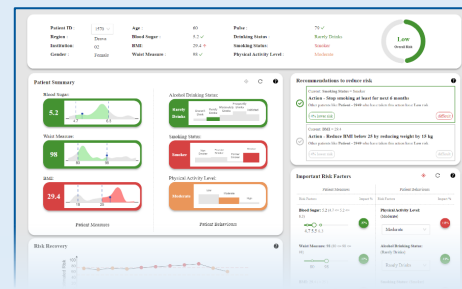
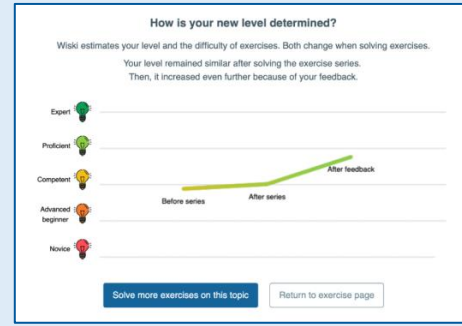
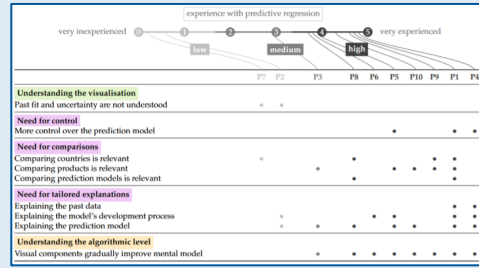
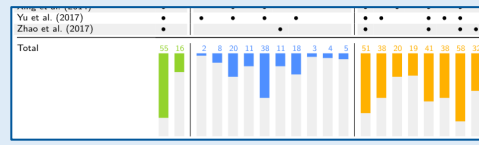
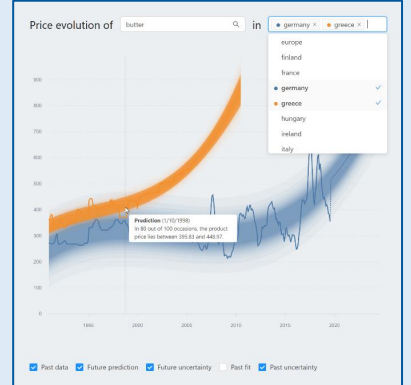
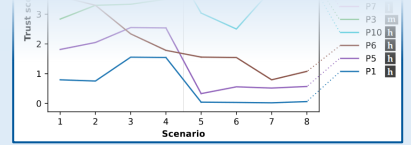
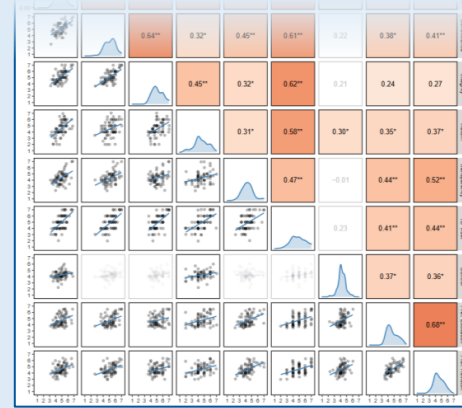


Interactive Visualisations for Explainable AI Research and Research Practices



Jeroen Ooge
jeroenooge.be



Maak een aangeraden oefening van hetzelfde hoofdstuk

Aangeraden

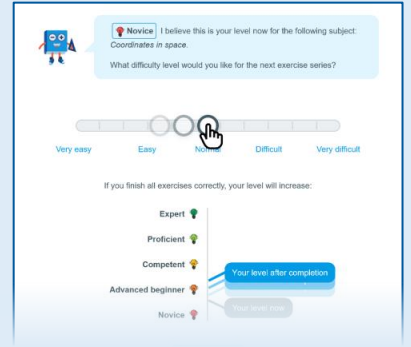
- Defening 37
- Defening 26
- Defening 21

Aantal pogingen medeleeringen nodig hadden om oefening 21 juist op te lossen

Maak oefening 21

... of kies zelf je volgende oefening

Naar het oefeningenoverzicht



About Me



Jeroen Ooge
jeroenooge.be

“yeh-roon ow-geh” 😂

About Me

I am VIS



Jeroen Ooge
jeroenooge.be

About Me

*Human-centred explainable AI
with **visualisations***

My short life

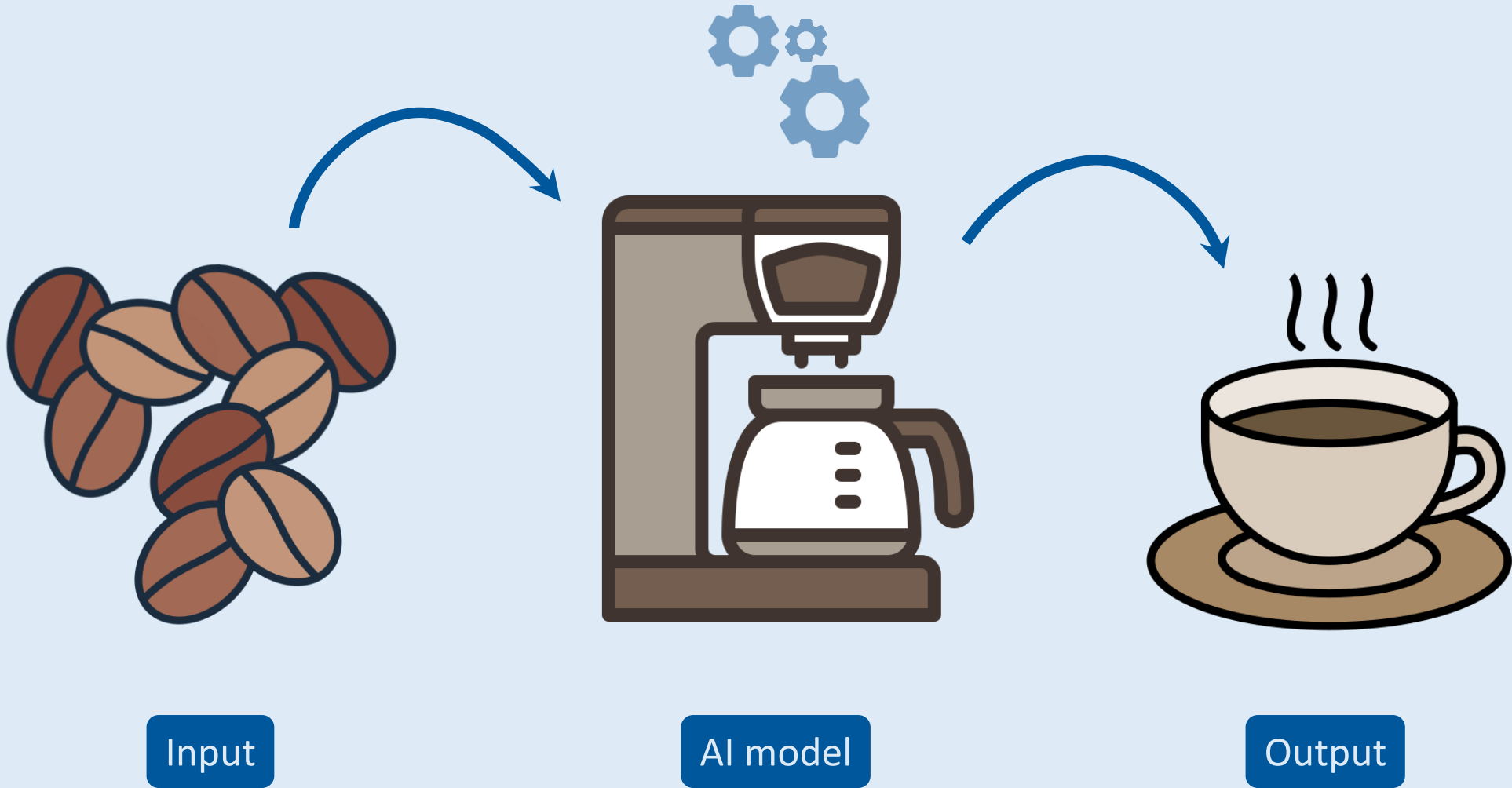
MSc fundamental mathematics
MSc applied informatics
2012-2019

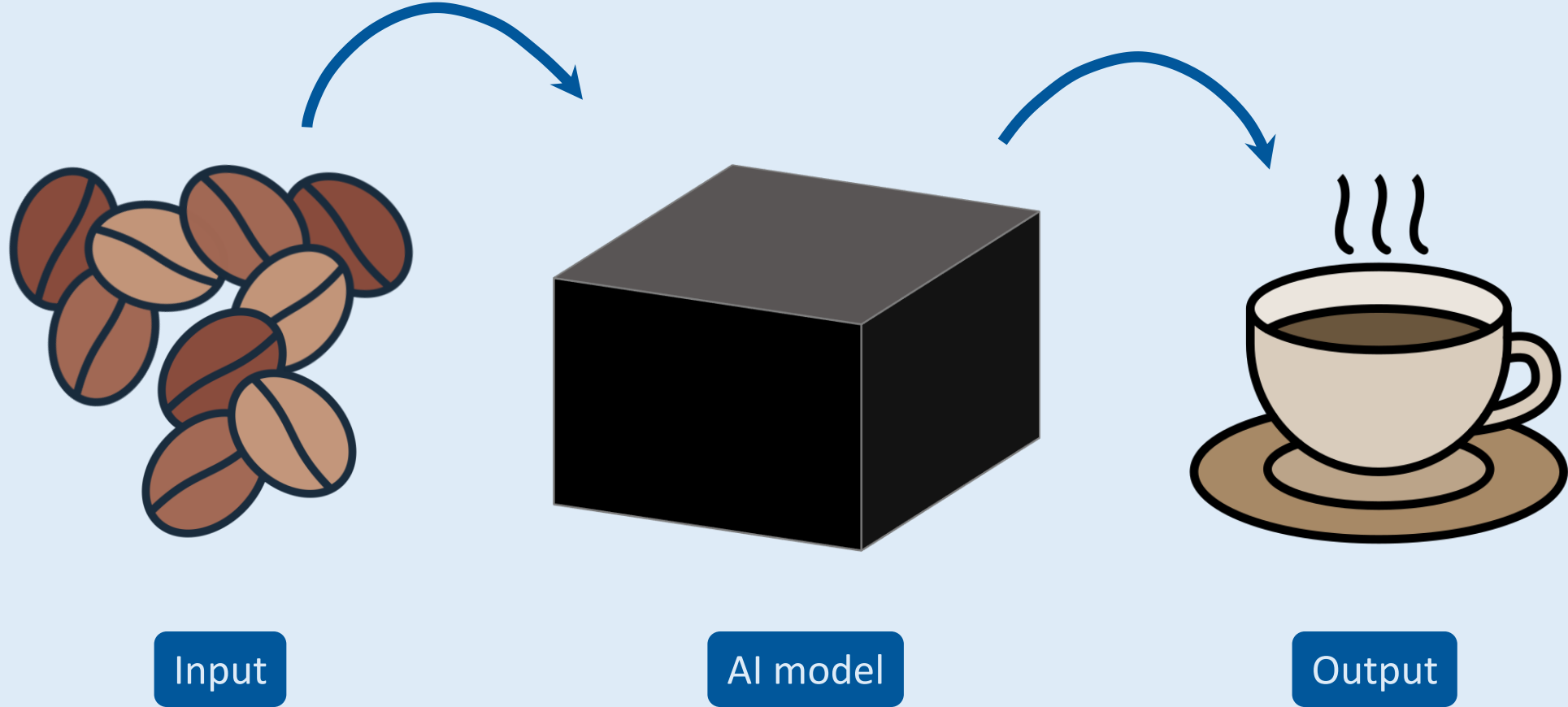
PhD researcher
at KU Leuven
2019-2023

Assistant professor
at Utrecht University
2023-now

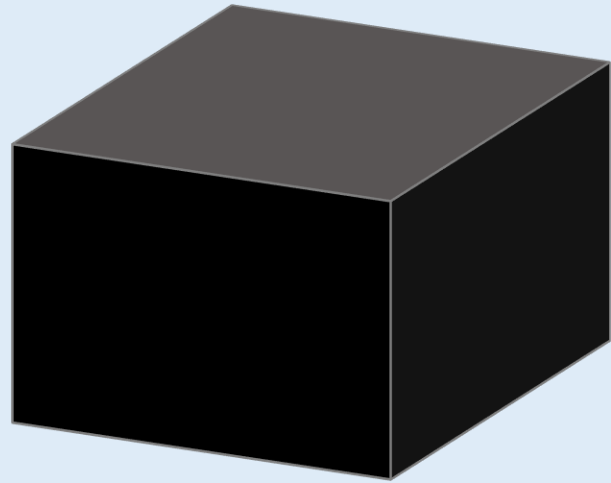


Jeroen Ooge
jeroenooge.be

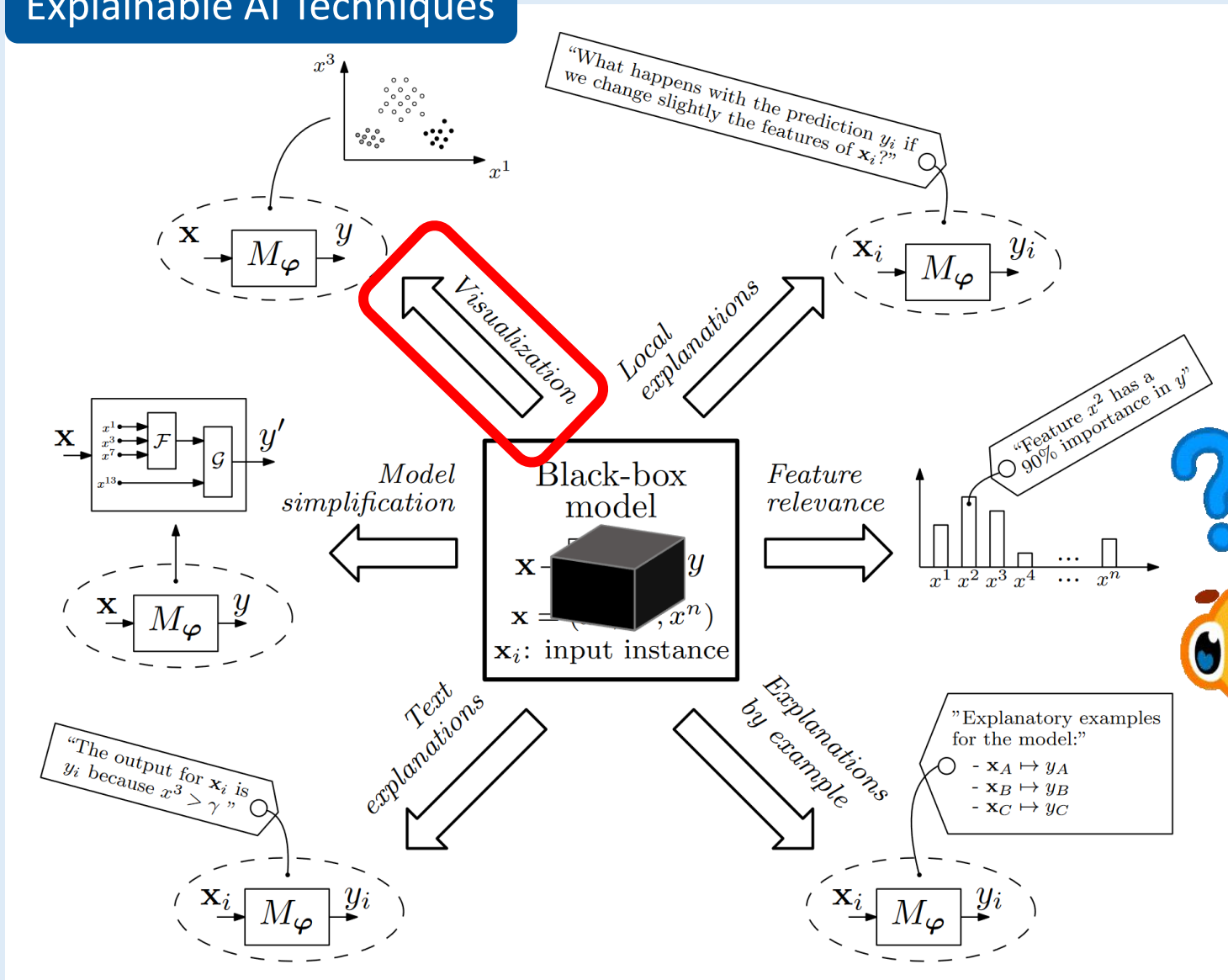




Can we “open”
this black box?



Explainable AI Techniques



Which explanation type do we use?

Who is the audience?

What is the context?





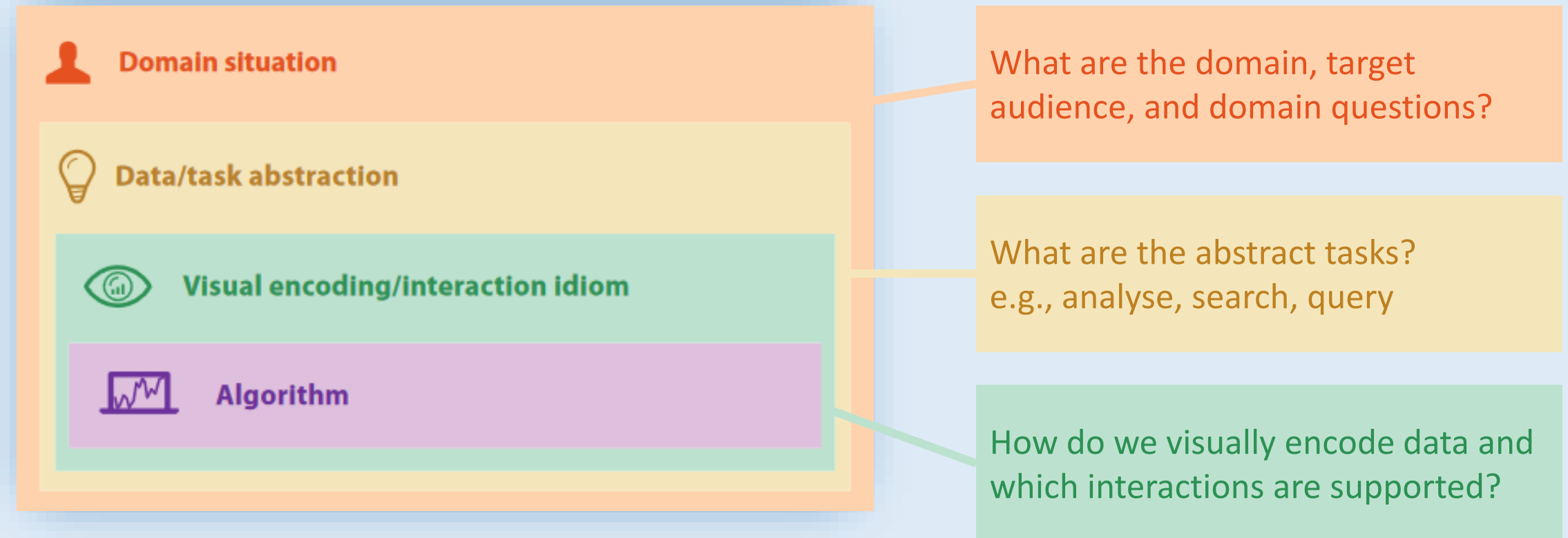
Human-Centred Explainable AI

Different people and contexts
need different explainability solutions

My goal: design visual explanations and
evaluate how they affect people's behaviour
(e.g., trust, understanding, motivation)

- I Complex Visual Dashboards**
- II Simple Visualisations**
- III Visualisations in Reports**

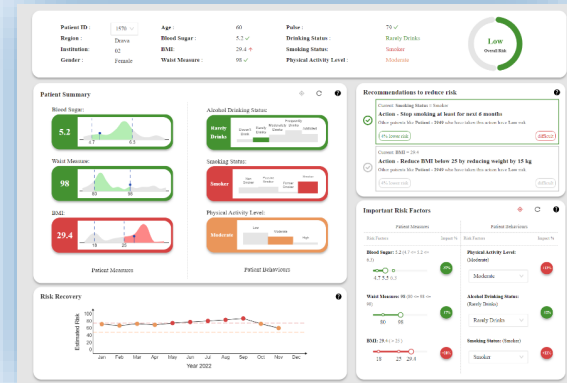
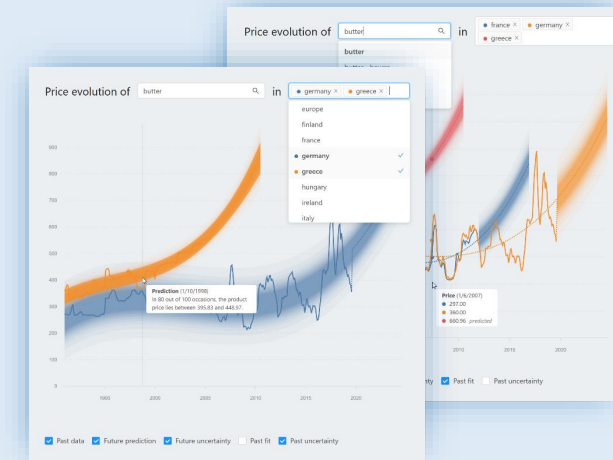
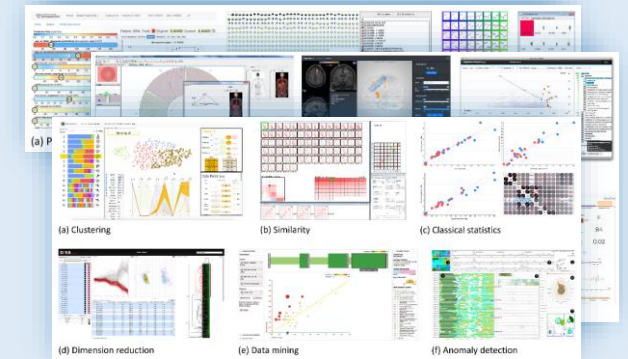
Quick Recap of the Nested Model



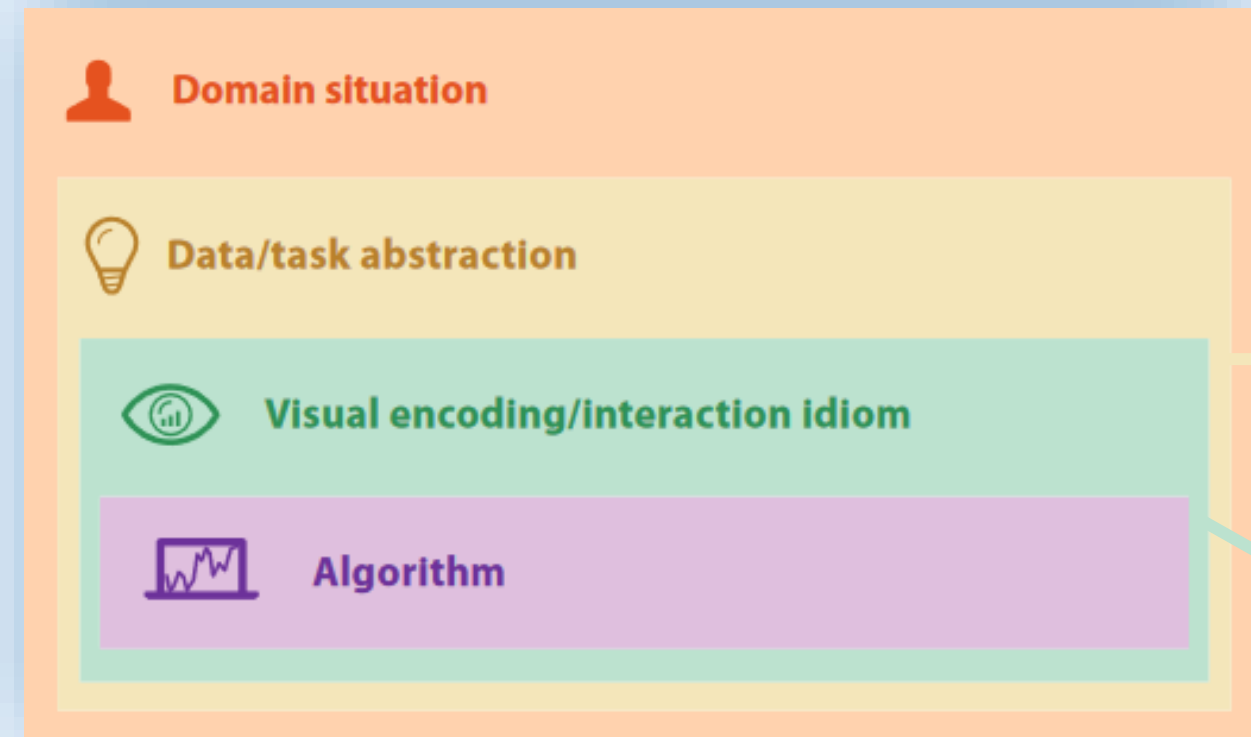
I Complex Visual Dashboards

II Simple Visualisations

III Visualisations in Reports



#1: Visual Analytics in Healthcare



Domain: healthcare

Audience: AI experts, domain experts, AI novices

Tasks: monitor, interpret, predict

Encoding: --

Interaction: abstract/elaborate, connect, encode, explore, filter, reconfigure, select, shepherd

