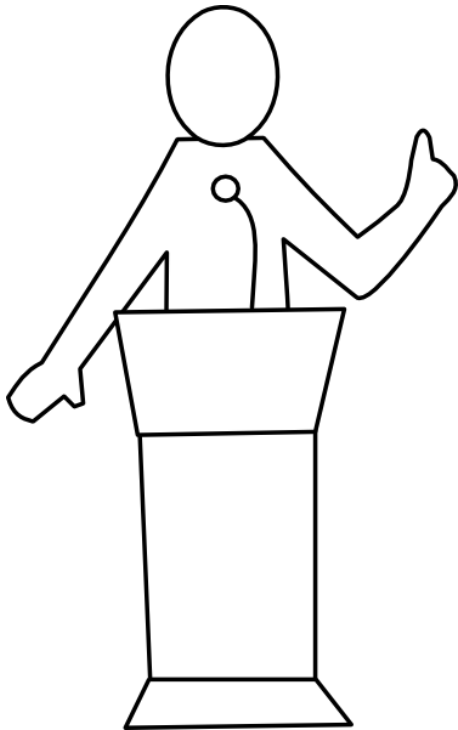


Scalable and Repeatable Extrinsic Evaluation for Pattern Discovery Systems

Mario Boley, Maïke Krause-Traudes, Bo Kang, Björn Jacobs
University of Bonn & Fraunhofer IAIS
mario@realKD.org

Recently at Q&A time...

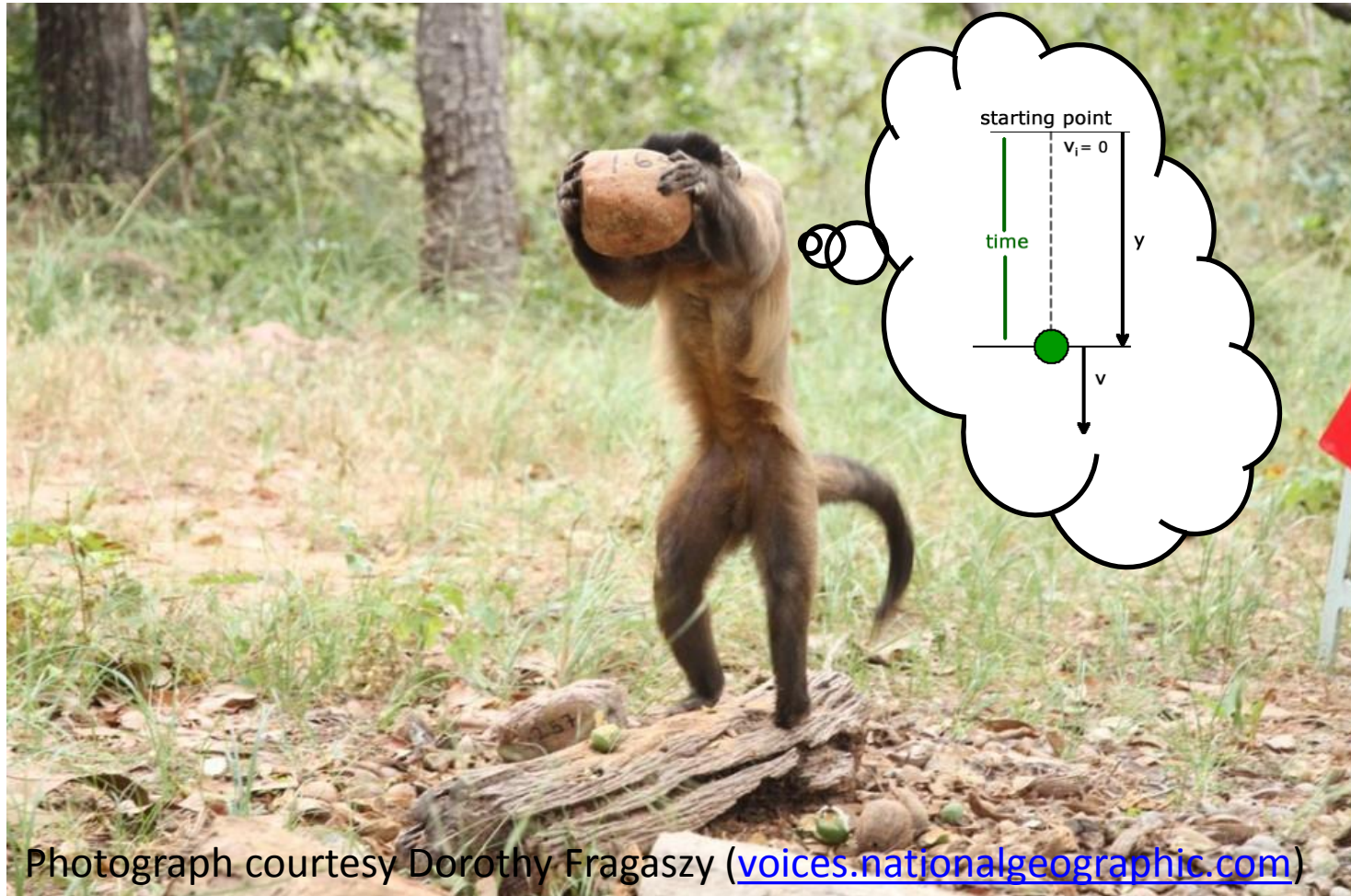


Q: This looks interesting, but is this really what *users* would want?

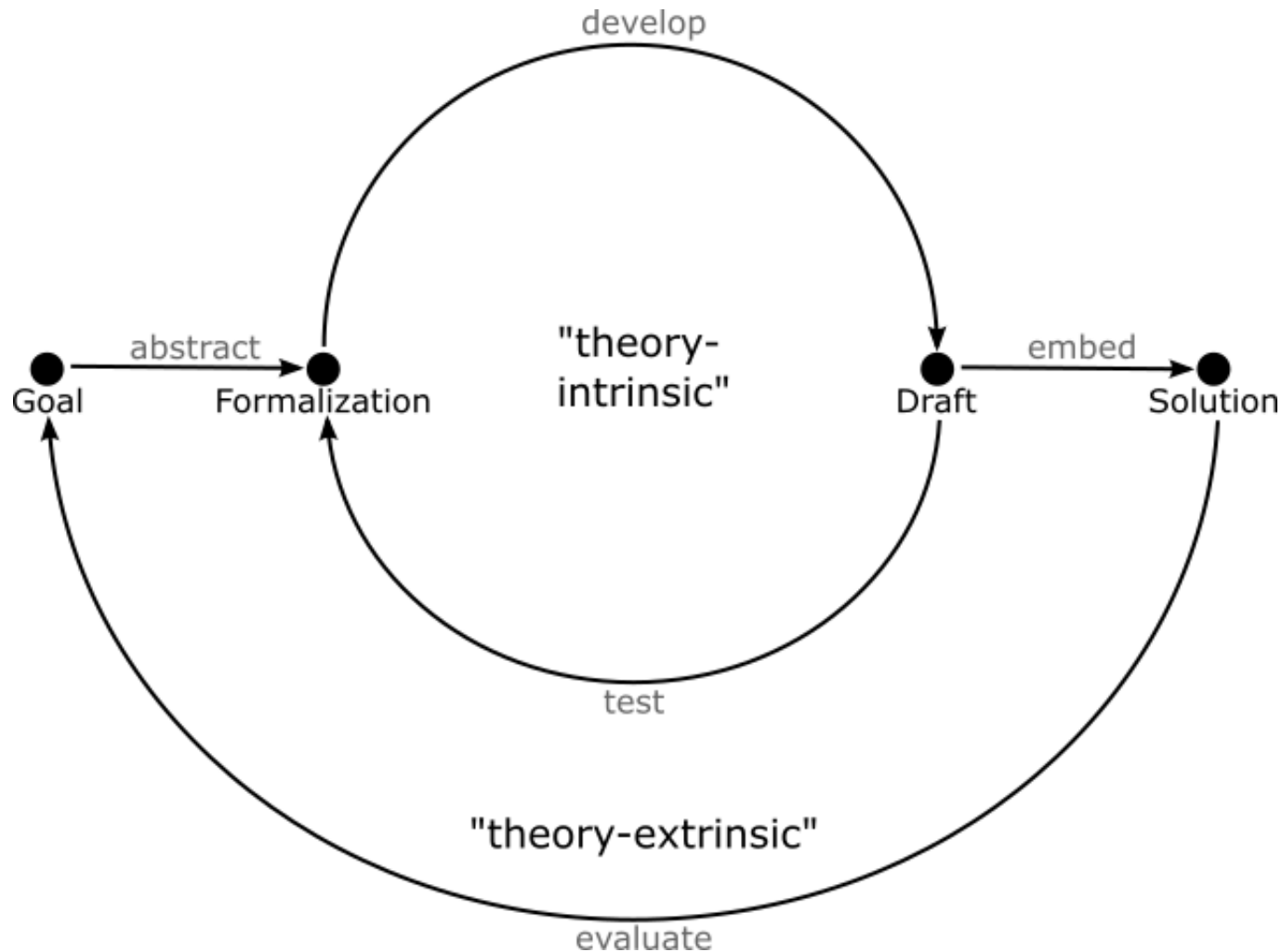
A: Well, I guess in order to really confirm that, we would need to test this *somehow* with real users.

