



From Data to Dialog: Visualization-empowered Human-AI Collaboration



Xinhuan Shu (<https://shuxinhuan.github.io>)

Lecturer @ Newcastle University

Data Visualization  / Human-Computer
Interaction  / Data-driven Storytelling 

hi buna ziua
hola hello
ciao ola tere
你好 Γεια σας hej
hallo witam أهلا
hei سلام 안녕하세요
bonjour sveiki
こんにちは
здравствуйте

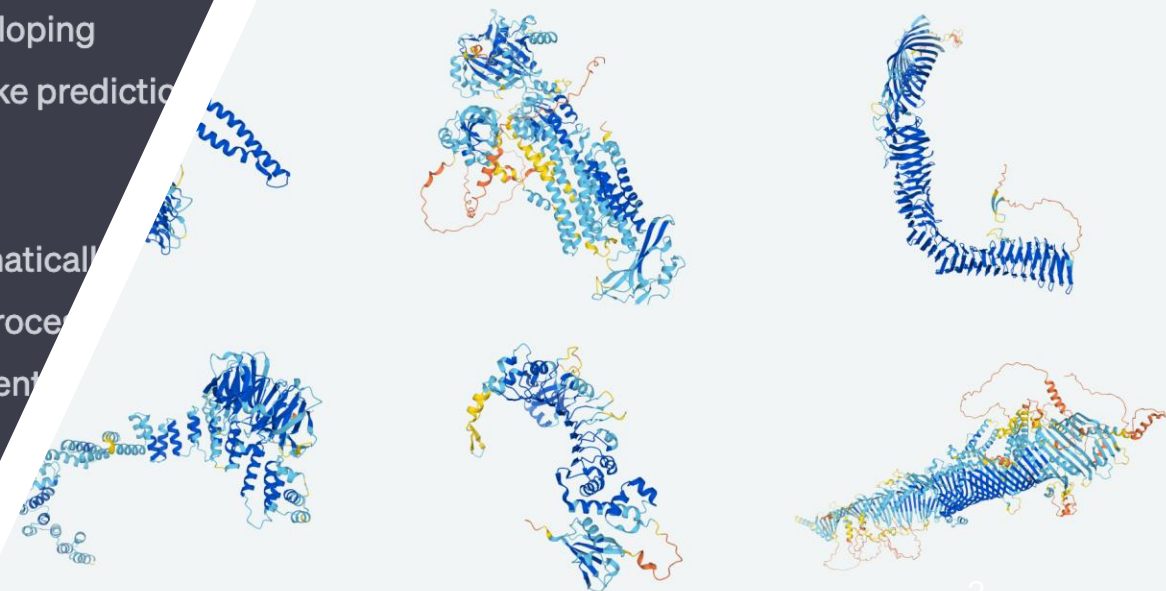




Machine learning is a subfield of artificial intelligence (AI) that focuses on developing algorithms and statistical models that enable computers to learn from and make predictions or decisions based on data, without being explicitly programmed.

In other words, machine learning involves the use of algorithms that can automatically improve their performance by learning from data. This is achieved through a process of training, in which a computer program is given a large dataset and uses it to identify patterns and relationships within the data.

Once the algorithm has been trained, it can be used to make predictions or decisions on new data that it has not seen before, based on the patterns and relationships it has learned.



How does AI work?

Input

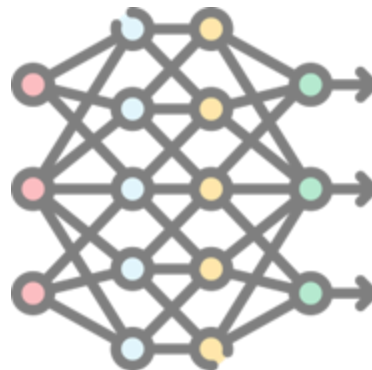
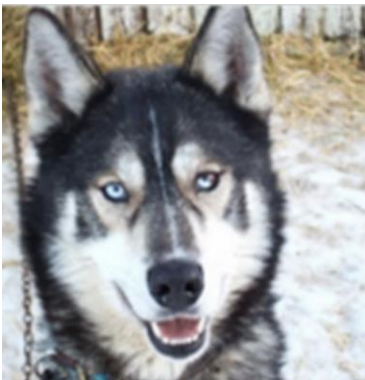


Model



Output

Predicted:
husky ✓



Predicted:
wolf ✗

Opinion

OP-ED CONTRIBUTOR

How to Make A.I. That's Good for People

By Fel-Fel Li

March 7, 2018

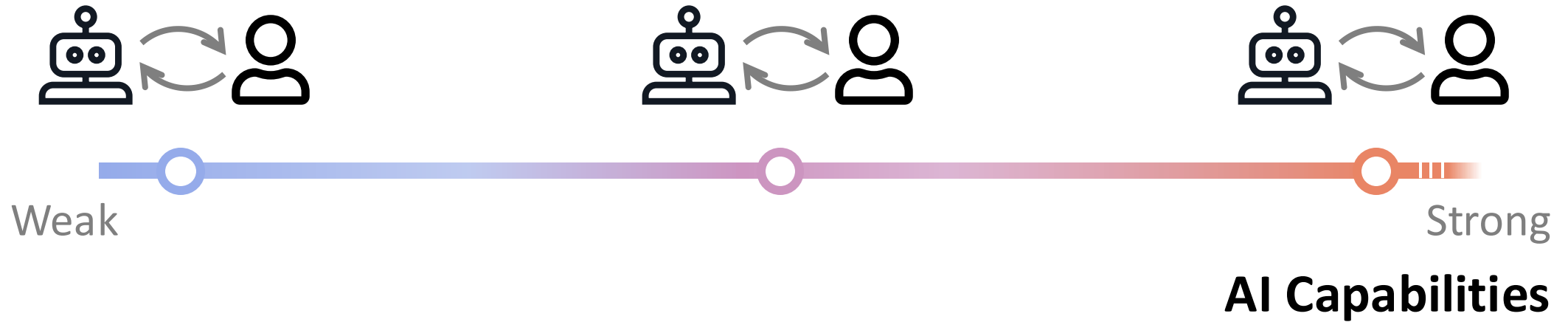
Give this article

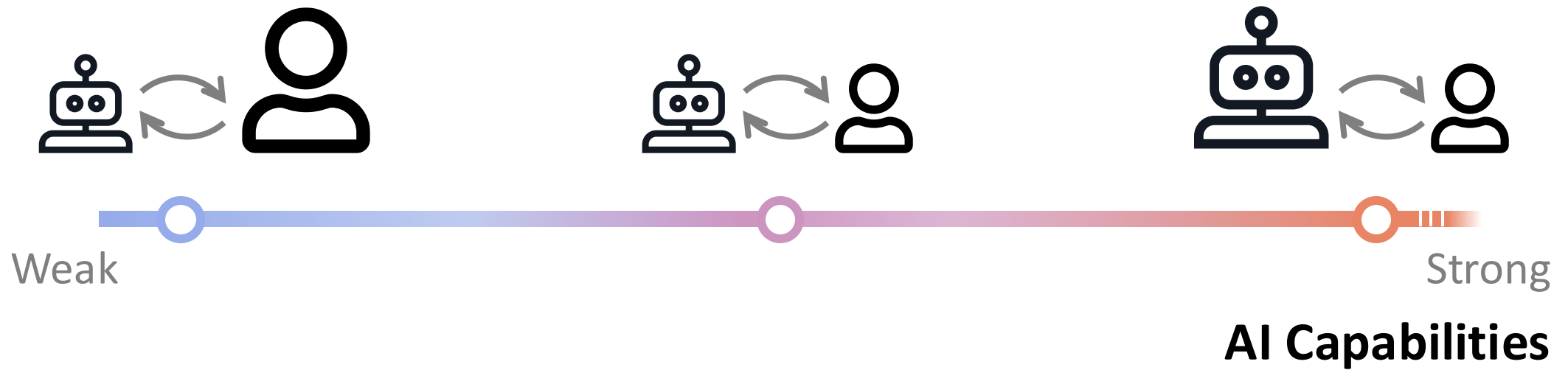


Elisa Macellari

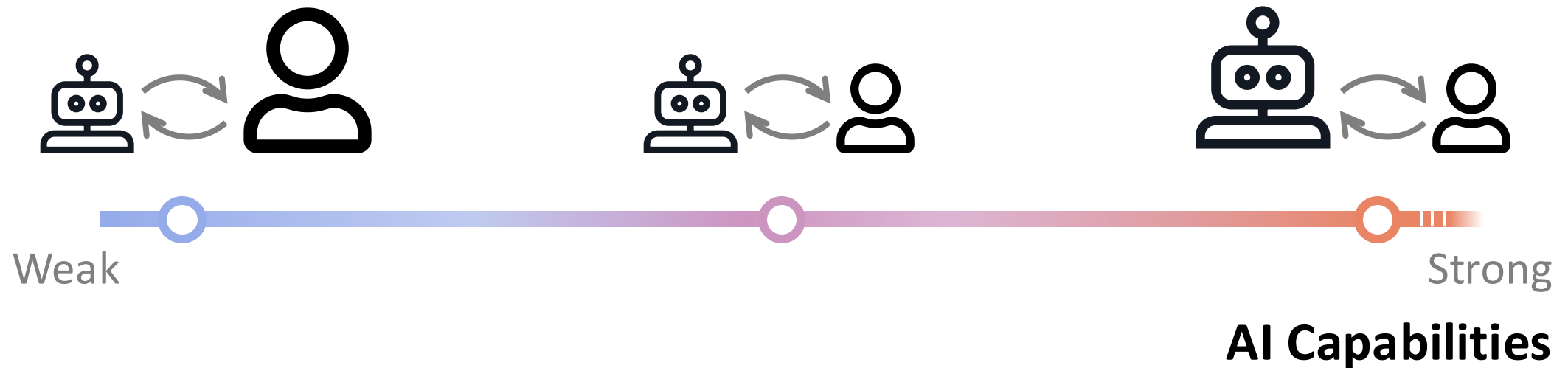
I worry, however, that enthusiasm for A.I. is preventing us from reckoning with its looming effects on society. Despite its name, there is nothing “artificial” about this technology — it is made by humans, intended to behave like humans and affects humans. So **if we want it to play a positive role in tomorrow’s world, it must be guided by human concerns.**

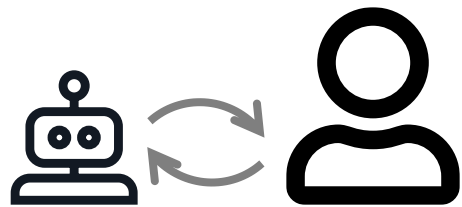
I call this approach **“human-centered A.I.”** It consists of three goals that can help responsibly guide the development of intelligent machines.