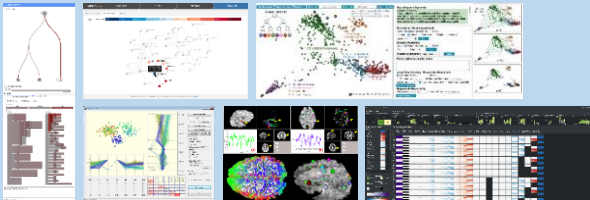
visualisation



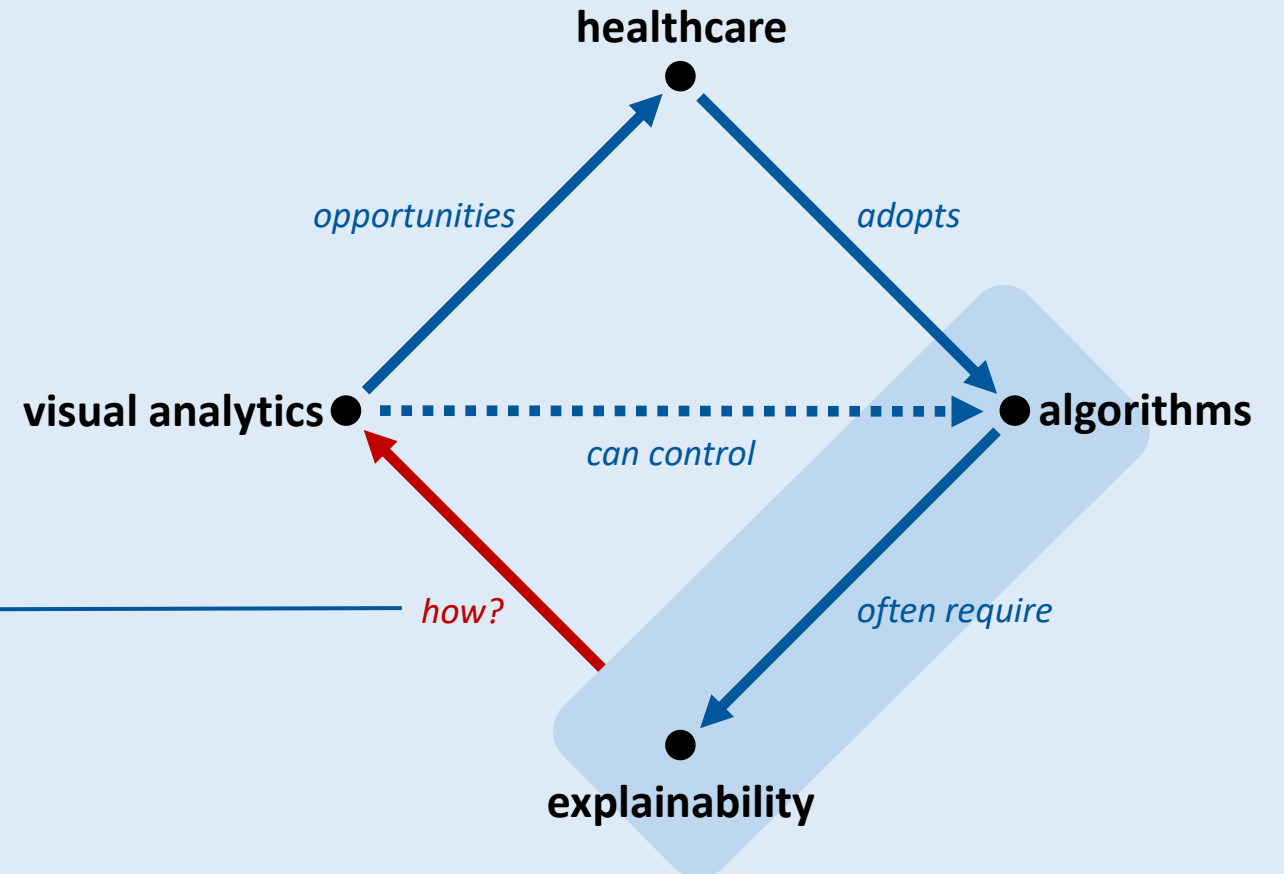
interaction



control



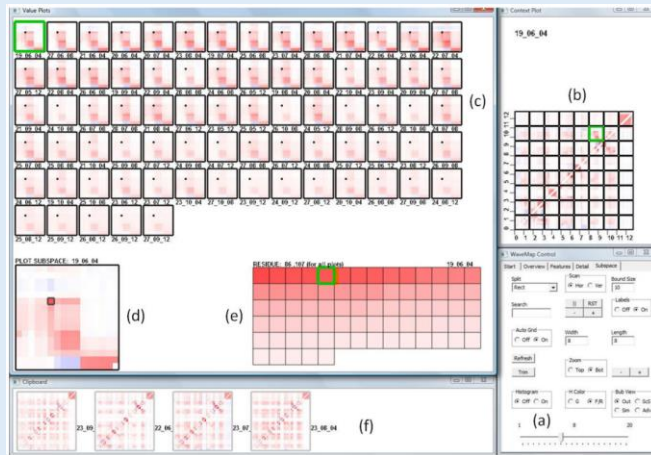
direct explanation



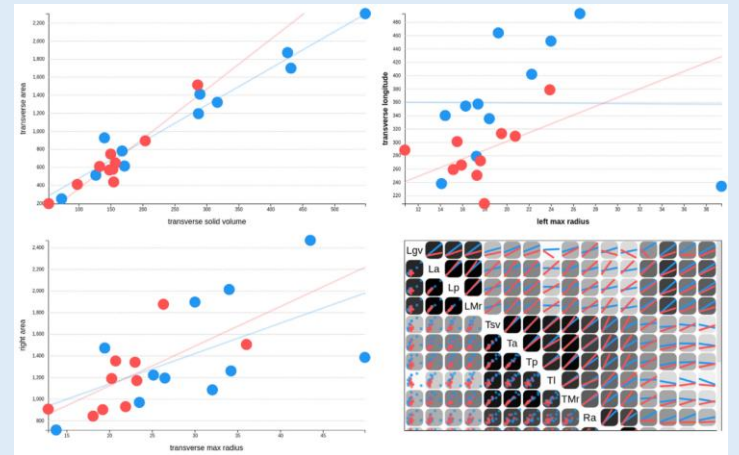
Visualisation



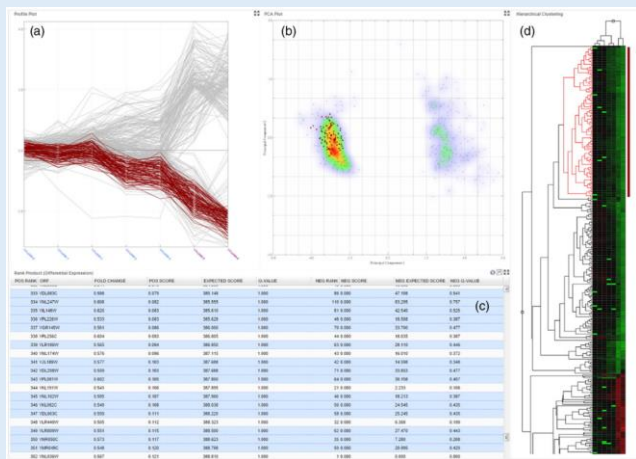
Clustering



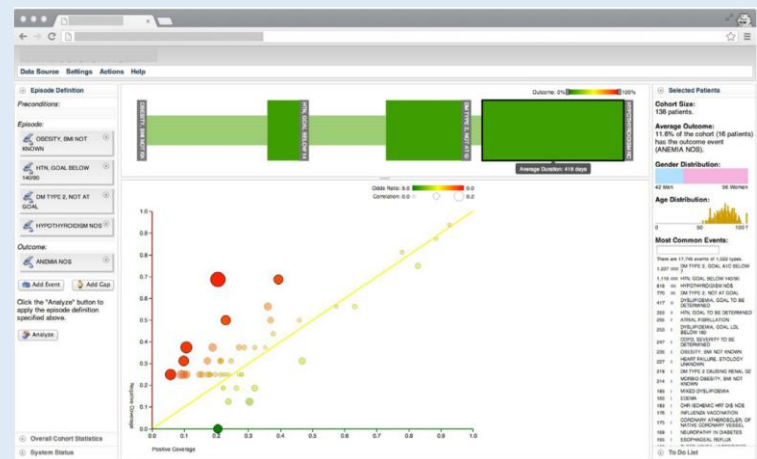
Similarity



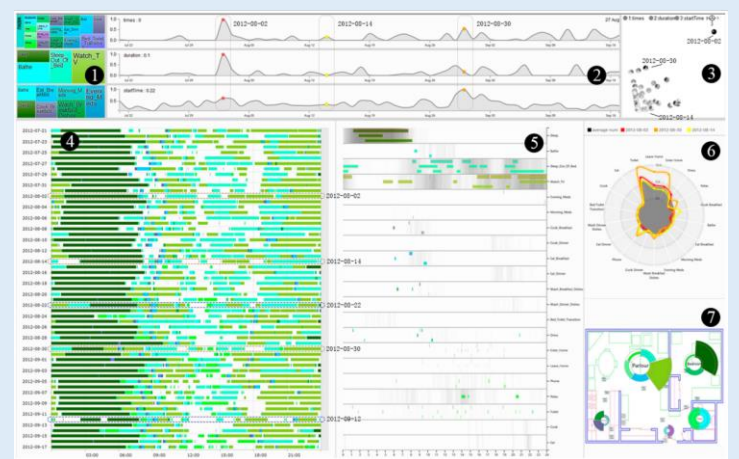
Classical statistics



Dimension reduction

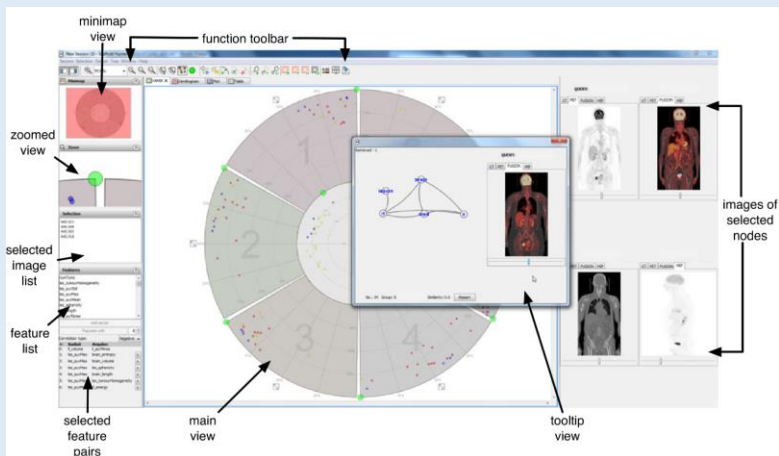


Data mining

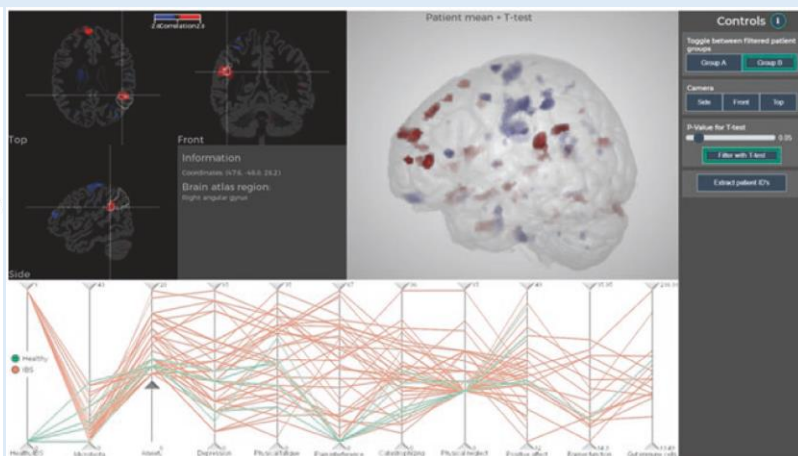


Anomaly detection

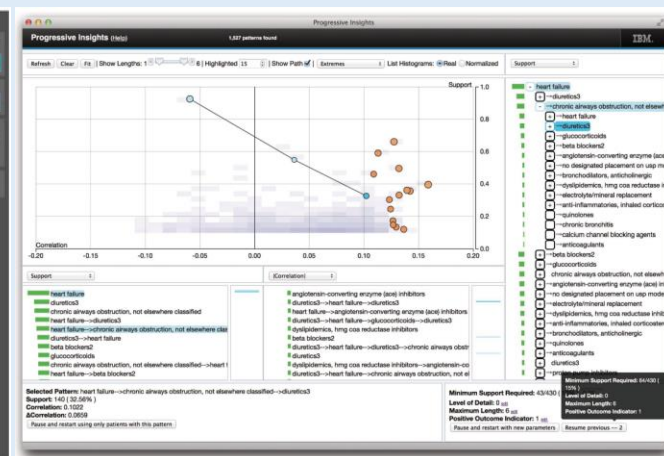
Interaction



Abstract/elaborate



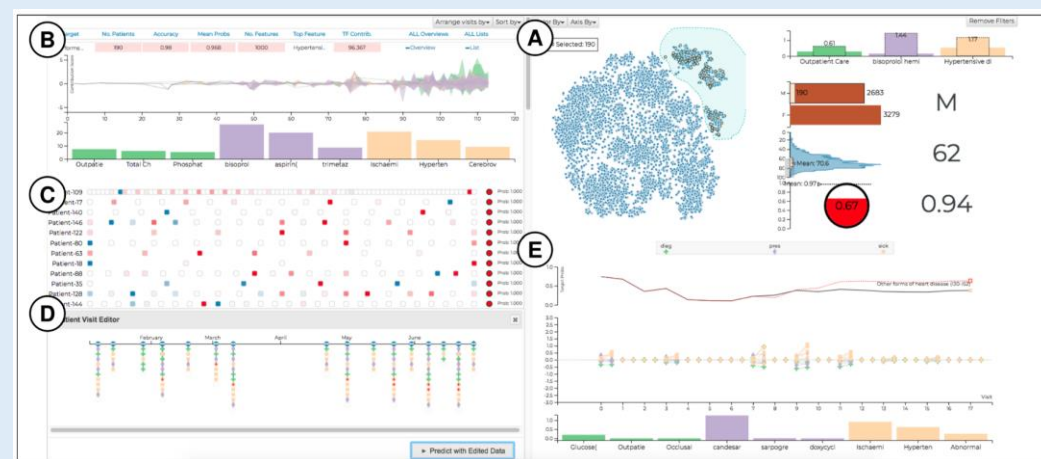
Filter



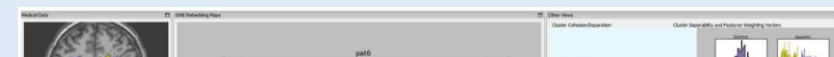
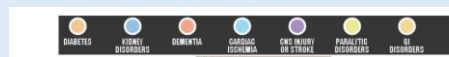
Reconfigure



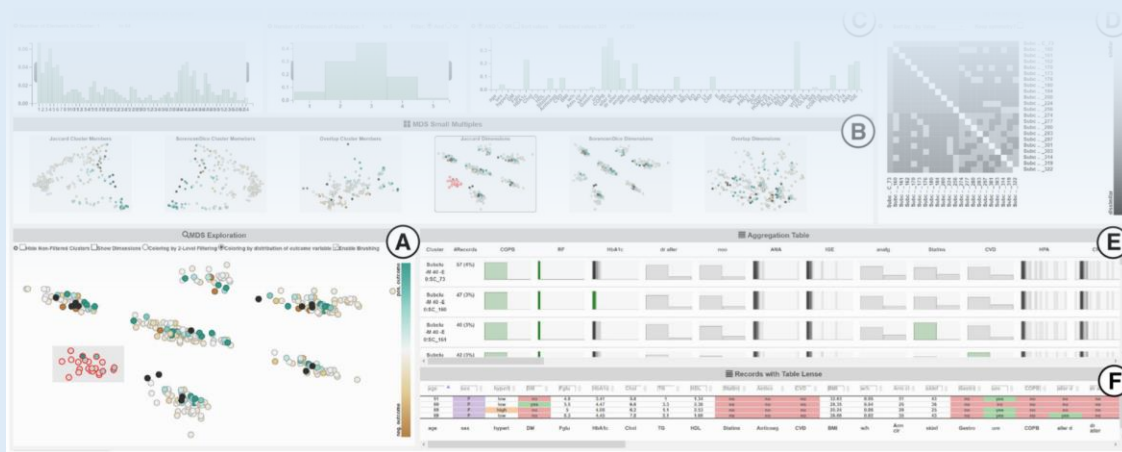
Encode



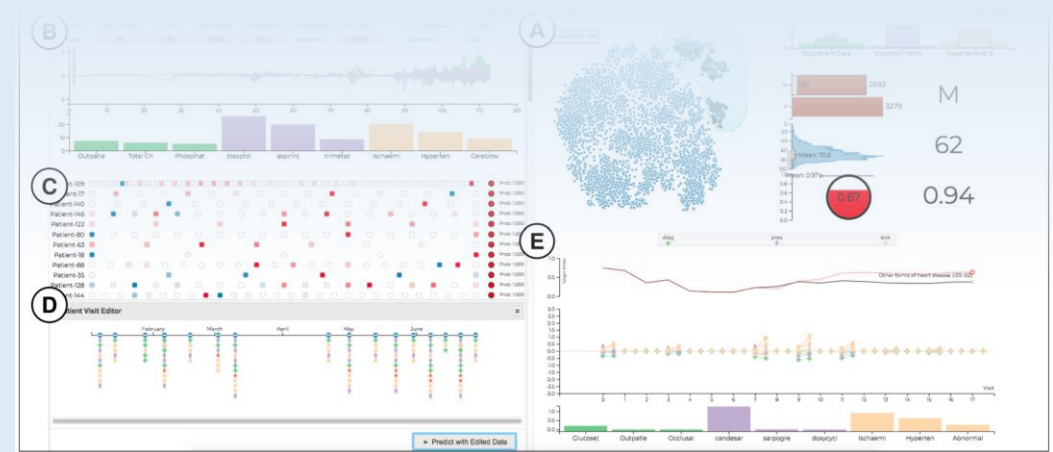
Select



Interaction



Encode



Select

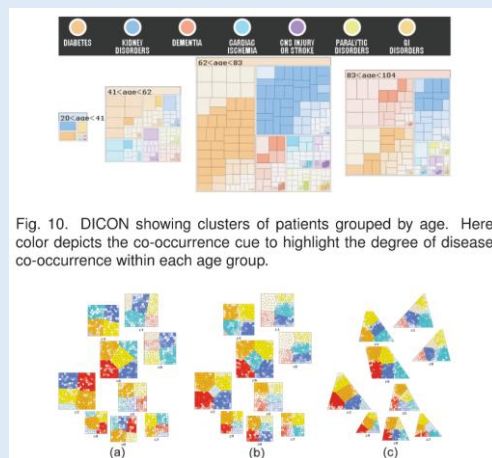
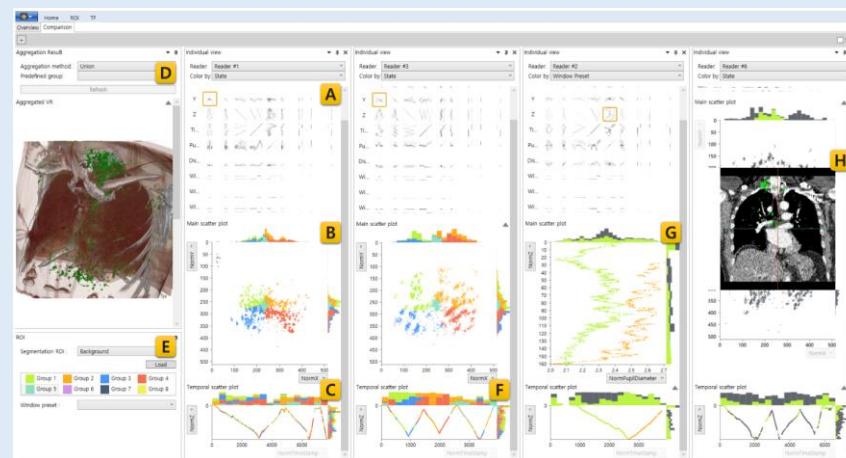
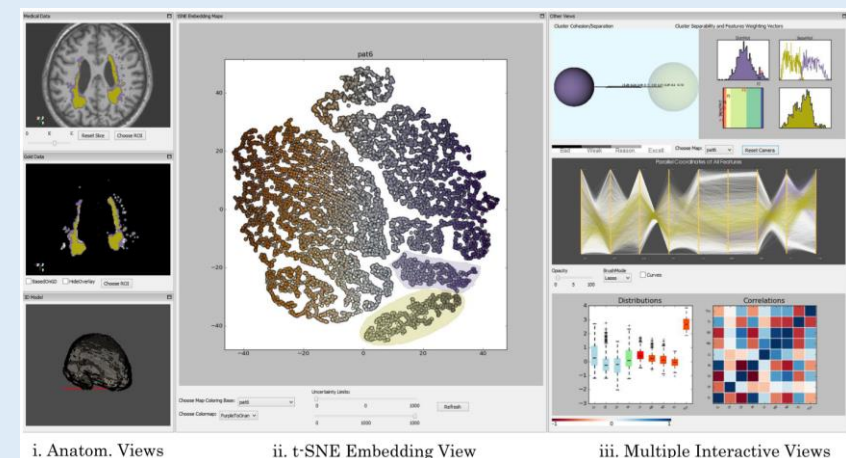


Fig. 10. DICON showing clusters of patients grouped by age. Here color depicts the co-occurrence cue to highlight the degree of disease co-occurrence within each age group.