Q: Yep, agreed. Thank you.

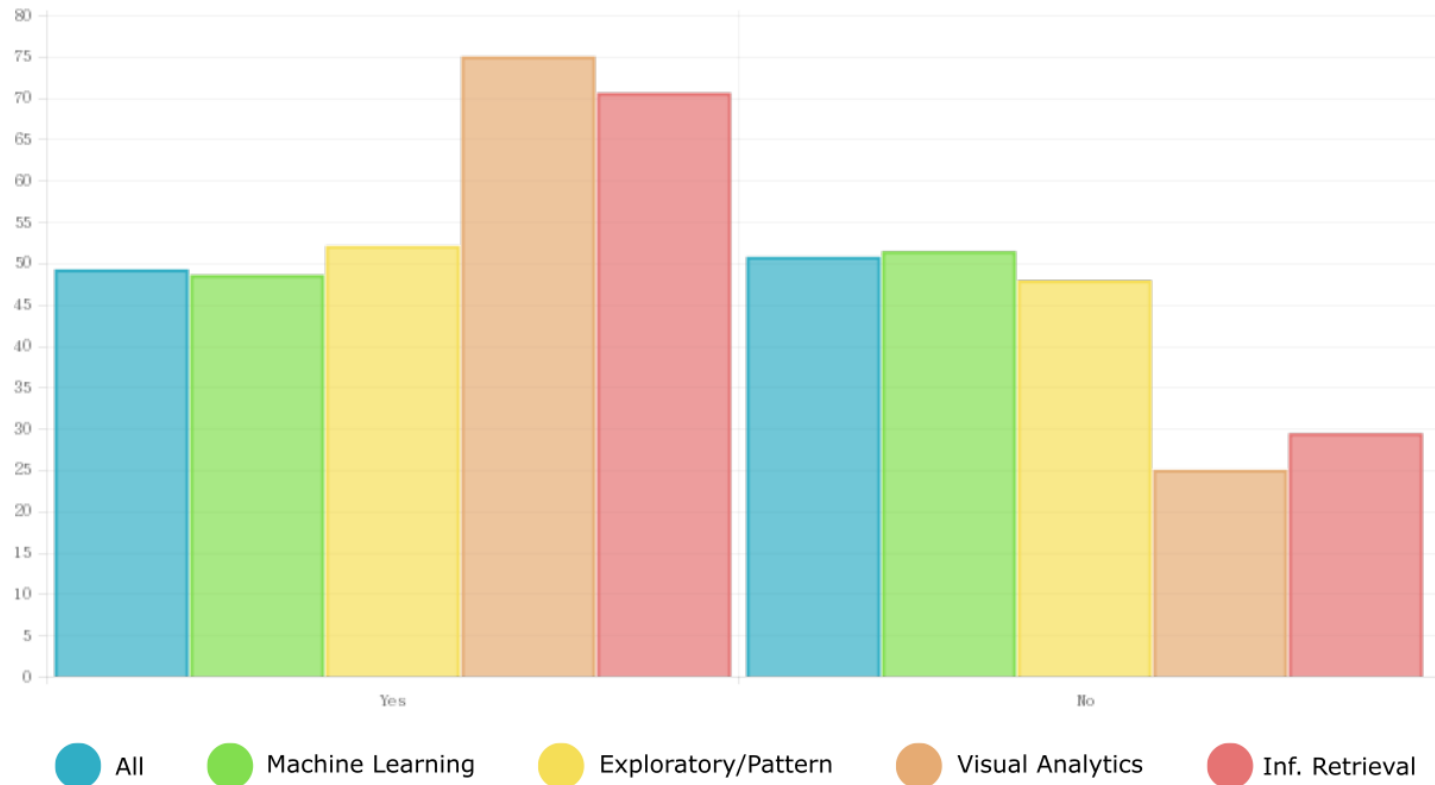
Extrinsic evaluation can support ultimate value of contributions



Extrinsic means: “not depending on theory used for development cycle”