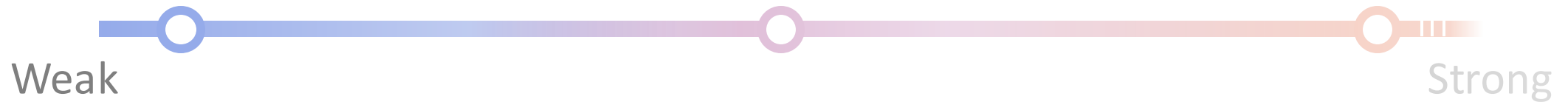
My research explores how we can design visualization-empowered Human-AI interactions that support diverse data tasks.





AI: distill knowledge

Humans: expert in making decisions

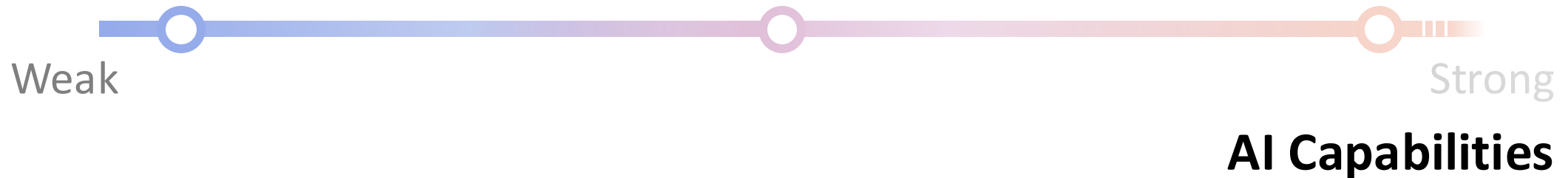
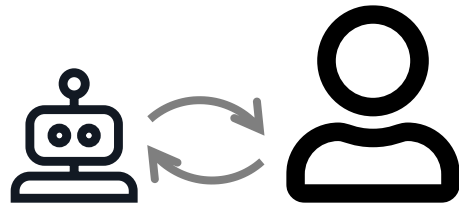


AI Capabilities

Emotion-Oriented Visual Summarization of Classroom Videos

TVCG'20

Haipeng Zeng, Xinhuan Shu, Yanbang Wang, Yong Wang, Liguozhang, Ting-Chuen Pong, Huamin Qu



Emotion analysis in classroom videos



Exploring students' emotion can help both teachers and parents know about students' learning status and further help teachers improve teaching.

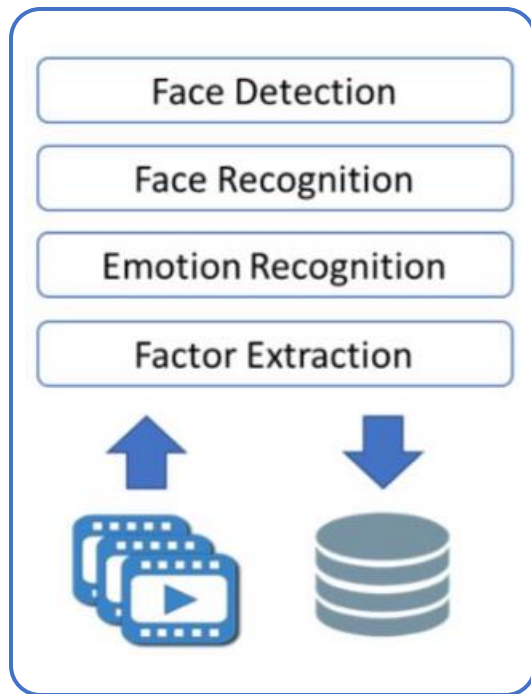
It is not easy for teachers to quickly capture and explore many students' emotions in the classroom.

How to carry out **efficient**, **informative**, **reliable** emotion analysis?

AI

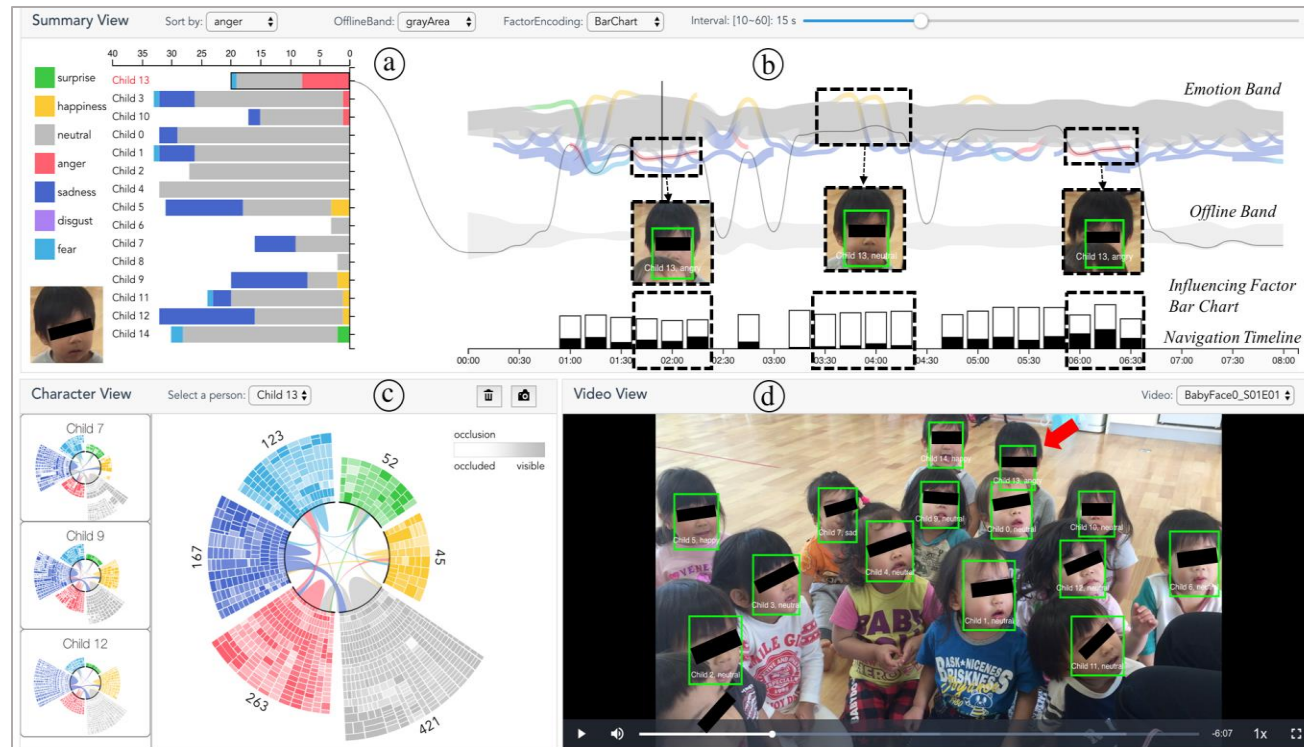
VIS

Human



AI

AI recognizes emotions of individual students, but the model output has **uncertainties**.