Connect

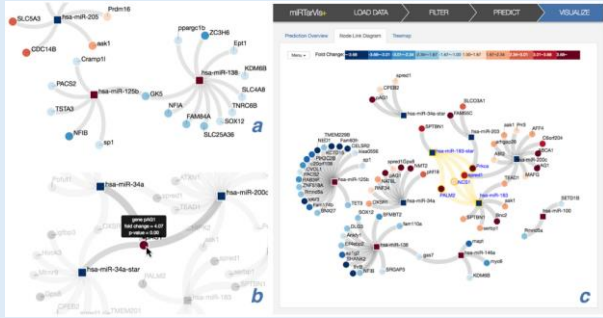


Explore

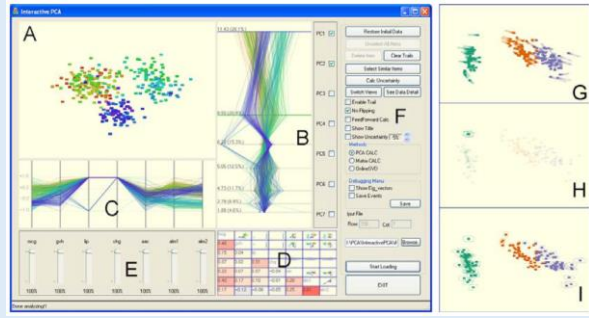


Select + connect

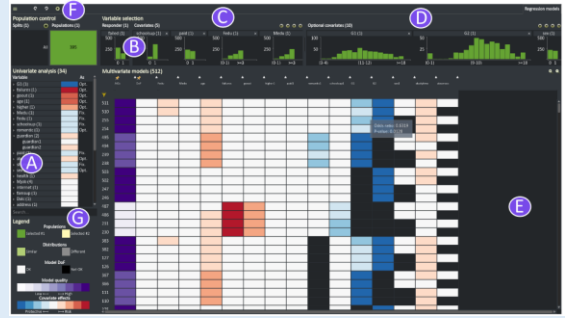
Control



Configuration window separate from visualisation

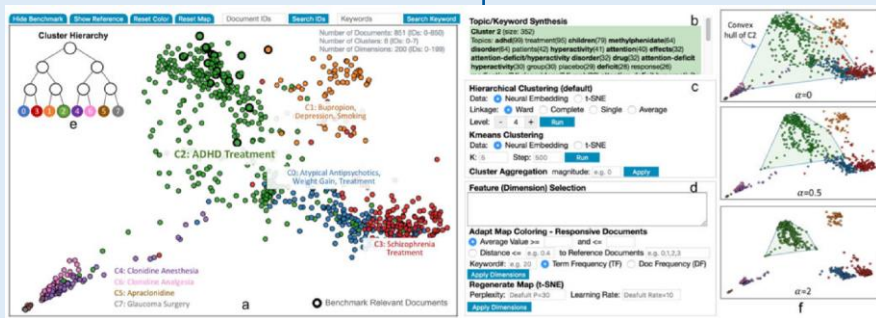


Fixed settings panel, automatically rerun algorithm



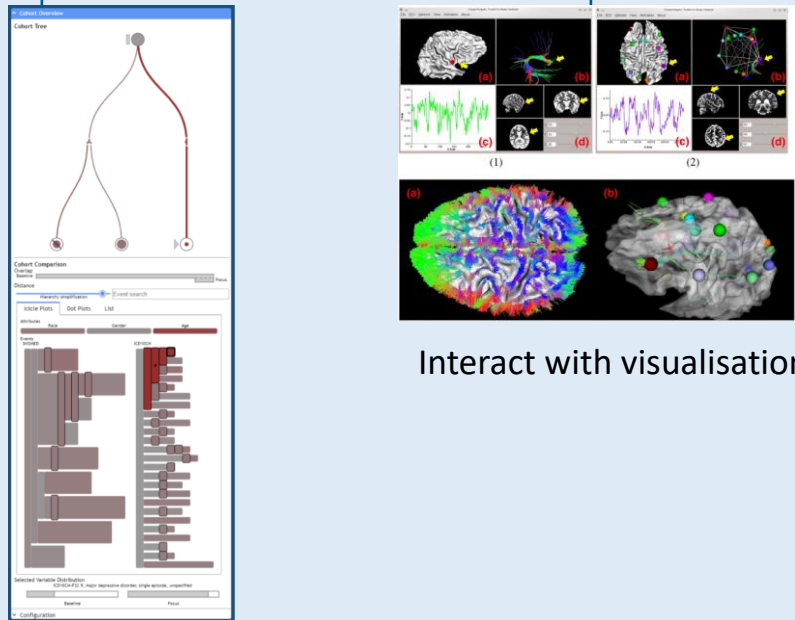
Automatically rerun algorithm when input features change

Semi-interactive



Fixed settings panel, manually rerun algorithm

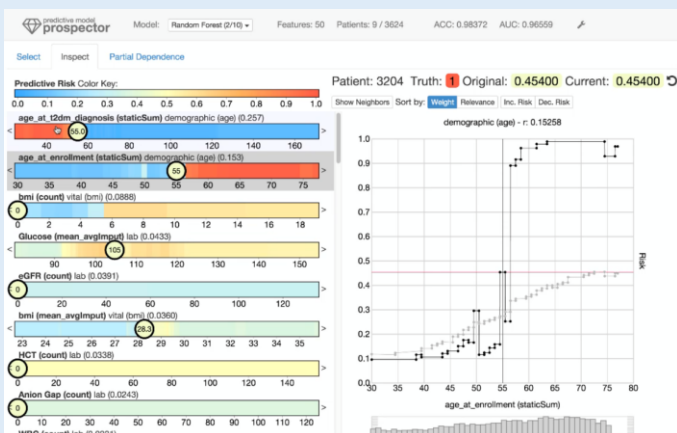
Tight integration



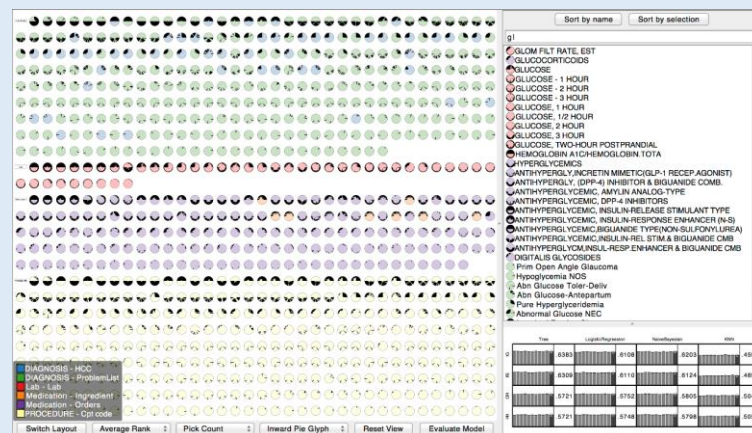
Interact with visualisation

Configuration in visualisation interface, automatically rerun algorithm

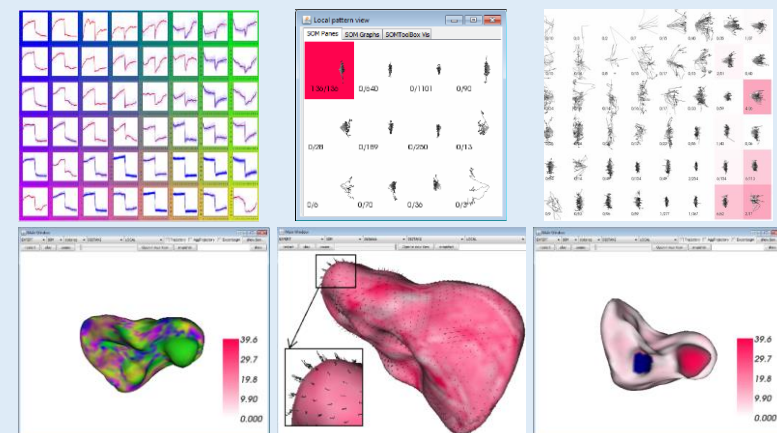
Direct explanation



Partial dependence plot

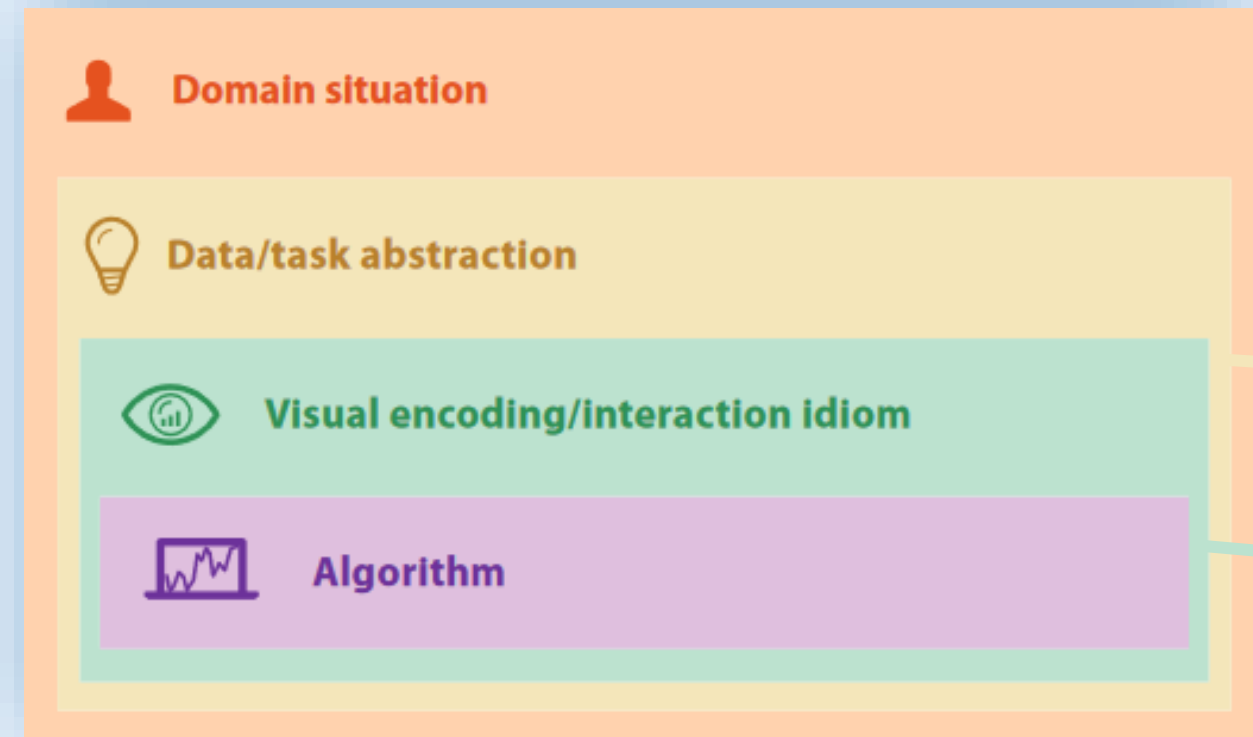


Feature importance



Visualisation

#2: Predicting Diabetes Onset



Domain: healthcare

Audience: healthcare professionals and patients

Tasks: compare to other patients, propose actions to minimise risks

Encoding: density plots, bar charts

Data-centric explanation

What-if explanation

Feature importance

Patient ID: 1570 **Age:** 60 **Pulse:** 79 ✓ **Region:** Drava **Blood Sugar:** 5.2 ✓ **Drinking Status:** Rarely Drinks **Institution:** 02 **BMI:** 29.4 ↑ **Smoking Status:** Smoker **Gender:** Female **Waist Measure:** 98 ✓ **Physical Activity Level:** Moderate

Overall Risk: Low

Patient Summary

Blood Sugar: 5.2 (4.7 - 6.3)

Alcohol Drinking Status: Rarely Drinks (Uncont. Drink, Heavy Drink, Moderately Drink, Addicted)

Waist Measure: 98 (80 - 98)

Smoking Status: Smoker (Non Smoker, Former Smoker, Smoker)

BMI: 29.4 (18 - 25)

Physical Activity Level: Moderate (Low, Moderate, High)

Recommendations to reduce risk

- Current: Smoking Status = Smoker
Action - Stop smoking at least for next 6 months
Other patients like Patient - 2949 who have taken this action have Low risk.
4% lower risk (difficult)
- Current: BMI = 29.4
Action - Reduce BMI below 25 by reducing weight by 15 kg
Other patients like Patient - 2949 who have taken this action have Low risk.
4% lower risk (difficult)

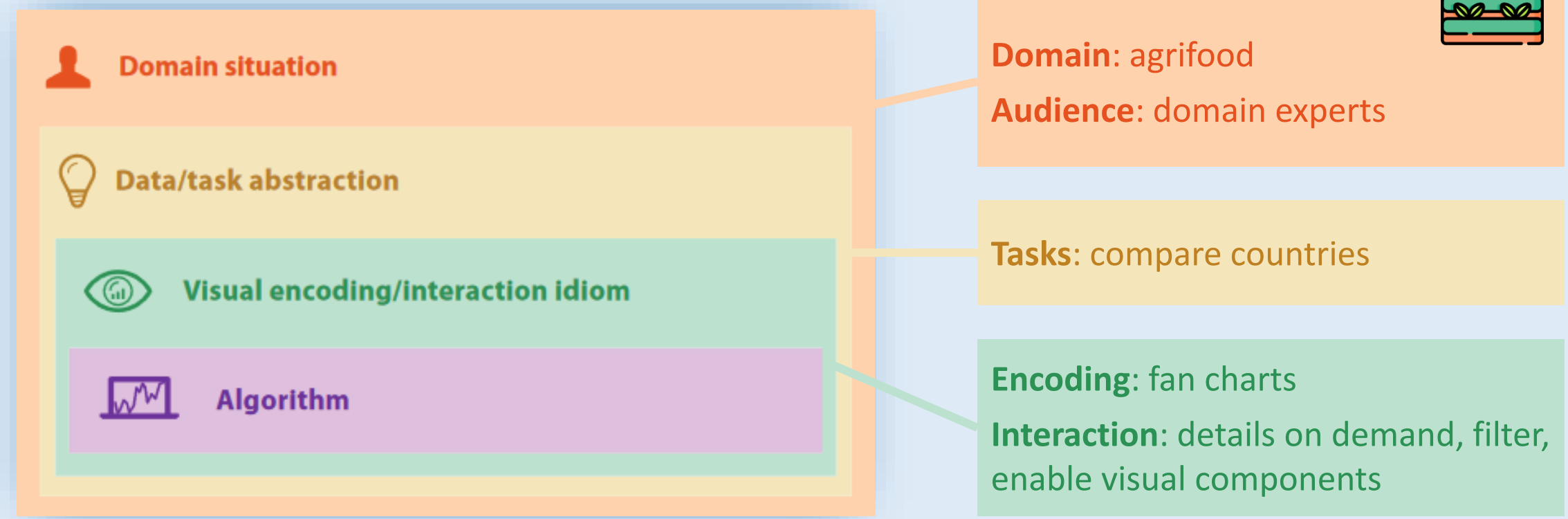
Important Risk Factors

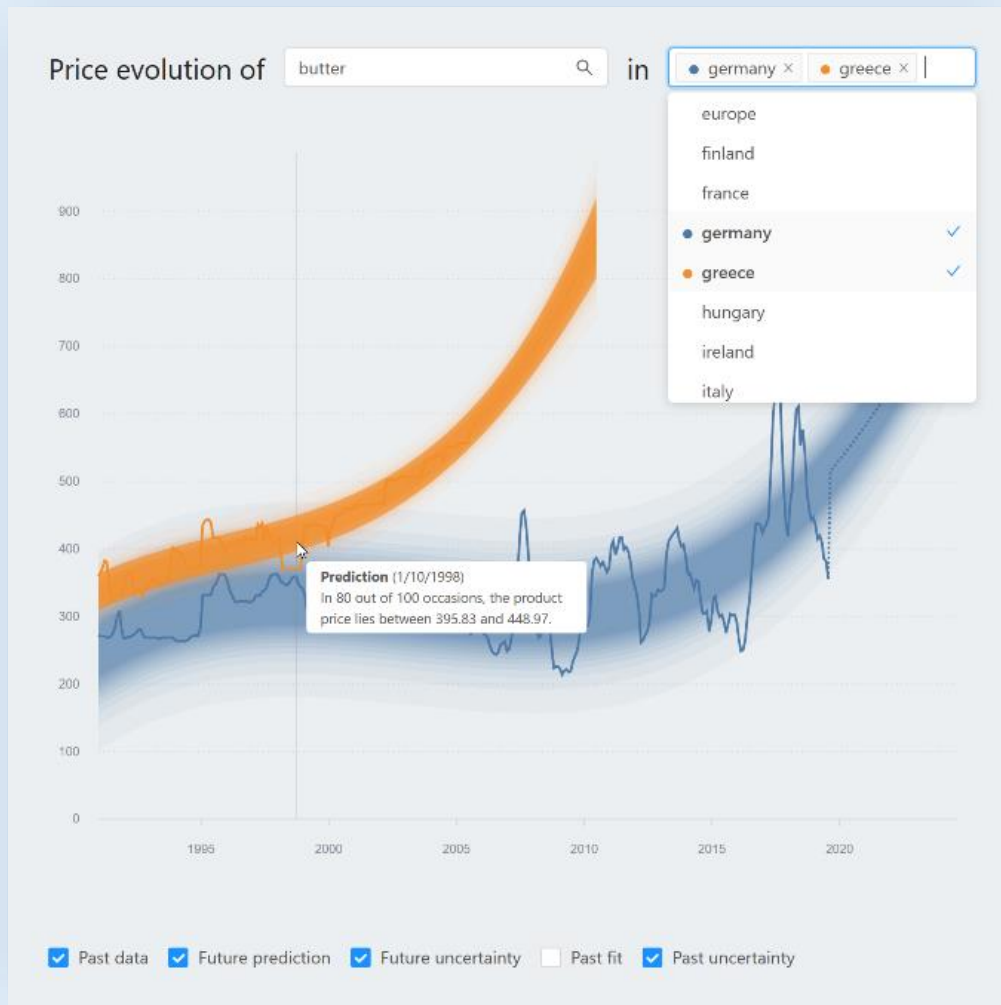
Patient Measures		Patient Behaviours	
Risk Factors	Impact %	Risk Factors	Impact %
Blood Sugar: 5.2 (4.7 <= 5.2 <= 6.3)	+37%	Physical Activity Level: (Moderate)	+13%
Waist Measure: 98 (80 <= 98 <= 98)	+17%	Alcohol Drinking Status: (Rarely Drinks)	+12%
BMI: 29.4 (> 25)	+10%	Smoking Status: (Smoker)	+11%

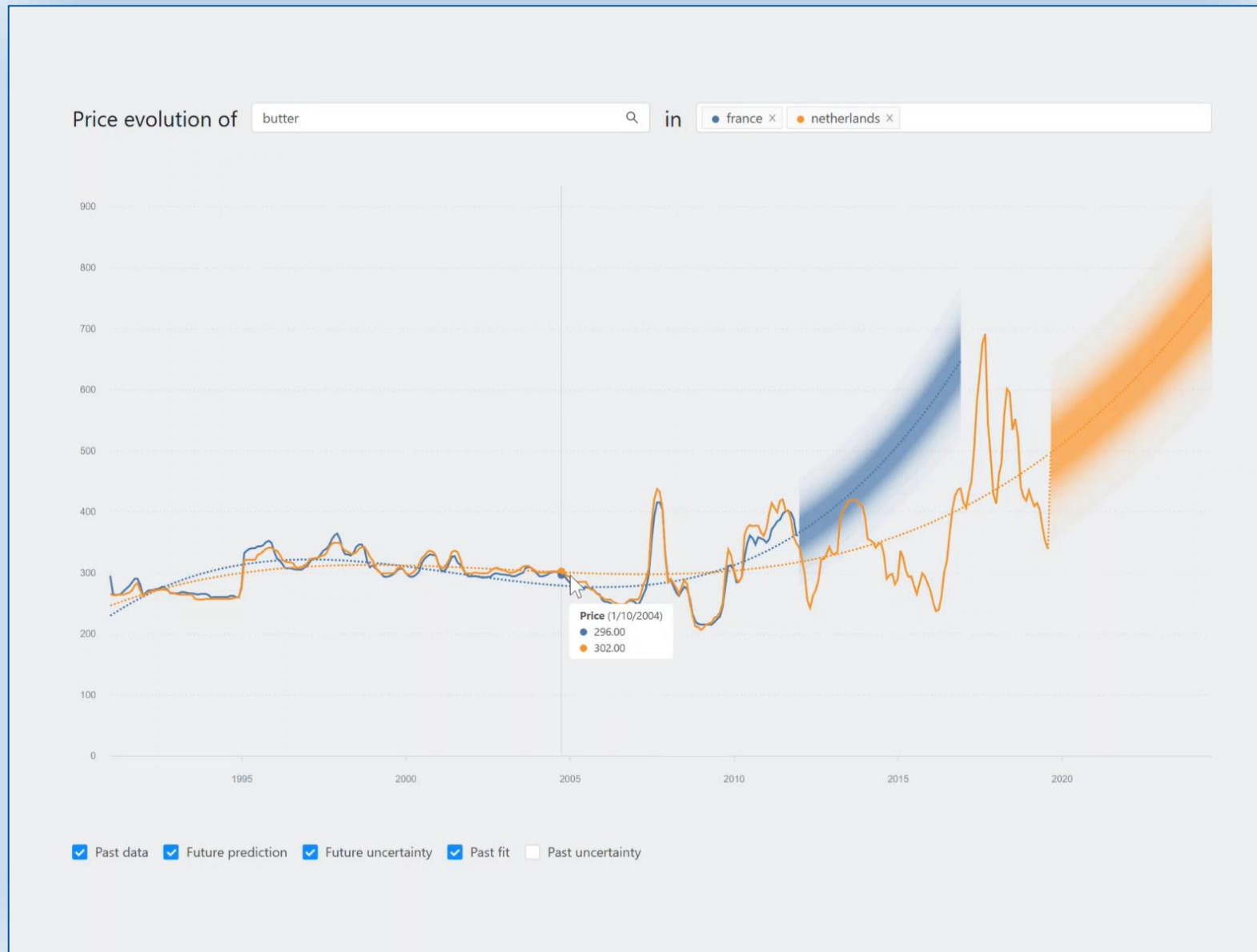
Risk Recovery

Estimated Risk vs Year 2022 (Jan to Dec)

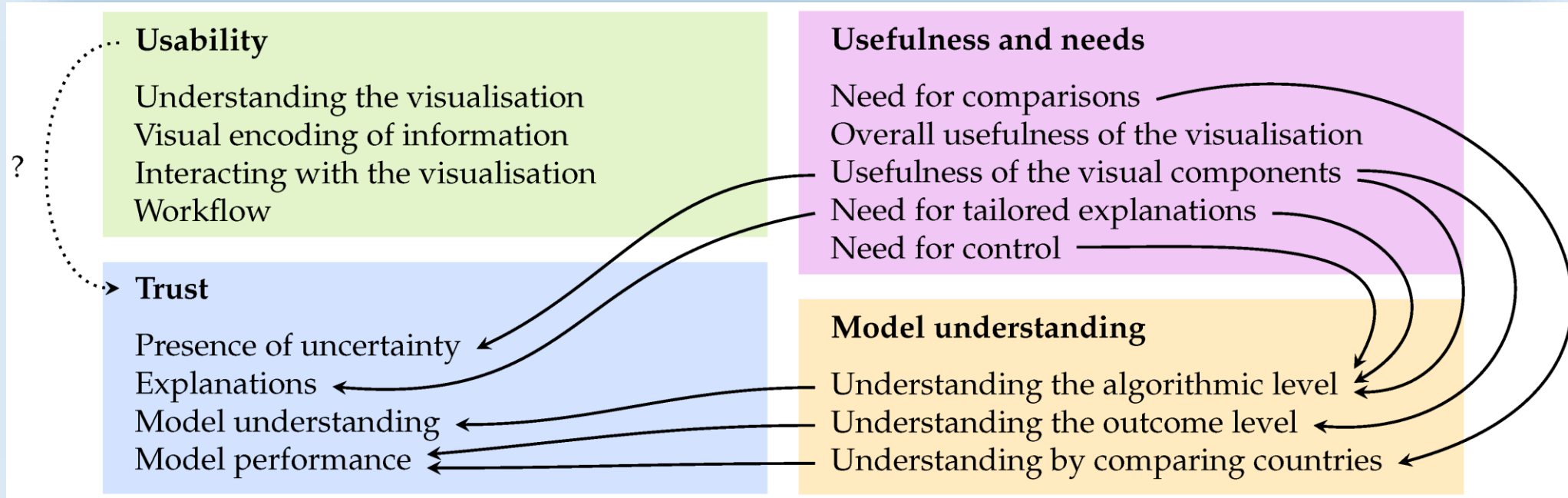
#3: Uncertain Predictions in Agrifood





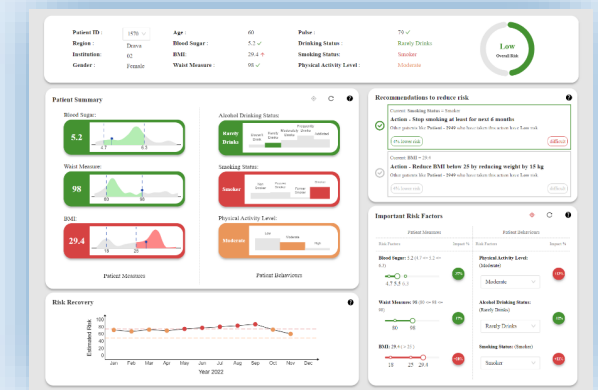
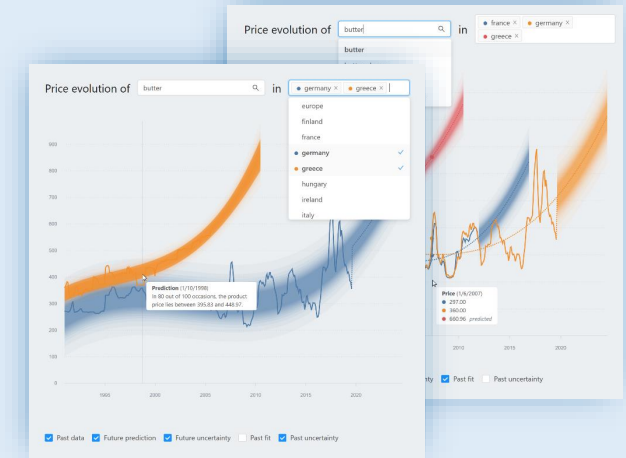
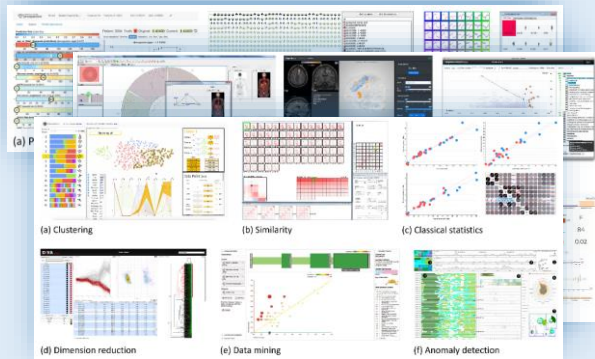