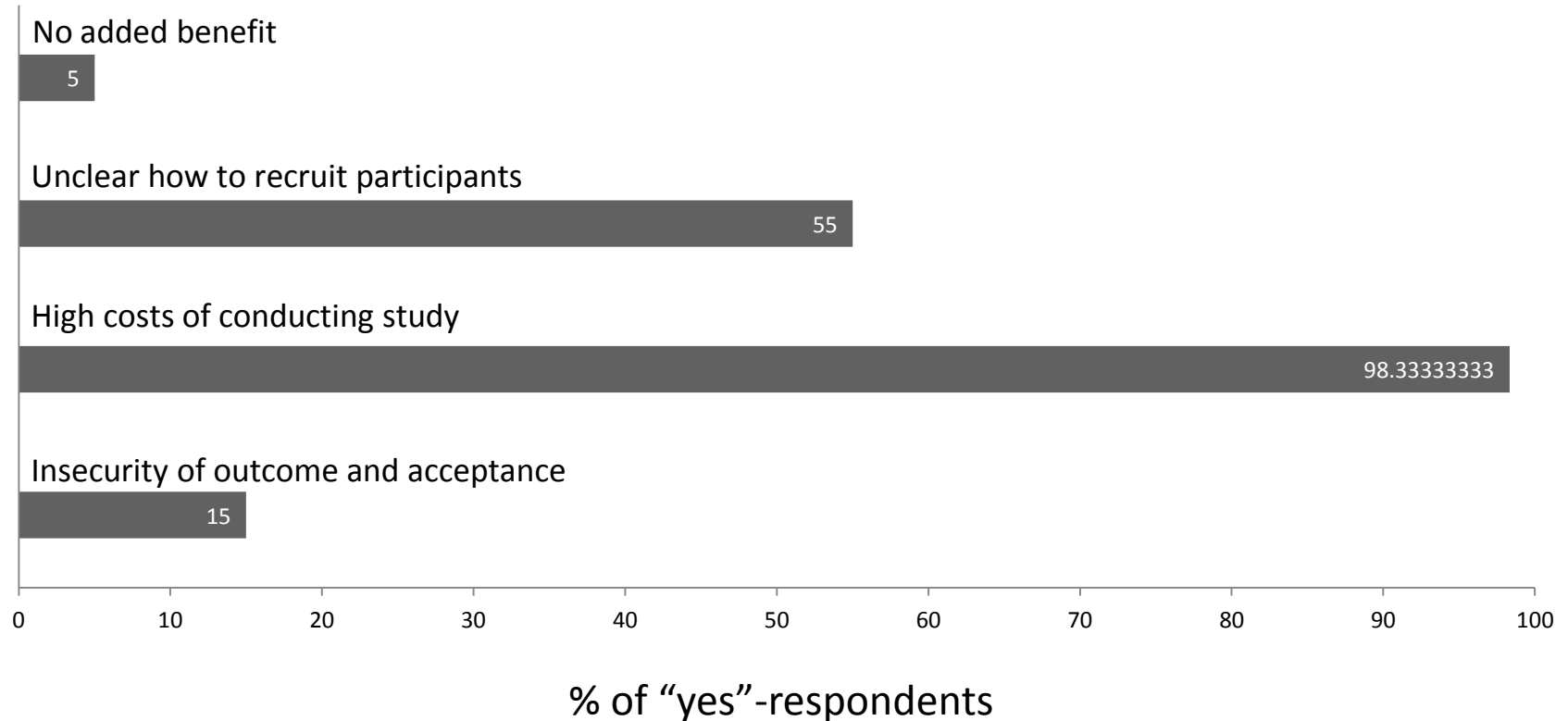


Poll among ECMLPKDD authors: *half* skipped potentially useful studies

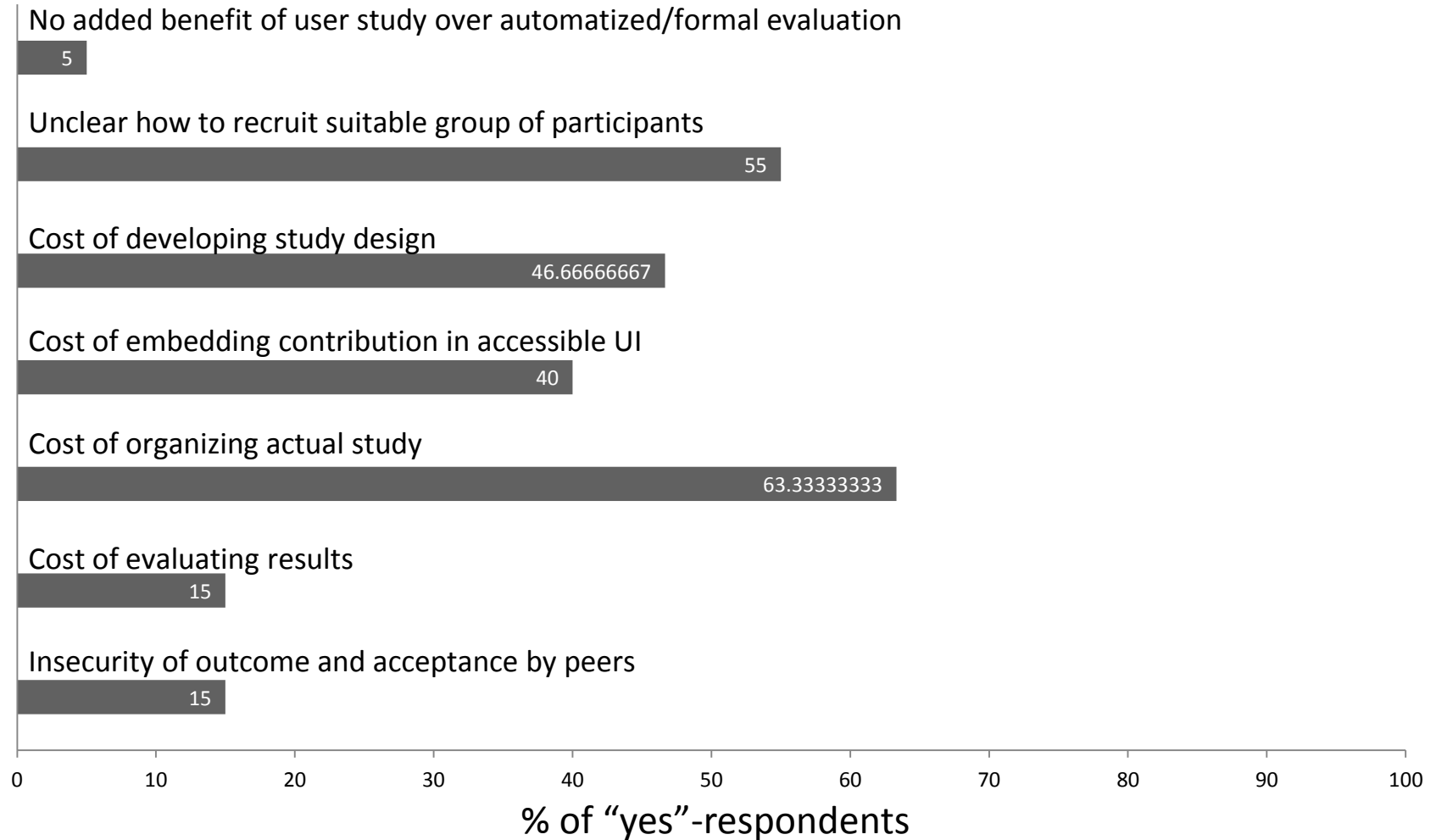


Details at <http://www.realkd.org/dm-userstudies/ecmlpkdd-authorpoll-march2015/>

High costs are dominant reason for skipping on “study opportunity”



High costs are dominant reason for skipping on “study opportunity”



Creedo's major contributions are...

- Allows definition of **reusable study designs**
- Elements focus on **scalable evaluation in application context**
- **Automatizes** process

A study is a process for providing evidence in favor or against...

Hypothesis:

“Users can solve a certain class of analysis tasks better with a specific target system than with other control systems.”

A study is a process for providing evidence in favor or against...

Hypothesis:

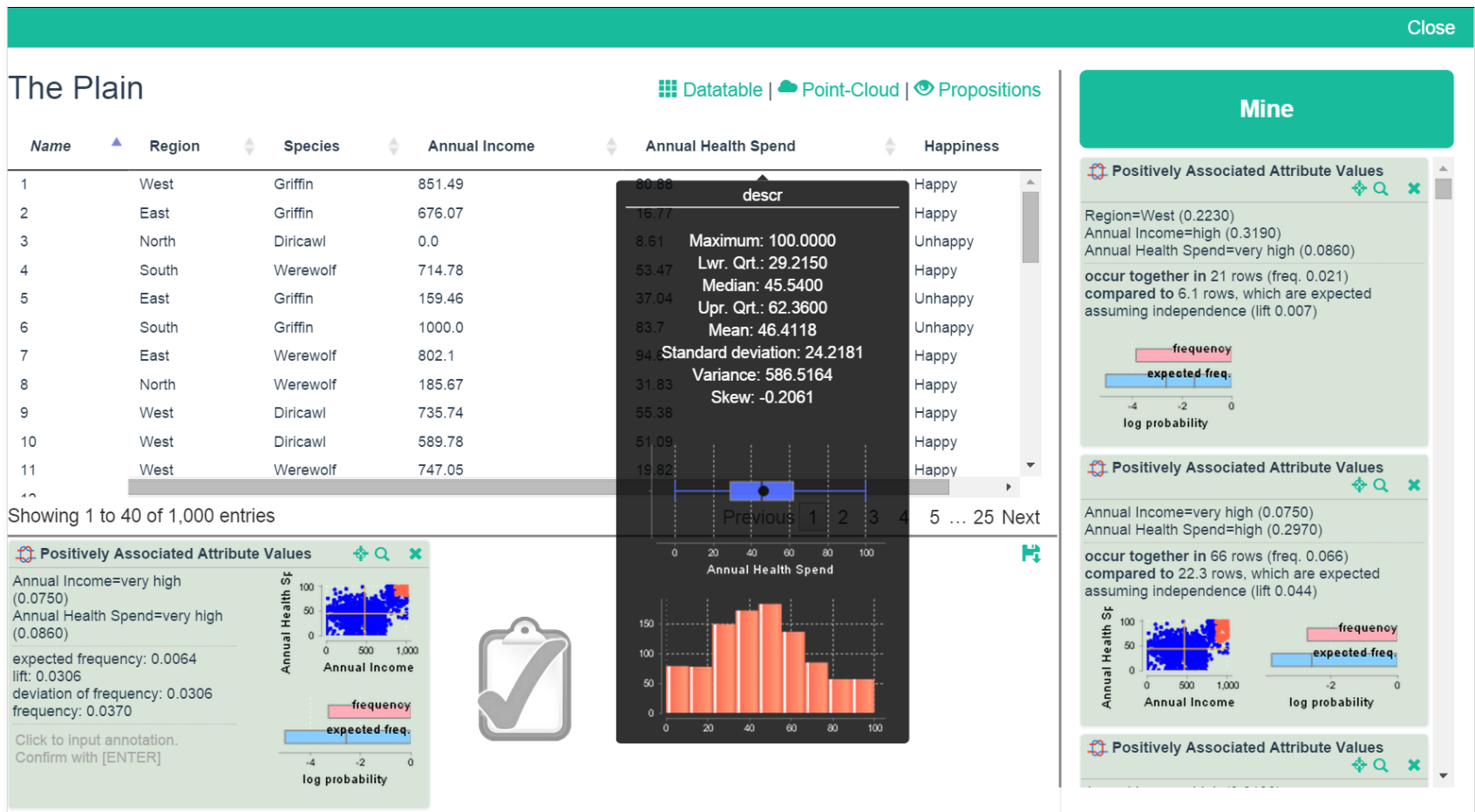
“Users can solve a certain class of analysis tasks better with a specific target system than with other control systems.”

Example:

“Users can discover a set of interesting patterns faster using a FORSIED-based association discovery process than when using a conventional association discovery process.”*

*based on a static interestingness measure that is oblivious to prior and gained knowledge

Data analysis systems are represented by Creed analytics dashboards



Algorithms can be integrated via the realKD library

One Click Mining - Socio-Economics of Germany - Google Chrome
localhost:8080/SpringOCM/dashboard.htm

Socio-Economics of Germany

Name	Type	State
Aachen	Urban	NRW
Ahrweiler	Rural	Rhineland-Palatina
Aichach-Friedberg	Rural	Bavaria
Alb-Donau-Kreis	Rural	Baden-Wuerttemb
Altenburger Land	Rural	Thuringia
Altenkirchen (Westerwald)	Rural	Rhineland-Palatina
Altmarkkreis Salzwedel	Rural	Saxony-Anhalt
Altoetting	Rural	Bavaria
Alzey-Worms	Rural	Rhineland-Palatina
Amberg	Urban	Bavaria
Amberg-Weizsach	Rural	Bavaria
Ammerland	Rural	Lower Saxony
Anhalt-Bitterfeld	Rural	Saxony-Anhalt
Ansbach	Urban	Bavaria
Ansbach	Rural	Bavaria
Aschaffenburg	Urban	Bavaria
Aschaffenburg	Rural	Bavaria

Showing 1 to 40 of 412 entries

Previous 1 2 3 4 5 ... 11 Next

Algorithm execution

Category: Exceptional Model Mining

Algorithm: EMM BeamSearch

EMM optimization function: freq*deviation_order

Number of results: 10


Beam width: 0

Target attributes: CDU 2005, SPD 2005

Model Class: Cosine distance between regression models

Cancel Execute!