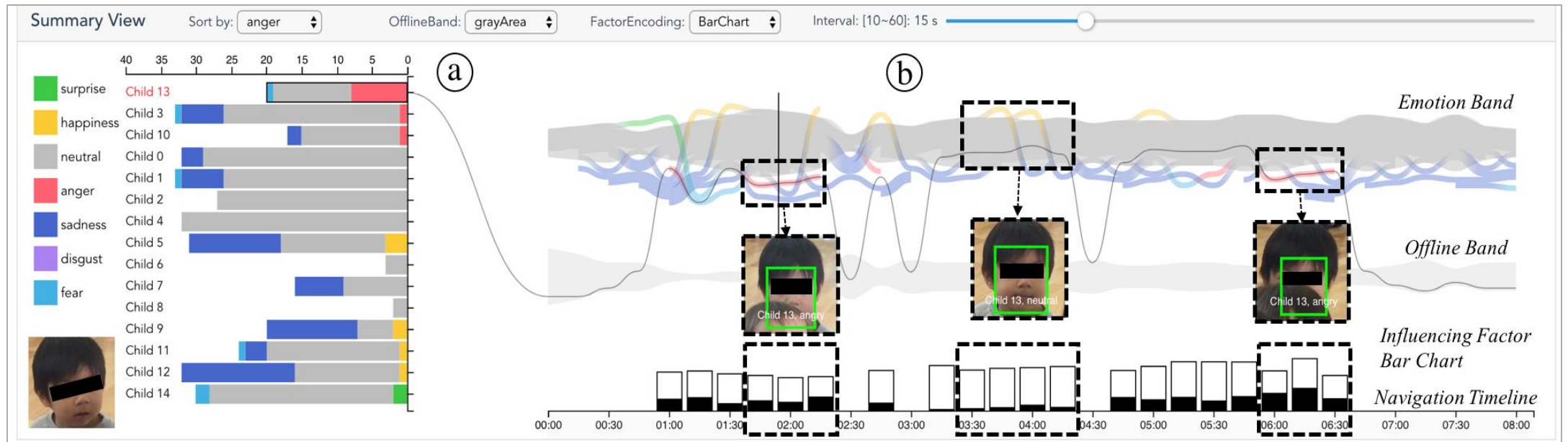
VIS

Visualization enhances understanding for insight discovery

Human

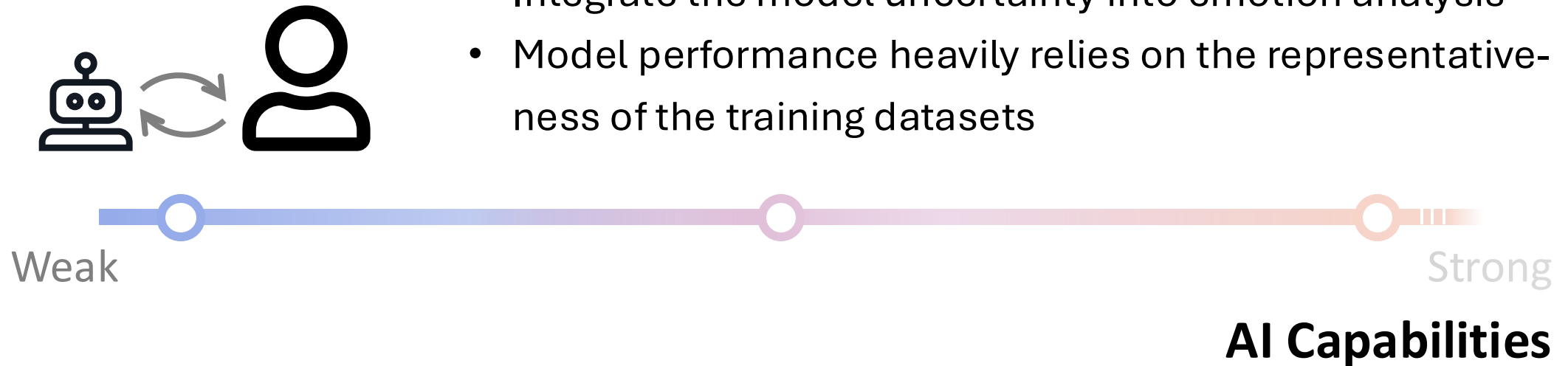
Teachers explore students' engagement status

Misidentified anger emotion

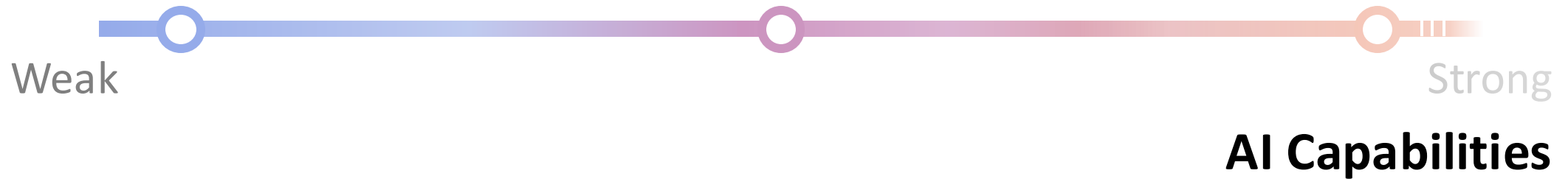


Lessons Learned

- Integrate the model uncertainty into emotion analysis
- Model performance heavily relies on the representativeness of the training datasets



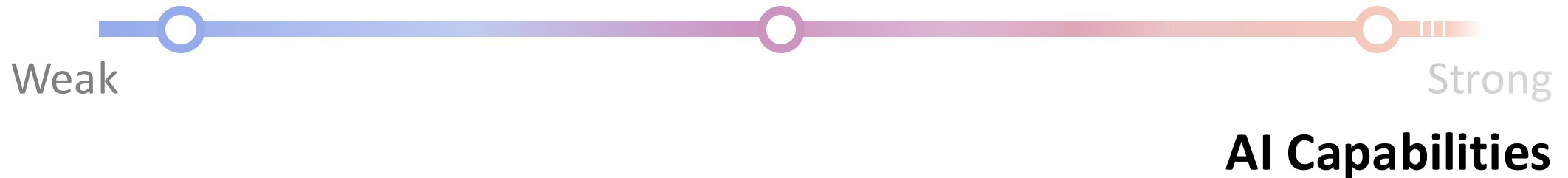
Human-AI Co-creation of visualizations



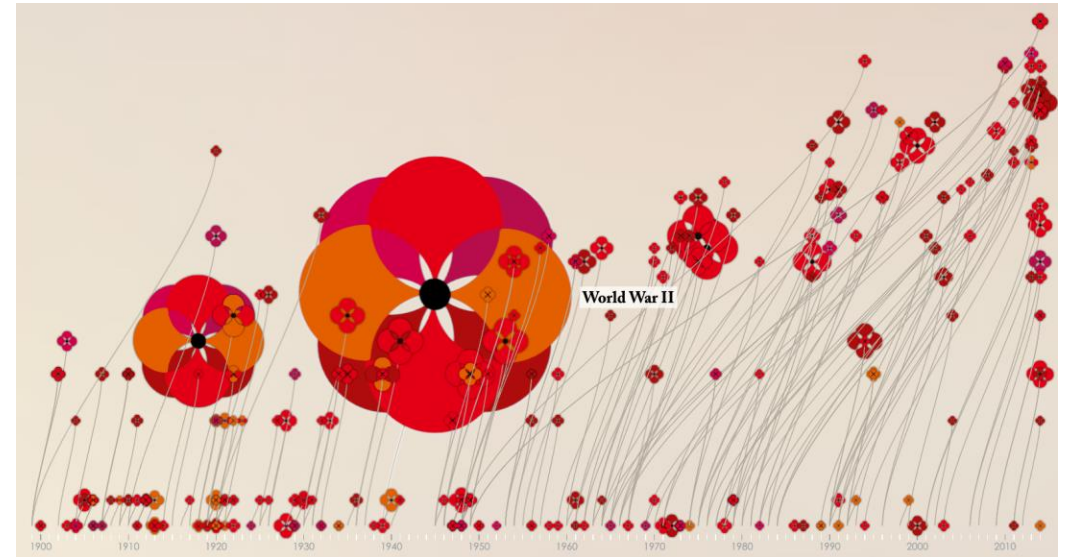
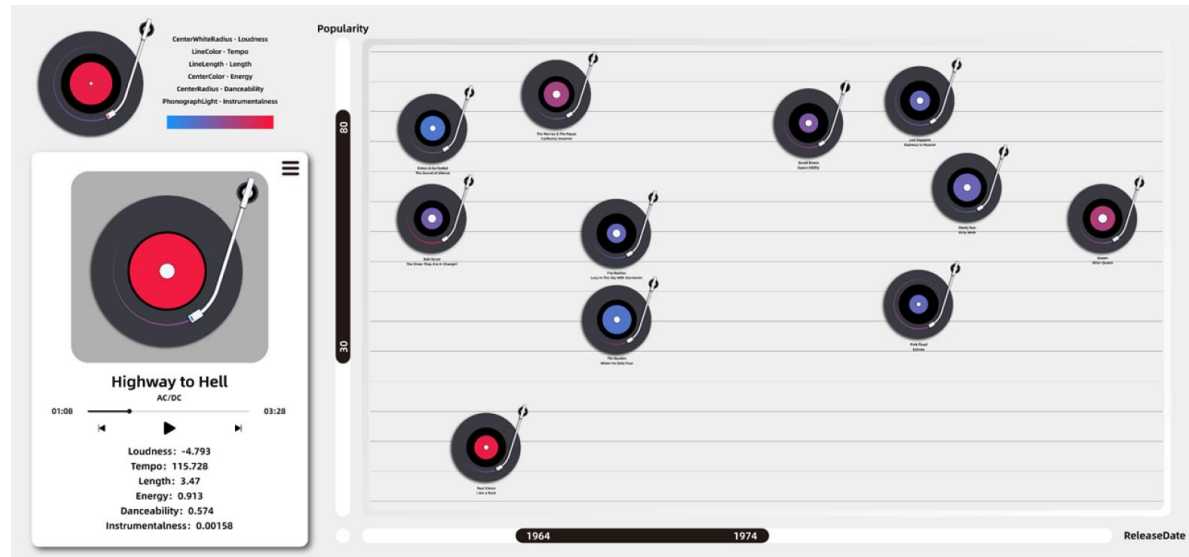
Automatic Generation of Metaphoric Glyph-based Visualization



Lu Ying, Xinhuan Shu, Dazhen Deng, Yuchen Yang, Tan Tang, Lingyun Yu, Yingcai Wu



Metaphoric Glyph-based Visualization (MGV)



How to Create MGVs?

