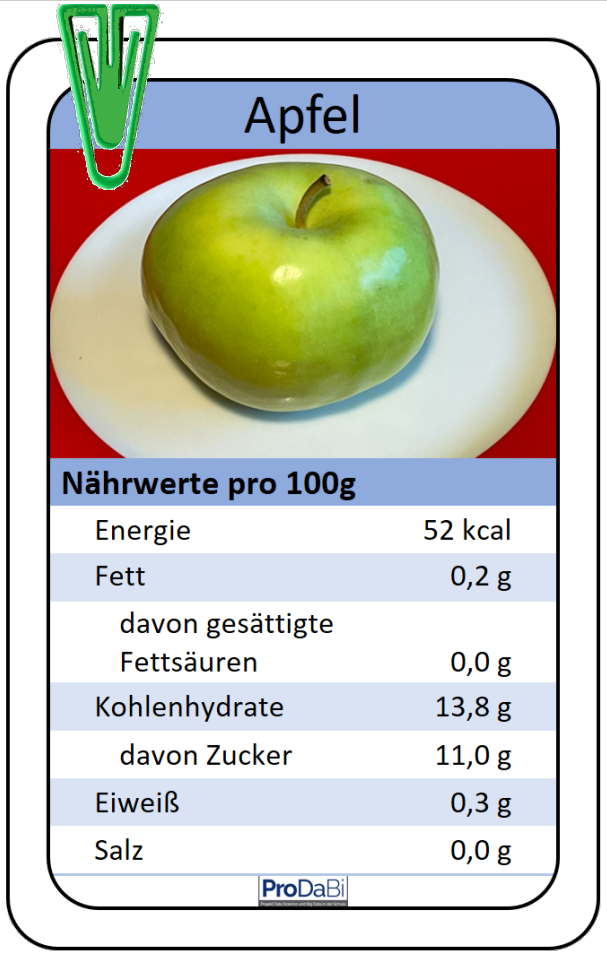
https://en.wikipedia.org/wiki/Food_Allergen_Labeling_and_Consumer_Protection_Act

Summary and Reflection

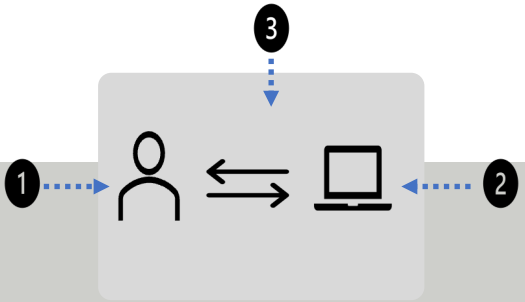
- Why these label?



- bias in label
- proxies



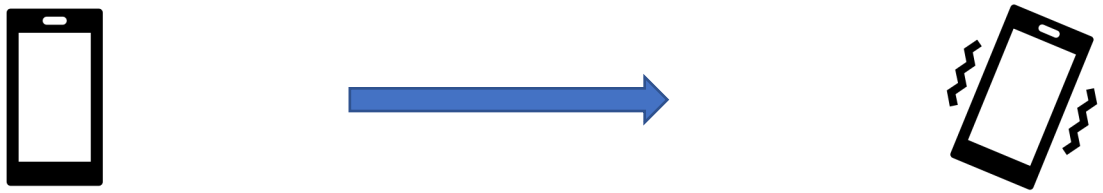
Proxies: O'NEIL, Cathy, 2017. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. 01. London: Penguin. ISBN 978-0-14-198541-1.