Mine



Creedo *tasks* bridge formal abstraction and application context

1. Introduction

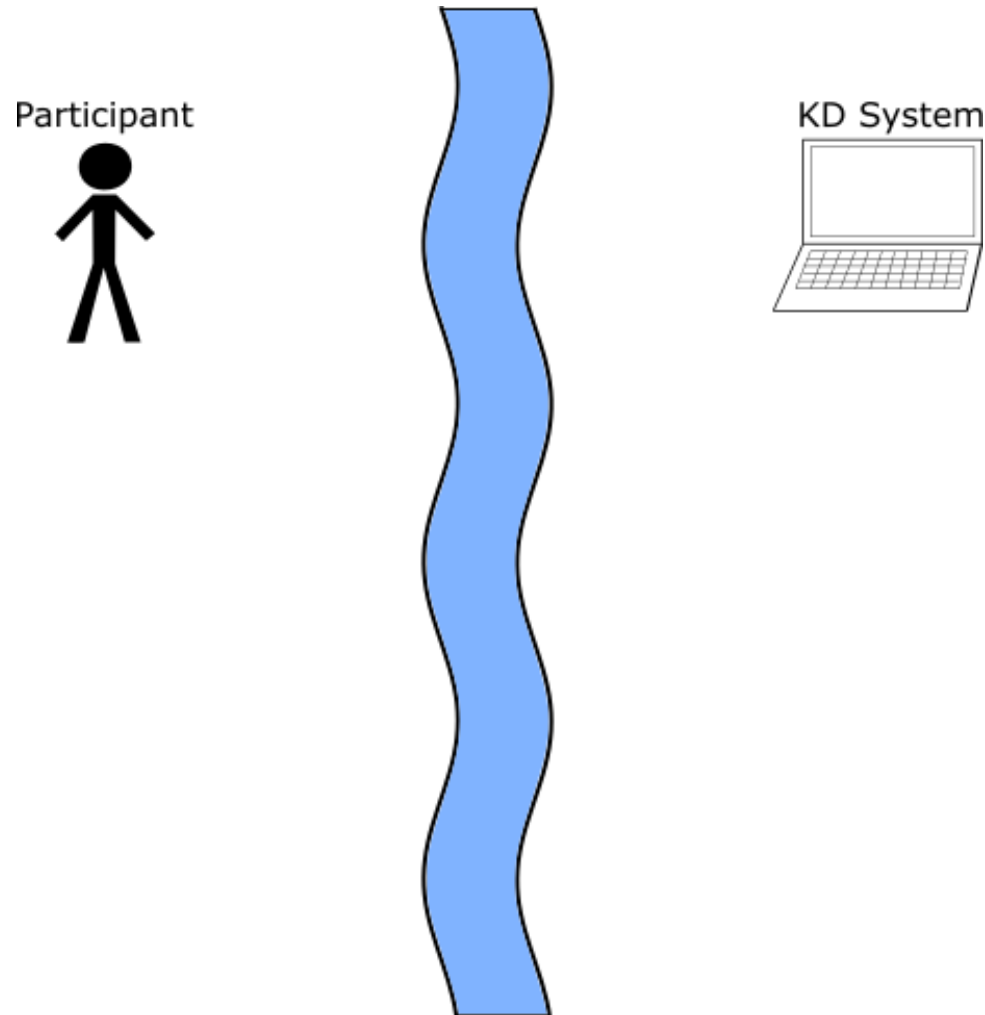
In this paper, we tackle the important problem of discovering interesting patterns from a given input dataset.

...

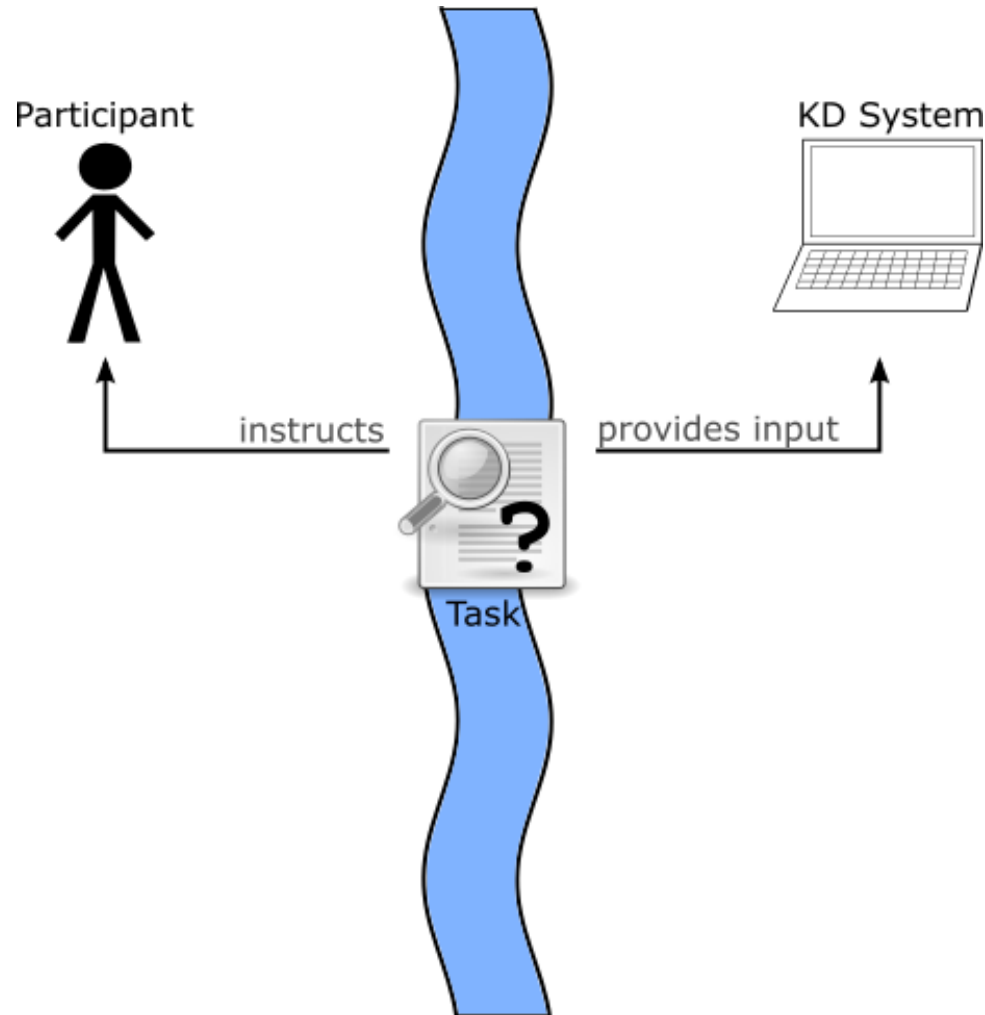
$$q(x) = \frac{1}{|D(x)|}(p_0 - p_x)^2$$

```
for each  $d \in D$ 
  if  $x \in D$  then
     $\hat{D}(x) \leftarrow \hat{D}(x) + 1$ 
  ...
```

Creedo *tasks* bridge formal abstraction and application context



Creedo *tasks* bridge formal abstraction and application context



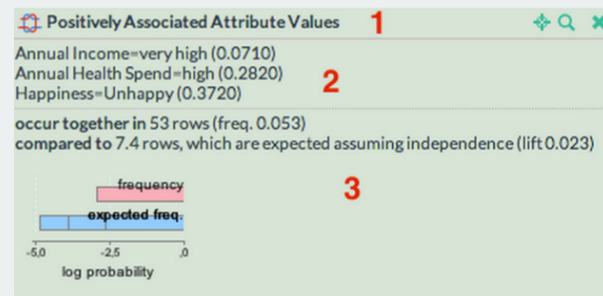
User perspective on task are natural language instructions

