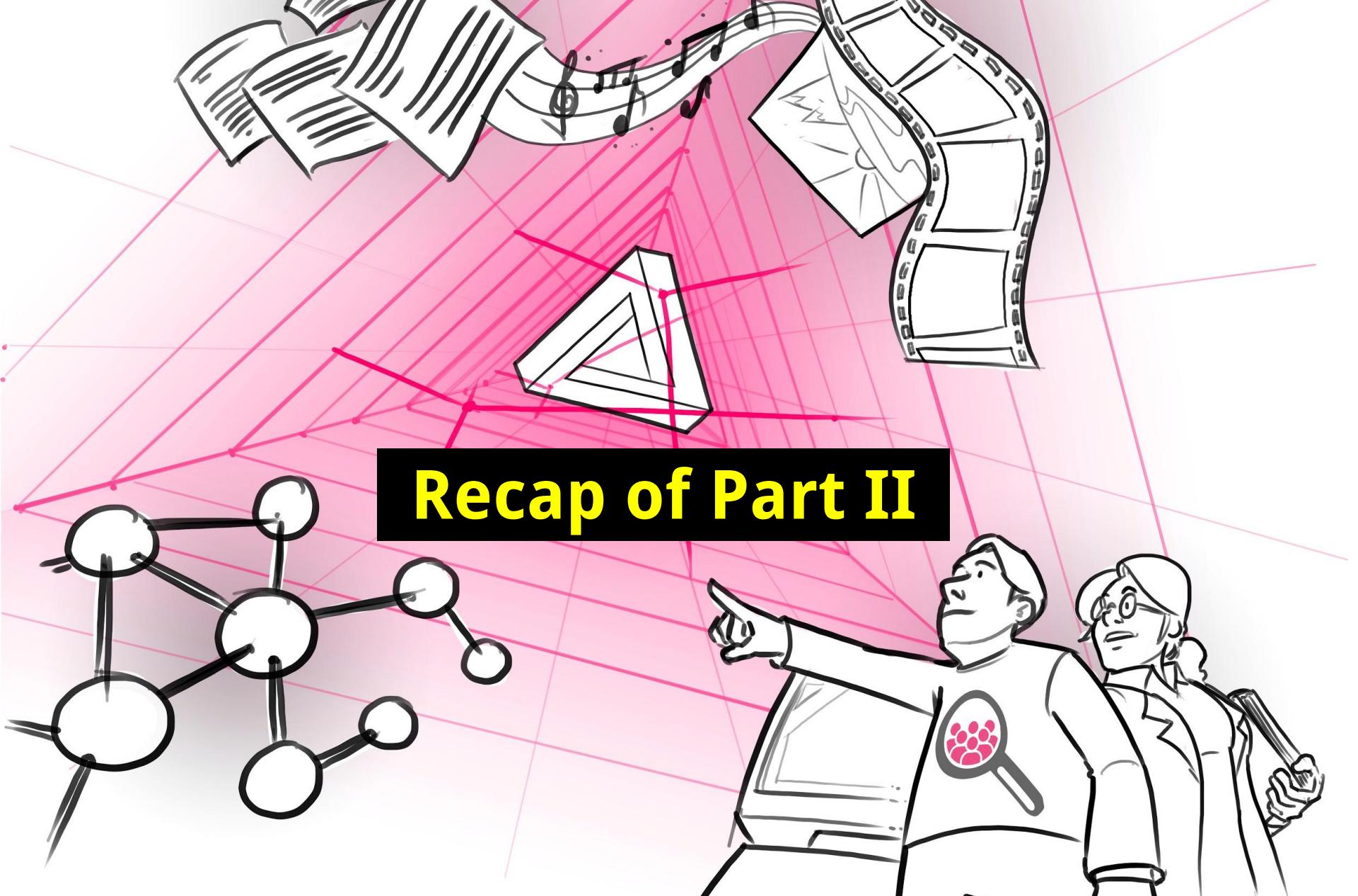


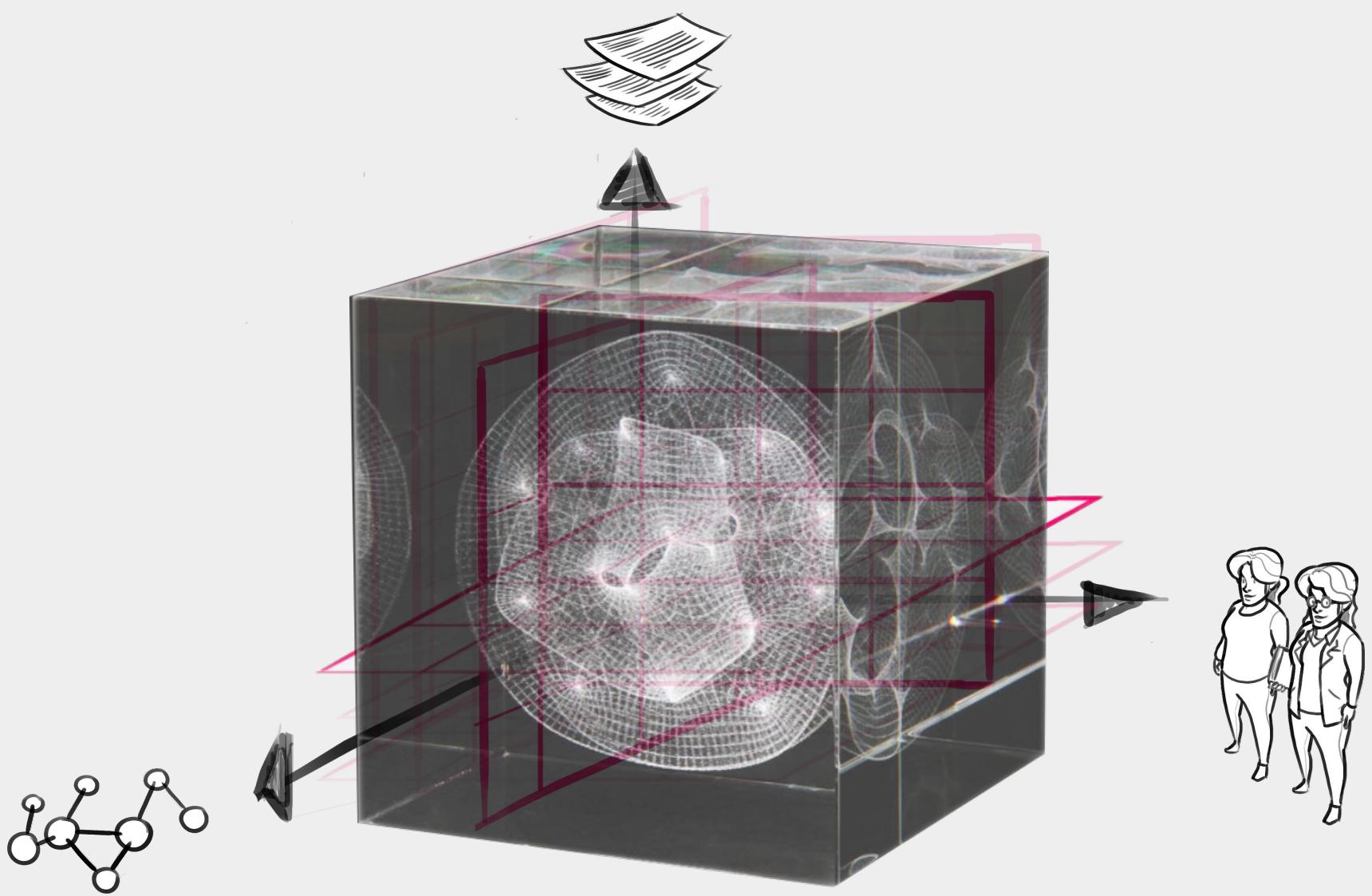
# CROWDTRUTH TUTORIAL

## Part III: Quality Assessment Data Processing & Metrics

Lora Aroyo, Anca Dumitrache, Oana Inel, Chris Welty

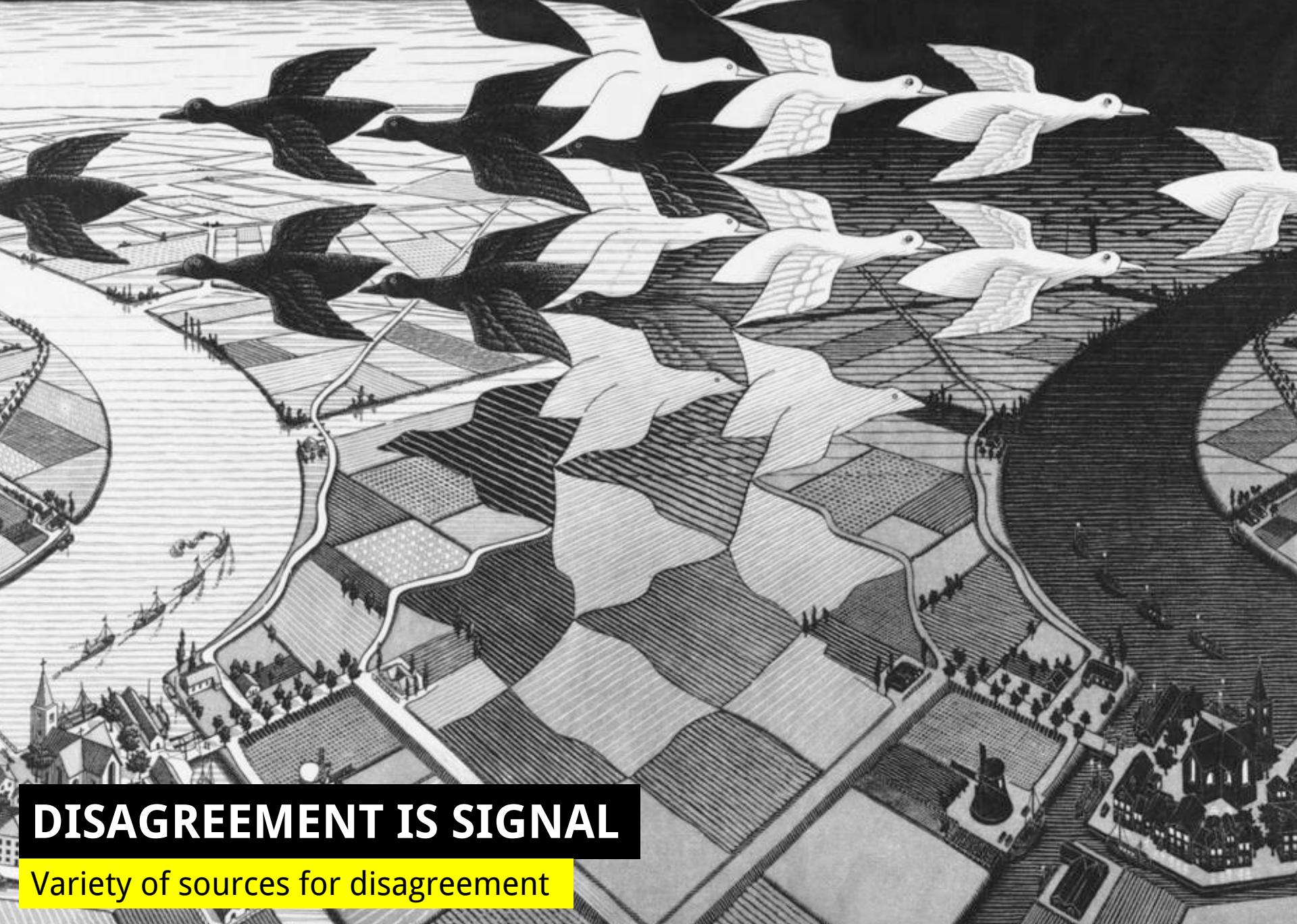


## Recap of Part II



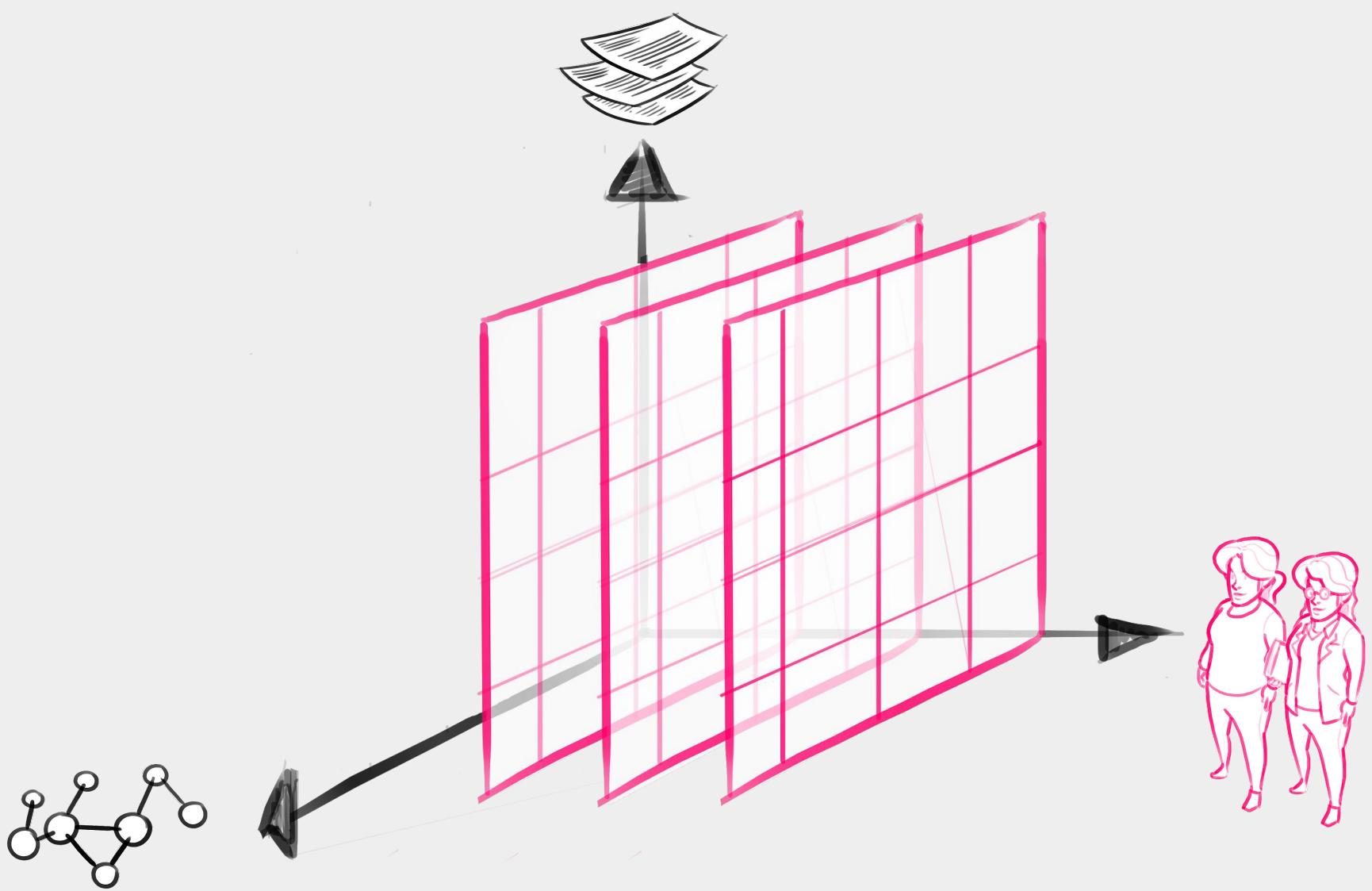
# HYPER-DIMENSIONAL SPACE

3 Sentence Slices



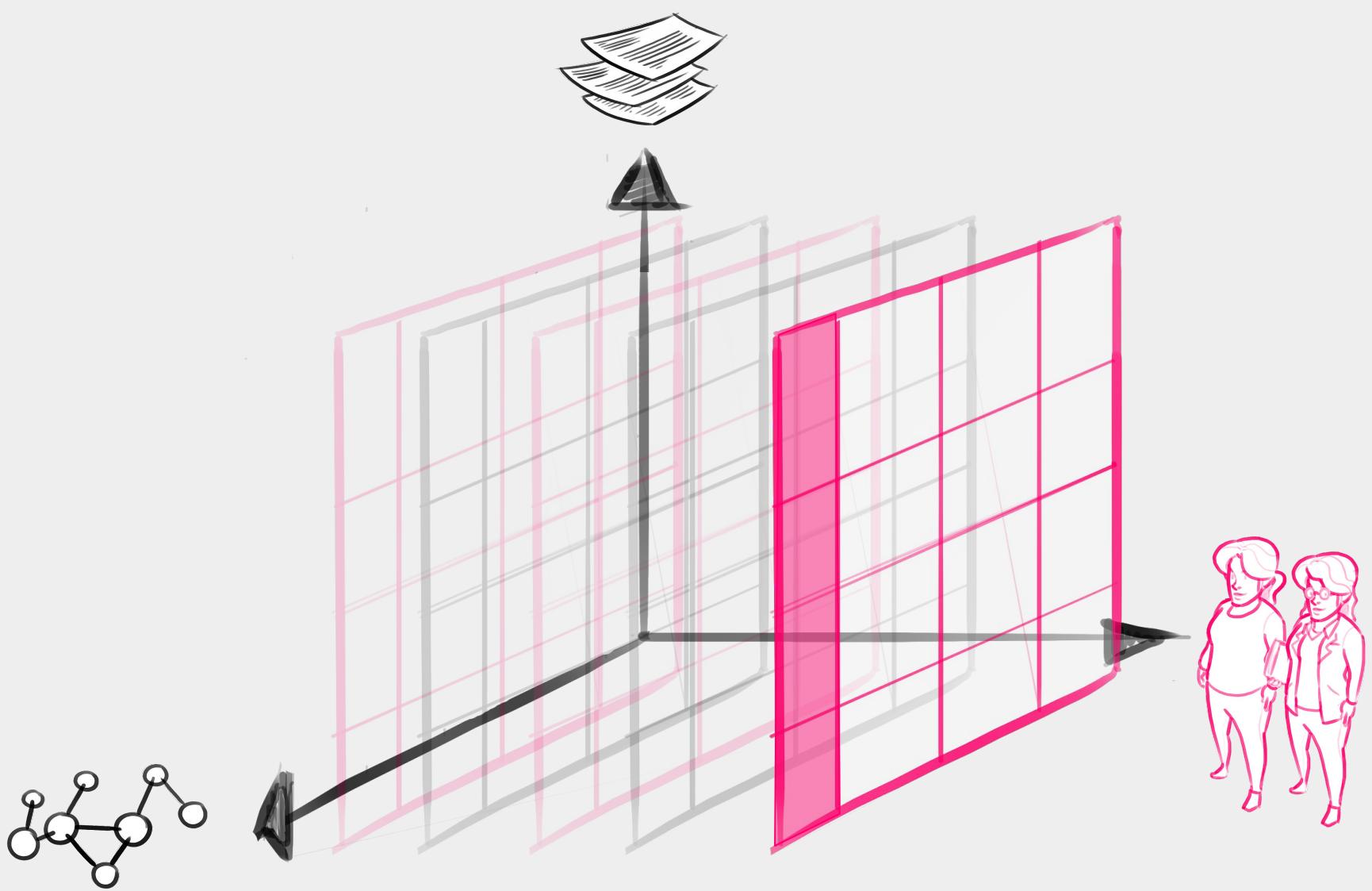