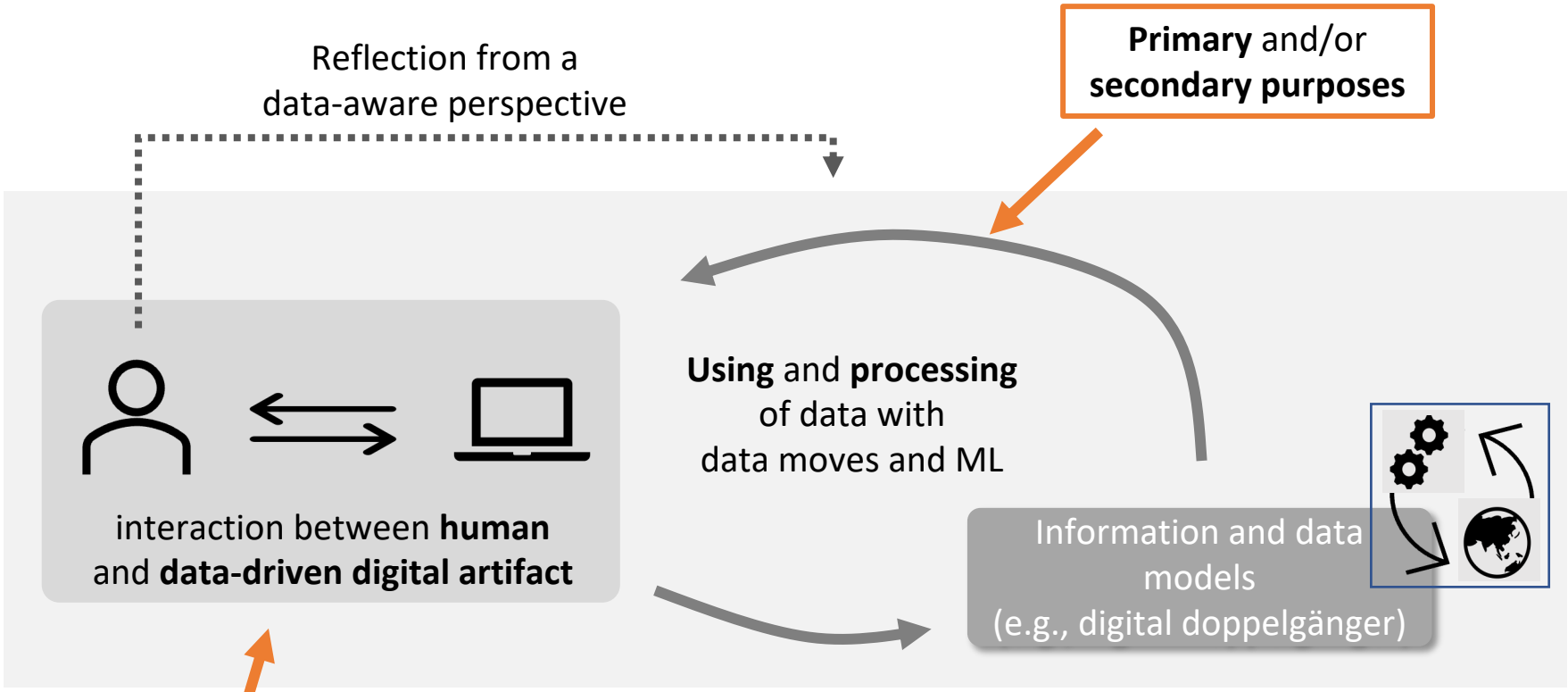
Summary: data cards

1 The role of the human: “program or be programmed”	“Trainer”: provide & model suitable data & suitable attributes / characteristics
2 The role of the artifact: replacement, augmentation, symbiosis	(Automated construction of a tree)
3 The role of the hybrid system: shaping and being shaped	Reflection on nutrition (healthy vs. advisable) ‘code is cruel’
4 The role of data:	Data quality == model quality Data as model
5 Paradigm change in teaching?	Reflect on the data model; not only on the algorithm

Example 3: Data Awareness



The concept *data awareness*



Based on ideas of [Bode & Kristensen, 2015; OECD, 2014; Pangrazio & Selwyn, 2019; Schulte & Budde, 2018; Tedre et al., 2020; Zuboff, 2019]

Using and processing the collected data

The purposes of using and processing data can essentially be divided into two areas:

1. data are used or processed to operate features of the data-driven digital artifact (**primary purpose**)
2. to investigate developments of the data-driven digital artifact or to achieve additional purposes (**secondary purpose**).

Construction of a digital doppelgänger

- Construction of a digital doppelgänger [Bode & Kristensen, 2015] (or digital self [Tedre et al., 2020]) which...
  - is constructed by processing the collected data – especially implicitly collected data
  - relates to a real person
  - represents the person within specific and limited proxies – no full representation
  - should be understand within digital doppelgänger of other users
  - can be used for different primary and secondary purposes