Next

Analysis task instructions

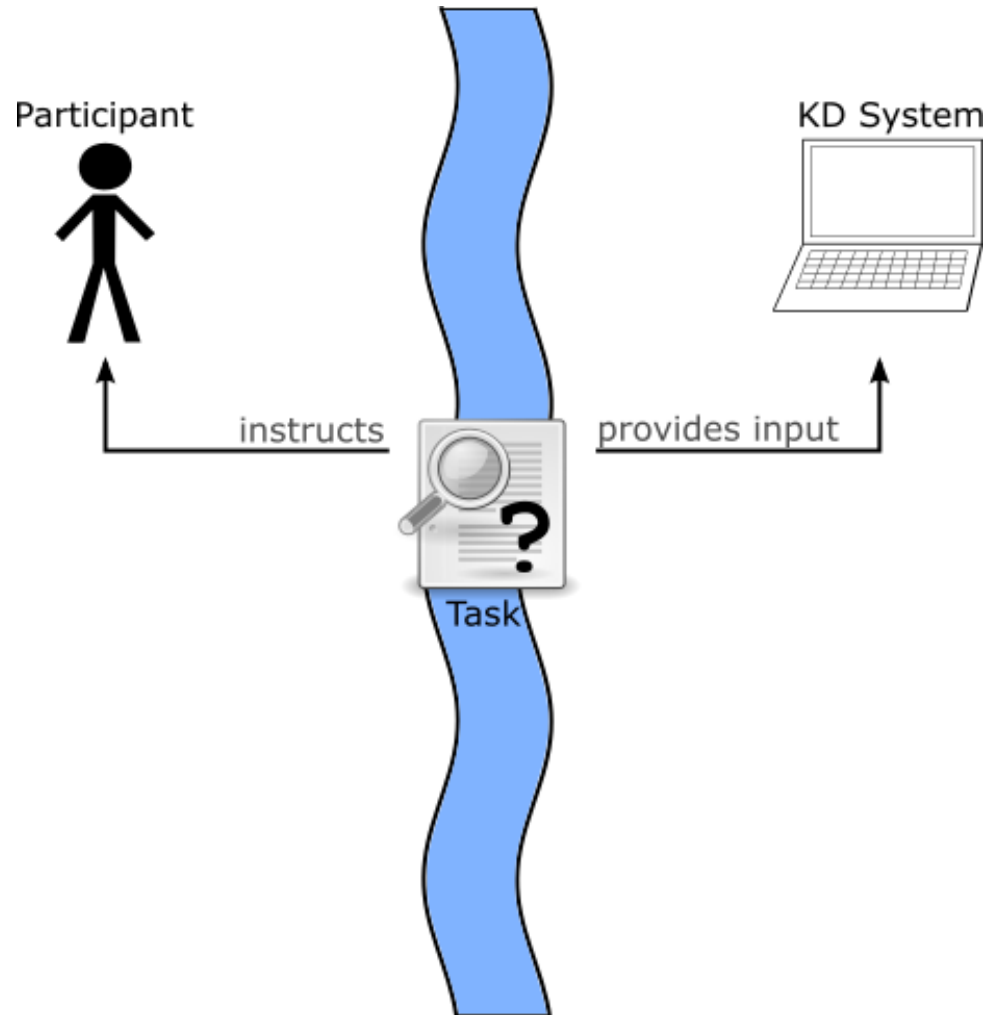
As a new employee of the Data Science Department of the government of **The Plain**, you have to get familiar with the socio-economics status of your country. Go on and use our data mining tool to discover key phenomena from **The Plain**'s socio-economic data. The data consists of the socio-economic records of 1000 representative inhabitant samples of your country.

The data mining tool will propose statements about the data, and measures associativity among the statements. Such information is summarized in graphic representations (**patterns**) like the figure below:

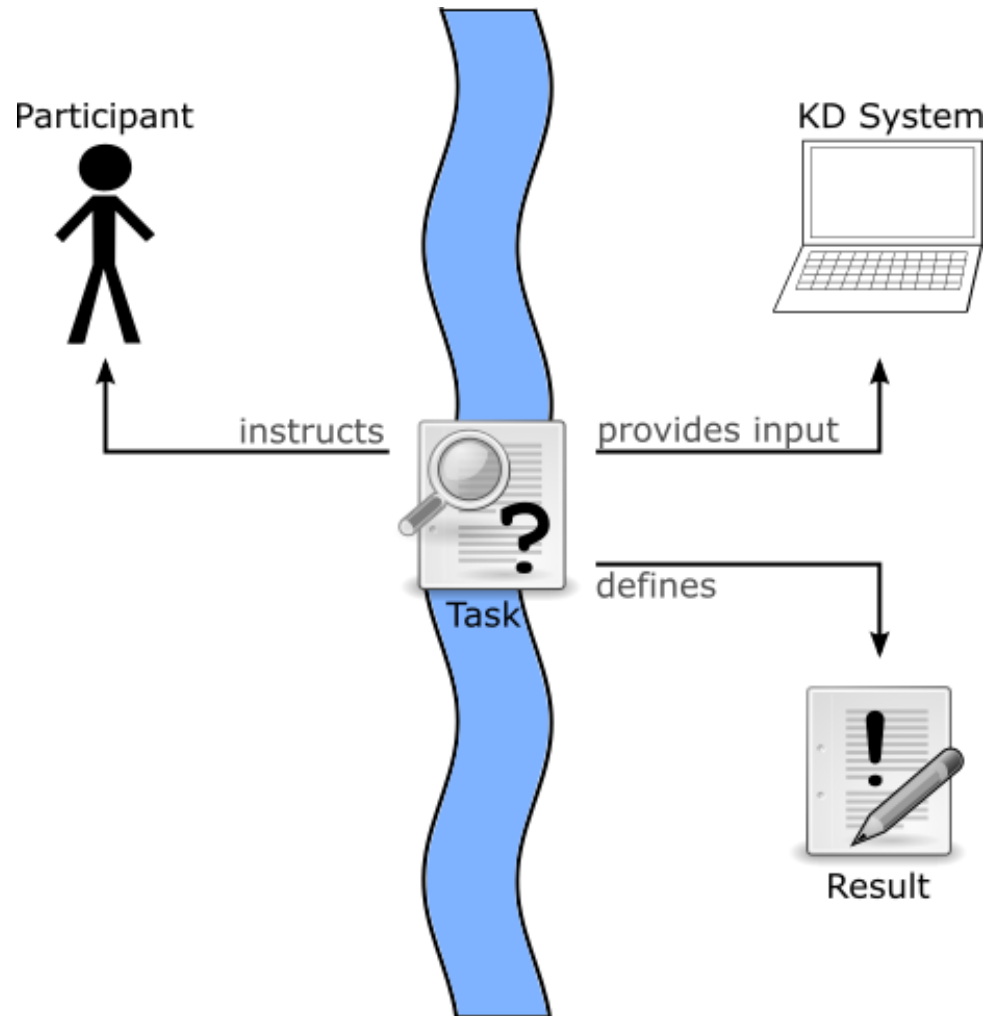


- Sec.1 states whether the statements are positively or negatively associated.
- Sec.2 lists the considered statements, along with the frequency (proportion of inhabitants) that each individual statement holds true.
- Sec.3 visualizes the difference between the expected frequency (blue bar) and the actual frequency (red bar) of the statements. The larger the difference the stronger the positive/negative association is.