## DISAGREEMENT IS SIGNAL

Variety of sources for disagreement



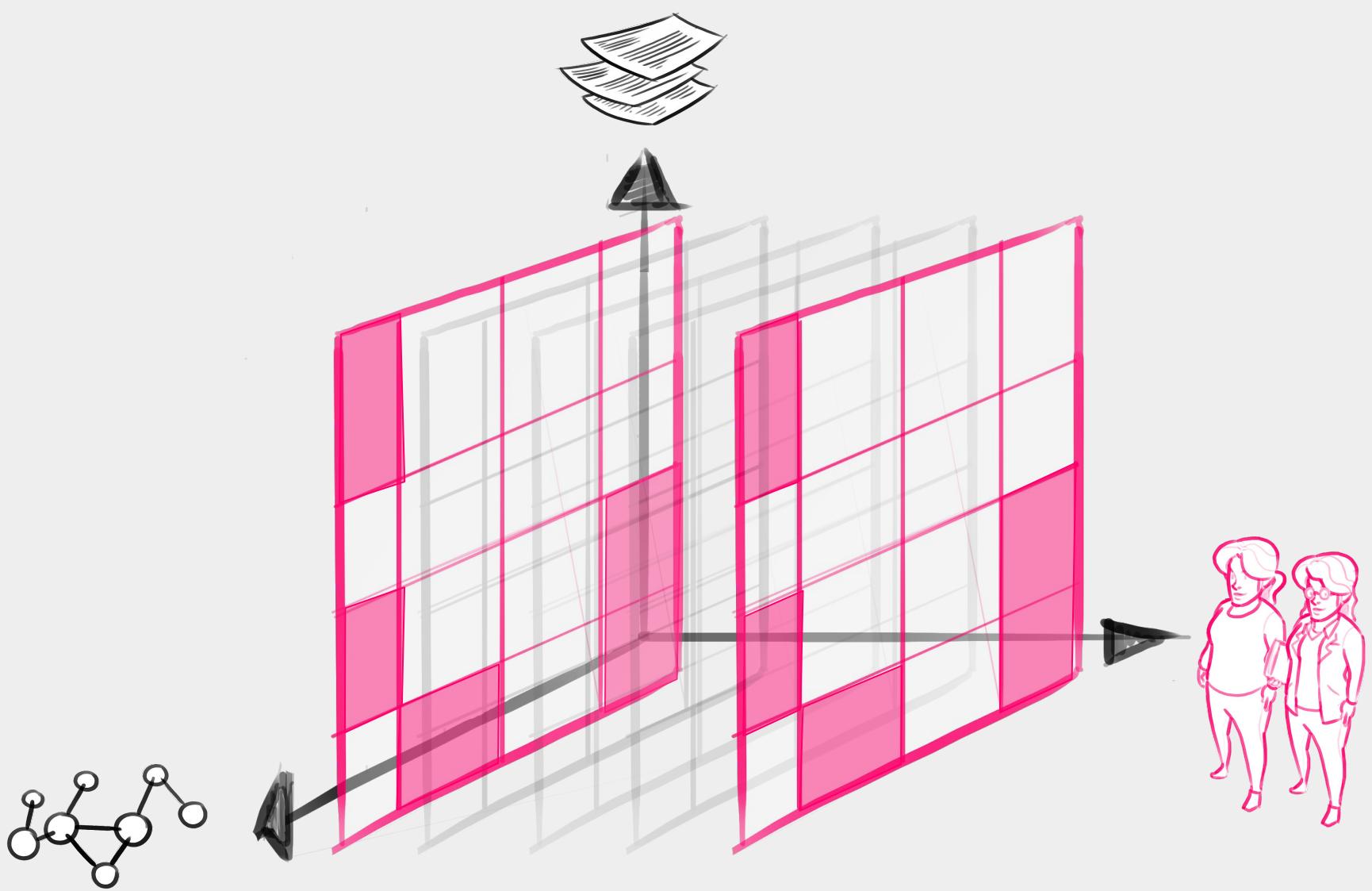
## DISAGREEMENT IS SIGNAL

Source 1: People's bias & perspective



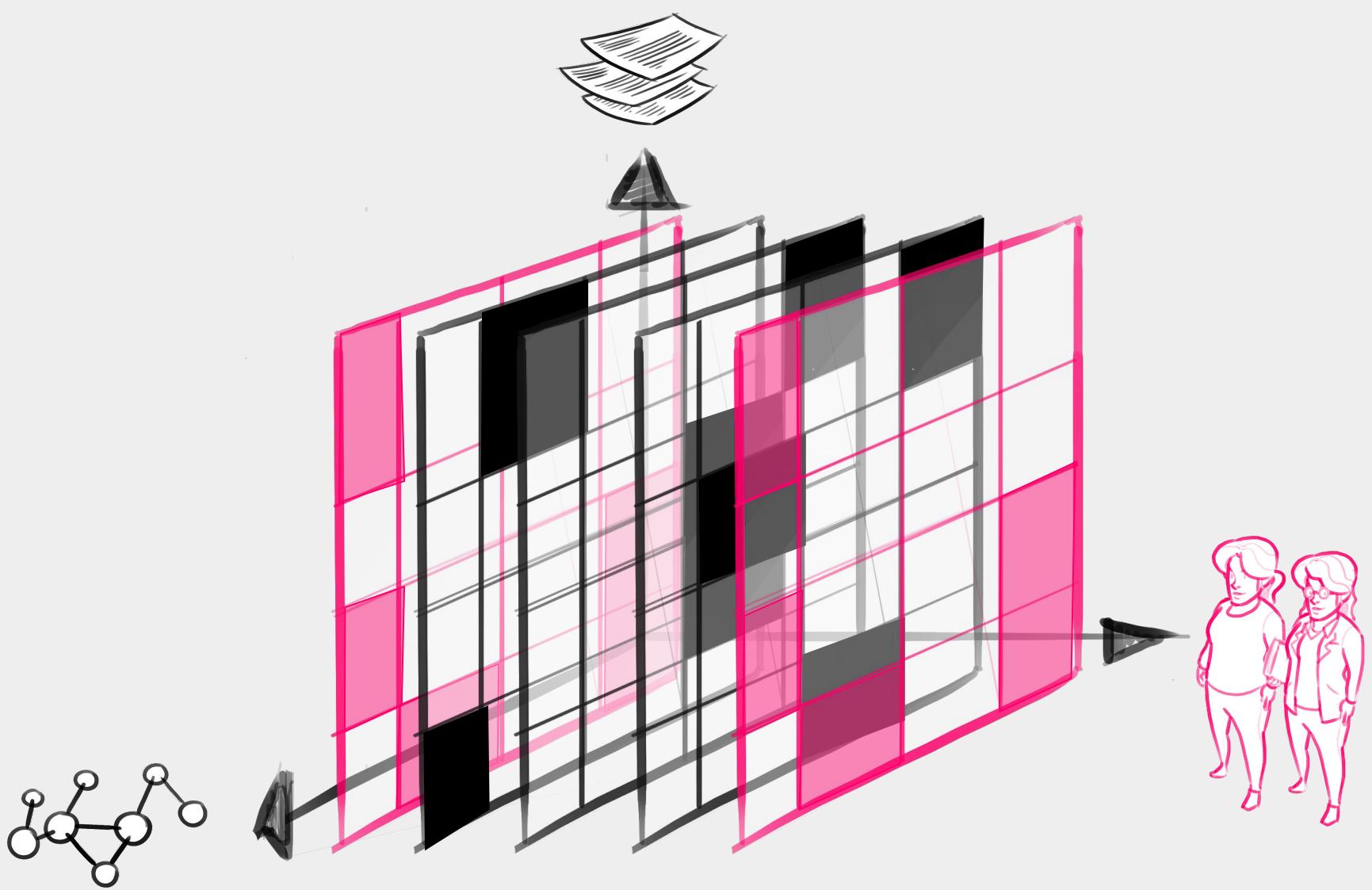
## DISAGREEMENT IS SIGNAL

Source 1: Worker systematically give same answer



## DISAGREEMENT IS SIGNAL

Source 1: Worker systematically give same answer



## DISAGREEMENT IS SIGNAL

Source 1: Worker systematically give same answer