Jeroen Ooge and Katrien Verbert. 2022. Visually Explaining Uncertain Price Predictions in Agrifood: A User-Centred Case-Study. *Agriculture* 12, 7: 1024. <https://doi.org/10.3390/agriculture12071024>



Usability, usefulness and needs, and model understanding
can directly and indirectly affect trust

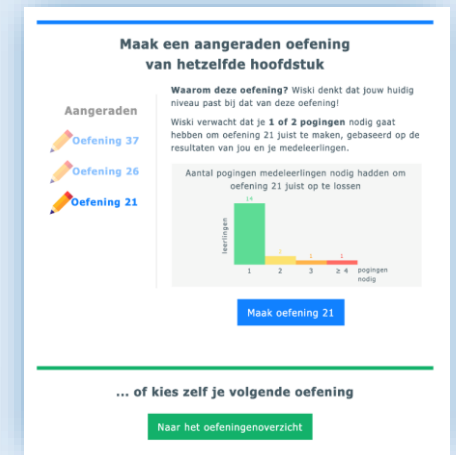
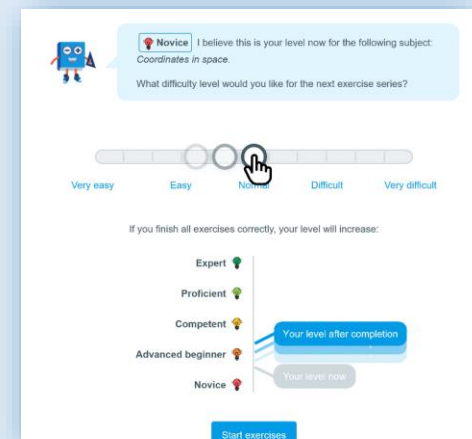
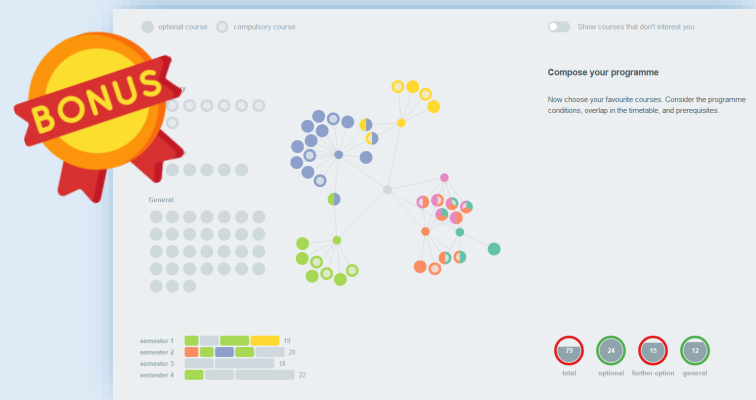
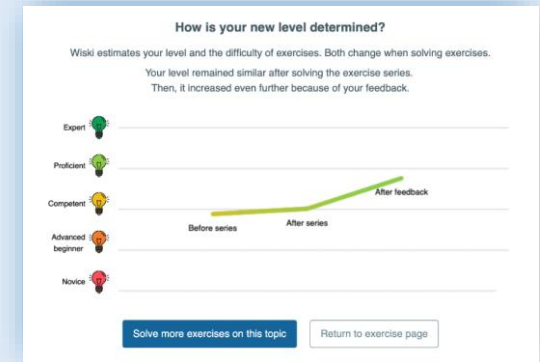
Break for Questions



I Complex Visual Dashboards

II Simple Visualisations

III Visualisations in Reports





expert

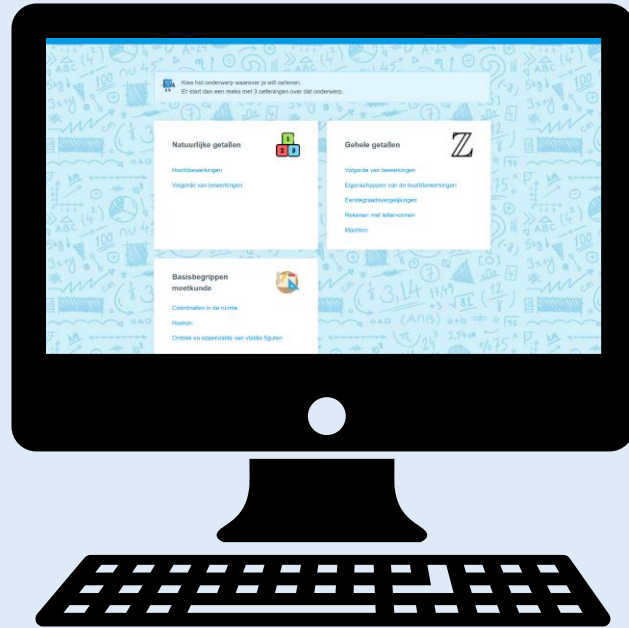


competent

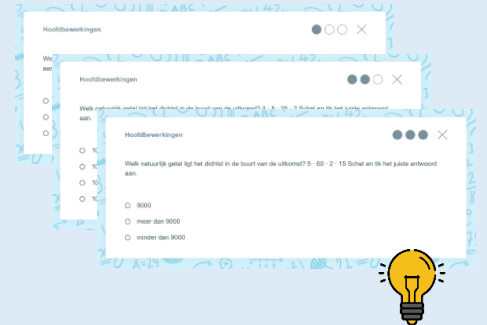
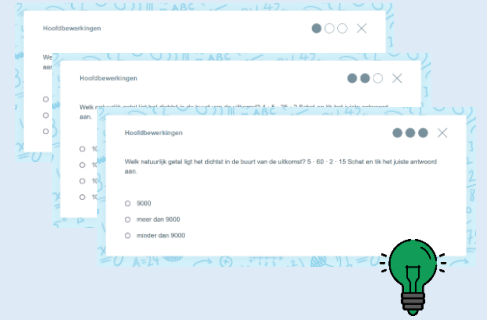


beginner

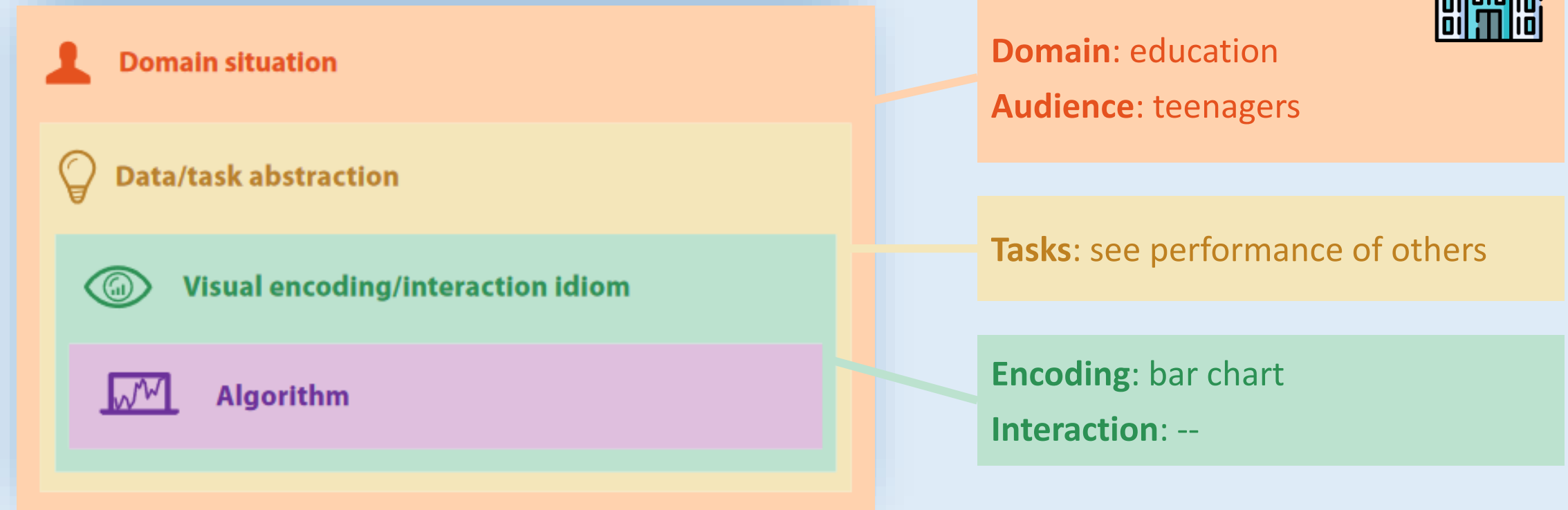
$$\sqrt{x} + \int dx + \pi$$



Recommender system



#4: Justifying Recommendations



Maak een aangeraden oefening van hetzelfde hoofdstuk

Aangeraden



Oefening 37



Oefening 26

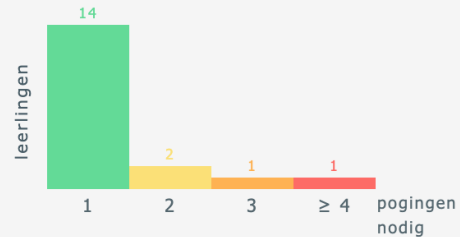


Oefening 21

Waarom deze oefening? Wiski denkt dat jouw huidig niveau past bij dat van deze oefening!

Wiski verwacht dat je **1 of 2 pogingen** nodig gaat hebben om oefening 21 juist te maken, gebaseerd op de resultaten van jou en je medeleerlingen.

Aantal pogingen medeleerlingen nodig hadden om oefening 21 juist op te lossen



Maak oefening 21

... of kies zelf je volgende oefening

Naar het oefeningenoverzicht

Textual explanation

Visual explanation

Maak een aangeraden oefening van hetzelfde hoofdstuk

Aangeraden

 Oefening 37

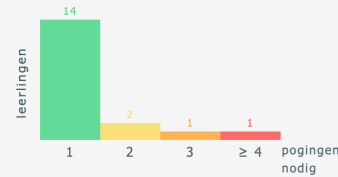
 Oefening 26

 Oefening 21

Waarom deze oefening? Wiski denkt dat jouw huidig niveau past bij dat van deze oefening!

Wiski verwacht dat je **1 of 2 pogingen** nodig gaat hebben om oefening 21 juist te maken, gebaseerd op de resultaten van jou en je medeleerlingen.

Aantal pogingen medeleerlingen nodig hadden om oefening 21 juist op te lossen



Maak oefening 21

Maak een aangeraden oefening van hetzelfde hoofdstuk

Aangeraden

 Oefening 27

 Oefening 40

 Oefening 45

Waarom deze oefening?

Oefening 27 is aangeraden omdat het algoritme van Wiski dat zo heeft berekend.



Maak oefening 27

... of kies zelf je volgende oefening

Naar het oefeningenoverzicht

Kies je volgende oefening

oefeningenoverzicht

Maak een aangeraden oefening van hetzelfde hoofdstuk

Aangeraden

 Oefening 27

 Oefening 40

 Oefening 45

Wiski raadt de volgende oefening aan



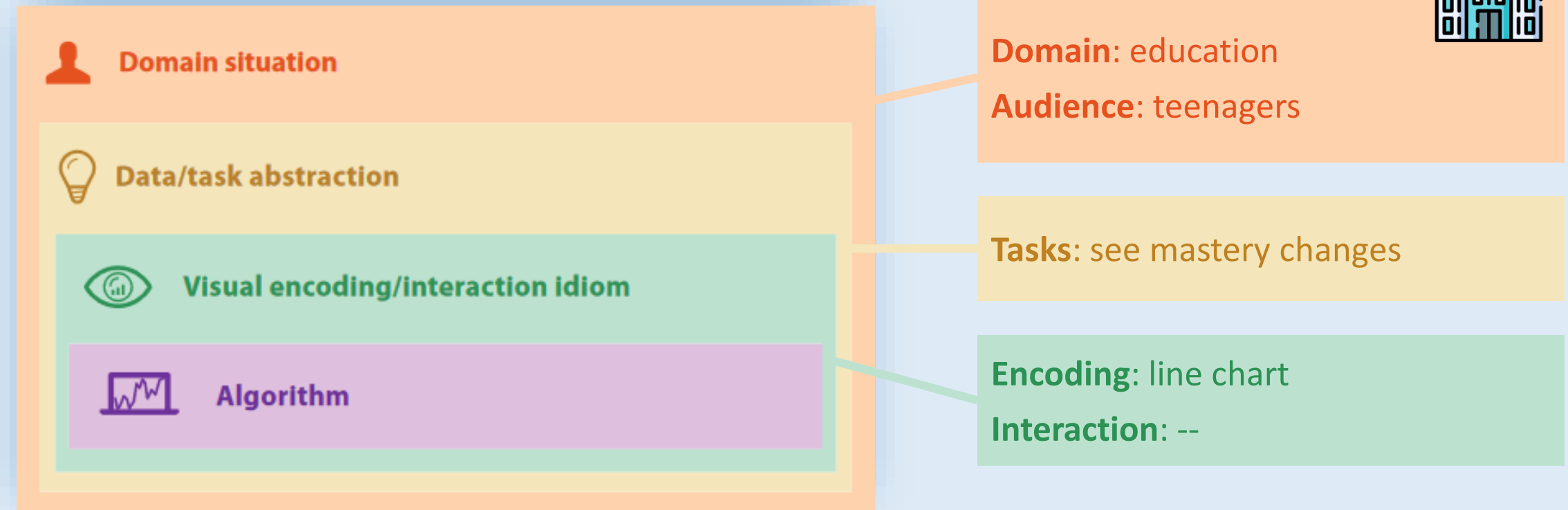
Maak oefening 27

... of kies zelf je volgende oefening

Naar het oefeningenoverzicht

Visual explanations can increase initial trust but may not be the most important factor for building it


#5: Visualise the Effect of Control





Hoe is je nieuw niveau bepaald?


Wiski schat jouw niveau en de moeilijkheid van oefeningen in. Beide veranderen bij het oplossen van oefeningen.


Je niveau is gestegen na het maken van de reeks oefeningen.
Daarna is het nog extra gestegen door je feedback.

Expert  _____

Bedreven  _____

Competent  _____

Gevorderde beginner  _____

Beginner  _____

Voor reeks Na reeks Na feedback

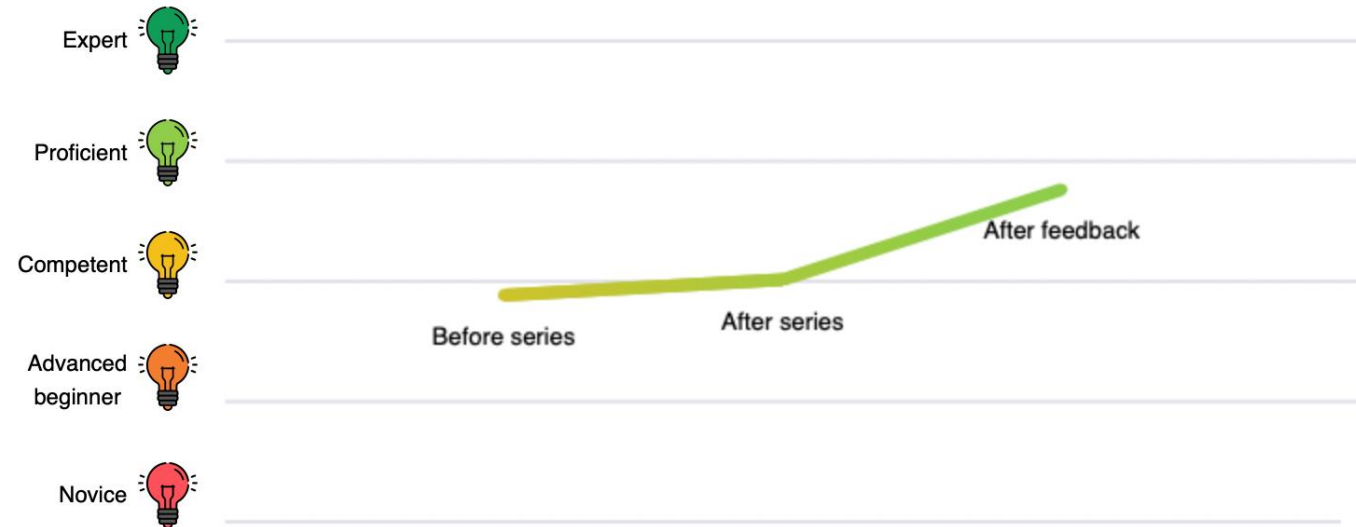
[Maak meer oefeningen over dit onderwerp](#) [Ga terug naar oefenpagina](#)

How is your new level determined?

Wiski estimates your level and the difficulty of exercises. Both change when solving exercises.

Your level remained similar after solving the exercise series.

Then, it increased even further because of your feedback.



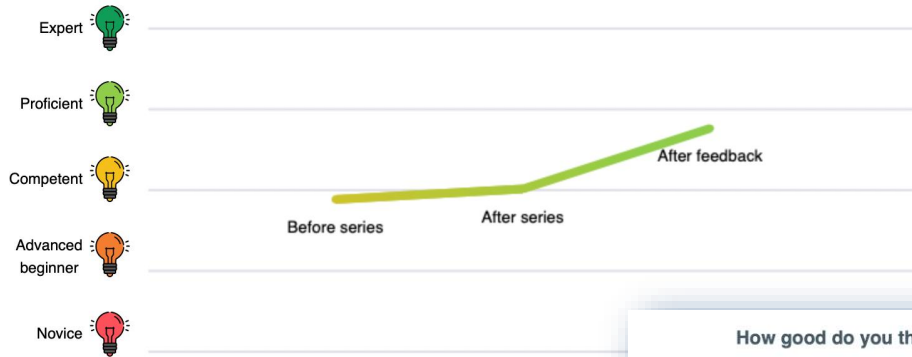
Solve more exercises on this topic

Return to exercise page

How is your new level determined?

Wiski estimates your level and the difficulty of exercises. Both change when solving exercises.

Your level remained similar after solving the exercise series.
Then, it increased even further because of your feedback.



Solve more exercises on this topic

Return to

How good do you think you are at mathematics?

There is no right or wrong answer. Wiski uses your answer to find suitable exercises for you.

- Expert:** mathematics holds no secrets for you.
- Proficient:** you score better than average on mathematics.
- Competent:** you score average on mathematics.
- Advanced beginner:** basic exercises are not a problem for you.
- Novice:** you often have a hard time understanding mathematics.

Submit

Wiski would like additional information from you

You solved a complete series of recommended exercises, congratulations! For the next series, you can give Wiski additional information so that Wiski knows better how you feel.

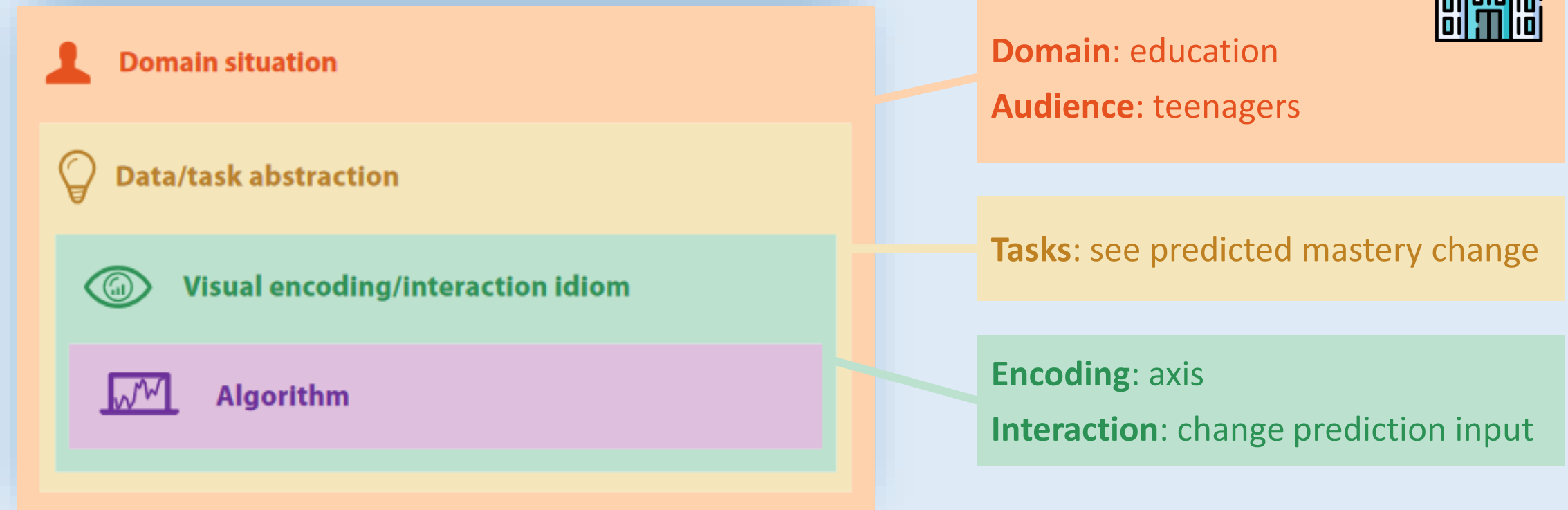
What difficulty of exercises would you like?





Submit feedback

Seeing the impact of control can increase initial trust, but control mechanisms by themselves do not necessarily

#6: Visual What-If Analysis








WISKI Oefenen 

 **gevorderde beginner** Volgens mij is dit nu je level voor het onderwerp *Volgorde van bewerkingen*

Welke moeilijkheidsgraad wil je voor de volgende oefeningensreeks?

Heel makkelijk Makkelijk Gewoon Moelijk Heel moeilijk

Als je alle oefeningen in de reeks juist oplost, dan stijgt je level:

- Expert 
- Bedreven 
- Competent 
- Gevorderde beginner  **Je level na de reeks**
- Beginner  Je level nu

Start de reeks

Bronvermeldingen Contact

Impact on level if series is solved correctly

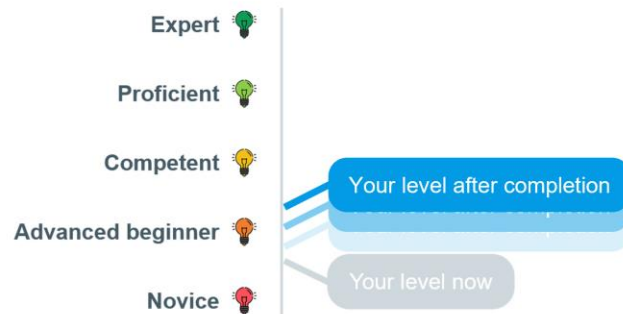


Novice I believe this is your level now for the following subject:
Coordinates in space.

What difficulty level would you like for the next exercise series?