Entry number	Primary Type	Secondary Type	Base Attack	Base Defense	Base HP	Special Attack	Special Defense	Base Speed	Height of Objects	Weight of Objects
70	grass	poison	90	50	65	85	45	55	6.4	1
49	bug	poison	65	60	70	90	75	90	12.5	1.5
83	normal	flying	90	55	52	58	62	60	15	0.8
71	grass	poison	105	65	80	100	70	70	15.5	1.7
45	grass	poison	80	85	75	110	90	50	18.6	1.2
48	bug	poison	55	50	60	40	55	45	30	1
12	bug	flying	45	50	60	90	80	70	32	1.1
22	normal	flying	90	65	65	61	61	100	38	1.2
84	normal	flying	85	45	35	35	35	75	39.2	1.4
94	ghost	poison	65	80	60	170	95	130	40.5	1.5
72	water	poison	40	35	40	50	100	70	45.5	0.9
62	water	fighting	95	95	90	70	90	70	54	1.3
42	poison	flying	80	70	75	65	75	90	55	1.6
34	poison	ground	102	77	81	85	75	85	62	1.4

Pokemon data

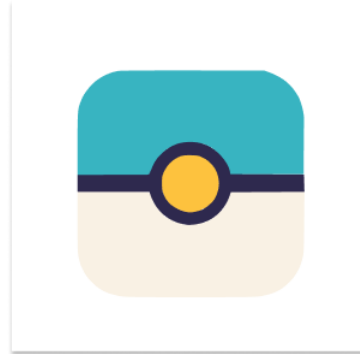


How to Create MGVs?

Challenge #1: *How to select an appropriate visual metaphor?*

Entry number	Primary Type	Secondary Type	Base Attack	Base Defense	Base HP	Special Attack	Special Defense	Base Speed	Height of Objects	Weight of Objects
70	grass	poison	90	50	65	85	45	55	6.4	1
49	bug	poison	65	60	70	90	75	90	12.5	1.5
83	normal	flying	90	55	52	58	62	60	15	0.8
71	grass	poison	105	65	80	100	70	70	15.5	1.7
45	grass	poison	80	85	75	110	90	50	18.6	1.2
48	bug	poison	55	50	60	40	55	45	30	1
12	bug	flying	45	50	60	90	80	70	32	1.1
22	normal	flying	90	65	65	61	61	100	38	1.2
84	normal	flying	85	45	35	35	35	75	39.2	1.4
94	ghost	poison	65	80	60	170	95	130	40.5	1.5
72	water	poison	40	35	40	50	100	70	45.5	0.9
62	water	fighting	95	95	90	70	90	70	54	1.3
42	poison	flying	80	70	75	65	75	90	55	1.6
34	poison	ground	102	77	81	85	75	85	62	1.4

Pokemon data



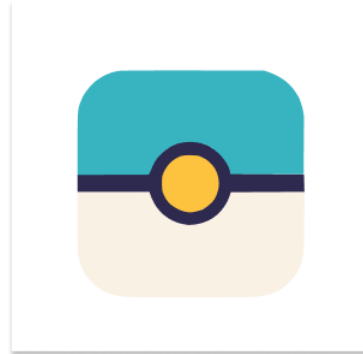
How to Create MGVs?

Challenge #2: *How to embed metaphors into glyph-based visualization design?*

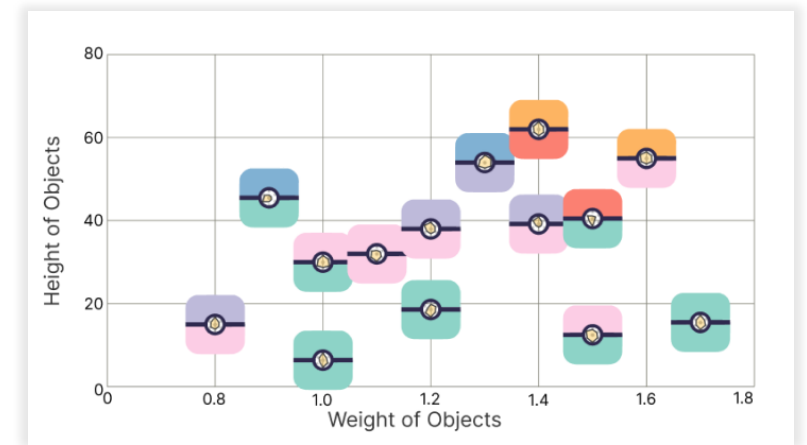
Entry number	Primary Type	Secondary Type	Base Attack	Base Defense	Base HP	Special Attack	Special Defense	Base Speed	Height of Objects	Weight of Objects
70	grass	poison	90	50	65	85	45	55	6.4	1
49	bug	poison	65	60	70	90	75	90	12.5	1.5
83	normal	flying	90	55	52	58	62	60	15	0.8
71	grass	poison	105	65	80	100	70	70	15.5	1.7
45	grass	poison	80	85	75	110	90	50	18.6	1.2
48	bug	poison	55	50	60	40	55	45	30	1
12	bug	flying	45	50	60	90	80	70	32	1.1
22	normal	flying	90	65	65	61	61	100	38	1.2
84	normal	flying	85	45	35	35	35	75	39.2	1.4
94	ghost	poison	65	80	60	170	95	130	40.5	1.5
72	water	poison	40	35	40	50	100	70	45.5	0.9
62	water	fighting	95	95	90	70	90	70	54	1.3
42	poison	flying	80	70	75	65	75	90	55	1.6
34	poison	ground	102	77	81	85	75	85	62	1.4

Pokemon data

+



=



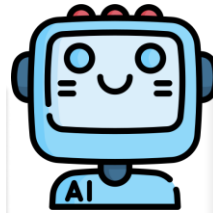
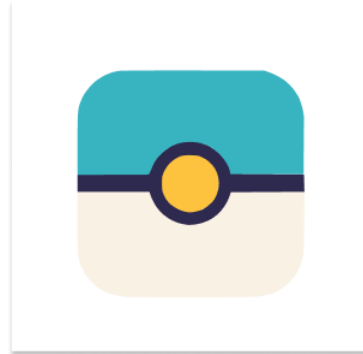
How to Create MGVs?

Challenge #2: *how to embed metaphors into glyph-based visualization design?*

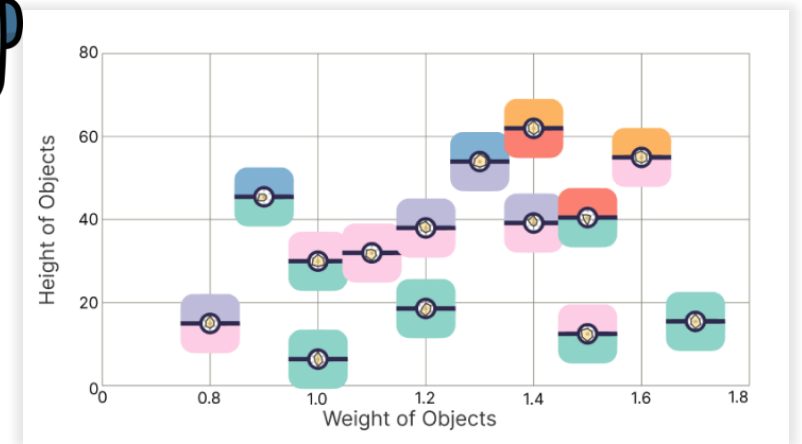
Entry number	Primary Type	Secondary Type	Base Attack	Base Defense	Base HP	Special Attack	Special Defense	Base Speed	Height of Objects	Weight of Objects
70	grass	poison	90	50	65	85	45	55	6.4	1
49	bug	poison	65	60	70	90	75	90	12.5	1.5
83	normal	flying	90	55	52	58	62	60	15	0.8
71	grass	poison	105	65	80	100	70	70	15.5	1.7
45	grass	poison	80	85	75	110	90	50	18.6	1.2
48	bug	poison	55	50	60	40	55	45	30	1
12	bug	flying	45	50	60	90	80	70	32	1.1
22	normal	flying	90	65	65	61	61	100	38	1.2
84	normal	flying	85	45	35	35	35	75	39.2	1.4
94	ghost	poison	65	80	60	170	95	130	40.5	1.5
72	water	poison	40	35	40	50	100	70	45.5	0.9
62	water	fighting	95	95	90	70	90	70	54	1.3
42	poison	flying	80	70	75	65	75	90	55	1.6
34	poison	ground	102	77	81	85	75	85	62	1.4