# DO THE SENTENCES EXPRESS A TREAT RELATION?

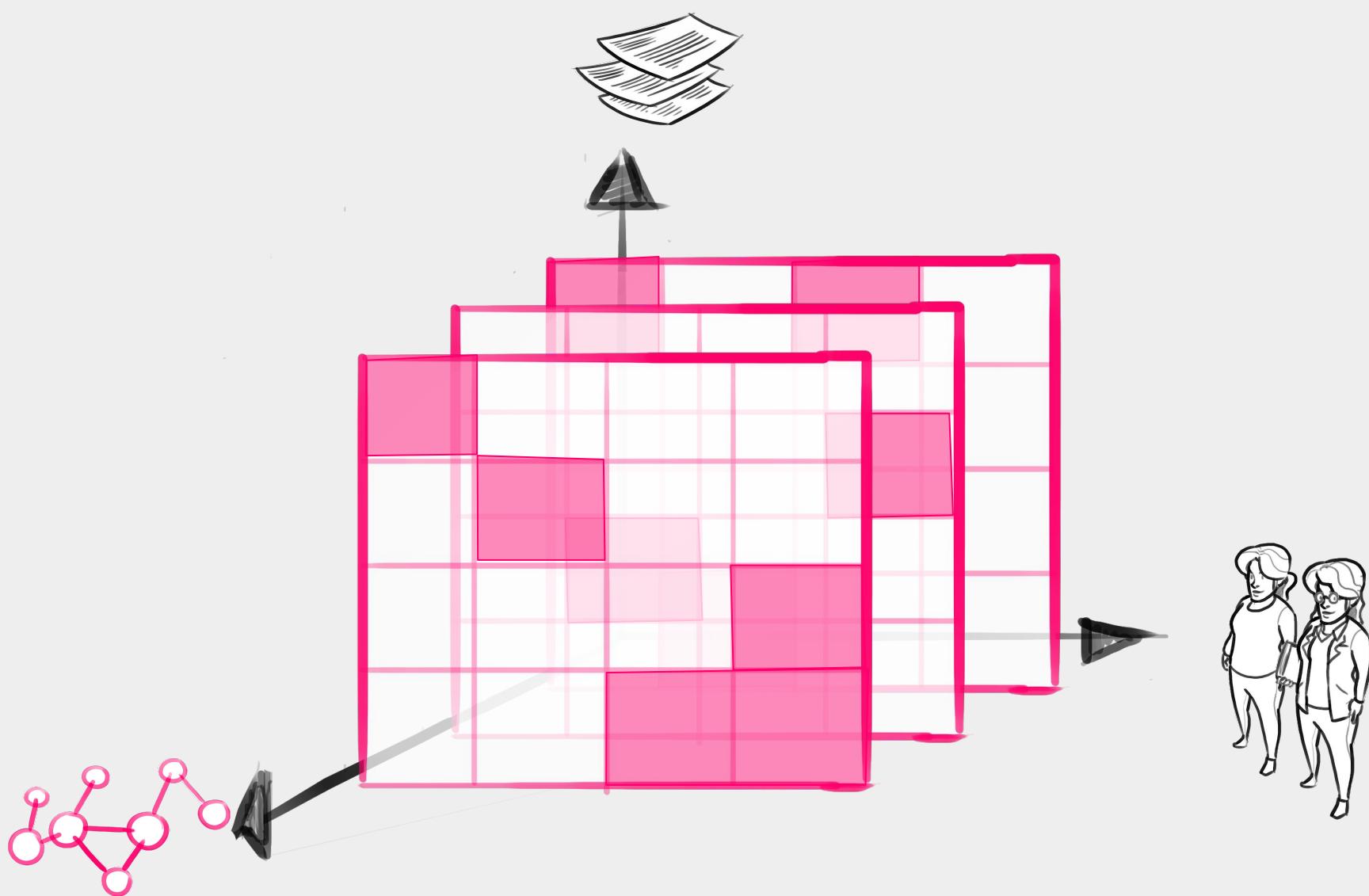
**ANTIBIOTICS** are the first line treatment for vectors of **TYPHUS**.

**QUININE** is not a reliable cure for **MALARIA**.

Worker	Sentence (1)	Sentence (2)
Worker 1	yes	no
Worker 2	yes	no
Worker 3	yes	
Worker 4		no
Worker 5	no	yes

## DISAGREEMENT CAUSE

Source 1: Low Quality Workers



## DISAGREEMENT IS SIGNAL

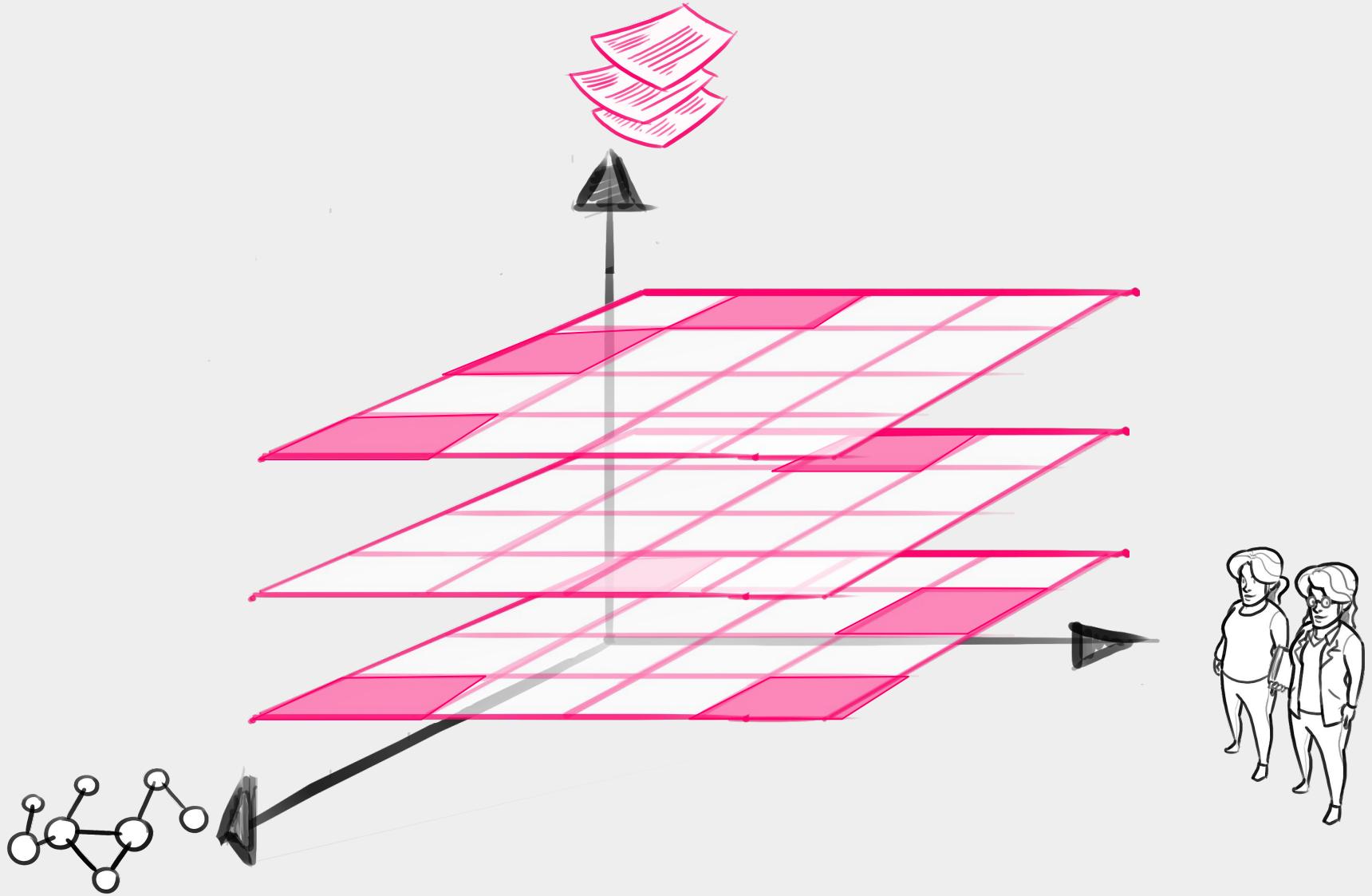
Source 2: Target semantics

# WHAT IS THE RELATION EXPRESSED? TREAT OR PREVENT?

**METHYLERGOMETRINE** is a blood vessel constrictor most commonly used to control excessive **BLEEDING**.

## DISAGREEMENT CAUSE

Source 2: Target semantics



## DISAGREEMENT IS SIGNAL

Source 3: Sentence Clarity

# DO THE SENTENCES EXPRESS A TREAT RELATION? BETWEEN CHLOROQUINE AND MALARIA ?

Rheumatoid arthritis and **MALARIA** have been treated with **CHLOROQUINE** for decades.