Teaching unit for grade 5 to 6

1st phase: experiment to introduce to the interaction system

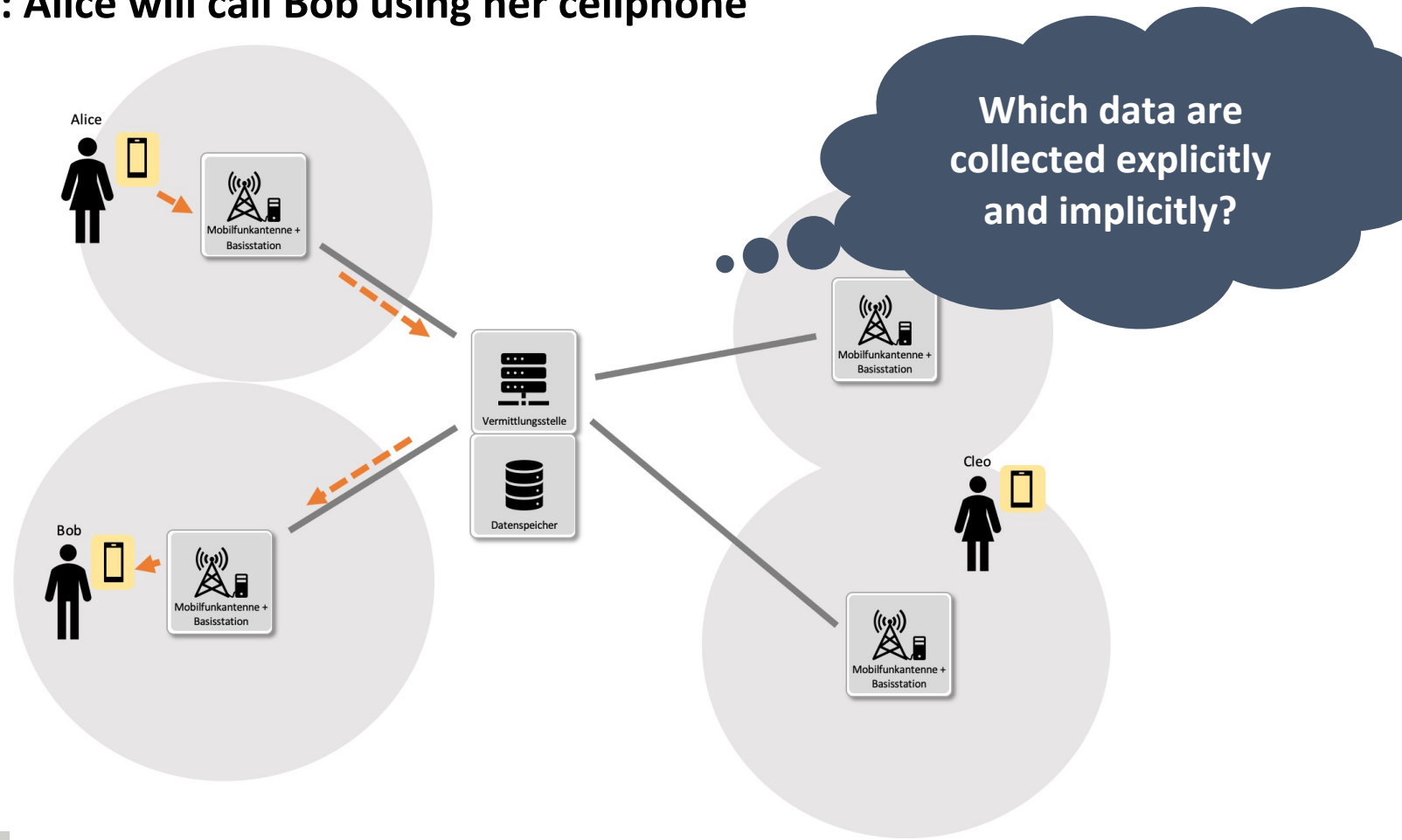
2nd phase: architecture of the mobile network (explicitly and implicitly collected data and primary purpose of using and processing these data)

3rd phase: exploring location data and constructing a digital doppelgänger (secondary purpose)

4th phase: reflection and discussion of the data-driven digital artifact and the interactionsystem