Pokemon data

+



=



MGV Generation Model



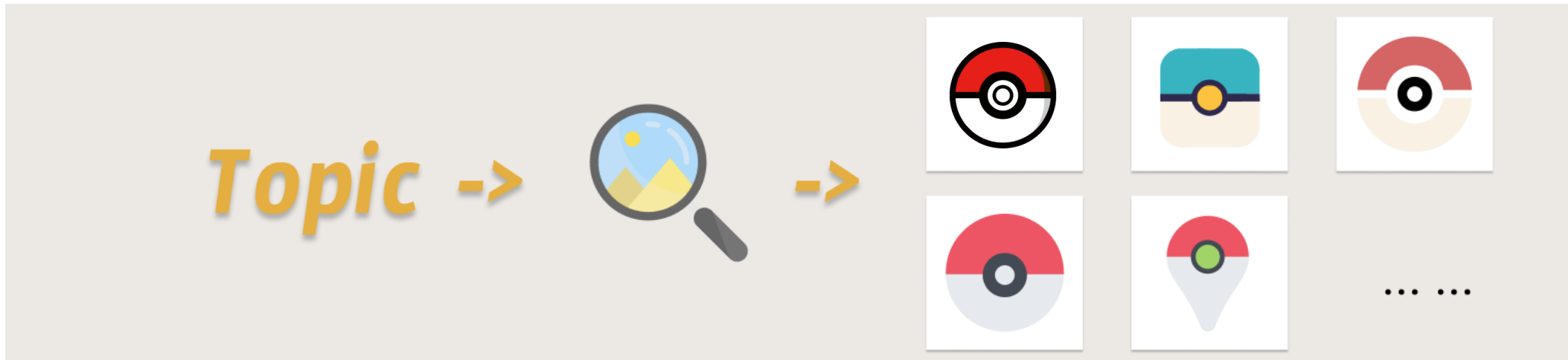
MGV Generation Model

Select Metaphoric Images

Construct

Render

1. It is **semantically** related to the data.
2. It is a **vector image** with a relatively **simple** structure.



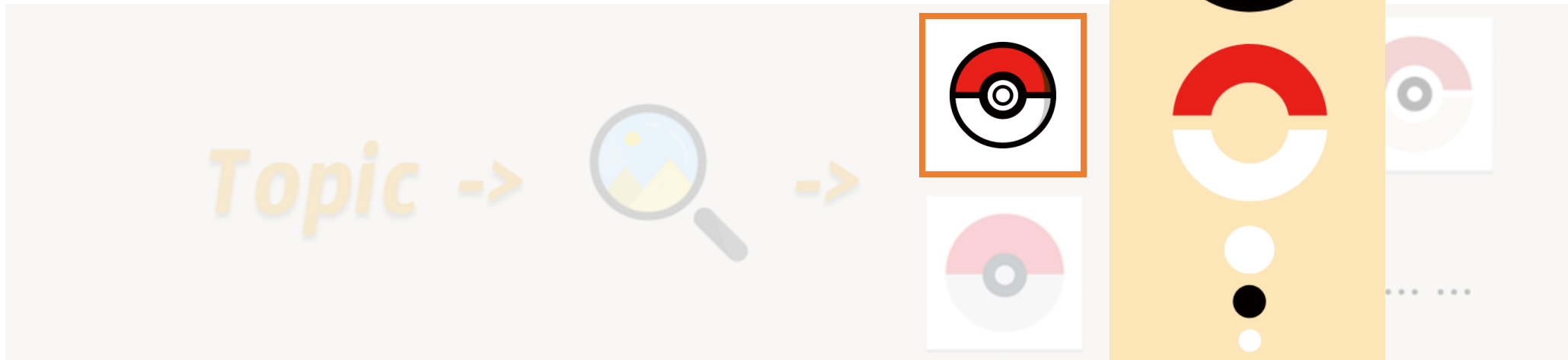
MGV Generation Model

Select Metaphoric Images

Construct

Render

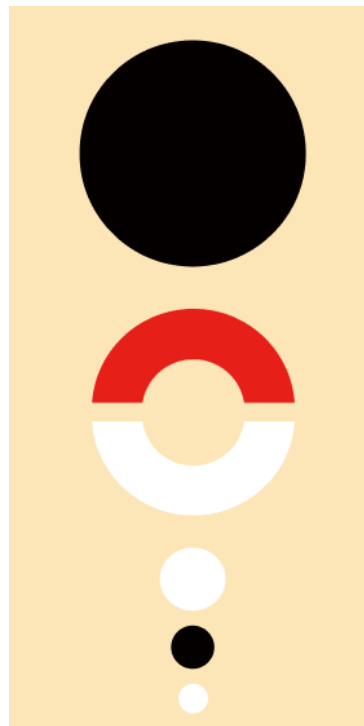
1. It is **semantically** related to the data.
2. It is a **vector image** with a relatively **simple** structure.



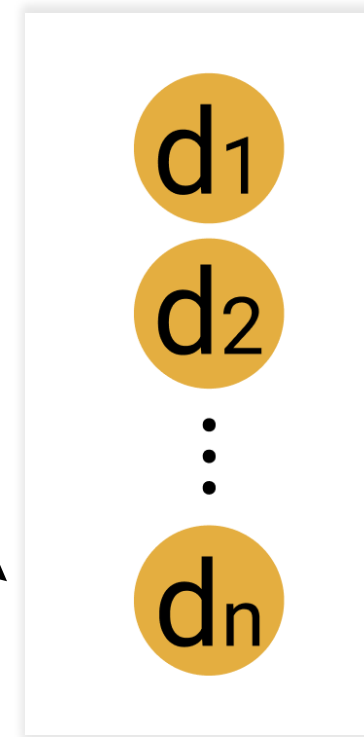
MGV Generation Model



Visual Elements and Layout



Monte Carlo
Tree Search



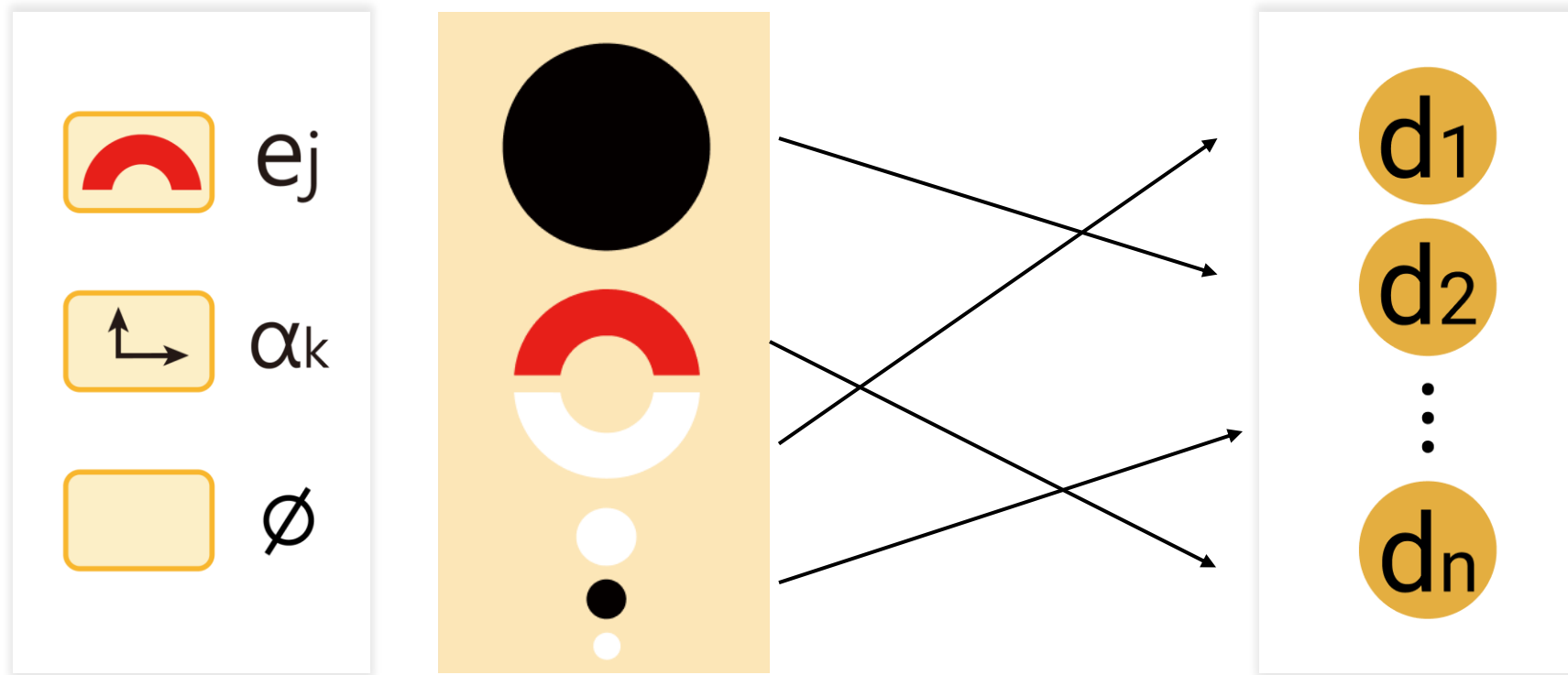
Data elements

MGV Generation Model

Select Metaphoric Images

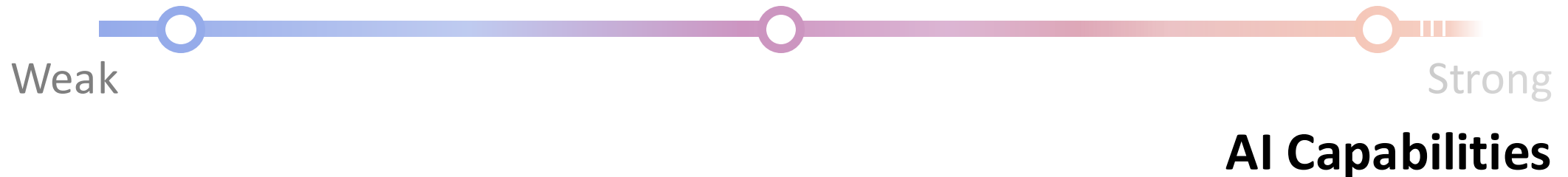
Construct

Render

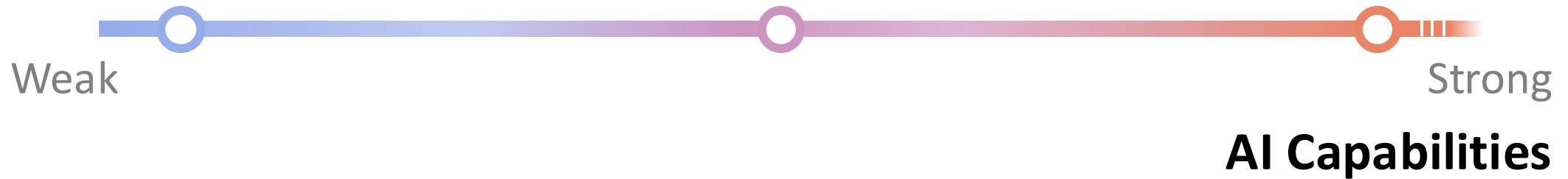
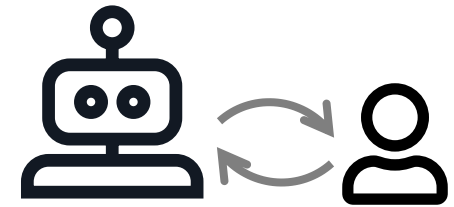


Lessons Learned

- Assigning the labor-intensive part to the machine and providing the connectors of subjective decisions to human beings.
- AI can enumerate possibilities, which serendipitously result in promising design.



