Disaster Monitoring

A novel Human Computation System based on Game with A Purpose

Changkun Ou
Yifei Zhan
Zhe Li

Recap: Functionalities

Game With A Purpose

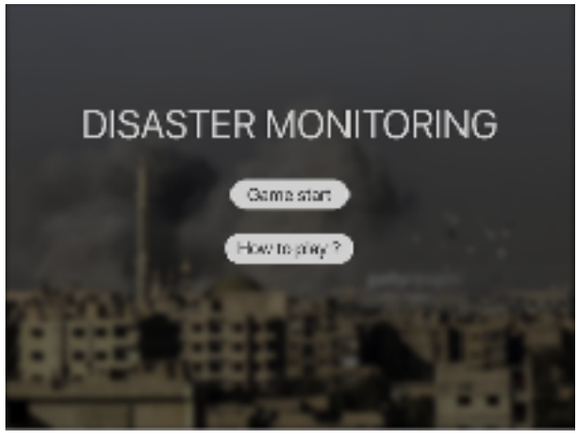
- Image tagging game

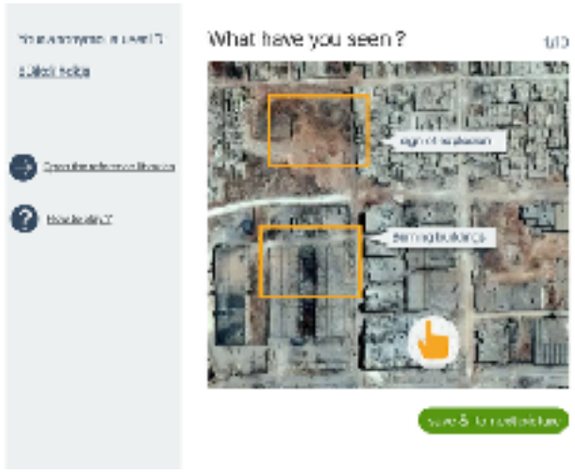
- Single Players

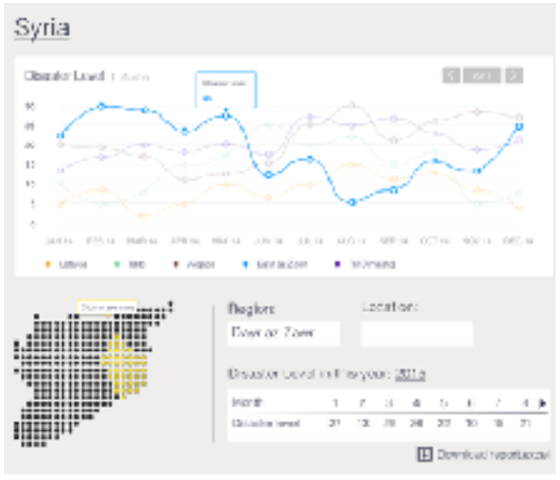
- Motivation: altruism

Review: Prototype

https://invis.io/WQCKJRPJK#/243555585_home-Page







Game with a purpose - GWAP

- Potential Problem : Information leakage

- Possible Solution : Image segmentation

Implementation Requirements: Web-based

- Front-end

  Polymer

- Back-end