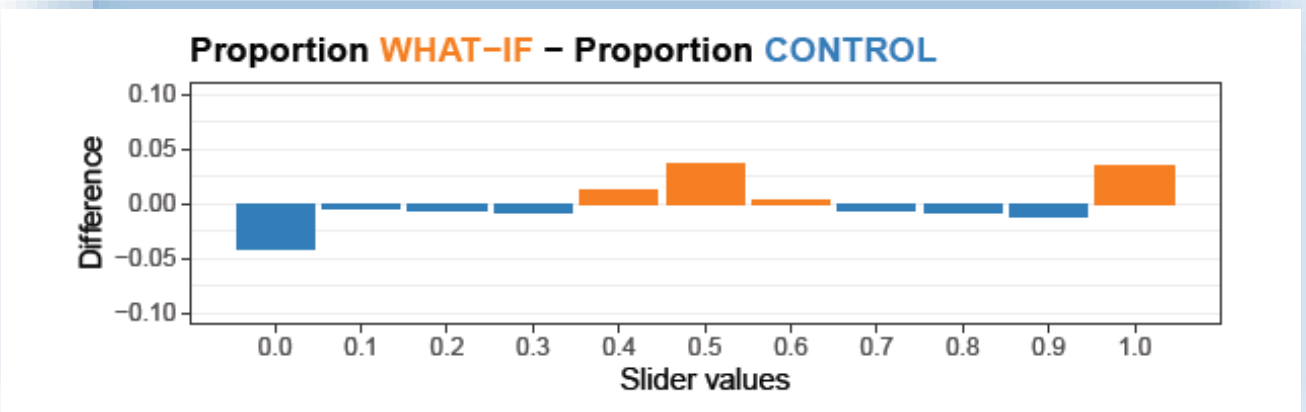



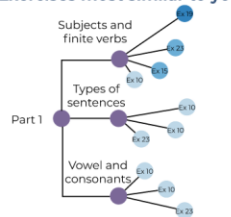
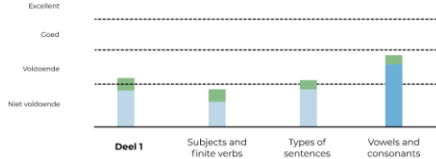
If you finish all exercises correctly, your level will increase:



Start exercises

What-if explanations can encourage teenagers to try harder exercises, but do not necessarily affect initial trust, metacognition, motivation, or performance



Recommended series			General	Subjects and finite verbs	Types of sentences	Vowels and consonants
Exercise 15 Part 1 Subjects and finite verbs	Exercise level Easy	Completed before? No	<p>Why this exercise series?</p> <p>The system looks for exercises adapted to your mastery level so you can make the most progress. Your mastery level is estimated based on exercises you have solved previously.</p> <p>1. Your level and why:</p> <p>PART 1 Passing</p> <p>Negative impact Positive impact</p>  <p>2. Exercises most similar to your level:</p>  <p>3. What if you correctly finish this series?</p> <p>Excellent Goed Voldoende Niet voldoende</p>  <p>Deel 1 Subjects and finite verbs Types of sentences Vowels and consonants</p> <p>By completing the next exercise series, you will level up for the topics 'Subjects and finite verbs', 'Types of sentences' and 'Vowels and consonants'. Keep going!</p>			
Exercise 23 Part 1 Types of sentences	Exercise level Easy	Completed before? No				
Exercise 12 Part 1 Subjects and finite verbs	Exercise level Easy	Completed before? No				
Exercise 35 Part 1 Vowels and consonants	Exercise level Easy	Completed before? No				
Exercise 10 Part 1 Types of sentences	Exercise level Easy	Completed before? No				

Start sequence



Exercise
Exercise 15 Part 1 Subjects and finite verbs
Exercise 23 Part 1 Types of sentences
Exercise 12 Part 1 Subjects and finite verbs
Exercise 35 Part 1 Vowels and consonants
Exercise 10 Part 1 Subjects and finite verbs



Group discussions with teachers and people developing educational technologies

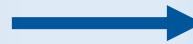
Recommended series		Explanation
Exercise	Exercise level	
Exercise 15 Part 1 Subjects and finite verbs	Easy	<p>Why this exercise series? The system recommends exercises that are closest to Your level. Your level is estimated by looking at exercises you've solved previously.</p> <p>What if you correctly solve this series? This is how the system estimates your mastery to change if you correctly solve the exercise series.</p> <p>Part 1</p> <p>● Non-recommended exercise ● Recommended exercise</p>
Exercise 23 Part 1 Types of sentences	Medium	
Exercise 12 Part 1 Subjects and finite verbs	Easy	
Exercise 35 Part 1 Vowels and consonants	Hard	
Exercise 10 Part 1 Subjects and finite verbs	Easy	



Focus groups with teachers and educational experts
Think-aloud sessions with adolescents

What if you correctly solve this series?

How the system estimates your mastery to change if you correctly solve the exercise series.



I believe this is your level now for the subject *Coordinates in space*:
💡 **Competent**

What difficulty level would you like for the next exercise series?

Easier Same Harder

If you finish all exercises correctly, your level will increase:

- Expert 💡
- Proficient 💡 — **Your level after completion**
- Competent 💡 — **Your level now**
- Advanced beginner 💡
- Novice 💡

Start exercises



I believe this is your level now for the subject *Coordinates in space*:
💡 **Novice**

What difficulty level would you like for the next exercise series?

Very easy Easy Medium

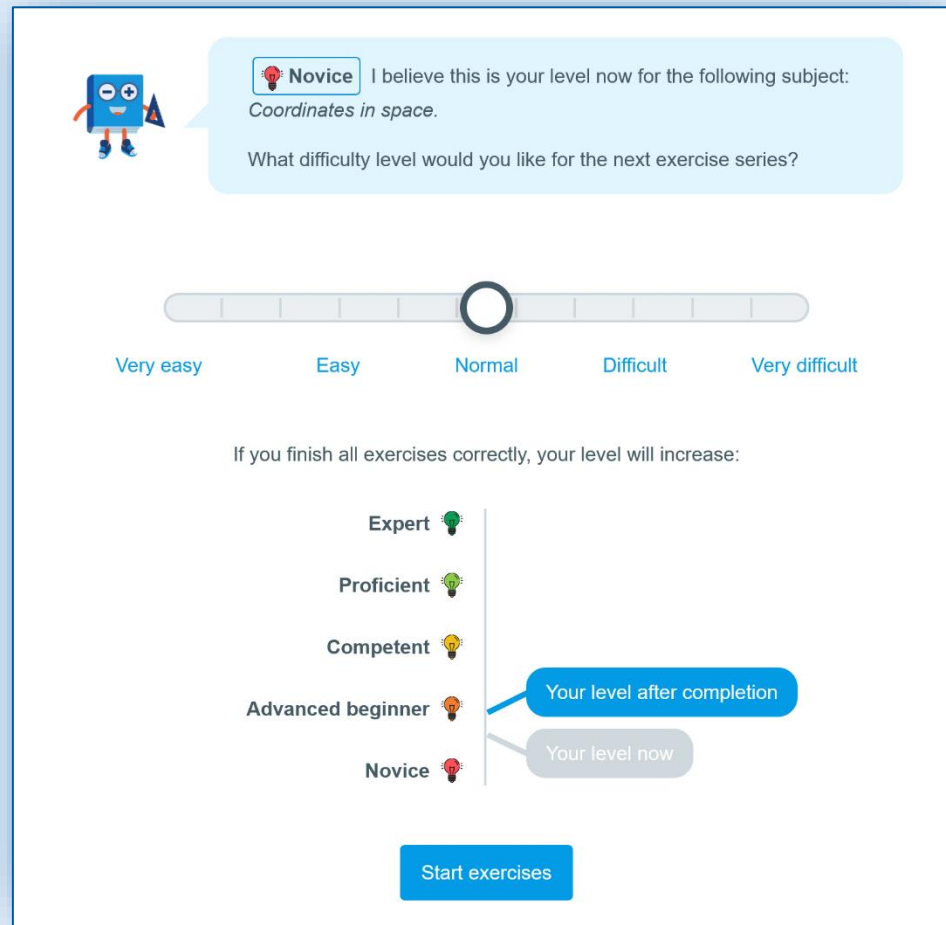
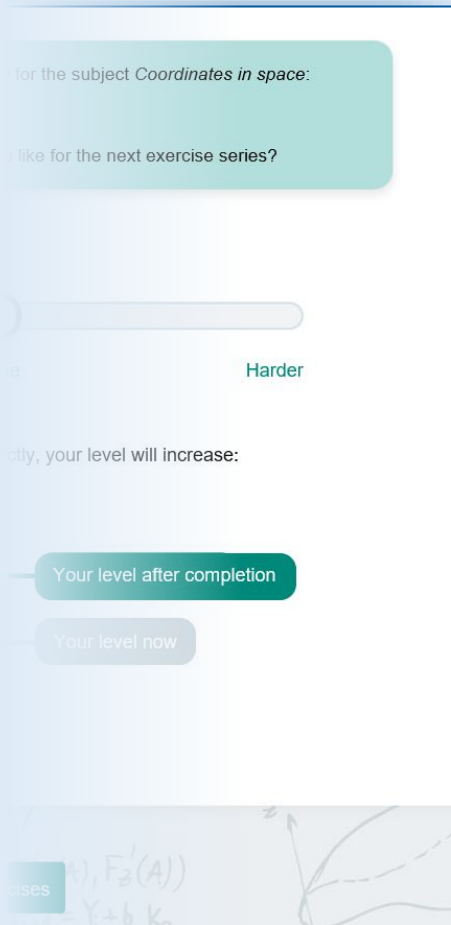
If you finish all exercises correctly, your level will increase:

- Expert 💡
- Proficient 💡
- Competent 💡
- Advanced beginner 💡
- Novice 💡

Start exercises



More think-aloud sessions with adolescents



Final prototype for evaluation!

#6: Visually Compose a University Programme



Domain situation

Domain: education

Audience: adult students



Data/task abstraction



Visual encoding/interaction idiom

Tasks: see connections between courses, compose programme



Algorithm

Encoding: network, bar chart, gauges

Interaction: filter, connect, select

● optional course ○ compulsory course

Show courses that don't interest you

Compulsory



More option

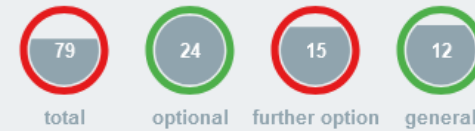


General



Compose your programme

Now choose your favourite courses. Consider the programme conditions, overlap in the timetable, and prerequisites.



Break for Questions

Novice | I believe this is your level now for the following subject: Coordinates in space. What difficulty level would you like for the next exercise series?

Very easy Easy Novice Difficult Very difficult

If you finish all exercises correctly, your level will increase:

- Expert
- Proficient
- Competent
- Advanced beginner
- Novice

Your level after completion

Your level now

Start exercises

How is your new level determined?

Wiski estimates your level and the difficulty of exercises. Both change when solving exercises. Your level remained similar after solving the exercise series. Then, it increased even further because of your feedback.

Expert

Proficient

Competent

Advanced beginner

Novice

Before series After series

After feedback

Solve more exercises on this topic Return to exercise page



optional course compulsory course Show courses that don't interest you

Compose your programme

Now choose your favourite courses. Consider the programme conditions, overlap in the timetable, and prerequisites.

semester 1 19

semester 2 20

semester 3 19

semester 4 22

total optional further option general

Maak een aangeraden oefening van hetzelfde hoofdstuk

Aangeraden

- Oefening 37
- Oefening 26
- Oefening 21

Waarom deze oefening? Wiski denkt dat jouw huidige niveau past bij dat van deze oefening!

Wiski verwacht dat je 1 of 2 pogingen nodig gaat hebben om oefening 21 juist te maken, gebaseerd op de resultaten van jou en je medeleerlingen.

Aantal pogingen medeleerlingen nodig hadden om oefening 21 juist op te lossen

Maak oefening 21

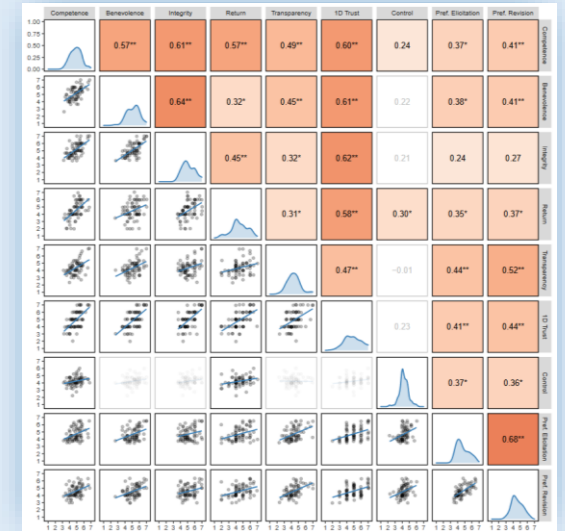
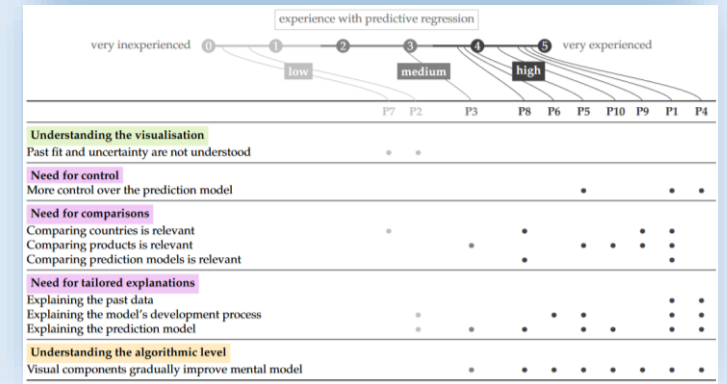
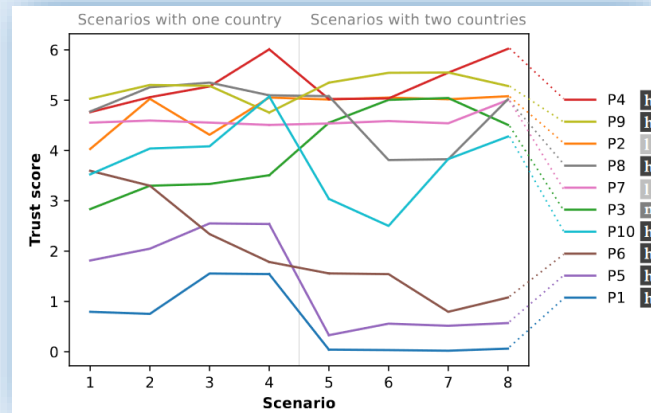
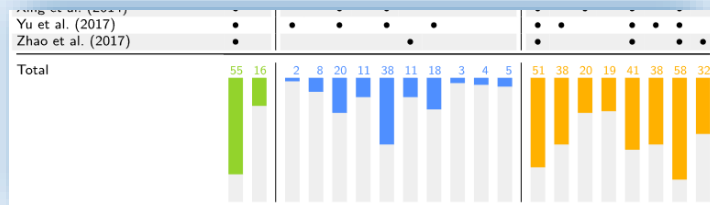
... of kies zelf je volgende oefening

Naar het oefeningenoverzicht

I Complex Visual Dashboards

II Simple Visualisations

III Visualisations in Reports



#1: Make Tables Insightful

	ACTIVITY		ALGORITHM								INTERACTION										
	Interpretation	Prediction	Anomaly detection	Artificial neural network	Classical statistics	Classification	Clustering/similarity	Data mining	Dimension reduction	Feature selection	Segmentation	Other	Abstract/elaborate	Connect	Encode	Explore	Filter	Reconfigure	Select	Shepherd	
Abbasloo et al. (2019)	•			•									•	•	•					•	•
Abdullah et al. (2020)	•						•		•				•	•			•	•	•	•	•
Afzal et al. (2011)		•										•	•							•	•
Abbasloo et al. (2019)	•			•									•	•						•	•
Yu et al. (2017)	•		•	•			•		•				•	•			•	•	•	•	•
Zhao et al. (2017)	•							•					•				•		•	•	•

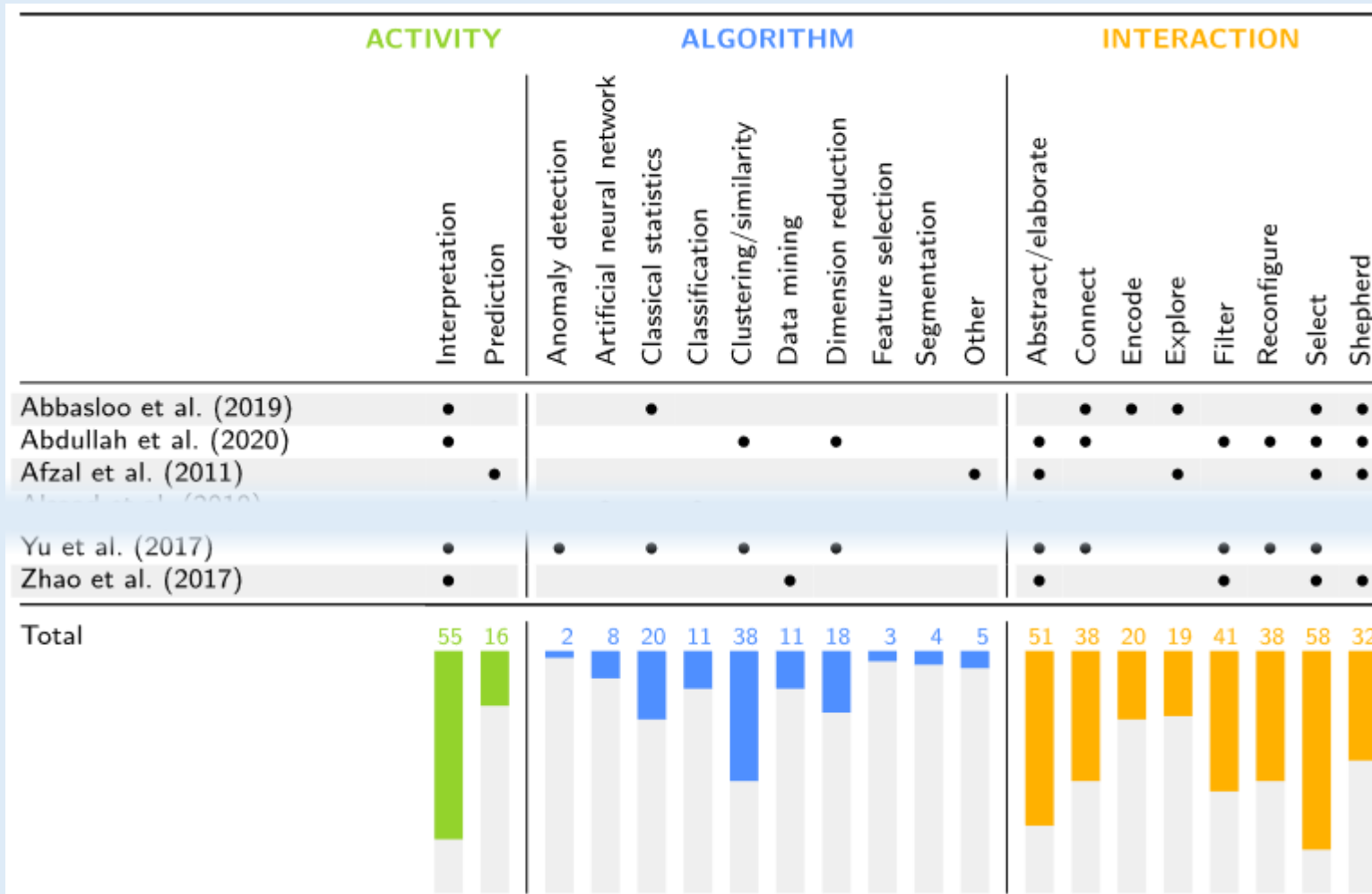
	ACTIVITY	ALGORITHM										INTERACTION								
	Interpretation Prediction	Anomaly detection	Artificial neural network	Classical statistics	Classification	Clustering/similarity	Data mining	Dimension reduction	Feature selection	Segmentation	Other	Abstract/elaborate	Connect	Encode	Explore	Filter	Reconfigure	Select	Shepherd	
Abbasloo et al. (2019)	•		•									•	•	•					•	•
Abdullah et al. (2020)	•											•	•						•	•
Afzal et al. (2011)		•									•	•							•	•
Abbasloo et al. (2019)	•											•	•						•	•
Alsaad et al. (2019)				•								•	•	•					•	•
Barlowe et al. (2013)												•	•						•	•
Behrisch et al. (2018)	•											•	•						•	•
Borland et al. (2020)	•											•	•						•	•
Brunker et al. (2019)	•											•	•						•	•
Cao et al. (2011)	•											•	•						•	•
Clark et al. (2017)	•											•	•						•	•
Dang et al. (2015)	•											•	•						•	•
Dingen et al. (2019)	•											•	•						•	•
Dixit et al. (2017)	•											•	•						•	•
Fang et al. (2017)	•											•	•						•	•
Farag et al. (2015)	•											•	•						•	•
Feller et al. (2018)	•											•	•						•	•
Geurts et al. (2015)	•											•	•						•	•
Gotz et al. (2011)	•											•	•						•	•
Gotz et al. (2014)	•											•	•						•	•
Gotz et al. (2020)	•											•	•						•	•
Guo et al. (2020)	•											•	•						•	•
Guo et al. (2018)	•											•	•						•	•
Herold et al. (2010)	•											•	•						•	•
Hinterberg et al. (2015)	•											•	•						•	•
Huang et al. (2015)	•											•	•						•	•
Huang et al. (2019)	•											•	•						•	•
Hund et al. (2016)	•											•	•						•	•
Hur et al. (2020)	•											•	•						•	•
Ji et al. (2017)	•											•	•						•	•
Ji et al. (2019a)	•											•	•						•	•
Ji et al. (2019b)	•											•	•						•	•
Jönsson et al. (2019)	•											•	•						•	•
Kakar et al. (2019)	•											•	•						•	•
Klemm et al. (2014)	•											•	•						•	•
Klimov et al. (2015)	•											•	•						•	•
Kovalerchuk et al. (2012)	•											•	•						•	•
Krause et al. (2014)	•											•	•						•	•
Krause et al. (2016)	•											•	•						•	•
Krause et al. (2018)	•											•	•						•	•
Kumar et al. (2015)	•											•	•						•	•
Kwon et al. (2018)	•											•	•						•	•
Kwon et al. (2019)	•											•	•						•	•
Kwon et al. (2020)	•											•	•						•	•
L'Yi et al. (2015)	•											•	•						•	•
L'Yi et al. (2017)	•											•	•						•	•
Lamy and Tsopra (2019)	•											•	•						•	•
Li et al. (2012)	•											•	•						•	•
Li et al. (2020)	•											•	•						•	•
Liao et al. (2017)	•											•	•						•	•
Males et al. (2020)	•											•	•						•	•
Malik et al. (2015)	•											•	•						•	•
Moschonas et al. (2016)	•											•	•						•	•
Müller et al. (2020)	•											•	•						•	•
Nauta et al. (2020)	•											•	•						•	•
Nguyen et al. (2011)	•											•	•						•	•
Nguyen et al. (2012)	•											•	•						•	•
Raidou et al. (2016a)	•											•	•						•	•
Raidou et al. (2016b)	•											•	•						•	•
Riegler et al. (2016)	•											•	•						•	•
Santamaría et al. (2008)	•											•	•						•	•
Santamaría et al. (2019)	•											•	•						•	•
Seo and Shneiderman (2002)	•											•	•						•	•
Song et al. (2017)	•											•	•						•	•
Spitz et al. (2020)	•											•	•						•	•
Stolper et al. (2014)	•											•	•						•	•
Verma et al. (2017)	•											•	•						•	•
Von Landesberger et al. (2013)	•											•	•						•	•
Widnangamachchi et al. (2017)	•											•	•						•	•
Xing et al. (2014)	•											•	•						•	•
Yu et al. (2017)	•											•	•						•	•
Zhao et al. (2017)	•											•	•						•	•

#1: Make Tables Insightful

	ACTIVITY		ALGORITHM								INTERACTION									
	Interpretation	Prediction	Anomaly detection	Artificial neural network	Classical statistics	Classification	Clustering/similarity	Data mining	Dimension reduction	Feature selection	Segmentation	Other	Abstract/elaborate	Connect	Encode	Explore	Filter	Reconfigure	Select	Shepherd
Abbasloo et al. (2019)	•			•									•	•	•	•			•	•
Abdullah et al. (2020)	•						•		•				•	•			•	•	•	•
Afzal et al. (2011)		•										•	•			•			•	•
Abbasloo et al. (2019)	•												•	•	•	•			•	•
Yu et al. (2017)	•		•	•			•		•				•	•			•	•	•	•
Zhao et al. (2017)	•							•					•				•		•	•
Total	55	16	2	8	20	11	38	11	18	3	4	5	51	38	20	19	41	38	58	32