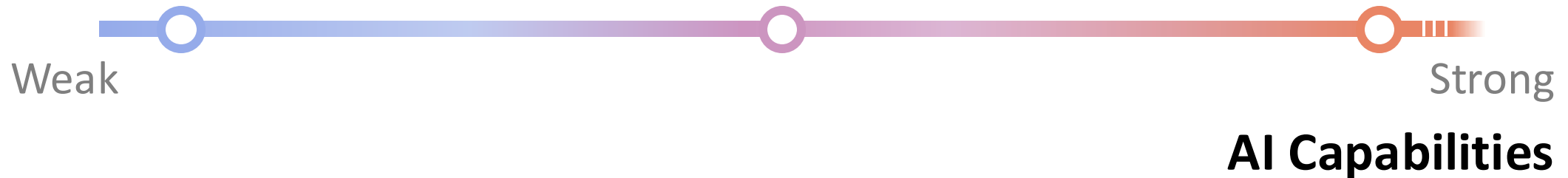
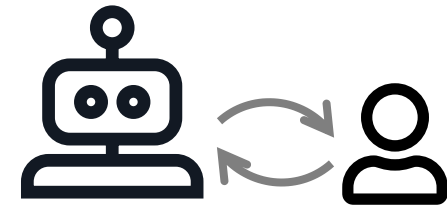
Humans may learn
from **AI**

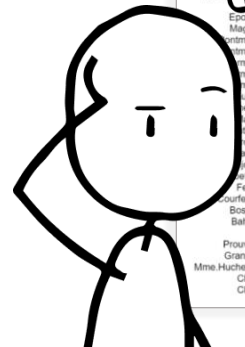
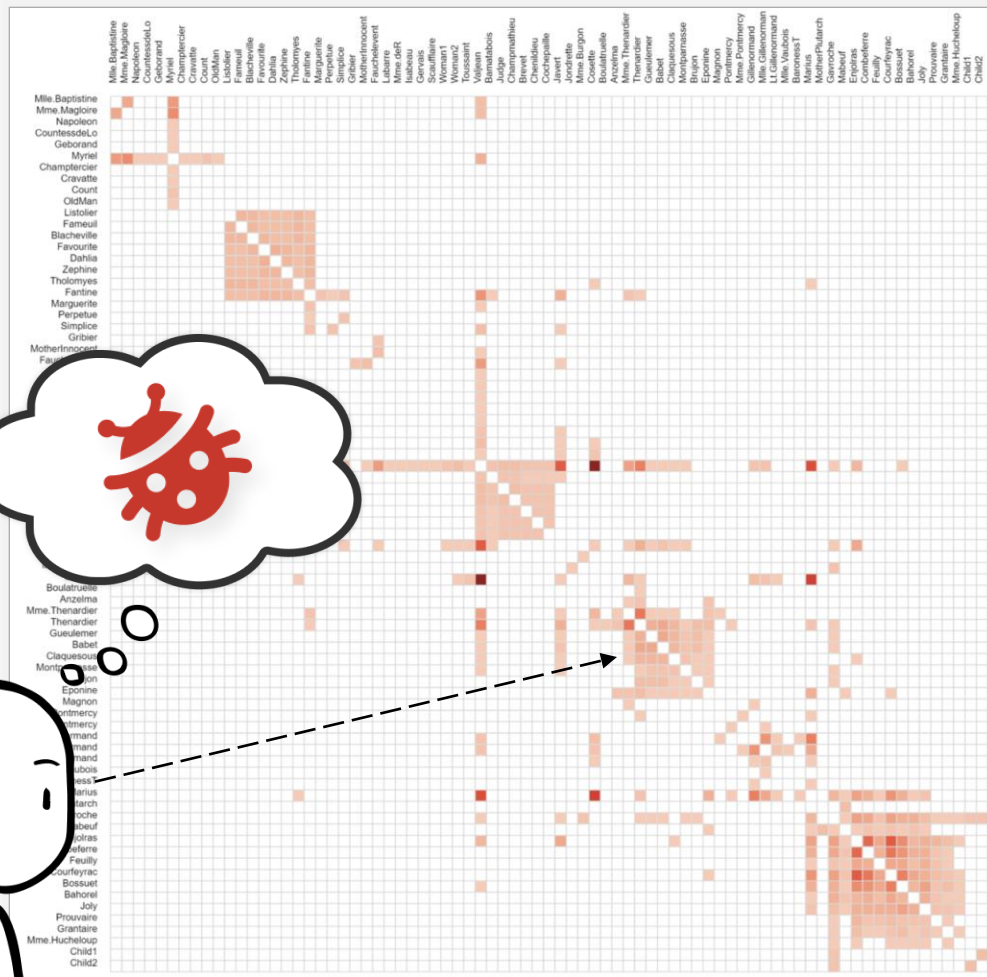


Interactive Pattern Explanation for Network Visualizations

 VIS 2024

Xinhuan Shu, Alexis Pister, Junxiu Tang,
Fanny Chevalier, Benjamin Bach





Adjacency Matrix

Block

Node Clique

Cluster

Diagonals

Off-diagonal cells

Dense row / column

Visual Patterns

Node Clique: Set of nodes where every node is connected to every other node.

Cluster: Set of nodes where almost all nodes are connected, or at least would be present, the cluster would be a clique.

Diagonals: Self-links are the first cells along the matrix diagonal. Self-links are links that connect a node to itself. Examples include self-citations in citation networks.

Off-diagonal cells: Connectors indicate links between two cliques or clusters (A and B).

Dense row / column: Highly connected nodes are visible by rows and columns with many cells. Cells do not need to be adjacent.

critical Attributes

data characteristics

accuracy of perception

columns, whereas rows is encoded in the

suited for several

enous node or edge

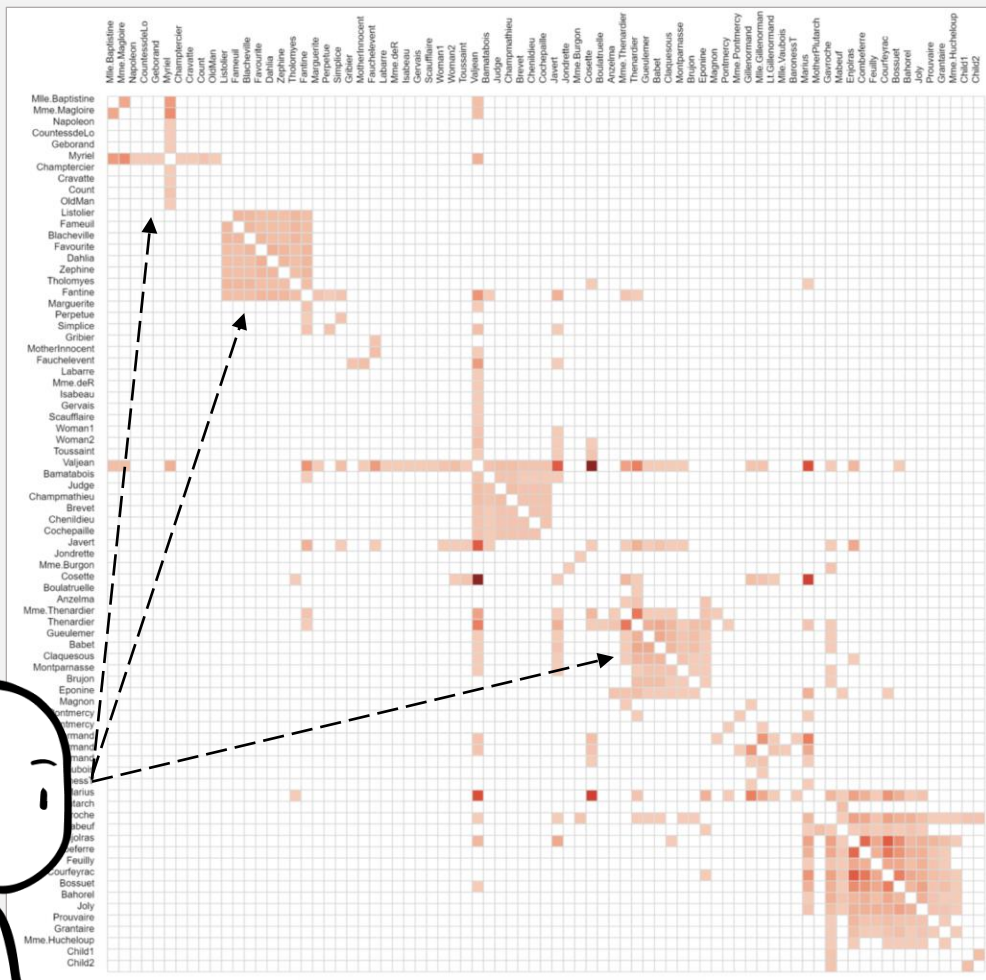
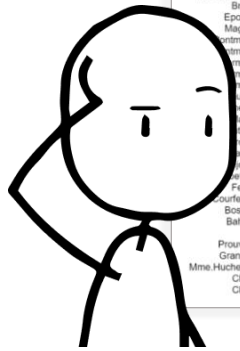
orks.

trees.

from Data to Viz

1. Data Collection
2. Data Cleaning
3. Data Analysis

<https://visualizations.github.io/>



Adjacency Matrix Visual Patterns

Block
Node Clique
Set of nodes where every node is connected to every other node.
It depicts a value as complete blocks without missing cells inside.