95%

For prevention of malaria, use only in individuals traveling to malarious areas where **CHLOROQUINE** resistant P. falciparum **MALARIA** has not been reported.

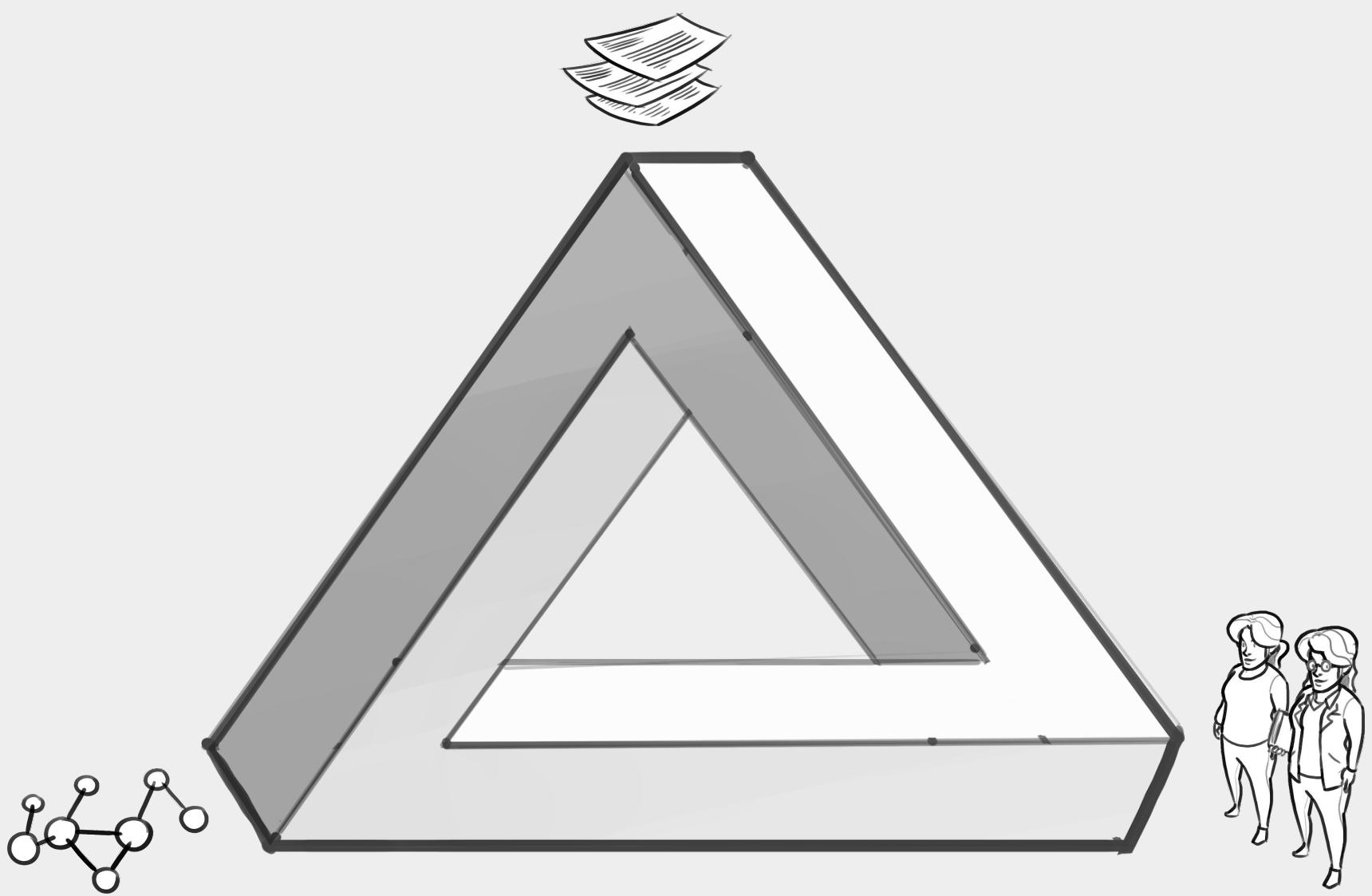
75%

Among 56 subjects reporting to a clinic with symptoms of **MALARIA** 53 (95%) had ordinarily effective levels of **CHLOROQUINE** in blood.