	ACTIVITY	ALGORITHM	INTERACTION
	Interpretation Prediction	Anomaly detection Artificial neural network Classical statistics Classification Clustering/similarity Data mining Dimension reduction Feature selection Segmentation Other	Abstract/elaborate Connect Encode Explore Filter Reconfigure Select Shepherd
Abbasloo et al. (2019)	•	•	•
Abdullah et al. (2020)	•		•
Afzal et al. (2011)	•		•
Abbasloo et al. (2019)	•	•	•
Behrisch et al. (2018)	•	•	•
Borland et al. (2020)	•	•	•
Brunker et al. (2019)	•	•	•
Cao et al. (2011)	•	•	•
Clark et al. (2017)	•	•	•
Dang et al. (2015)	•	•	•
Dingen et al. (2019)	•	•	•
Dixit et al. (2017)	•	•	•
Fang et al. (2017)	•	•	•
Farag et al. (2015)	•	•	•
Feller et al. (2018)	•	•	•
Geurts et al. (2015)	•	•	•
Gotz et al. (2011)	•	•	•
Gotz et al. (2014)	•	•	•
Gotz et al. (2020)	•	•	•
Guo et al. (2020)	•	•	•
Guo et al. (2018)	•	•	•
Herold et al. (2010)	•	•	•
Hinterberg et al. (2015)	•	•	•
Huang et al. (2015)	•	•	•
Huang et al. (2019)	•	•	•
Hund et al. (2016)	•	•	•
Hur et al. (2020)	•	•	•
Ji et al. (2017)	•	•	•
Ji et al. (2019a)	•	•	•
Ji et al. (2019b)	•	•	•
Jönsson et al. (2019)	•	•	•
Kakar et al. (2019)	•	•	•
Klemm et al. (2014)	•	•	•
Klimov et al. (2015)	•	•	•
Kovalerchuk et al. (2012)	•	•	•
Krause et al. (2014)	•	•	•
Krause et al. (2016)	•	•	•
Krause et al. (2018)	•	•	•
Kumar et al. (2015)	•	•	•
Kwon et al. (2018)	•	•	•
Kwon et al. (2019)	•	•	•
Kwon et al. (2020)	•	•	•
L'Yi et al. (2015)	•	•	•
L'Yi et al. (2017)	•	•	•
Lamy and Tsopra (2019)	•	•	•
Li et al. (2012)	•	•	•
Li et al. (2020)	•	•	•
Liao et al. (2017)	•	•	•
Males et al. (2020)	•	•	•
Malik et al. (2015)	•	•	•
Moschonas et al. (2016)	•	•	•
Müller et al. (2020)	•	•	•
Nauta et al. (2020)	•	•	•
Nguyen et al. (2011)	•	•	•
Nguyen et al. (2012)	•	•	•
Raidou et al. (2016a)	•	•	•
Raidou et al. (2016b)	•	•	•
Riegler et al. (2016)	•	•	•
Santamaría et al. (2008)	•	•	•
Santamaría et al. (2019)	•	•	•
Seo and Shneiderman (2002)	•	•	•
Song et al. (2017)	•	•	•
Spitz et al. (2020)	•	•	•
Stolper et al. (2014)	•	•	•
Verma et al. (2017)	•	•	•
Von Landesberger et al. (2013)	•	•	•
Widjagumachchi et al. (2017)	•	•	•
Xing et al. (2014)	•	•	•
Yu et al. (2017)	•	•	•
Zhao et al. (2017)	•	•	•

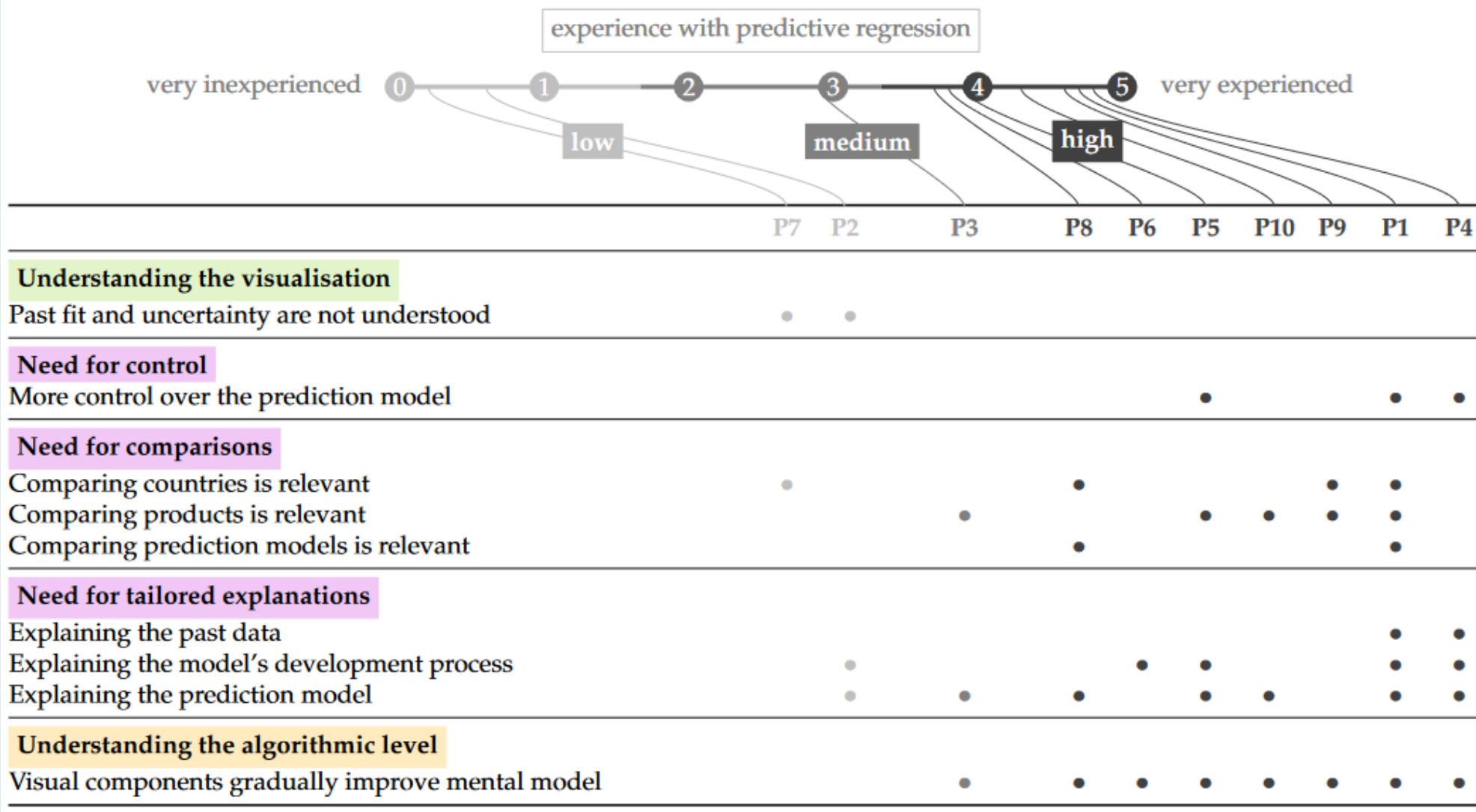
#1: Make Tables Insightful



	ACTIVITY	ALGORITHM	INTERACTION
	Interpretation Prediction	Anomaly detection Artificial neural network Classical statistics Classification Clustering/similarity Data mining Dimension reduction Feature selection Segmentation Other	Abstract/elaborate Connect Encode Explore Filter Reconfigure Select Shepherd
Abbasloo et al. (2019)	•	•	•
Abdullah et al. (2020)	•		•
Afzal et al. (2011)			•
Alsaad et al. (2019)		•	•
Barlowe et al. (2013)			•
Behrisch et al. (2018)	•		•
Borland et al. (2020)	•		•
Brunker et al. (2019)	•		•
Cao et al. (2011)	•		•
Clark et al. (2017)	•	•	•
Dang et al. (2015)	•		•
Dingen et al. (2019)	•		•
Dixit et al. (2017)	•		•
Fang et al. (2017)	•		•
Farag et al. (2015)	•		•
Feller et al. (2018)	•		•
Geurts et al. (2015)	•		•
Gotz et al. (2011)	•		•
Gotz et al. (2014)	•		•
Gotz et al. (2020)	•		•
Guo et al. (2020)	•	•	•
Guo et al. (2018)	•		•
Herold et al. (2010)	•		•
Hinterberg et al. (2015)	•	•	•
Huang et al. (2015)	•		•
Huang et al. (2019)	•		•
Hund et al. (2016)	•		•
Hur et al. (2020)	•		•
Ji et al. (2017)	•		•
Ji et al. (2019a)	•		•
Ji et al. (2019b)	•		•
Jönsson et al. (2019)	•		•
Kakar et al. (2019)	•		•
Klemm et al. (2014)	•		•
Klimov et al. (2015)	•		•
Kovalerchuk et al. (2012)	•		•
Krause et al. (2014)	•		•
Krause et al. (2016)	•		•
Krause et al. (2018)	•		•
Kumar et al. (2015)	•		•
Kwon et al. (2018)	•		•
Kwon et al. (2019)	•		•
Kwon et al. (2020)	•		•
L'Yi et al. (2015)	•		•
L'Yi et al. (2017)	•		•
Lamy and Tsopra (2019)	•		•
Li et al. (2012)	•		•
Li et al. (2020)	•		•
Liao et al. (2017)	•		•
Males et al. (2020)	•		•
Malik et al. (2015)	•		•
Moschonas et al. (2016)	•		•
Müller et al. (2020)	•		•
Nauta et al. (2020)	•		•
Nguyen et al. (2011)	•		•
Nguyen et al. (2012)	•		•
Raidou et al. (2016a)	•		•
Raidou et al. (2016b)	•		•
Riegler et al. (2016)	•		•
Santamaría et al. (2008)	•		•
Santamaría et al. (2019)	•		•
Seo and Shneiderman (2002)	•		•
Song et al. (2017)	•		•
Spitz et al. (2020)	•		•
Stolper et al. (2014)	•		•
Verma et al. (2017)	•		•
Von Landesberger et al. (2013)	•		•
Widanagamagechchi et al. (2017)	•		•
Xing et al. (2014)	•		•
Yu et al. (2017)	•		•
Zhao et al. (2017)	•		•

	P7	P2	P3	P8	P6	P5	P10	P9	P1	P4
Understanding the visualisation										
Past fit and uncertainty are not understood	•	•								
Need for control										
More control over the prediction model						•			•	•
Need for comparisons										
Comparing countries is relevant	•			•				•	•	
Comparing products is relevant			•			•	•	•	•	
Comparing prediction models is relevant				•					•	
Need for tailored explanations										
Explaining the past data									•	•
Explaining the model's development process		•			•	•			•	•
Explaining the prediction model		•	•	•		•	•		•	•
Understanding the algorithmic level										
Visual components gradually improve mental model			•	•	•	•	•	•	•	•

Table 2. Some topics raised by the participants, ordered by their experience with predictive regression (P2 and P7 have low experience; P3 has medium experience; others have high experience).



#2: Use Consistent Colours

prices and uncertainty in the predictions. We evaluated who are active in different parts of agrifood; collecting quantitative data. In particular, we focused on the following

RQ1 Usability : How user-friendly are the interaction in our visual DSS?

RQ2 Usefulness and needs : How useful is our visual the needs of people active in agrifood?

RQ3 Model understanding : How does visualising understanding of the prediction model underlying

RQ4 Trust : How does visualising uncertain prediction model underlying our visual DSS?

Our research contribution consists of extensively perspectives. First, considering our prototype as a pro

Article

Visually Explaining Uncertain Price Predictions in Agrifood: A User-Centred Case-Study

Jeroen Ooge* and Katrien Verbert

Department of Computer Science, KU Leuven, 3001 Leuven, Belgium; katrien.verbert@kuleuven.be
* Correspondence: jeroen.ooge@kuleuven.be; Tel.: +32-16-32-42-86

Abstract: The rise of ‘big data’ in agrifood has increased the need for decision support systems that harvest the power of artificial intelligence. While many such systems have been proposed, their uptake is limited, for example because they often lack uncertainty representations and are rarely designed in a user-centred way. We present a prototypical visual decision support system that incorporates price prediction, uncertainty, and visual analytics techniques. We evaluated our prototype with 10 participants who are active in different parts of agrifood. Through semi-structured interviews and questionnaires, we collected quantitative and qualitative data about four metrics: usability, usefulness and needs, model understanding, and trust. Our results reveal that the first three metrics can directly and indirectly affect appropriate trust, and that perception differences exist between people with diverging experience levels in predictive modelling. Overall, this suggests that user-centred approaches are key for increasing uptake of visual decision support systems in agrifood.

Keywords: visual analytics; visualisation; uncertainty; explainable artificial intelligence; decision support systems; mixed-methods; thematic analysis



Citation: Ooge, J.; Verbert, K. Visually Explaining Uncertain Price Predictions in Agrifood: A User-Centred Case-Study. *Agriculture* 2022, 12, 1024. <https://doi.org/10.3390/agriculture12071024>

Academic Editor: Maciej Zaborowicz

Received: 13 June 2022

Accepted: 9 July 2022

Published: 14 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

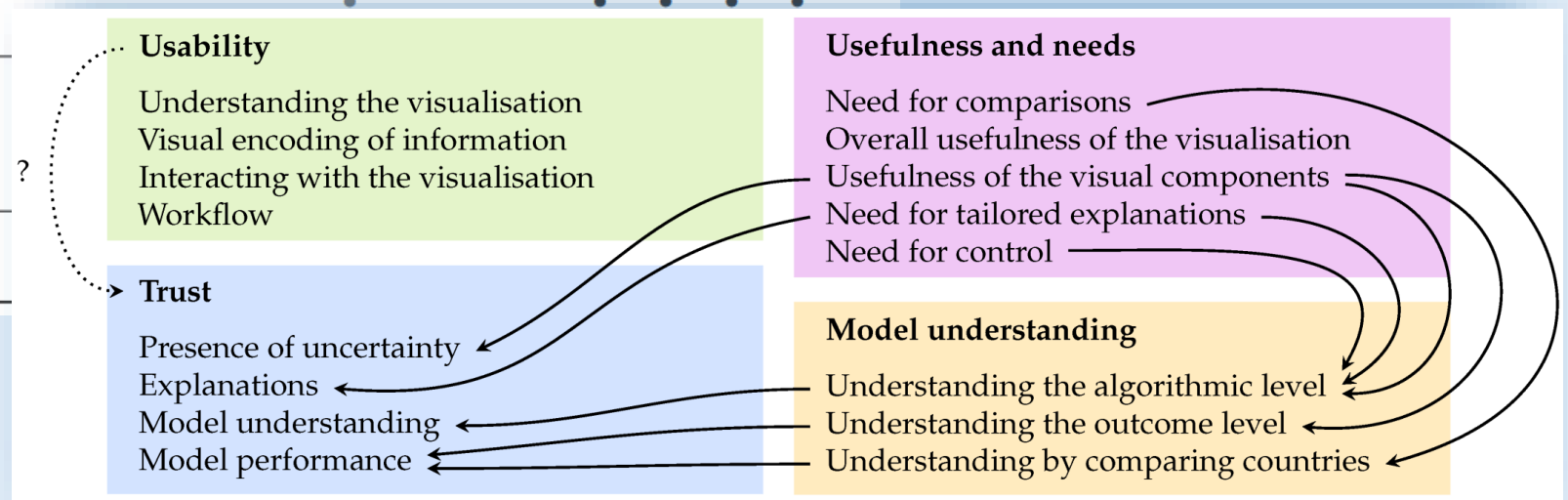
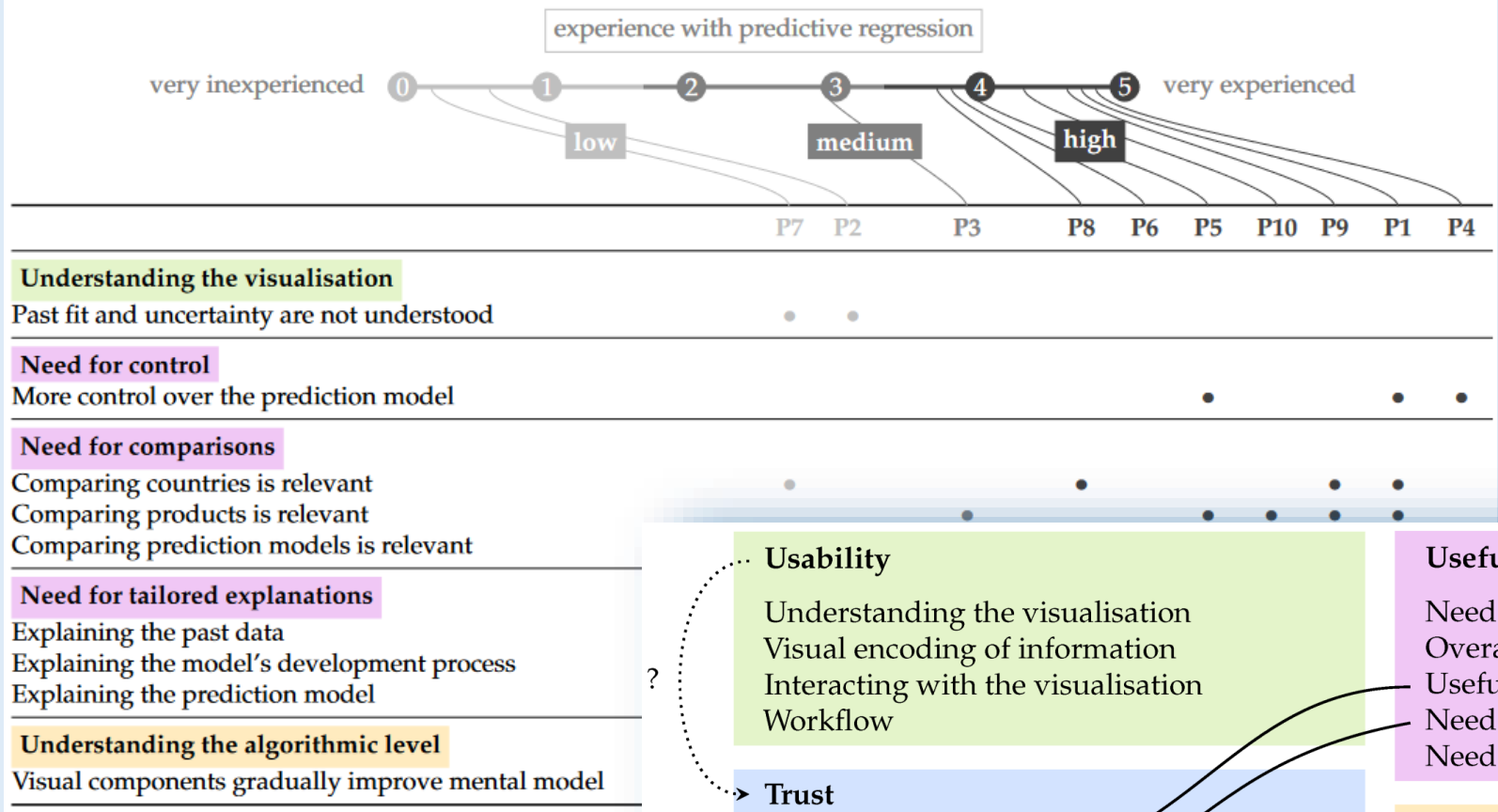
1. Introduction

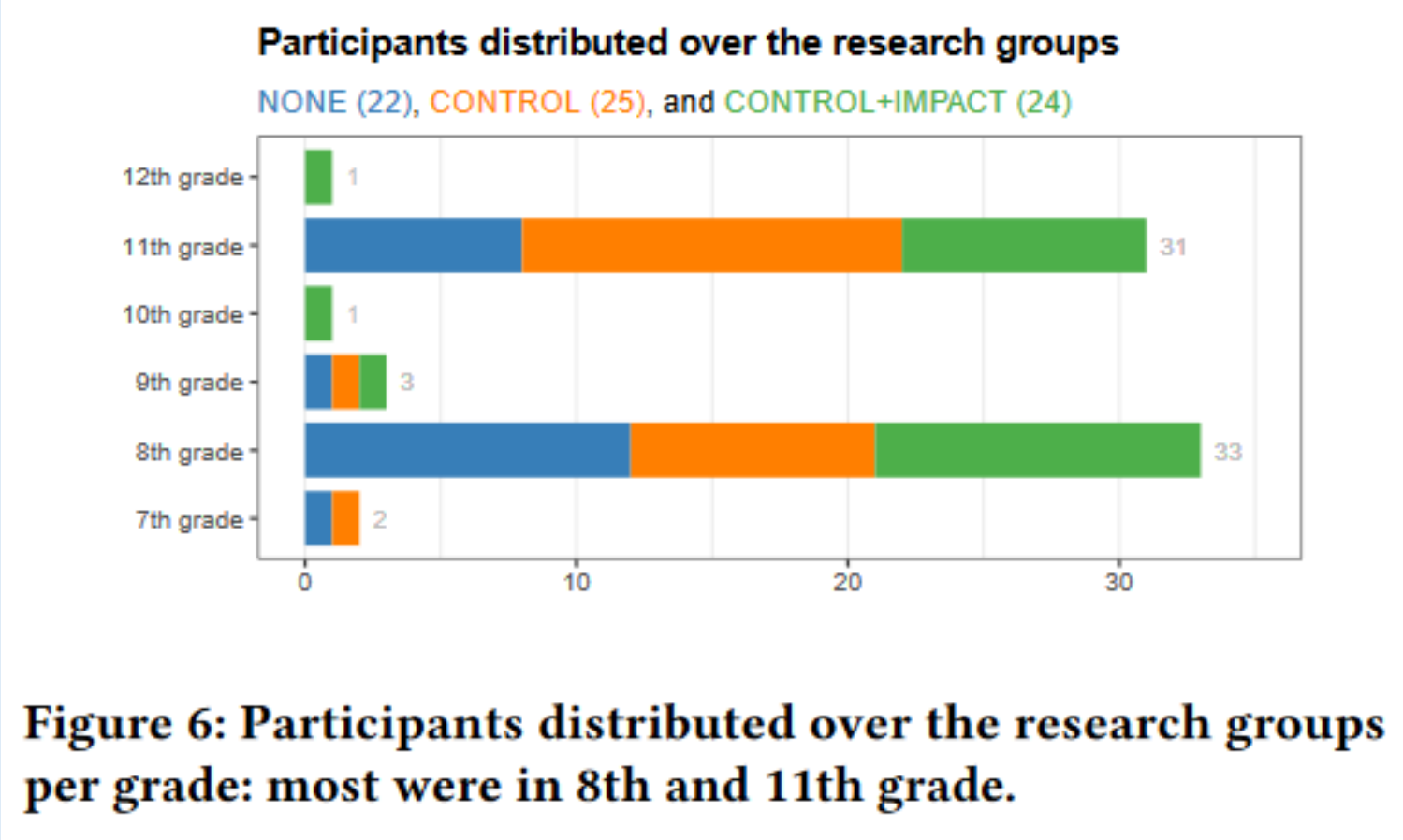
Under the impulse of success stories in other domains, artificial intelligence and ‘big data’ are on the rise in agrifood [1], leading to promising research directions such as *Agriculture 4.0* [2] and the broader *Agrifood 4.0* [3], *precision agriculture* [4–6], and *smart farming* [7–9]. While the adoption of such technologies is still modest in real-life agrifood applications [10], it is expected that the wide availability of cloud computing and remote sensing [11] will further boost their spread [12]. To process the explosive amount of information in this era of growing digitisation and to make data-grounded decisions, agrifood stakeholders increasingly need the assistance of *decision support systems* (DSSs) [2] that facilitate learning and allow to modify decision processes by integrating domain knowledge, rather than systems that merely prescribe actions [13,14].

Yet, even though the need for DSSs in agrifood has been acknowledged for over two decades [13] and many prototypes have been proposed [2,15], the uptake of these systems has been limited so far. Parker et al. [16,17], Zhai et al. [2], and Rose et al. [18] discussed several reasons for this low uptake: user interfaces of DSSs are not always user-friendly and lack visualisations, DSSs are not necessarily relevant when they do not meet end users’ needs or decision-making styles, outputs often miss uncertainty representations, and end users often distrust DSSs with opaque underlying algorithms. In other words, developers of DSSs for agrifood face important design challenges such as increasing usability, guarding usefulness for end users, and raising appropriate trust in underlying decision models.

Tackling these challenges requires human-centred approaches, which lie at the core of *human-computer interaction* (HCI), an interdisciplinary field that connects computer science, social sciences, and technology-applying domains such as agrifood. Specifically, HCI studies how interfaces can be designed and tailored to specific end users or application contexts to improve user experience, for example [19–21]. Two subdomains of HCI specialise in visualising complex information and explaining artificial intelligence, respectively.

Table 2. Some topics raised by the participants, ordered by their experience with predictive regression (P2 and P7 have low experience; P3 has medium experience; others have high experience).





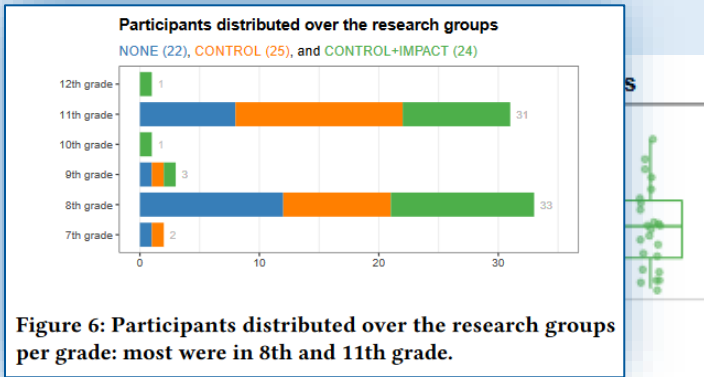


Figure 6: Participants distributed over the research groups per grade: most were in 8th and 11th grade.