Cluster
Node Cluster
Set of nodes where almost all nodes are connected, or all nodes would be present, the cluster would be a clique.
It cluster is value as a large "stamp" of cells but can contain empty cells.
Empty cells indicate unconnected nodes in the cluster.
Dense Sparse

Diagonal cells
Connectors indicate links between two clusters or nodes (i1 and i2).
This example has 2 links between i1 and i2.

Dense row / column
Hub nodes
Highly connected nodes are value by row and column with many cells.
Cells do not need to be adjacent.
More links Less links

Visual Patterns
Critical Attributes
data characteristics
accuracy of perception

It's essential to quickly spot and interpret Visual Patterns

of columns, whereas
all suited for several
enous node or edge
orks.
trees.

<https://visualizationcheatsheets.github.io>

Visualization Literacy

*“the ability to confidently use a given data visualization to translate questions specified in the **data domain** into visual queries in the **visual domain**, as well as **interpret Visual Patterns** in the visual domain as properties in the data domain.”*

A principled way of assessing visualization literacy [Boy et al. 2014]

*“the ability to make meaning from and **interpret patterns**, trends, and correlations in visual representations of data”*

Data visualization literacy [Börner et al. 2019]

*“detects salient **visual patterns**, translates them into conceptual information structures”*

Seeking patterns of visual pattern discovery for knowledge building [Andrienko et al. 2022]

Research problem

How to support people to quickly

spot and **interpret Visual Patterns**

Interactive Pattern Explanations

Explaining **visual patterns** and their **data patterns on-demand** in user-defined parts of visualization.

Explainer Pattern

les miserable Adjacency Matrix ?

[Return to data view](#)

Pattern Overview:

Filter:

Hub Bridge Clique Cluster
Fan Strong Link

>		Hub	3
>		Bridge	2
∨		Clique	16
>		Cluster	6
>		Fan	2
>		Strong Link	22

Cliques are groups of nodes where every node is connected to every other node of the clique.

Your selection has **4** network patterns, including **2 Clique**, **2 StrongLink**.

Clique (2) Strong Link (2)
Clique #1 Clique #2

The selected **Block** pattern represents a **Clique**.

A **Clique** is a **subgraph pattern**, where a group of nodes are connected to every other node of the clique.

Your selection is a clique with **8** nodes and **56** mutual links connecting each of the nodes to each other.

Block pattern:

◆ The selected pattern may visually vary, like:

This pattern is shown as a solid squared block of cells along the diagonal. In some cases, the block can be fragmented, i.e., the square is divided into full rectangles or arrays of cells. The larger the block, the more nodes are in the clique. The grey cells on the diagonal denote self connections.

Browse related Cliques and Blocks

◆ Similar instances in your network (ranked by size):

[Clique #1](#) [Clique #2](#) [Clique #3](#)


Interactive Pattern Explanations



Interactive Pattern Explanations

Pattern Explorer x
Pattern Explorer x
+

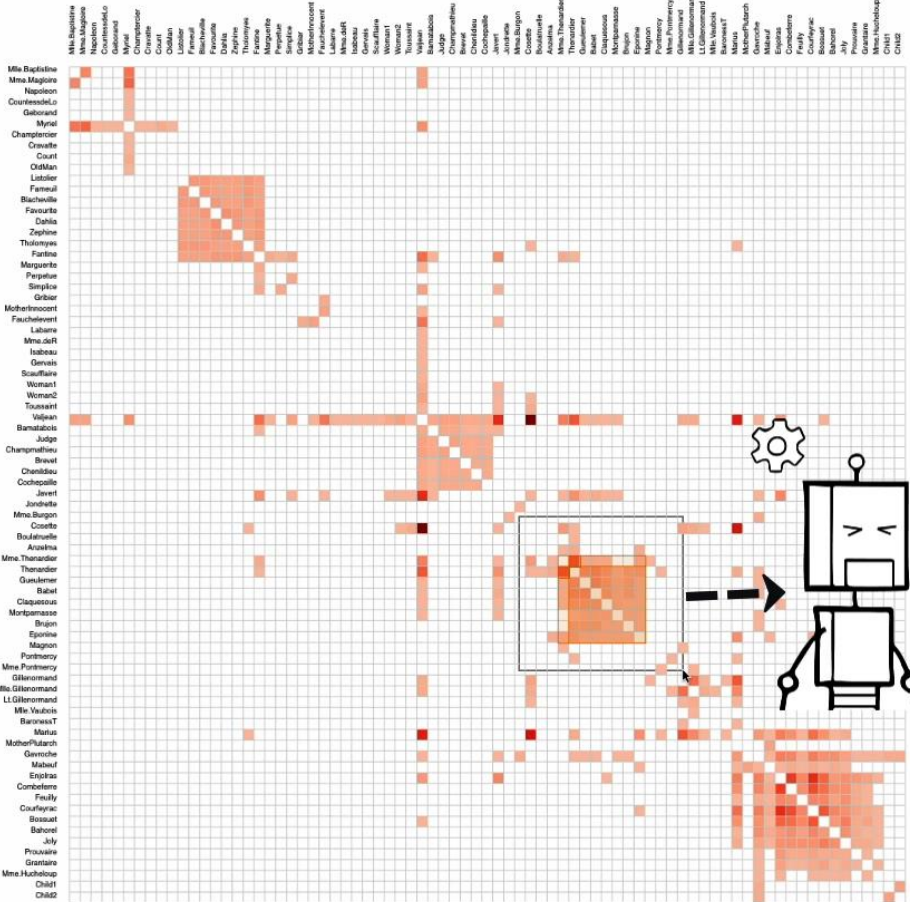
localhost:5173
Guest




Network: les-mis

Pattern Overview:

Seriation: barycentre



	NODE-LINK DIAGRAM	ADJACENCY MATRIX	TIME-ARCS
Link Patterns			
Node Patterns			
Subgraph Patterns			



Interactive Pattern Explanations

The screenshot displays the Explainer Pattern application interface. On the left, the 'Explainer Pattern' logo and 'Network: les-mis' are visible. The main area shows a network visualization with a grid of nodes and edges. A specific pattern is highlighted with a red box. On the right, a 'Network pattern explanation' panel is open, providing details about the selected pattern.

Network pattern explanation

Block pattern

◆ The selected pattern may visually vary.

Browse related Block and Clique

- ◆ Similar instances in your network (ranked by the size):
 - > Clique [0, 4]
 - > Clique [5, 9]
 - > Clique [10, 14]
 - > Clique [15, 19]

Visual pattern explanation

Interactive Pattern Explanations

The screenshot displays the Explainer Pattern application interface. On the left, the 'Network: les-mis' is shown with a 'Pattern Overview' section. The main area features a large grid visualization of the network, with a specific pattern highlighted in orange. A tooltip window is open, providing an explanation of the selected pattern, which is a 'Clique'.

Explainer Pattern
Network: les-mis
Pattern Overview: [toggle]

Serialiation: barycentre

Clique (2) Strong Link (2)
Clique #1: Clique #2

The selected pattern **Block** represents a **Clique**.

Clique
A **Clique** is a **subgraph pattern**, where a group of nodes are connected to every other node of the clique.

Your selection is a clique with **7 nodes** and **42 mutual links** connecting each of the nodes to each other.

Block pattern
The selected pattern may visually vary, like:

This pattern is shown as a solid squared block of cells along the diagonal. In some cases, the block can be fragmented, i.e., the square is divided into full rectangles or arrays of cells. The larger the block, the more nodes are.

Browse related Block and Clique
Similar instances in your network (ranked by the size):

- > Clique [0, 4]
- > Clique [5, 9]
- > Clique [10, 14]
- > Clique [15, 19]

[Browse other instances](#)

Evaluation

- Qualitative Study:
 - We ran a **within-subject design** with 12 participants, asking them to learn and interpret three network visualizations in an **open-ended way**.
- Quantitative Study:
 - We ran a **between-subject design** with 20 participants to measure **how many patterns people can accurately identify** after learning adjacency matrices.