50%

## DISAGREEMENT CAUSE

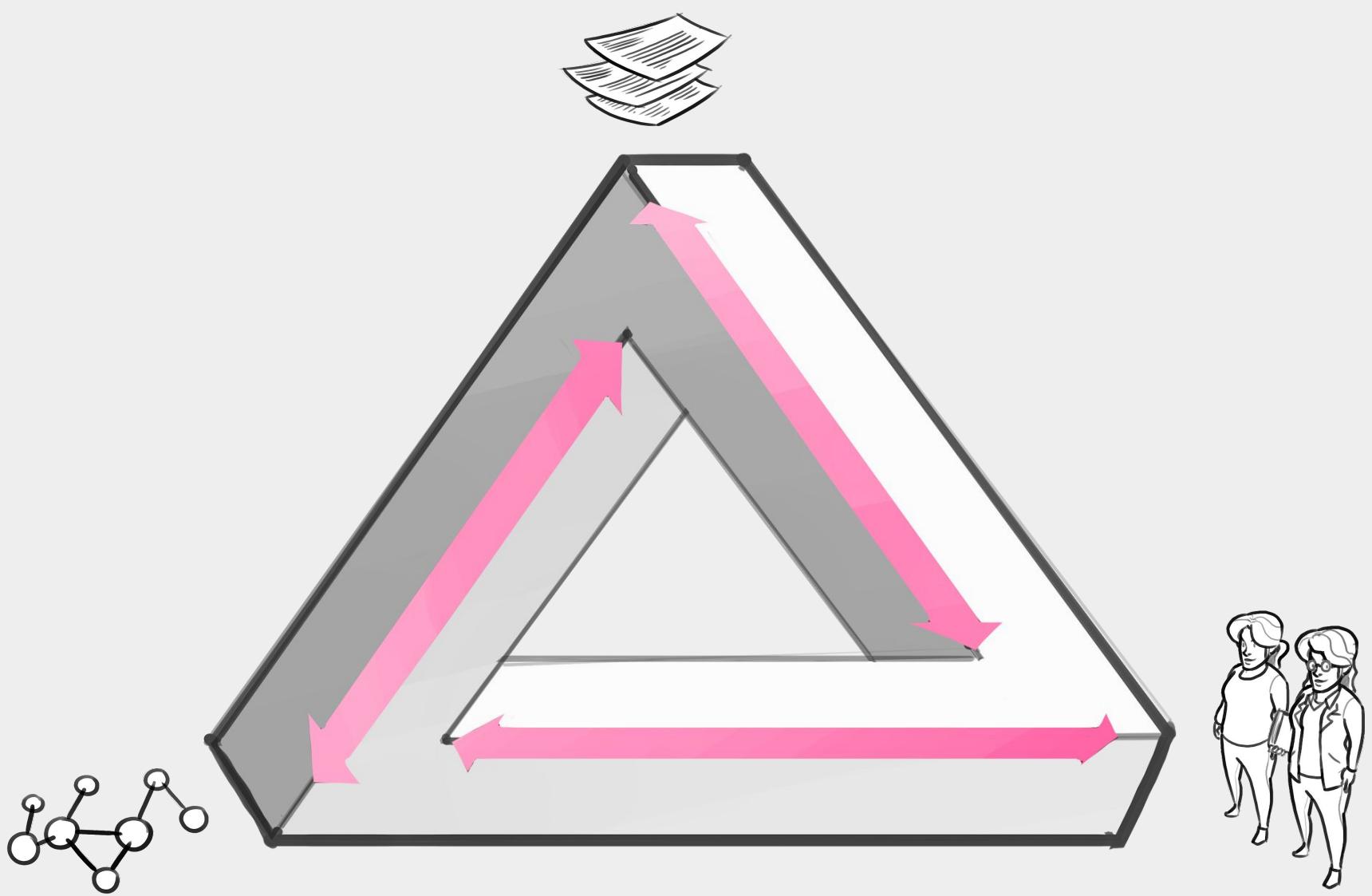
Source 3: Sentence Clarity



# TRIANGLE OF MEANING

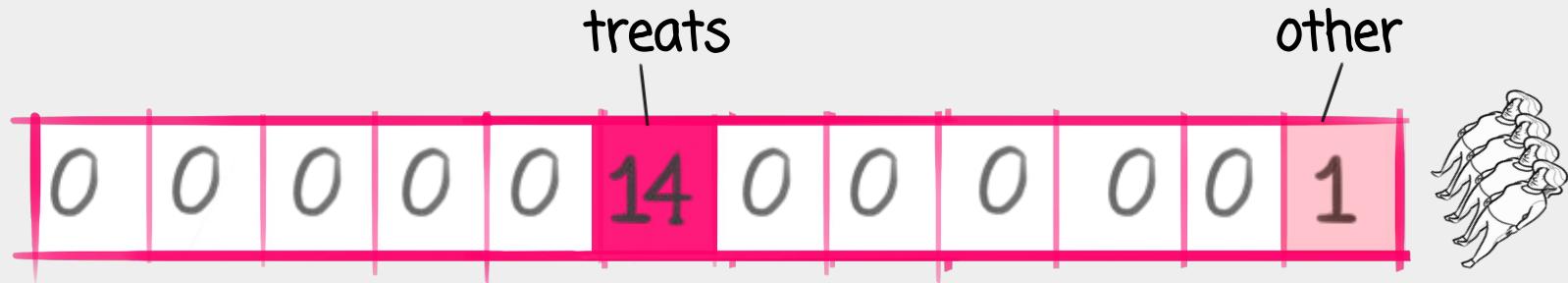
Model of semantic interpretation

*(Ogden & Richards, 1936)  
The Meaning of Meaning*



## TRIANGLE OF MEANING

Model of semantic interpretation



## TRIANGLE OF MEANING

CrowdTruth metrics for quality assessment

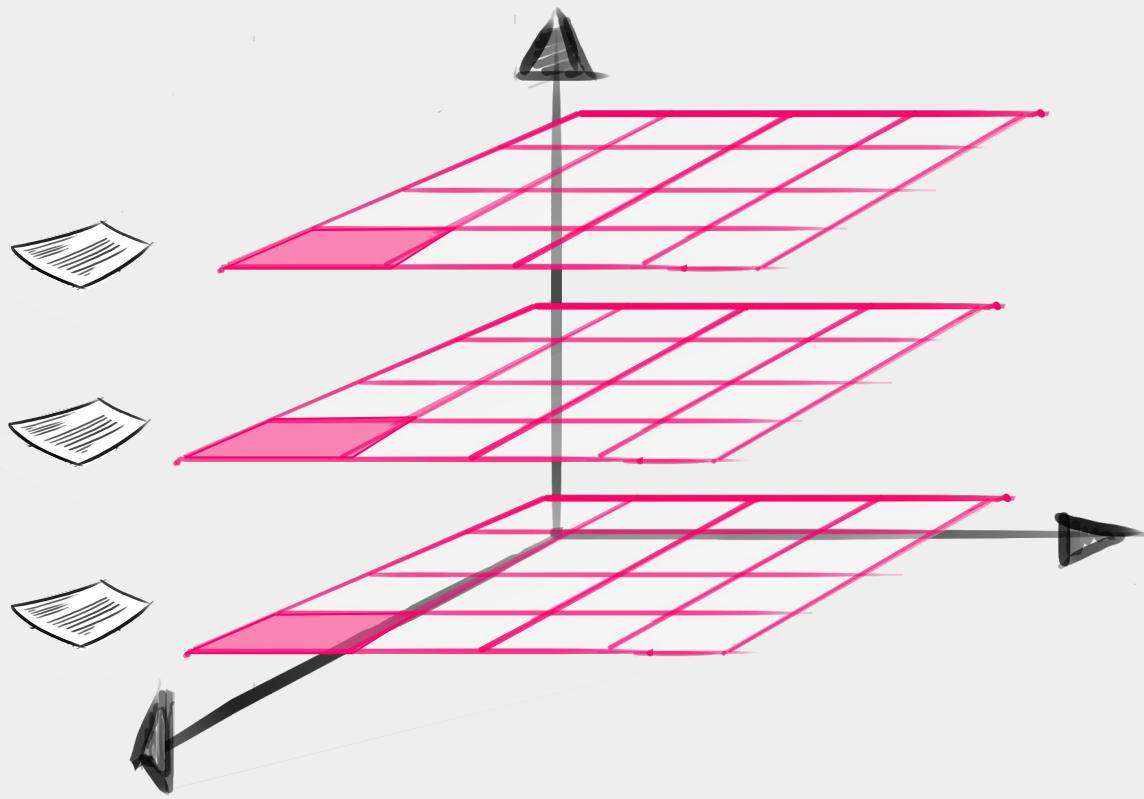


treats

0	0	0	0	0	14	0	0	0	0	0	1
---	---	---	---	---	----	---	---	---	---	---	---



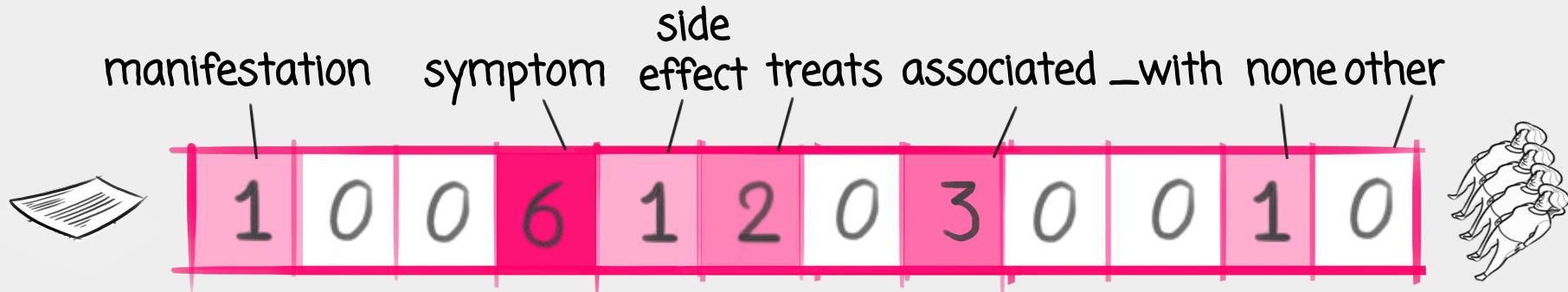
other



## QUALITY ASSESSMENT

Spam

Among 56 subjects reporting to a clinic with symptoms of **MALARIA** 53 (95%) had ordinarily effective levels of **CHLOROQUINE** in blood.



## QUALITY ASSESSMENT

Sentence ambiguity

manifestation contraindicates treats associated\_with none other



1	0	3	0	0	6	0	0	2	0	1	4
---	---	---	---	---	---	---	---	---	---	---	---



manifestation contraindicates treats associated\_with none other



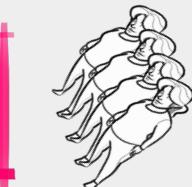
2	0	2	0	0	8	0	0	0	0	1	0
---	---	---	---	---	---	---	---	---	---	---	---



treats



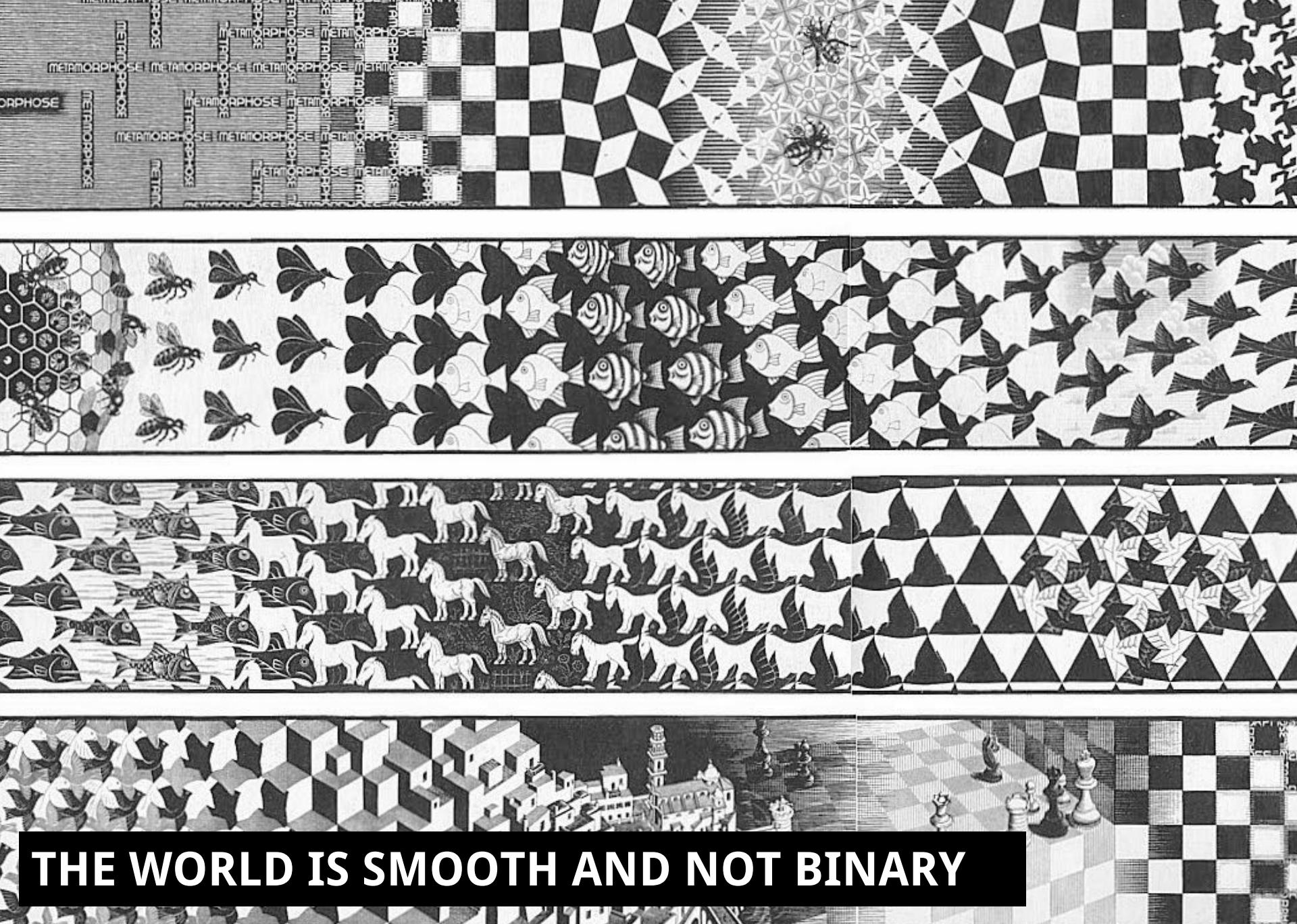
0	0	0	0	0	14	0	0	0	0	0	1
---	---	---	---	---	----	---	---	---	---	---	---