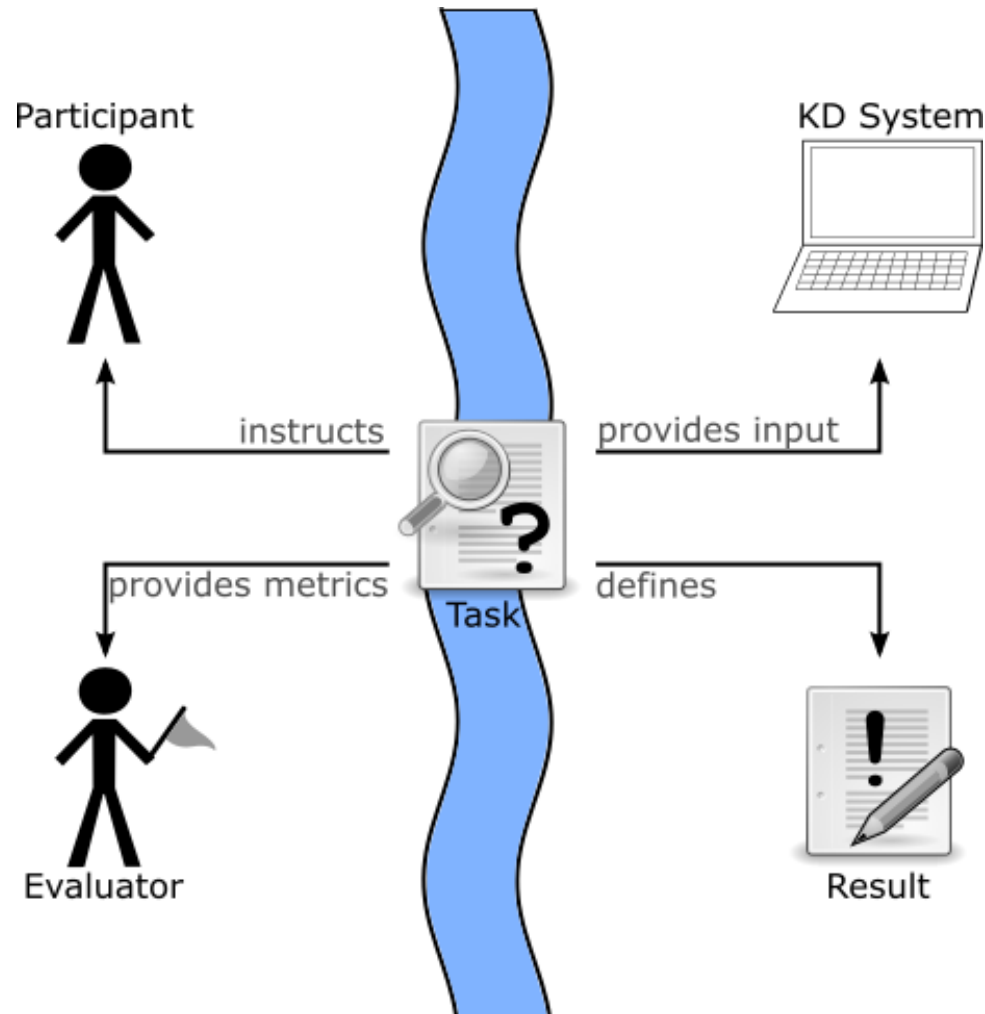
Creedo *tasks* bridge formal abstraction and application context



Creedo *tasks* bridge formal abstraction and application context



Creedo *tasks* bridge formal abstraction and application context



Task also defines elementary attributes of results

e. **Submit**

Rate result patterns

b.

Data-Table | Point-Cloud | Code-Table

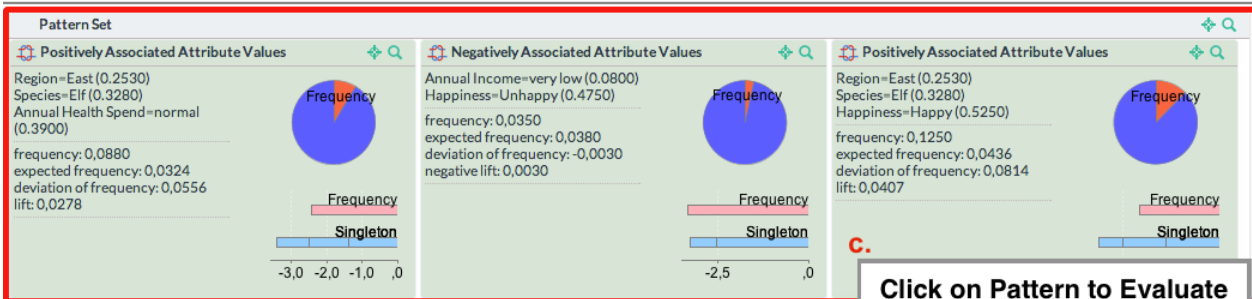
Name	Region	Species	Annual Income	Annual Health Spend	Happiness
1	East	Human	909.96	80.54	Unhappy
2	South	Human	858.36	37.74	Happy
3	South	Human	630.39	48.19	Unhappy
4	East	Orc	725.54	26.98	Happy
5	West	Human	934.14	53.01	Happy
6	South	Orc	399.12	39.38	Unhappy
7	East	Elf	143.22	75.7	Happy
8	West	Human	664.45	52.24	Unhappy
9	South	Elf	743.8	51.9	
10	North	Orc	514.84	33.0	
11	East	Elf	737.77	80.48	

a.

Data Area

Showing 1 to 40 of 1,000 entries

Previous 1 2 3 4 5 ... 25 Next



Ratings:

Do you consider this statement reveals diverse aspects in assisting to answer the question?

diverse (2) rather diverse (1)

undecided (0) rather not diverse (-1)

not diverse (-2)

Do you consider this statement completely answers the asked question?

complete (2) rather complete (1)

undecided (0) rather not complete (-1)

not complete (-2)

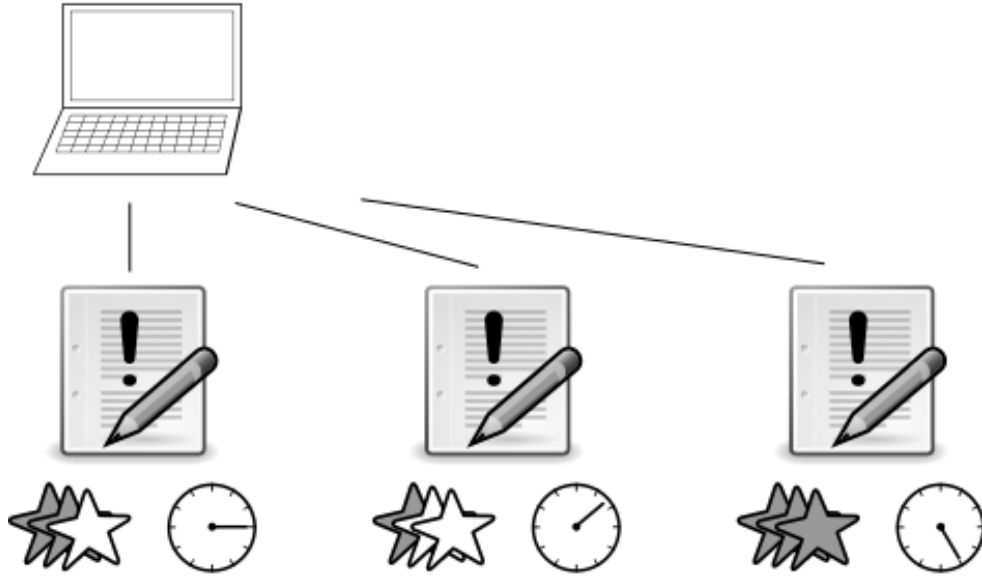
d.

Click on Value of Metrics to Evaluate

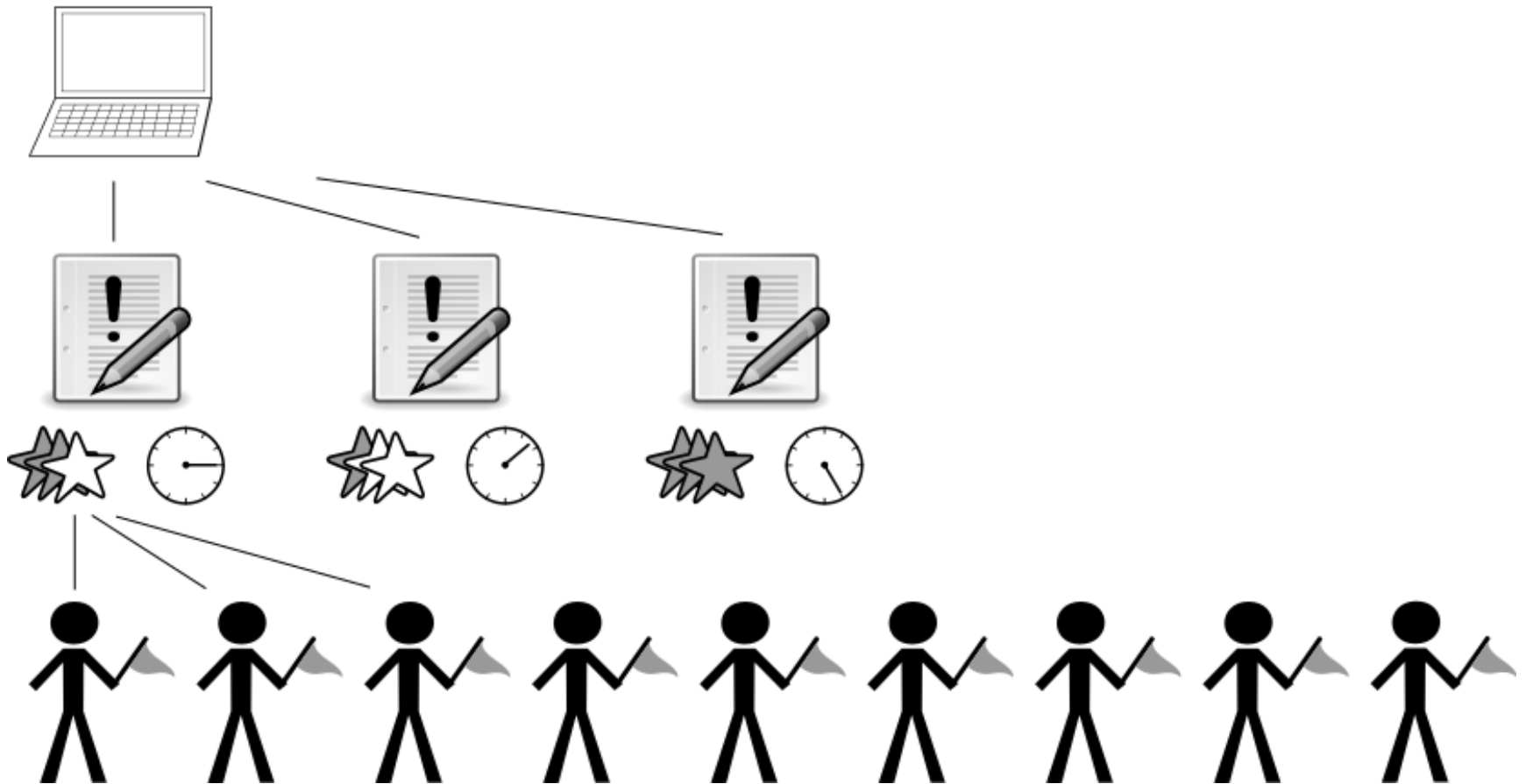
All measurements can be aggregated to *system performance measures*