Text-only



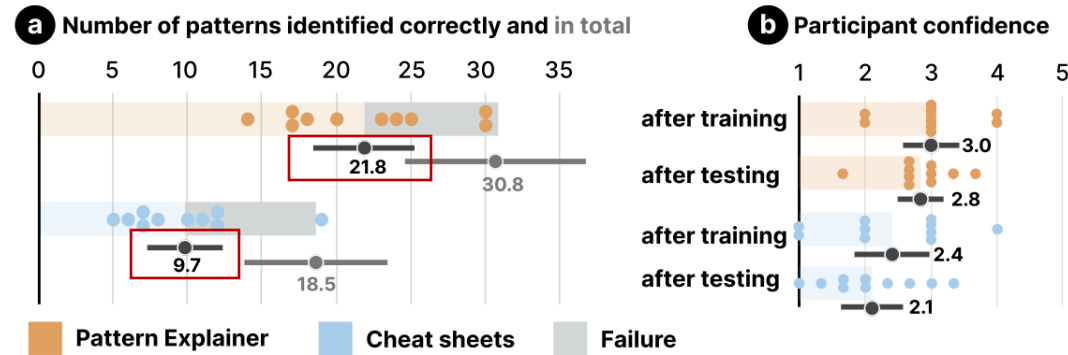
Cheatsheet



Pattern Explorer

Evaluation

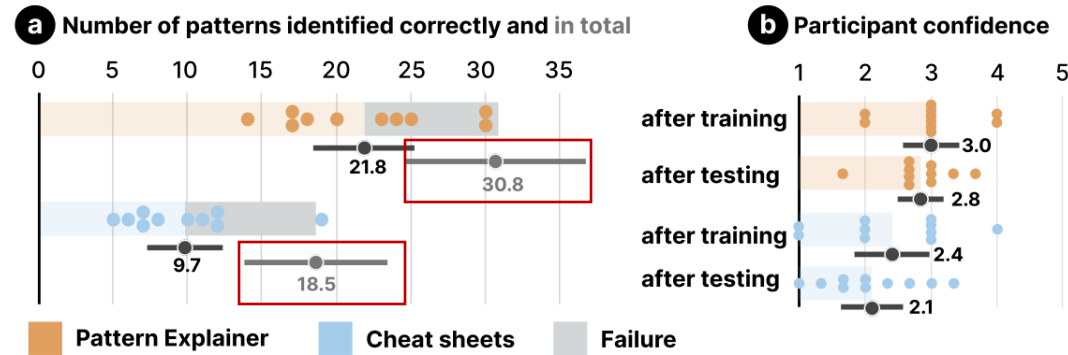
- Results:
 - Increase the number of patterns people identify correctly.



Condition	Sum	#SL	#Hub	#Bridge	#Fan	#Clique	#Cluster
Pattern Explainer	21.8	1.9	2.4	2.7	1.1	8.0	5.7
Cheat sheets	9.7	1.3	0.9	0.7	0.7	3.6	2.5

Evaluation

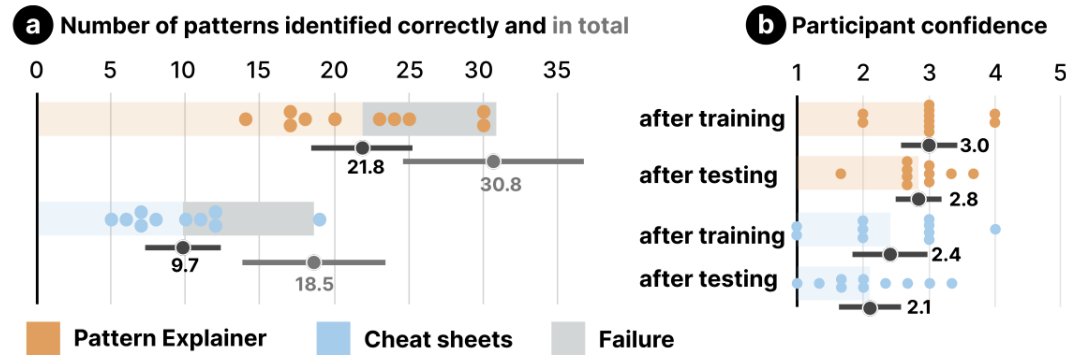
- Results:
 - Increase the number of patterns people identify correctly.



Condition	Sum	#SL	#Hub	#Bridge	#Fan	#Clique	#Cluster
Pattern Explainer	21.8	1.9	2.4	2.7	1.1	8.0	5.7
Cheat sheets	9.7	1.3	0.9	0.7	0.7	3.6	2.5

Evaluation

- Results:
 - Increase the number of patterns people identify correctly.



Condition	Sum	#SL	#Hub	#Bridge	#Fan	#Clique	#Cluster
Pattern Explainer	21.8	1.9	2.4	2.7	1.1	8.0	5.7
Cheat sheets	9.7	1.3	0.9	0.7	0.7	3.6	2.5

Evaluation

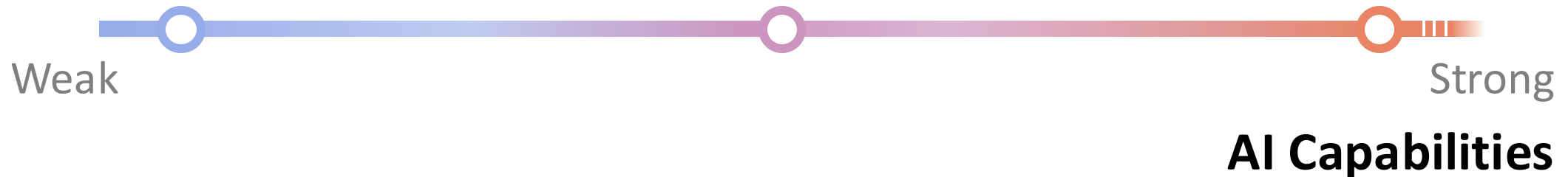
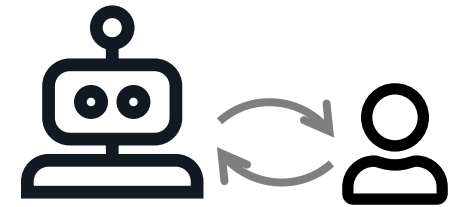
- Results:
 - Increase the number of patterns people identify correctly.
 - Appreciate the in-situ and on-the-fly explanations.

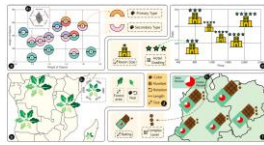
“Interactive pattern explanations put the theoretic concepts into practice.”

“Cheat sheets could not directly apply to the visualization. I got more confused about whether my understanding of this cheat sheet was right.”

Lessons Learned

- In-context and on-the-fly explanations can help novices improve visualization (data) literacy.
- Personalized, guided, and progressive tutorials are promising for AI-assisted education.

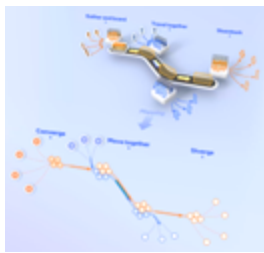




MetaGlyph
VIS'22



WonderFlow
TVCG'23



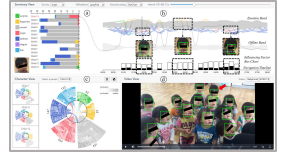
RouteFlow
CHI'25
🏆 Best Paper



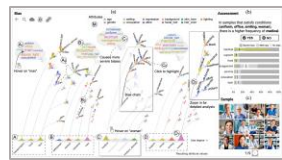
Pattern Explanation
VIS'24



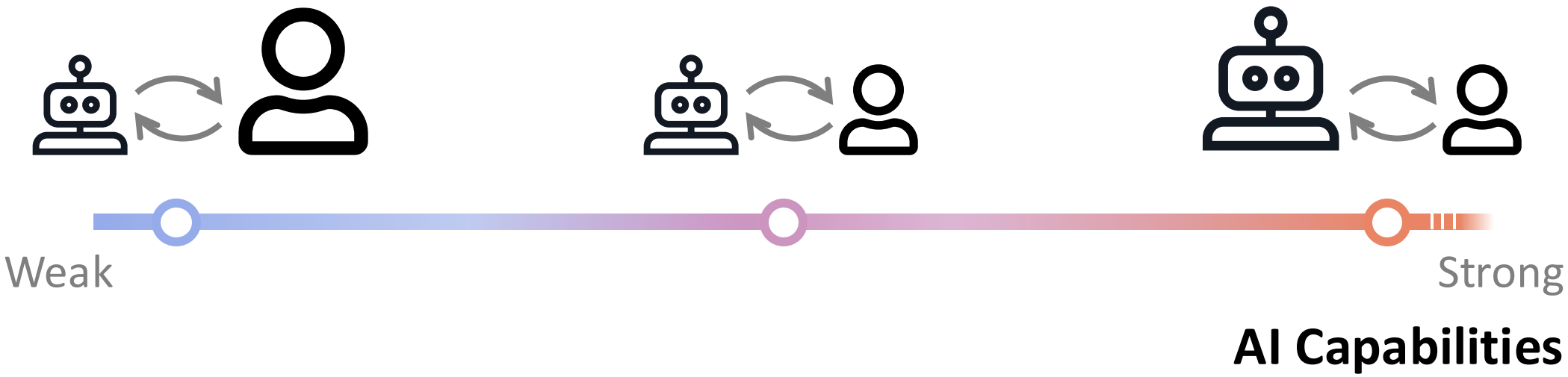
FretMate
IUI'24



EmotionCues
TVCG'20



BiasField
Under review





From Data to Dialog: Visualization-empowered Human-AI Collaboration



Xinhuan Shu (<https://shuxinhuan.github.io>)

Lecturer @ Newcastle University

Data Visualisation  / Human-Computer
Interaction  / Data-driven Storytelling 