other

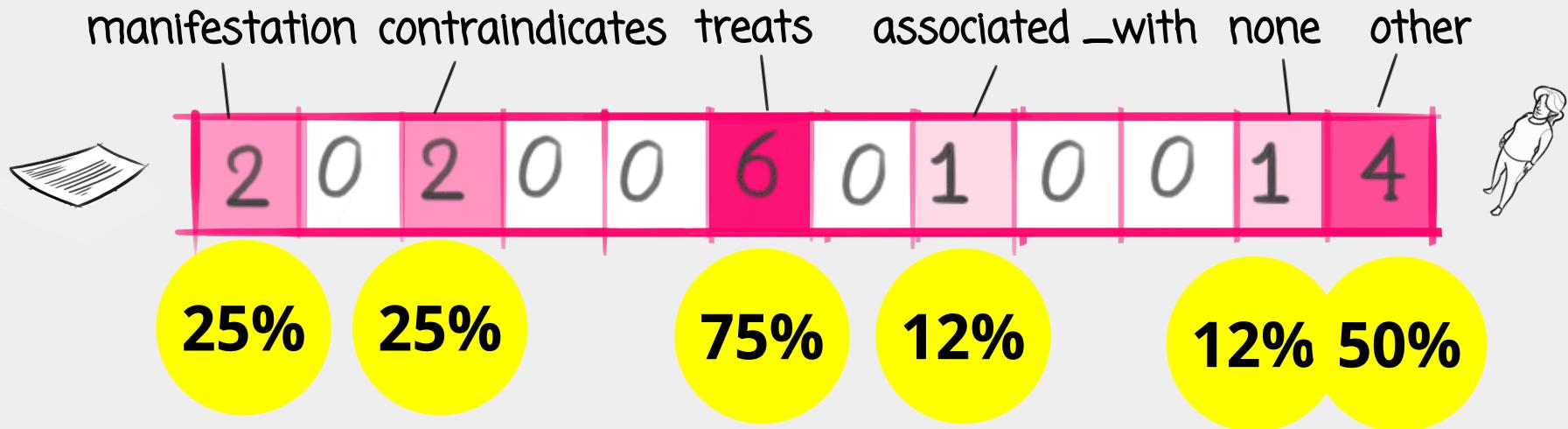
## TREATS RELATION

Yes or No?



**THE WORLD IS SMOOTH AND NOT BINARY**

For prevention of malaria, use only in individuals traveling to  
malarious areas where **CHLOROQUINE** resistant P.  
*falciparum* **MALARIA** has not been reported.



## CROWDTRUTH METRICS

Agreement as percentage

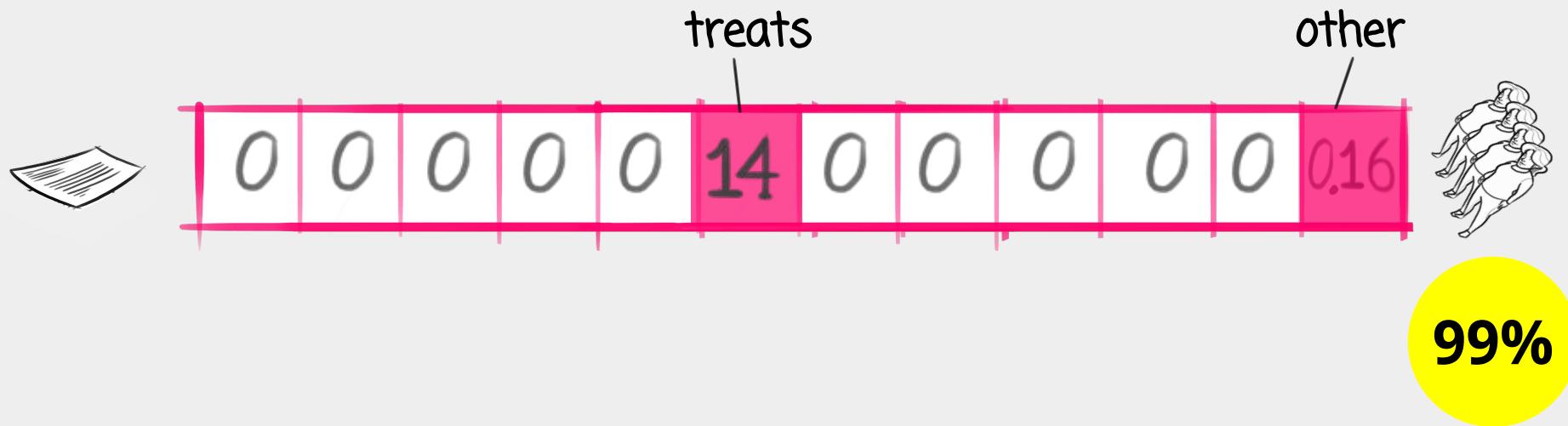
For prevention of malaria, use only in individuals traveling to  
malarious areas where **CHLOROQUINE** resistant P.  
*falciparum* **MALARIA** has not been reported.



## CROWDTRUTH METRICS

Applying all sides of the triangle

For prevention of malaria, use only in individuals traveling to  
malarious areas where **CHLOROQUINE** resistant P.  
*falciparum* **MALARIA** has not been reported.



## CROWDTRUTH METRICS

Applying all sides of the triangle

# APPLYING ALL SIDES OF THE TRIANGLE

## Sentence quality:

overall worker agreement over one sentence

## Worker quality:

overall agreement between one worker and the other workers

## Annotation quality:

overall worker agreement over one annotation

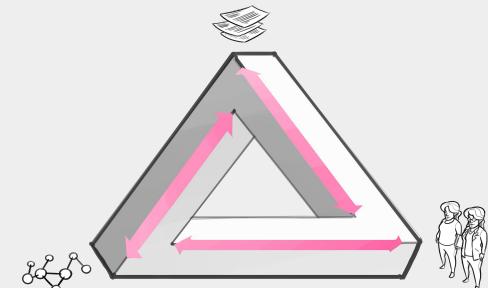
Calculated using **average cosine similarity** over **sentence vectors**.