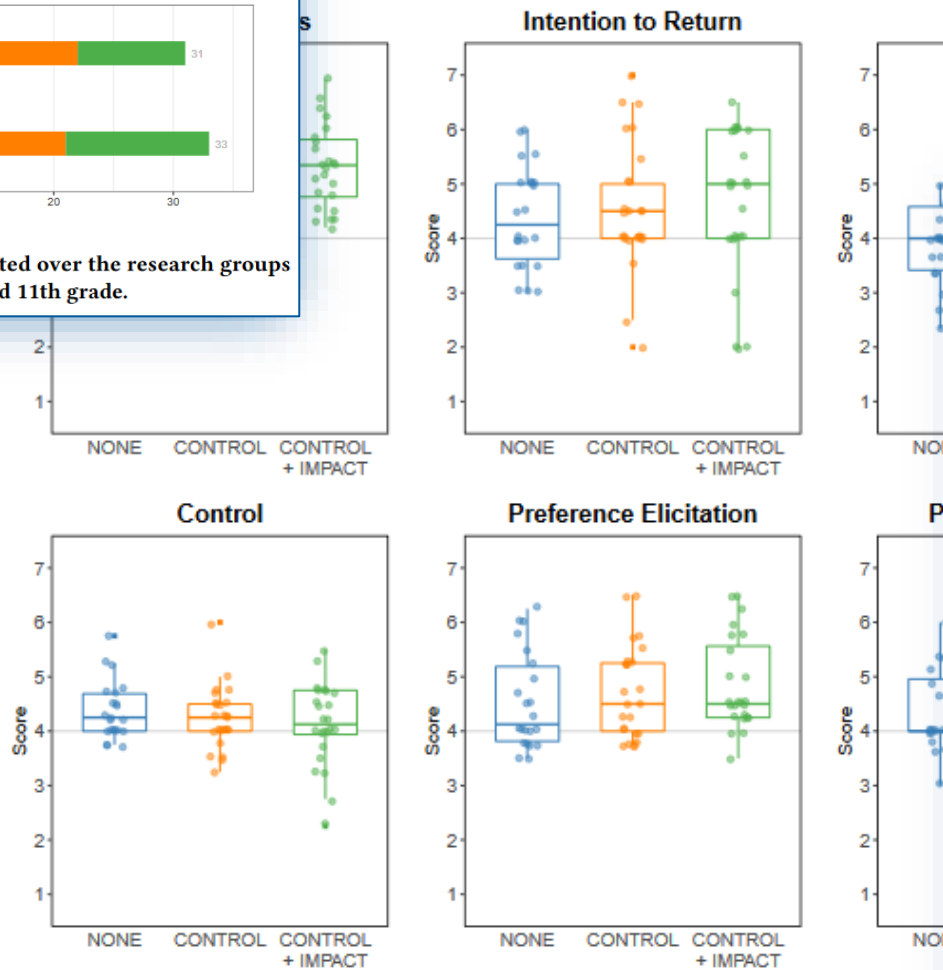


Figure 8: Box plots of the responses to the questionnaire in Table 2 for each group. The dot plots are slightly jittered horizontally and vertically.

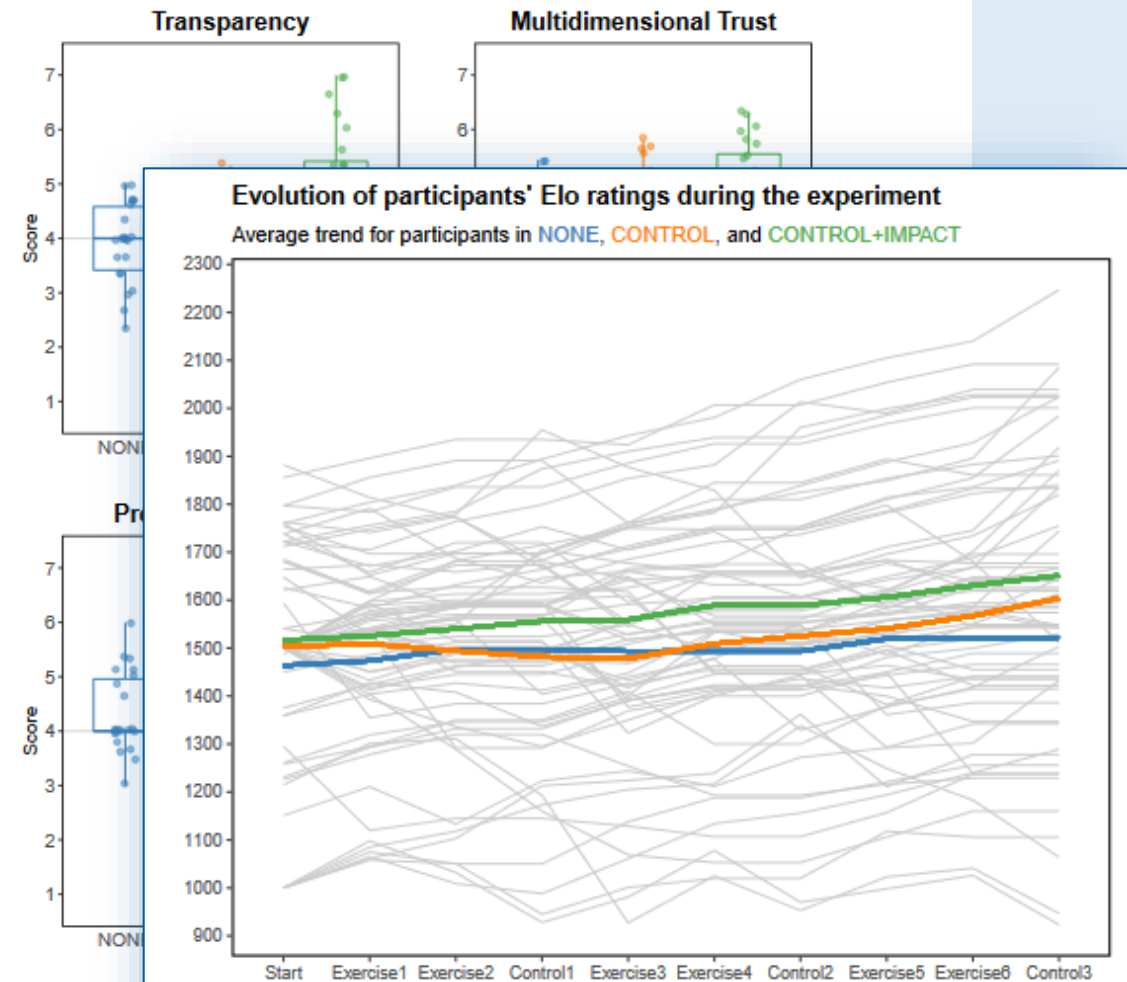


Figure 10: Evolution of participants' Elo ratings during the experiment and the average evolution per research group.

#3: Suppress Less Important Information

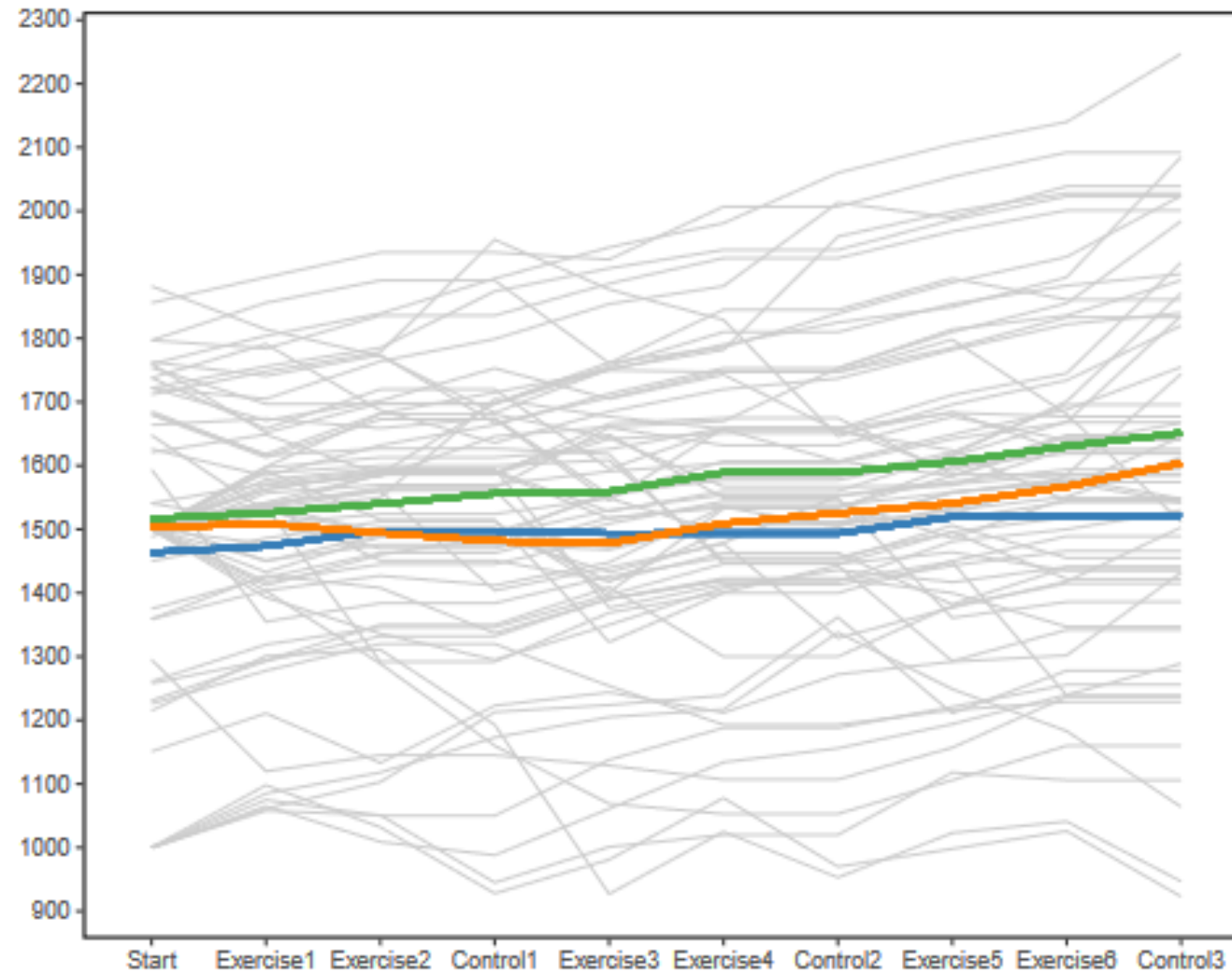
Table 1: Comparing the research groups with t-tests (Mann-Whitney U test for one-dimensional trust). Cells contain the effect sizes (second group mean minus first group mean).

	NONE VS. CONTROL	NONE VS. CONTROL+IMPACT	CONTROL VS. CONTROL+IMPACT
Benevolence	0.16 (<i>p</i> = 0.263)	0.61 (<i>p</i> = 0.011)	0.45 (<i>p</i> = 0.035)
Trusting beliefs	-0.01 (<i>p</i> = 0.529)	0.38 (<i>p</i> = 0.042)	0.40 (<i>p</i> = 0.030)
Transparency	0.29 (<i>p</i> = 0.068)	1.04 (<i>p</i> = 0.000)**	0.74 (<i>p</i> = 0.002)*
One-dimens. trust	0.00 (<i>p</i> = 0.504)	0.78 (<i>p</i> = 0.017)	0.78 (<i>p</i> = 0.020)
Multidimens. trust	0.15 (<i>p</i> = 0.207)	0.55 (<i>p</i> = 0.009)*	0.40 (<i>p</i> = 0.039)
Preference revision	0.33 (<i>p</i> = 0.080)	0.43 (<i>p</i> = 0.030)	0.10 (<i>p</i> = 0.325)

p* < 0.01, *p* < 0.001, non-significant results (*p* ≥ 0.5) are greyed out

Evolution of participants' Elo ratings during the experiment

Average trend for participants in NONE, CONTROL, and CONTROL+IMPACT



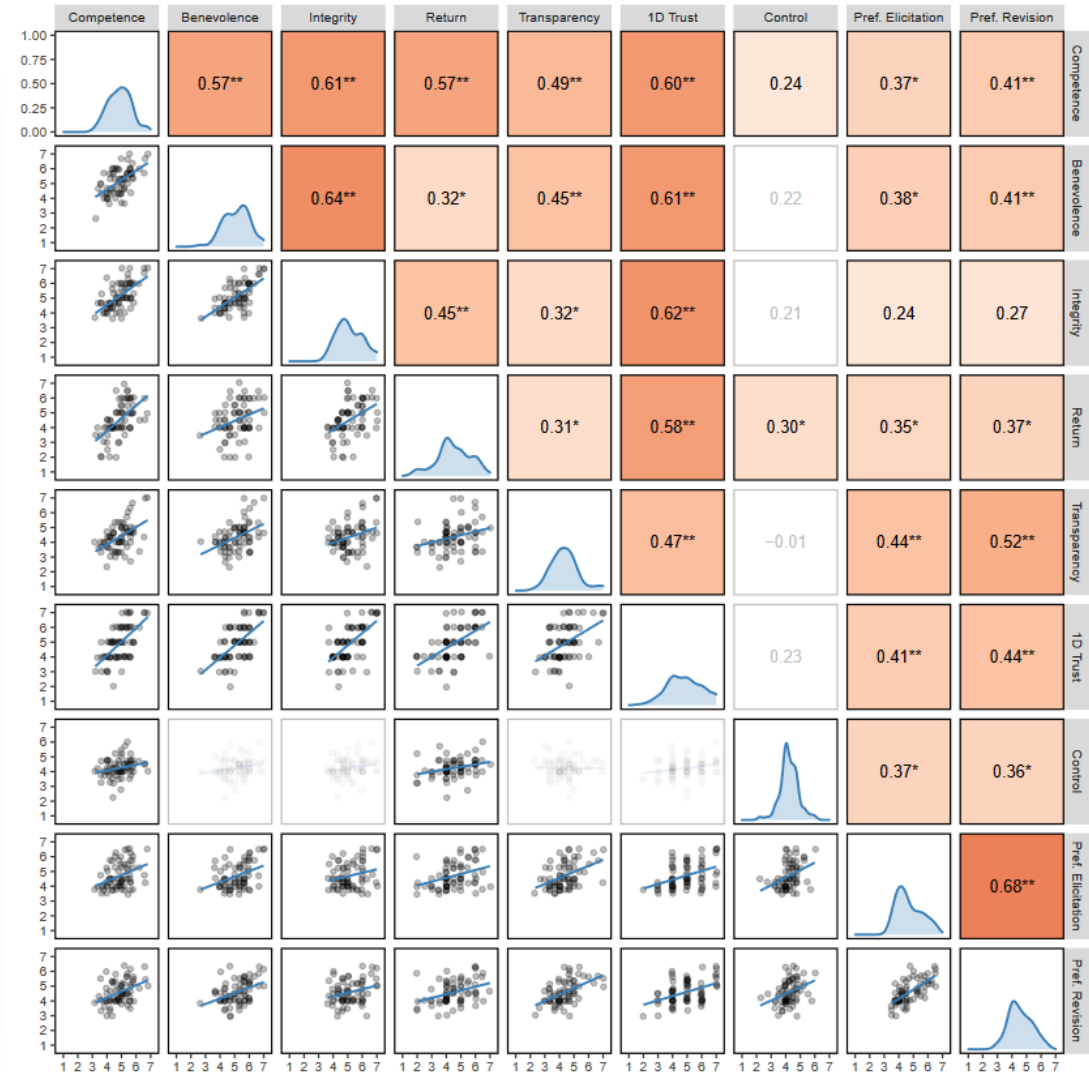
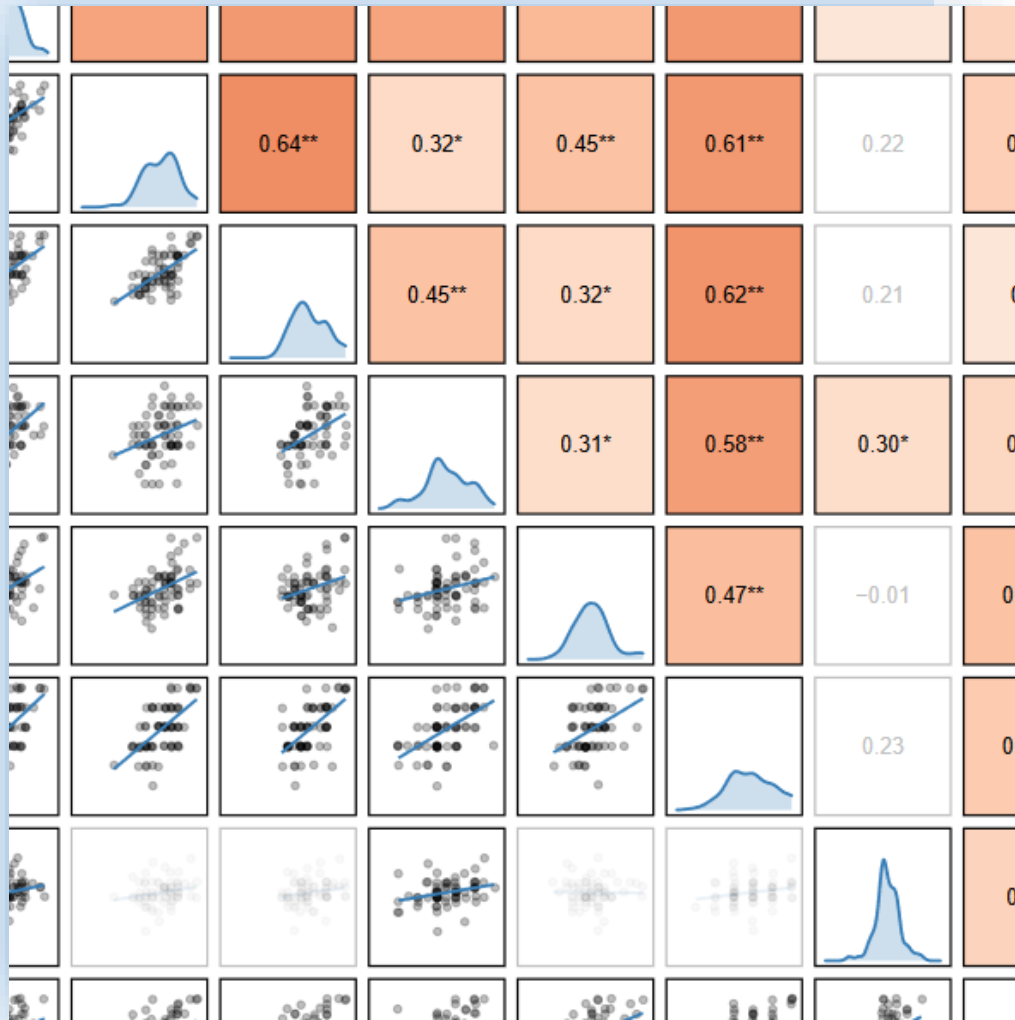


Figure 9: Relations between trust-related and control-related constructs. Lower triangle: dot plots with fitted regression lines. Diagonal: density plot of constructs. Upper triangle: correlations colour-coded by value (* $p < 0.01$, ** $p < 0.001$). Non-significant relations ($p \geq 0.05$) are greyed out.

#4: Show the Actual Data

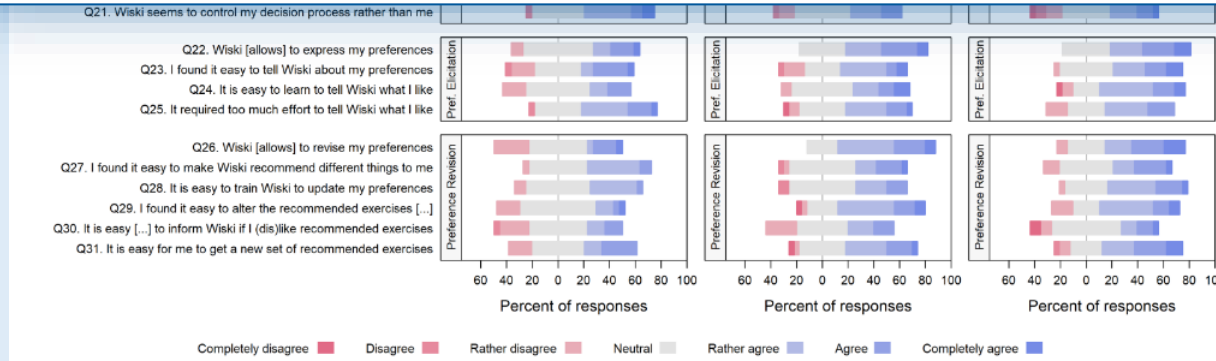
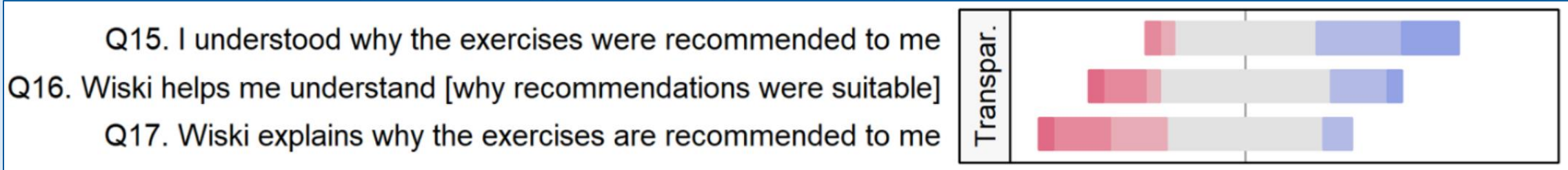
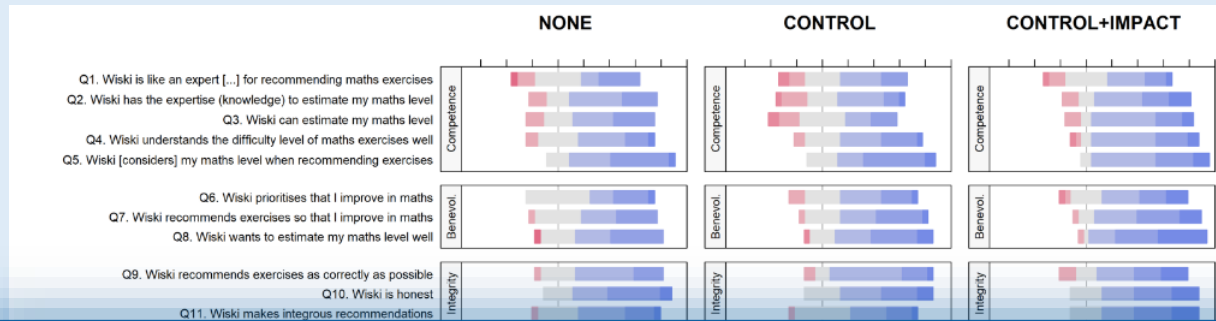


Figure 7: Diverging bar charts [36] of responses to the questionnaire in Table 2 after reverse-scoring, comparing the three research groups. Questions have been abbreviated for brevity and have been grouped per construct for clarity.