Phase 2: architecture of the mobile network

Situation: Alice will call Bob using her cellphone



Phase 2: collected data

Standortdaten einer Person aus dem Mobilfunknetz

	Beginn	Ende	Dienst	ein/ausgehend	Standort	
0	8/31/09 7:57	8/31/09 8:09	GPRS	ausgehend	13.39611111	52.52944444
1	8/31/09 8:09	8/31/09 8:09	GPRS	ausgehend	13.38361111	52.53
2	8/31/09 8:09	8/31/09 8:15	GPRS	ausgehend	13.37472222	52.53027778
3	8/31/09 8:15	8/31/09 8:39	GPRS	ausgehend	13.37472222	52.53027778
4	8/31/09 8:39	8/31/09 9:09	GPRS	ausgehend	13.37472222	52.53027778
5	8/31/09 9:09	8/31/09 9:39	GPRS	ausgehend	13.37472222	52.53027778
6	8/31/09 9:12	8/31/09 9:12	Telefonie	ausgehend	13.37472222	52.53027778
7	8/31/09 9:39	8/31/09 10:09	GPRS	ausgehend	13.37472222	52.53027778
8	8/31/09 10:09	8/31/09 10:39	GPRS	ausgehend	13.37472222	52.53027778
9	8/31/09 10:39	8/31/09 10:54	GPRS	ausgehend	13.37472222	52.53027778
10	8/31/09 10:55	8/31/09 11:25	GPRS	ausgehend	13.37472222	52.53027778
11	8/31/09 11:25	8/31/09 11:55	GPRS	ausgehend	13.37472222	52.53027778
12	8/31/09 11:55	8/31/09 12:25	GPRS	ausgehend	13.37472222	52.53027778
13	8/31/09 12:25	8/31/09 12:55	GPRS	ausgehend	13.37472222	52.53027778
14	8/31/09 12:30	8/31/09 12:31	Telefonie, CFNRy	eingehend	13.37472222	52.53027778
15	8/31/09 12:55	8/31/09 13:02	GPRS	ausgehend	13.37472222	52.53027778
16	8/31/09 13:02	8/31/09 13:02	GPRS	ausgehend	13.37583333	52.48972222
17	8/31/09 13:02	8/31/09 13:14	GPRS	ausgehend	13.37583333	52.48972222

Phase 3: Exploration of location data

Filter the location data and explore them on the map

Locations in the time range:

from: From hour

to: to hour

Locations on the weekday:

Tagwahl: Choose a weekday

Locations in the month:

Month choi… Choose a month

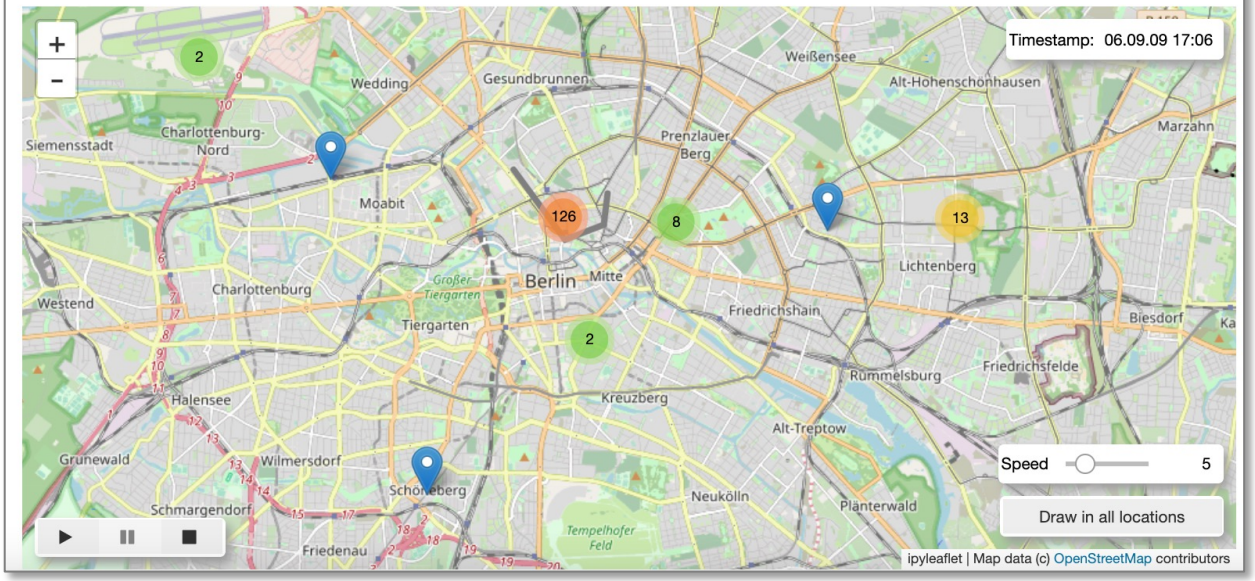
Reset or apply the filters:

Apply filters and update the map

Folgende Daten sind geladen:

	Begin	End	Service	in/outgoing	Longitude	Latitude
0	31.08.09 07:57	31.08.09 08:09	GPRS	ausgehend	13.396111	52.529444
1	31.08.09 08:09	31.08.09 08:09	GPRS	ausgehend	13.383611	52.530000
2	31.08.09 08:09	31.08.09 08:15	GPRS	ausgehend	13.374722	52.530278
...	...	...	...	...	...	...
21503	13.02.10 10:00	13.02.10 10:10	GPRS	ausgehend	13.404444	52.531667
21504	13.02.10 10:10	13.02.10 10:10	GPRS	ausgehend	13.404444	52.531667
21505	13.02.10 10:10	13.02.10 10:13	GPRS	ausgehend	13.404444	52.531667

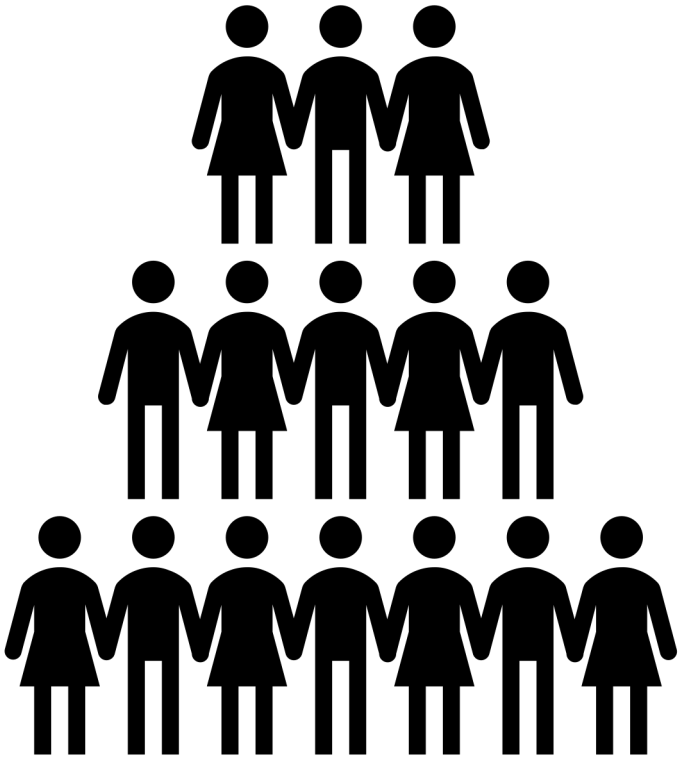
21506 rows × 6 columns



Application: go.upb.de/ExplorationLocationData

Phase 3: Digital Doppelgänger / Data Poster

- Male, 40 years old
- No kids
- Loves ice cream
- Works as politician, priest, trucker, lawyer…



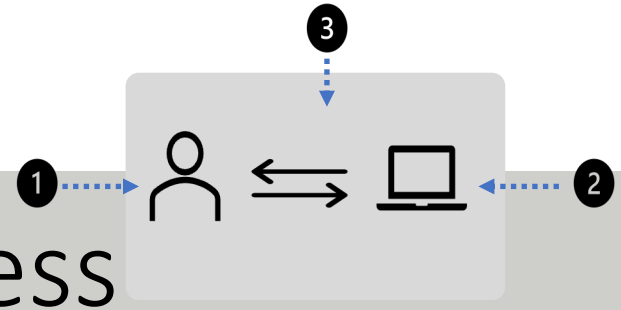
Phase 4: reflection and discussion

Within an interaction with a data-driven digital artifact, students should be aware of and reflect on

- the explicitly and implicitly collection of personal data.
- the using and processing of the personal data for primary and secondary purposes.
  - The possibility of a digital doppelgänger

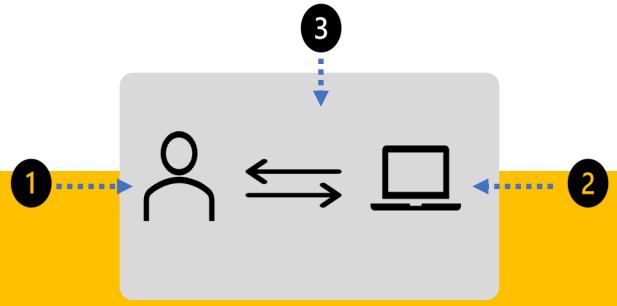
Transformation of self-view, world-view and habits (*Bildung*)