Quality metrics are **inter-dependent**:

e.g. ambiguous sentences count less when determining worker quality, because disagreement is expected

## CROWDTRUTH QUALITY METRICS

Ambiguity at any corner disseminates in the other corners



**Frame-Sentence Score (FSS):** the degree with which a particular frame matches the sense of the word in the sentence

$$FSS(s, f) = \frac{\sum \text{workers that picked } f \text{ in } s \cdot \text{worker quality}}{\sum \text{workers for } s \cdot \text{worker quality}}$$

**Sentence Quality Score (SQS):** overall worker agreement over one sentence, measured with cosine similarity

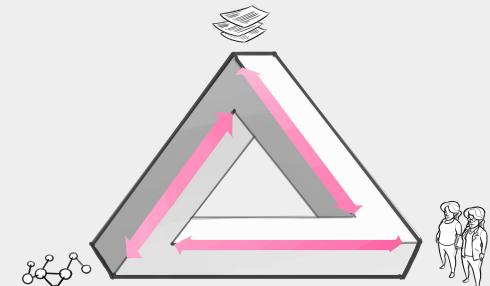
$$SQS(s) = \frac{\sum_{w1, w2 \in \text{workers for } s} \text{weighted cos sim}(\vec{w1}, \vec{w2}, FQS) \cdot \text{worker quality}(w1, w2)}{\sum_{w1, w2 \in \text{workers for } s} \text{worker quality}(w1, w2)}$$

**Frame Quality Score (FQS):** agreement over a frame in all sentences where the frame was picked at least once

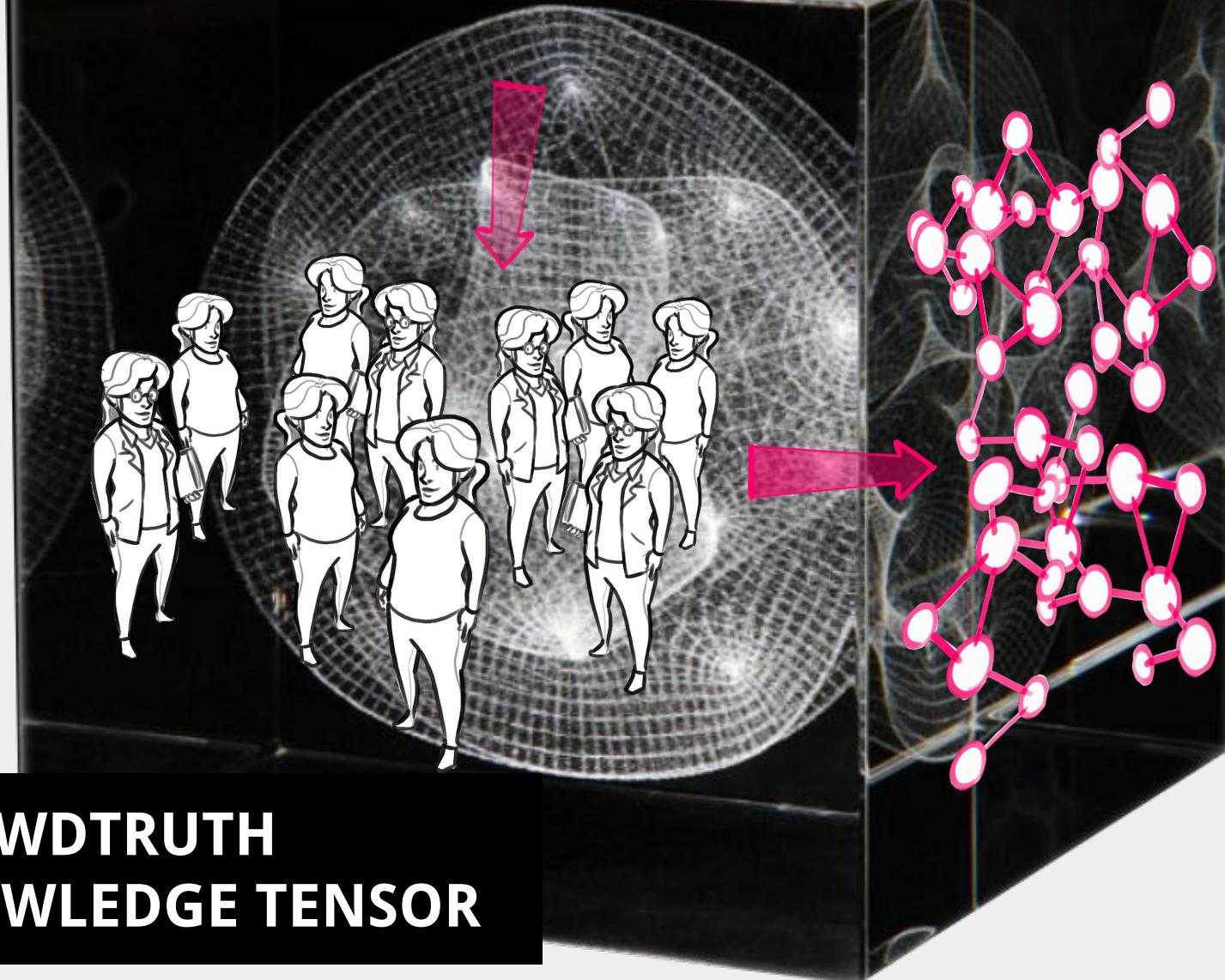
$$FQS(f) = \frac{\sum_{FSS(s, f) > 0} FSS(s, f) SQS(s)}{\sum_{FSS(s, f) > 0} SQS(s)}$$

## CROWDTRUTH QUALITY METRICS

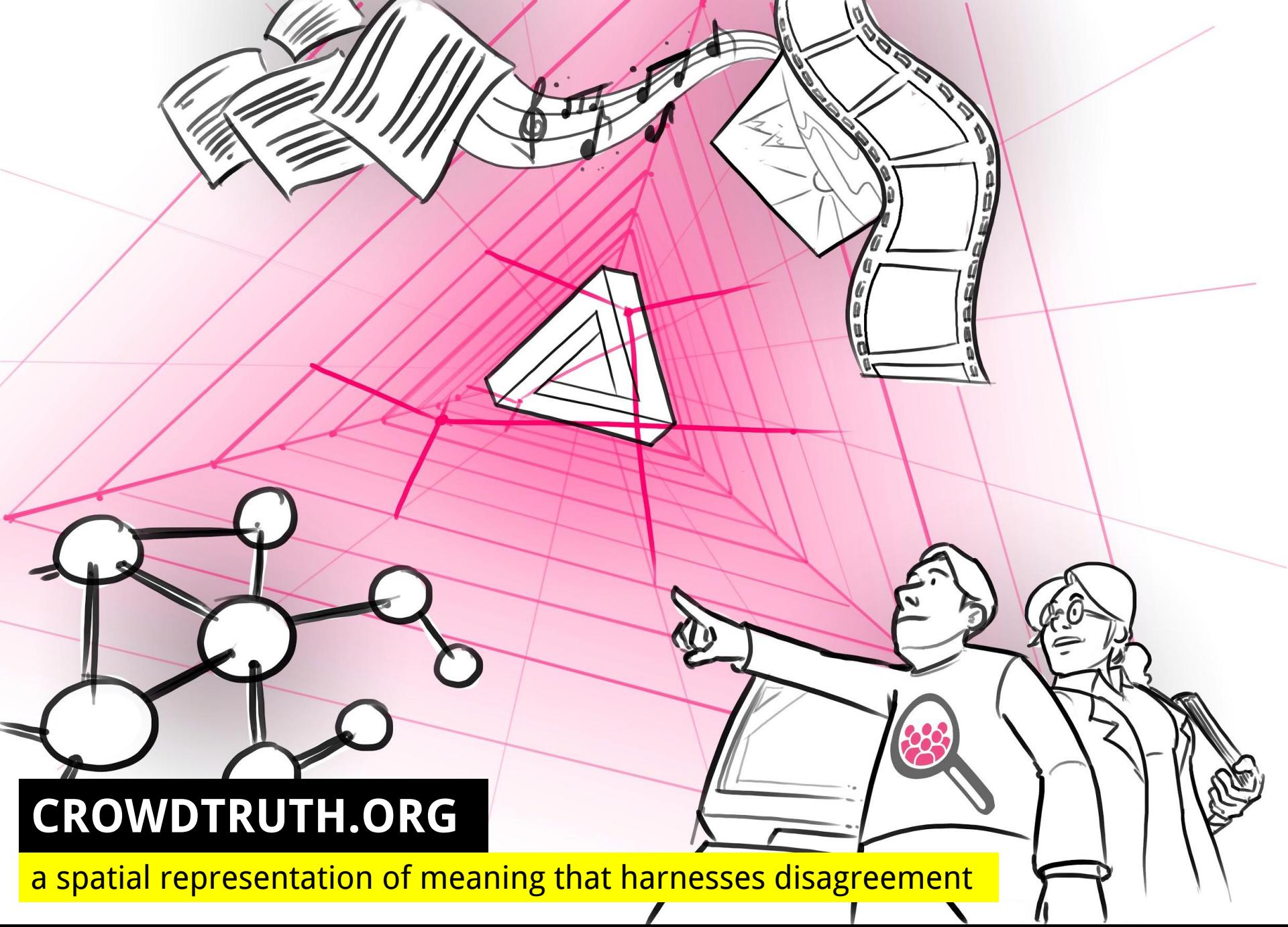
Ambiguity at any corner disseminates in the other corners



For prevention of malaria, use only in individuals traveling to malarious areas where **CHLOROQUINE** resistant *P. falciparum* **MALARIA** has not been reported.



# CROWDTRUTH KNOWLEDGE TENSOR



**CROWDTRUTH.ORG**

a spatial representation of meaning that harnesses disagreement

# CROWDTRUTH.ORG

GITHUB.COM/CROWDTRUTH/CROWDTRUTH-CORE  
PYPI.ORG/PROJECT/CROWDTRUTH

DATA.CROWDTRUTH.ORG

<http://lora-aroyo.org> @laroyo