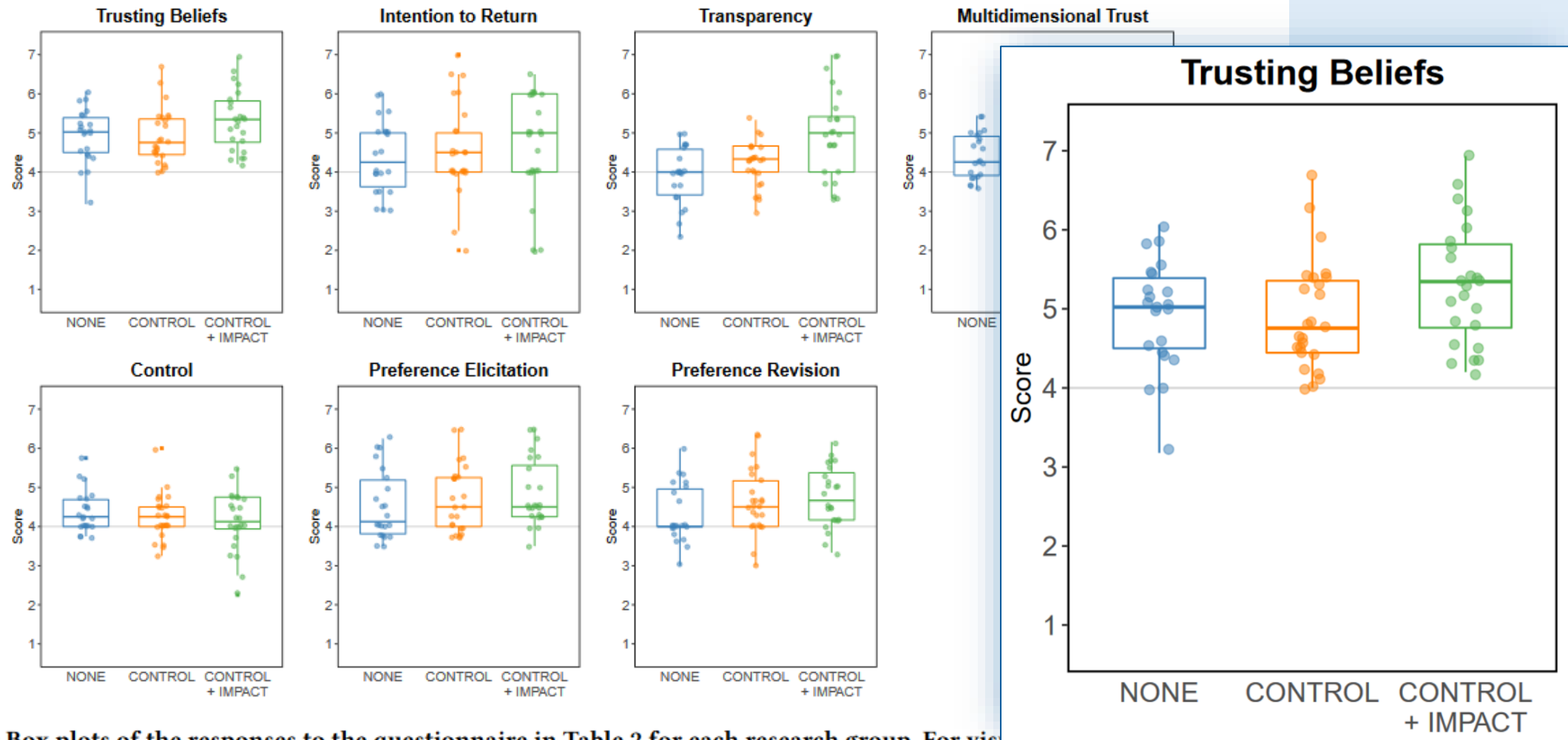
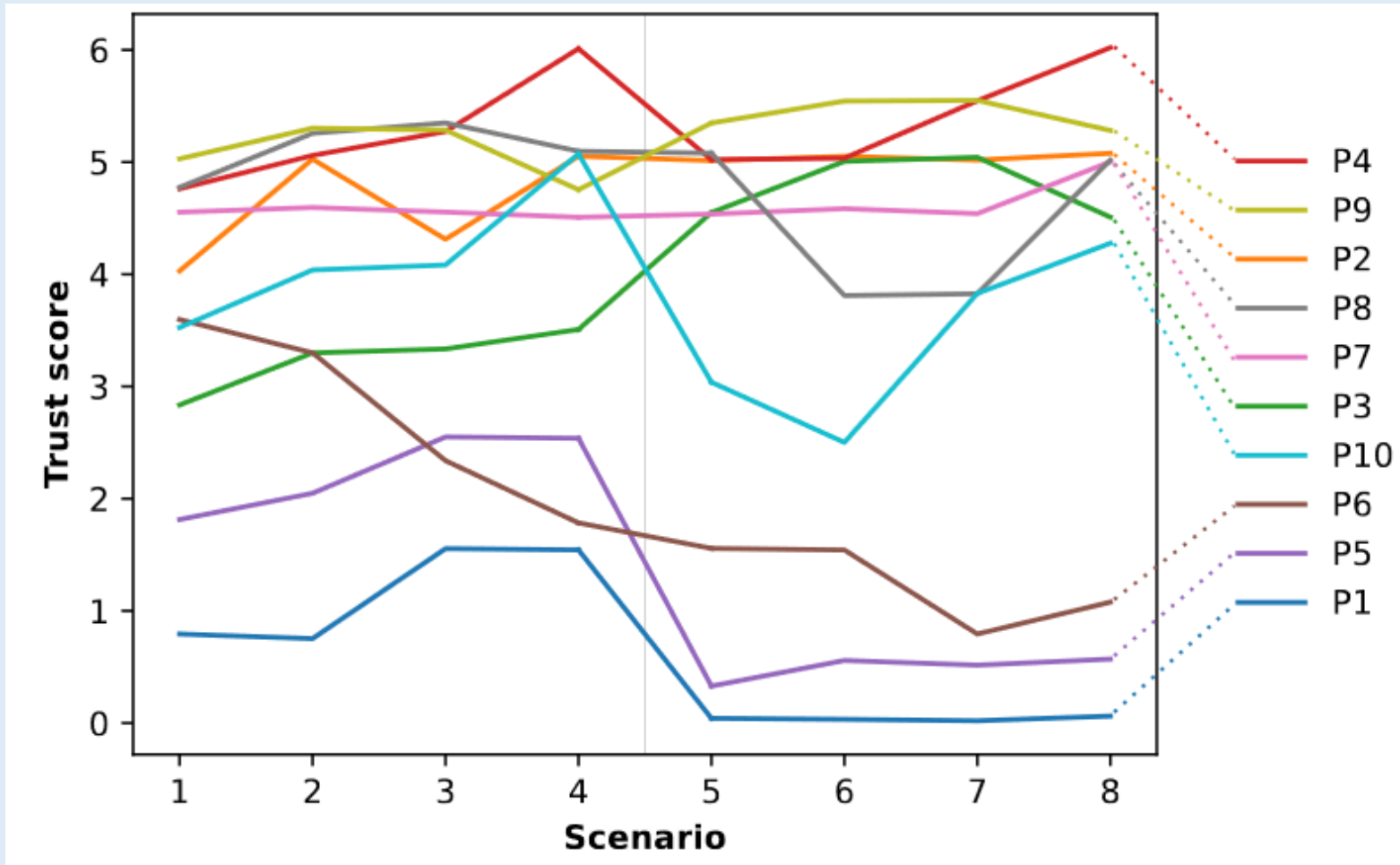


Figure 8: Box plots of the responses to the questionnaire in Table 2 for each research group. For visual clarity, the box plots and dot plots are slightly jittered horizontally and vertically.

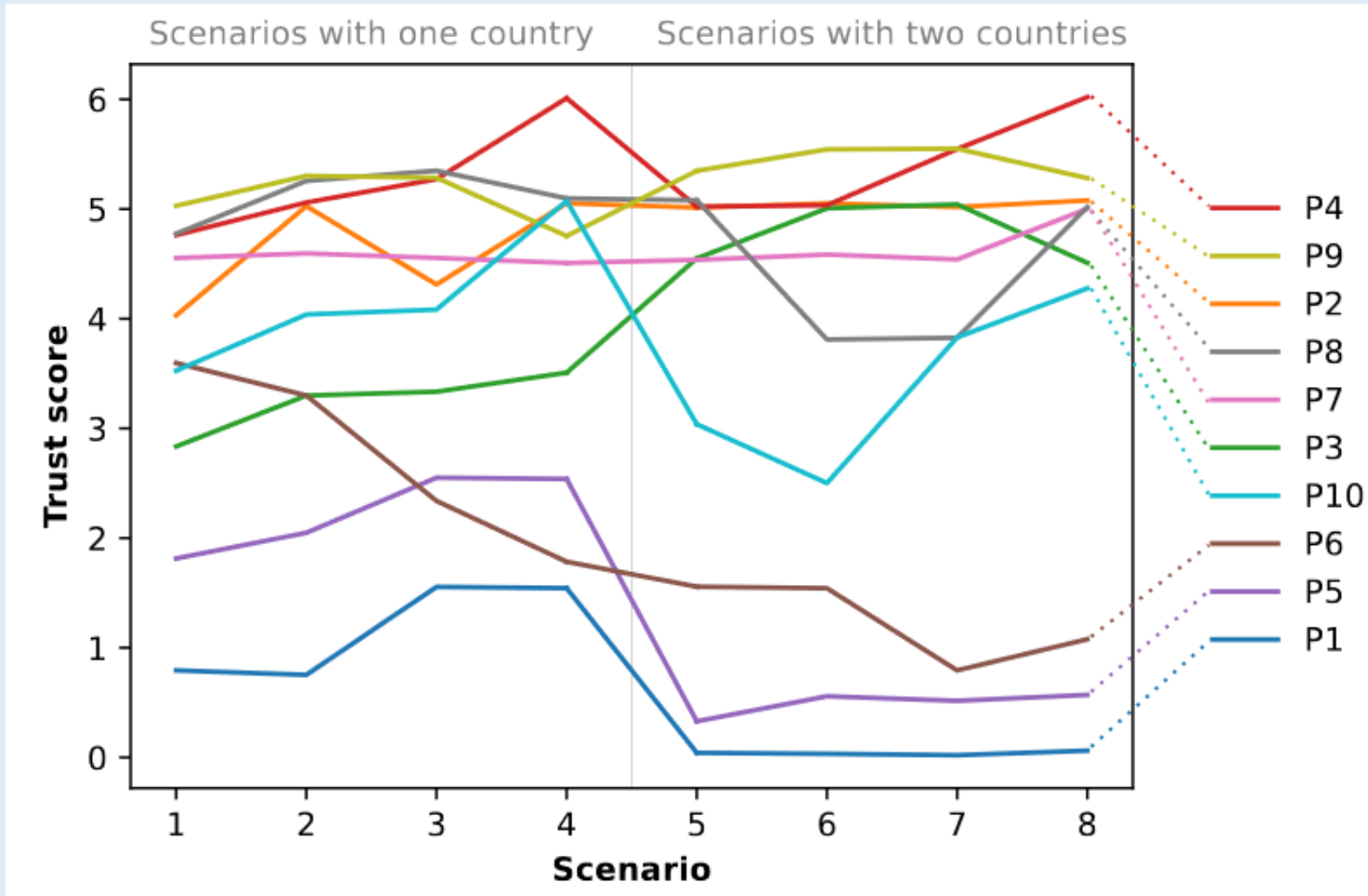
Jeroen Ooge, Leen Dereu, and Katrien Verbert. 2023. Steering Recommendations and Visualising Its Impact: Effects on Adolescents' Trust in E-Learning Platforms. In *Proceedings of the 28th International Conference on Intelligent User Interfaces (IUI '23)*, 156–170. <https://doi.org/10.1145/3581641.3584046>

Nick Desbarats. 2021. I've Stopped Using Box Plots. Should You? <https://nightingaledvs.com/ive-stopped-using-box-plots-should-you/>

#5: Enrich Visualisations Meaningfully



#5: Enrich Visualisations Meaningfully



#5: Enrich Visualisations Meaningfully

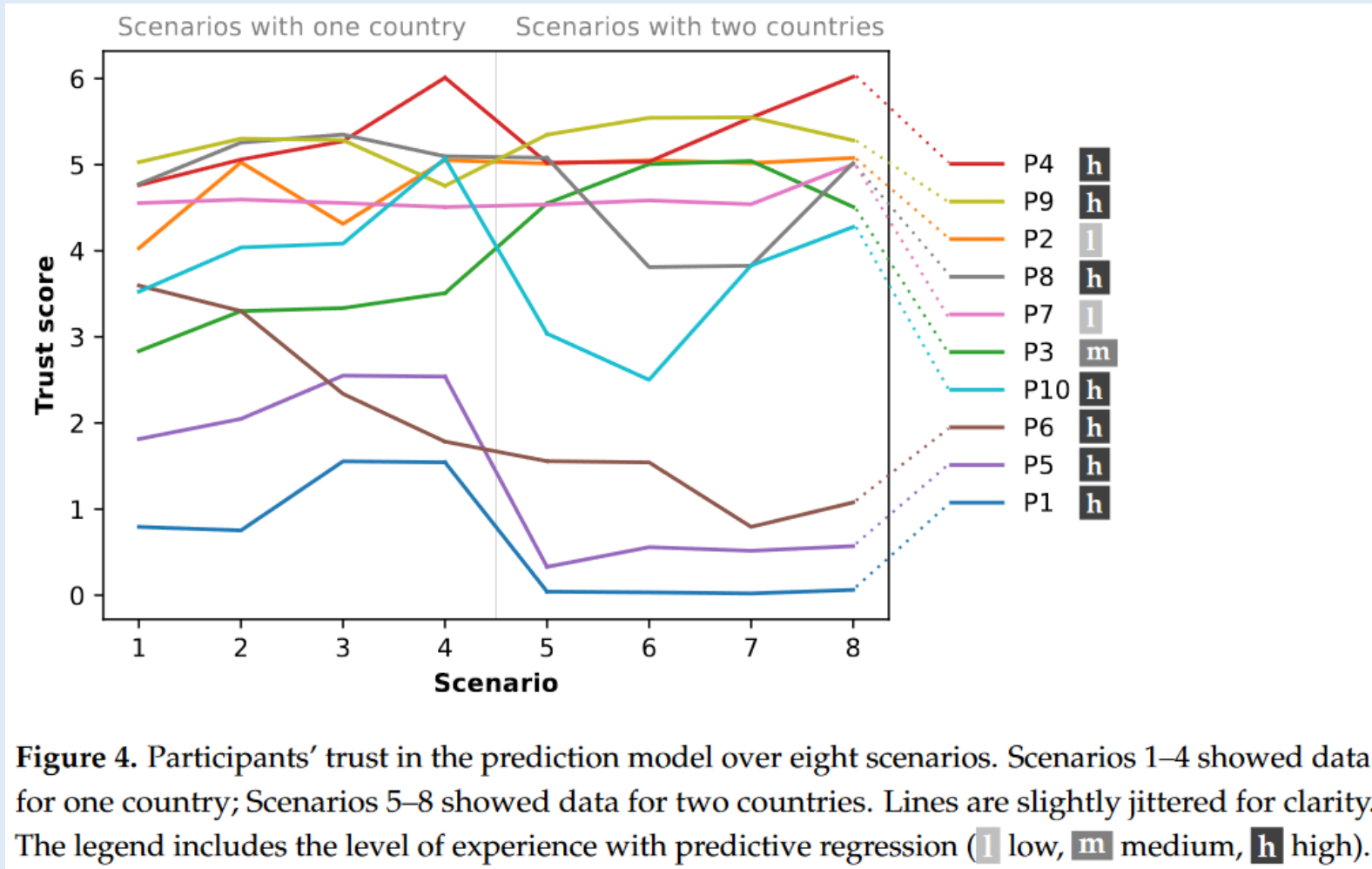
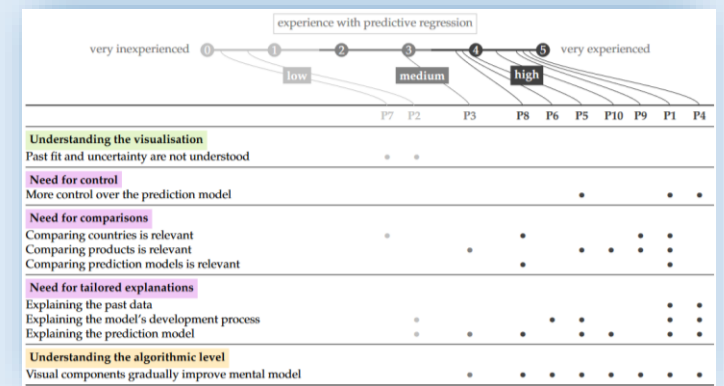
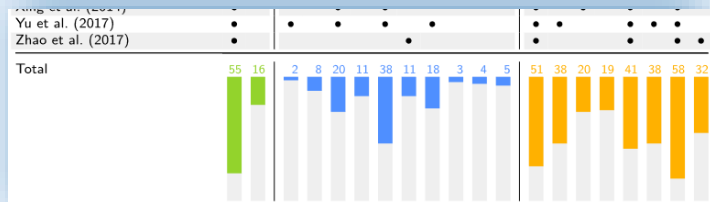
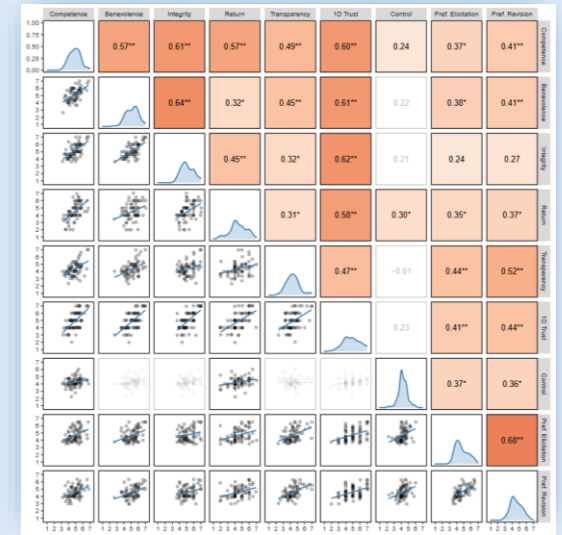
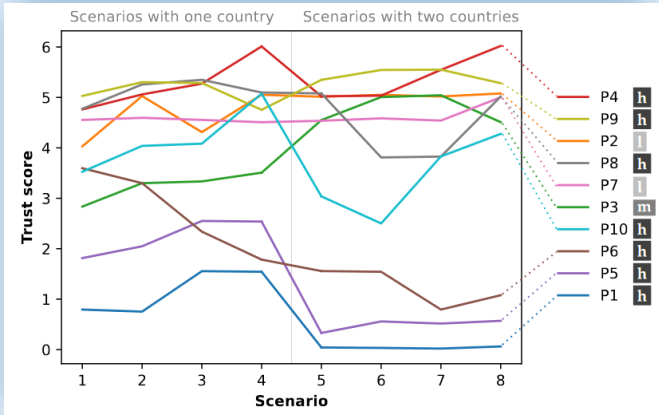
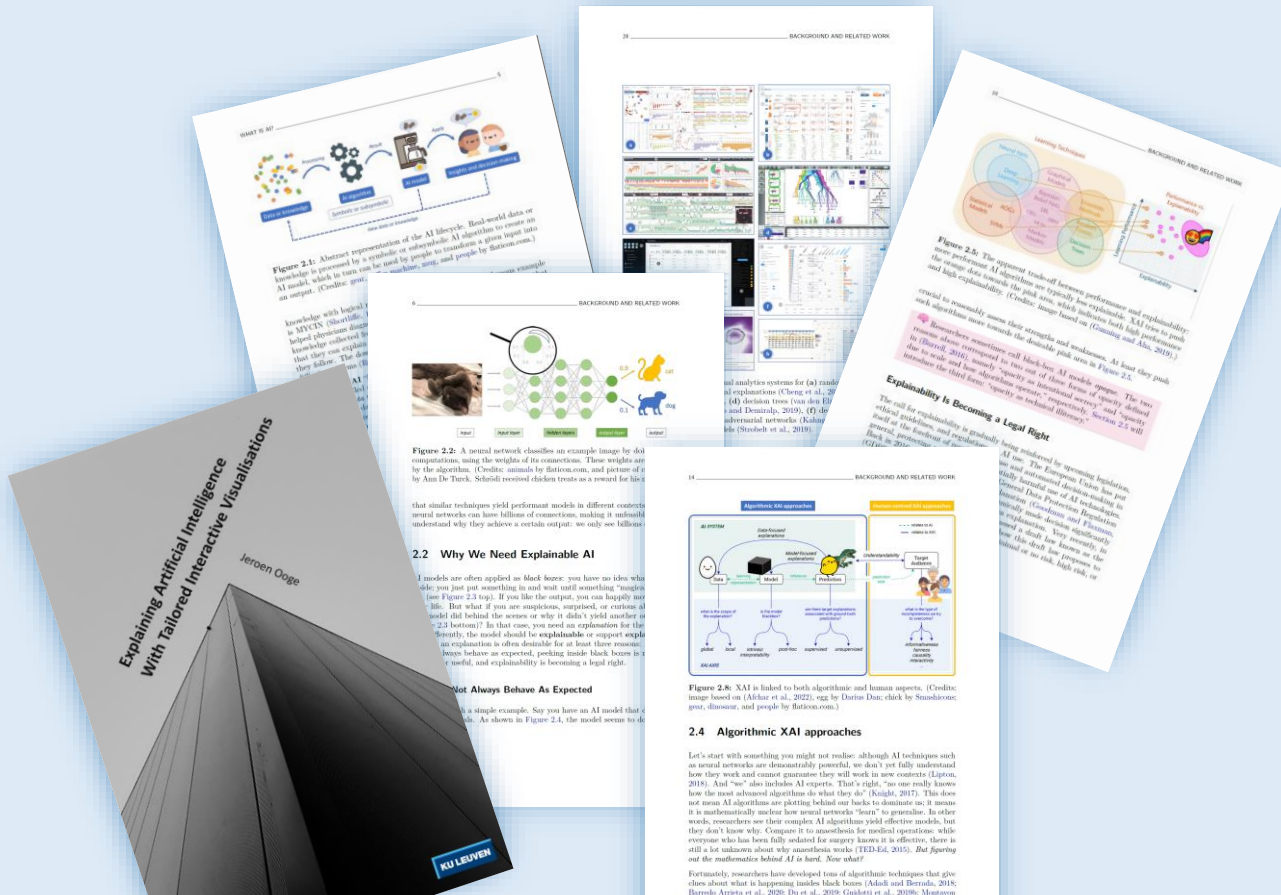


Figure 4. Participants' trust in the prediction model over eight scenarios. Scenarios 1–4 showed data for one country; Scenarios 5–8 showed data for two countries. Lines are slightly jittered for clarity. The legend includes the level of experience with predictive regression (l low, m medium, h high).

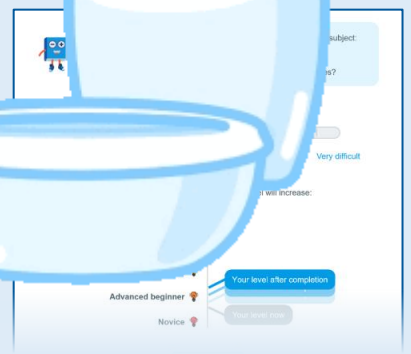
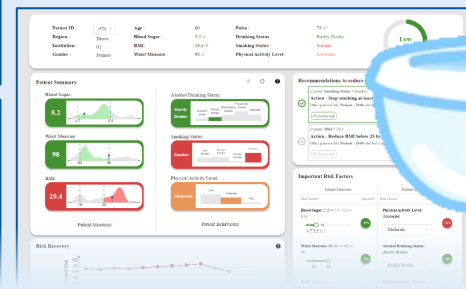
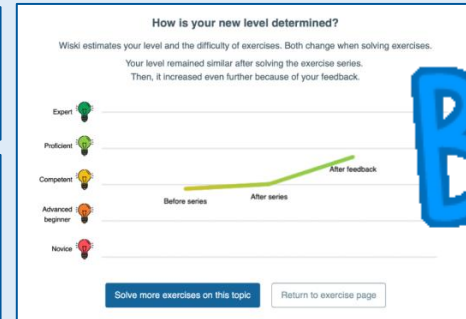
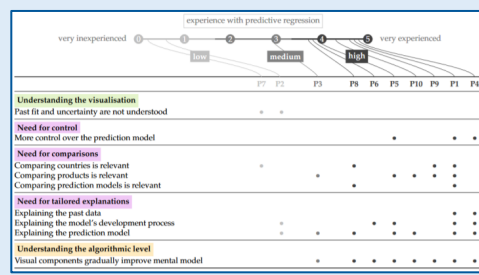
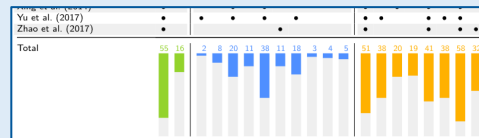
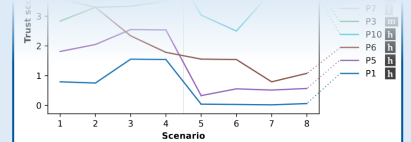
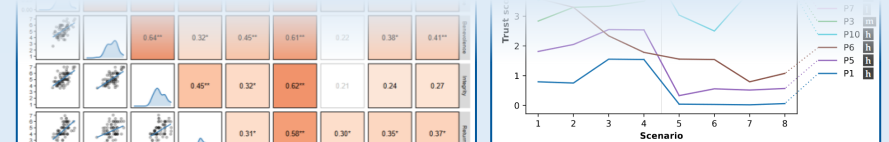
Break for Questions



Check out jeroenooge.be for regular research updates and more!



Interactive Visualisations for Explainable AI Research and Research Practices



Jeroen Ooge
jeroenooge.be
Universiteit Utrecht