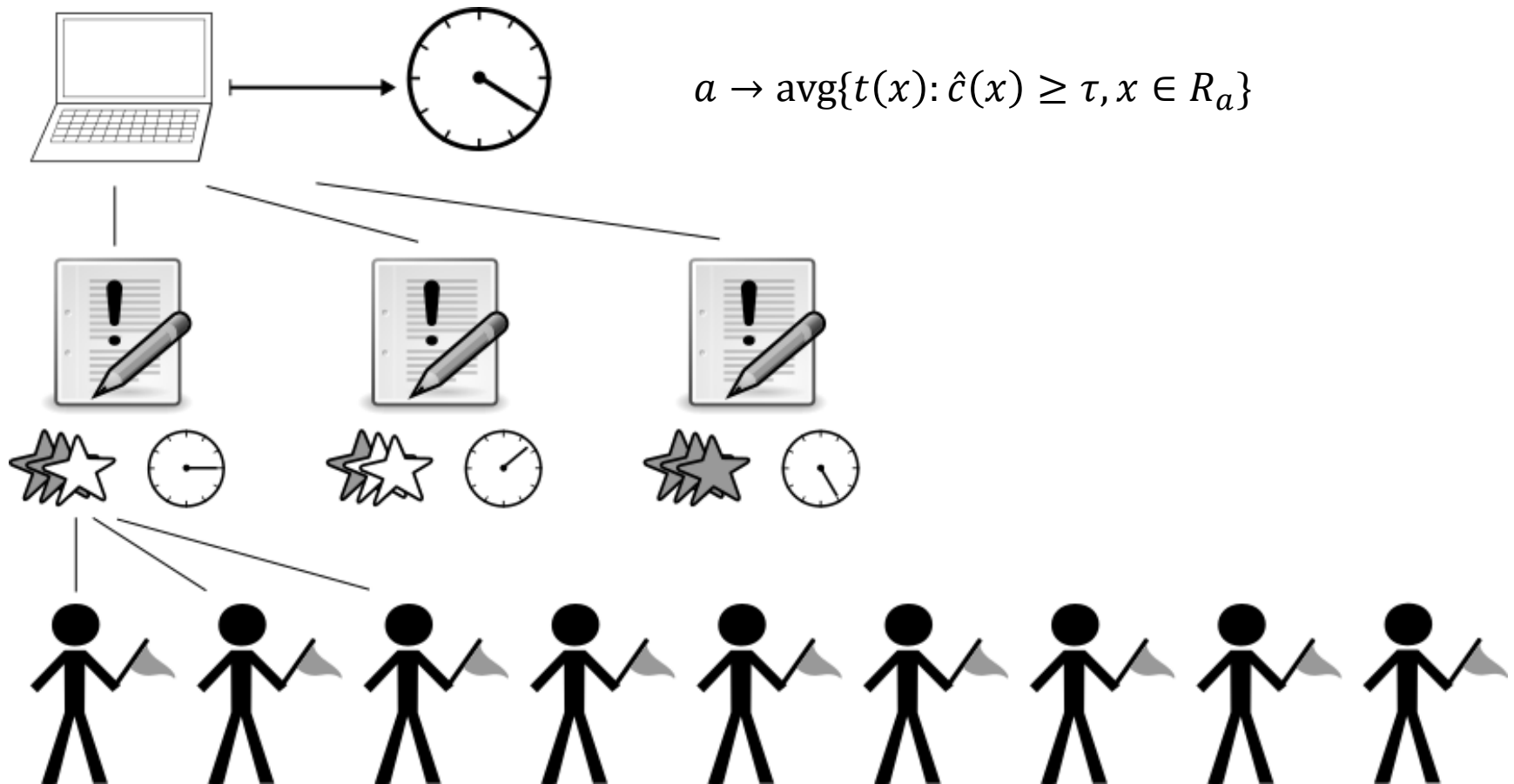
All measurements can be aggregated to *system performance measures*



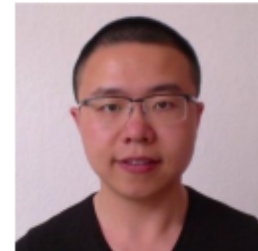
All measurements can be aggregated to *system performance measures*



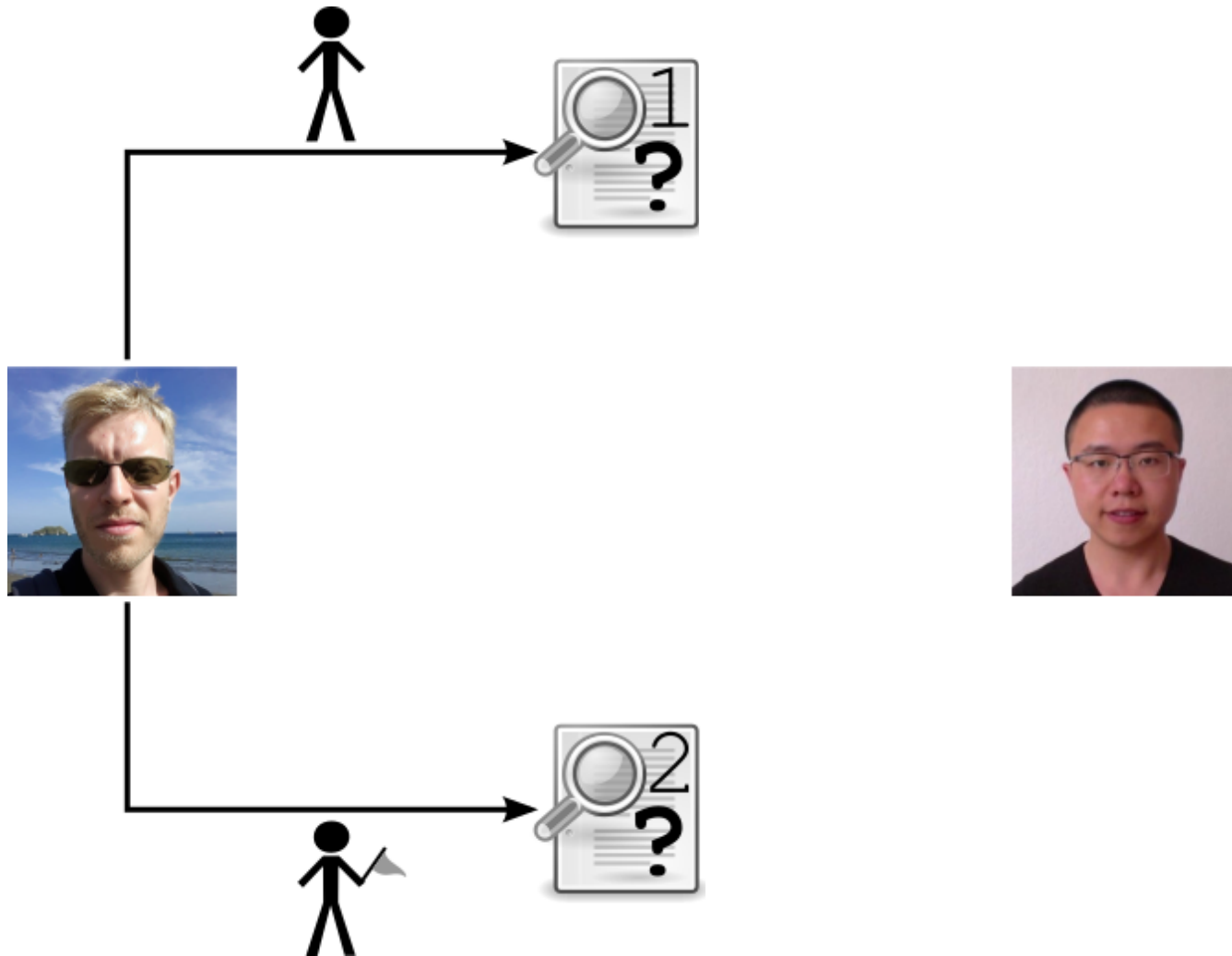
All measurements can be aggregated to *system performance measures*