Bildung: SCHULTE, Carsten and BUDDE, Lea, 2018. A Framework for Computing Education: Hybrid Interaction System: The need for a bigger picture in computing education. In: *18th Koli Calling International Conference on Computing Education Research (Koli Calling'18)*. Koli, Finland: ACM. 22 November 2018



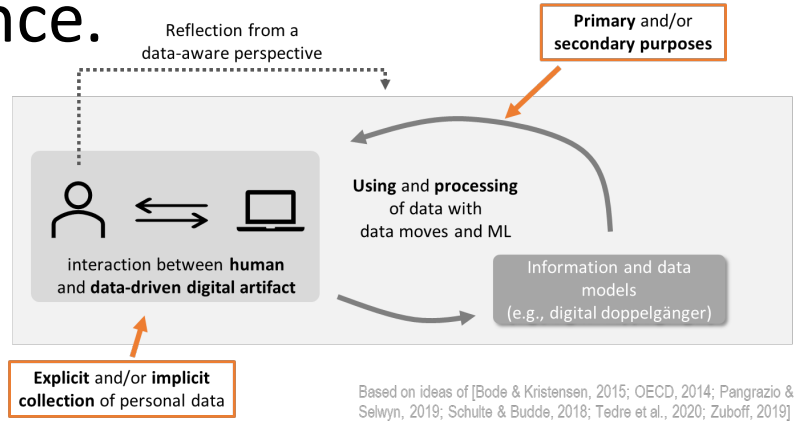
Summary: Data Awareness

1 The role of the human: “program or be programmed”	Personal data is collected and generated explicitly and implicitly during interaction
2 The role of the artifact: replacement, augmentation, symbiosis	Primary and secondary data processing
3 The role of the hybrid system: shaping and being shaped	Stop-moment, data awareness
4 The role of data:	(Implicit, explicit; primary, secondary) Digital Doppelgänger
5 Paradigm change in teaching?	Reflection on Architecture and Relevance of data in interaction with a dA



Summary / Questions

"AI education" requires developing an adequate picture of the hybrid interaction system – a kind of data-driven, emergent eco-system which needs to be made explicit to understand the transformative role as well as the technological basics of these artificial intelligence tools and how they are related to Data Science.



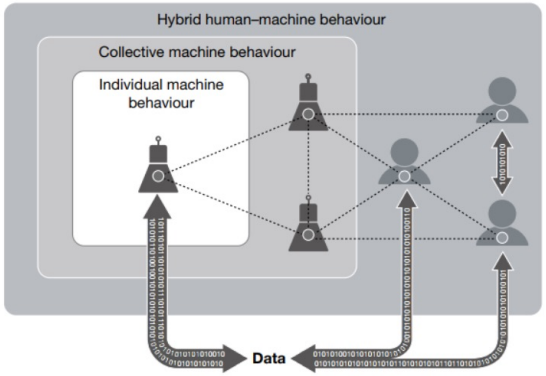


Fig. 4 | Scale of inquiry in the machine behaviour ecosystem. AI systems represent the amalgamation of humans, data and algorithms. Each of these domains influences the other in both well-understood and unknown ways. Data—filtered through algorithms created by humans—influences

PADERBORN UNIVERSITY
The University for the Information Society

Thanks for listening!

- Bode, M., & Kristensen, D. (2015). The digital doppelgänger within. A study on self-tracking and the quantified self-movement. In R. Canniford & D. Bajde (Eds.), Assembling Consumption. Researching actors, networks and markets (pp. 119–134). Routledge. https://doi.org/10.4324/9781315743608
- OECD. (2014). Summary of the OECD Privacy Expert Roundtable „Protecting Privacy in a Data-driven Economy: Taking Stock of Current Thinking". DSTI/ICCP/- REG(2014)3.
- Pangrazio, L., & Selwyn, N. (2019). 'Personal data literacies': A critical literacies approach to enhancing understandings of personal digital data. New Media & Society, 21(2), 419–437. https://doi.org/10.1177/1461444818799523
- Tedre, M., Vartiainen, H., Kahila, J., Toivonen, T., Jormanainen, I., & Valtonen, T. (2020). Machine Learning Introduces New Perspectives to Data Agency in K—12 Computing Education. 2020 IEEE Frontiers in Education Conference (FIE), 1–8.
- Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power (First edition). PublicAffairs.

Questions for Discussion

Q1: How should we design the relationship between “ideas and artefacts”?

Q2: In what aspects and how much do we need to reinforce the role of data?

Q3: (Where) Do we have to change typical teaching pattern?