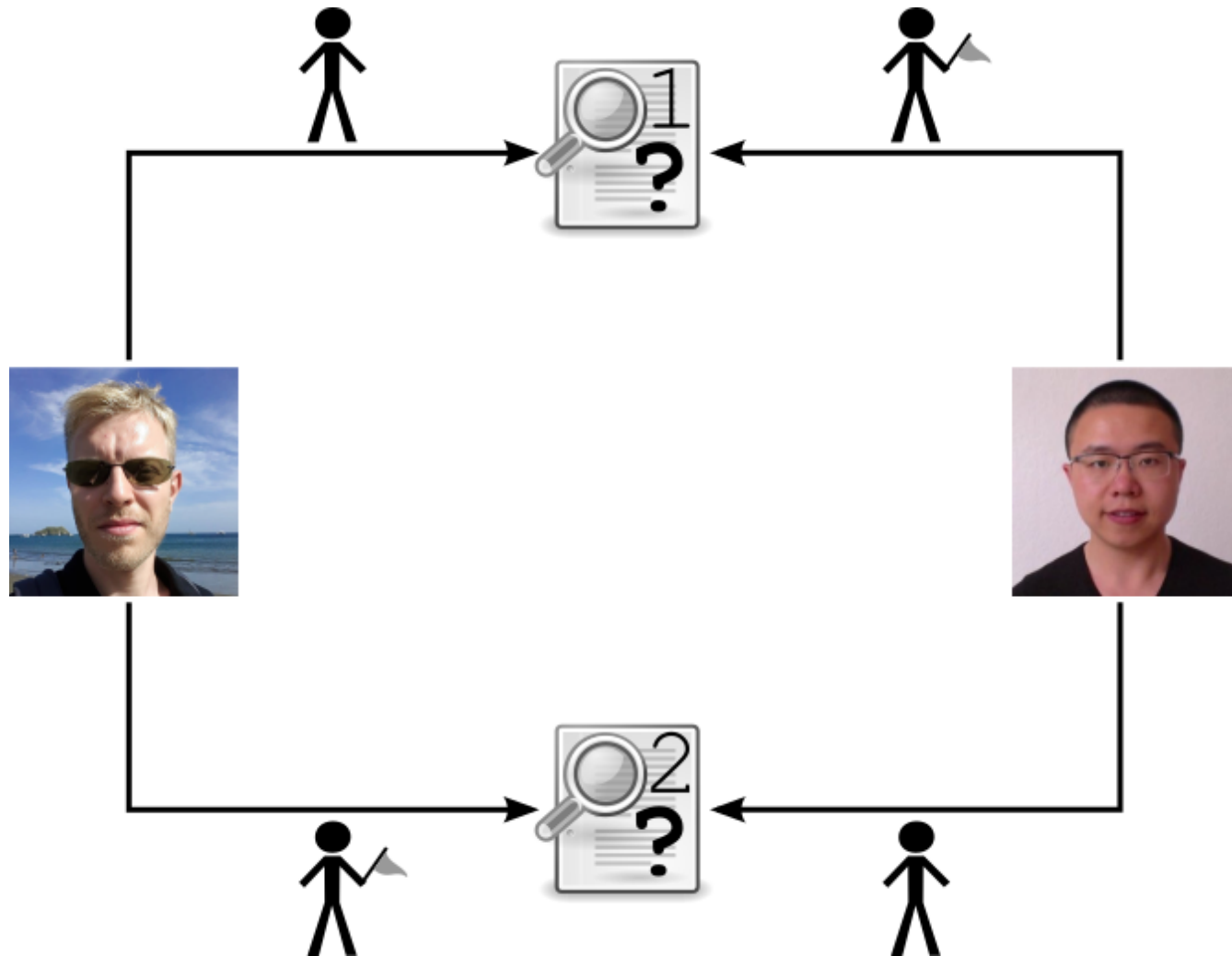
Assignment logic can control biases and balance confidence



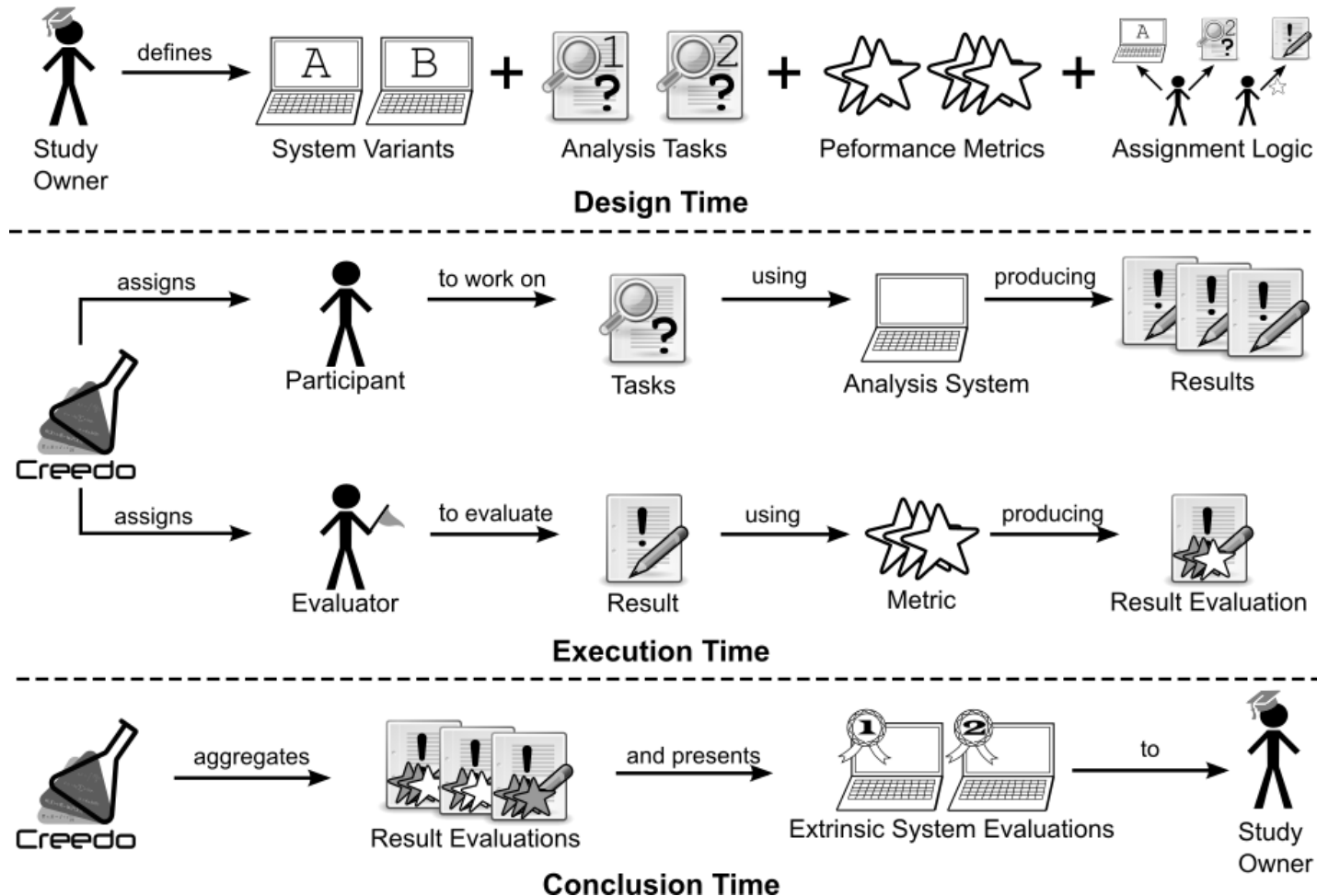
Assignment logic can control biases and balance confidence



Assignment logic can control biases and balance confidence



Creedo organizes study process



Yes, we can



mario@realKD.org