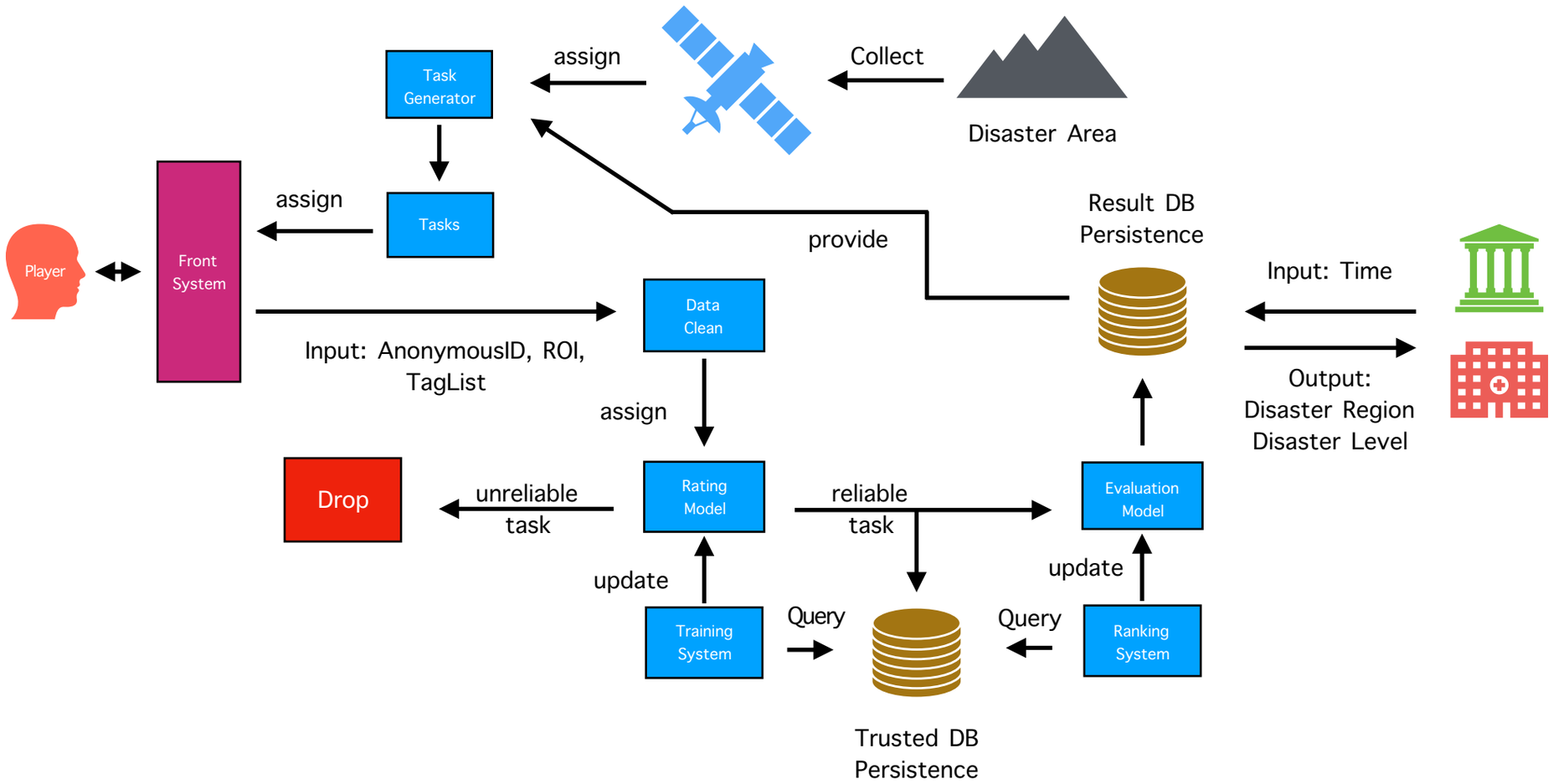
  Node.js / Python

- Database

  MongoDB

Recap: System Design

System Architecture (Recap)



- A novel Human Computation system for disaster monitoring

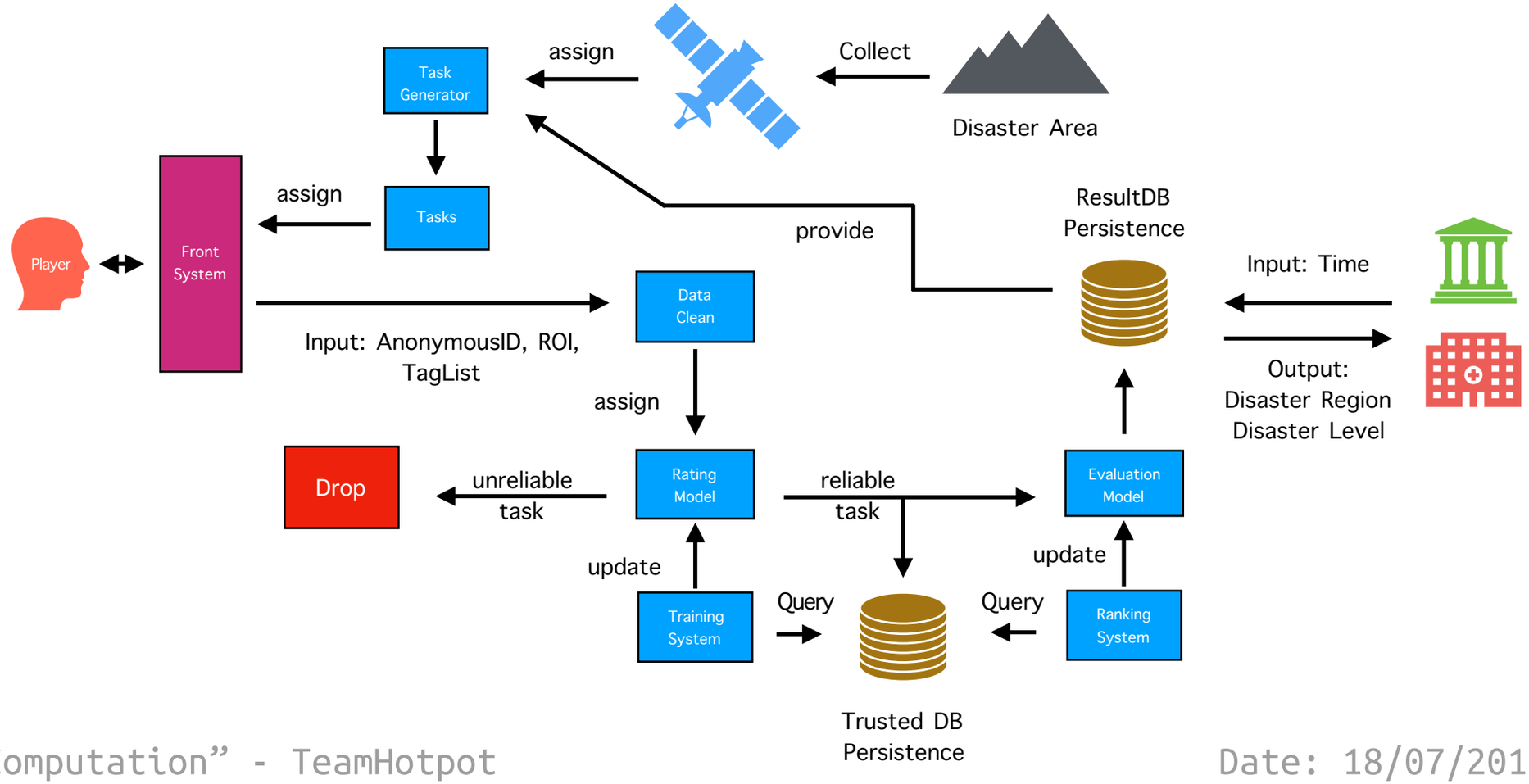
- Players select a ***ROI (Region of interests)*** on satellite images, then tag disaster keywords

Issues from Lab2 session…

- What is the reason to define the ***Player Rating Graph*** weight via **Matching Area Ratio** and **Pearson Correlation**? Any others?

- How to solve the ***graph initialization*** problem?

- How to ***handle new tags*** from players?



Databases

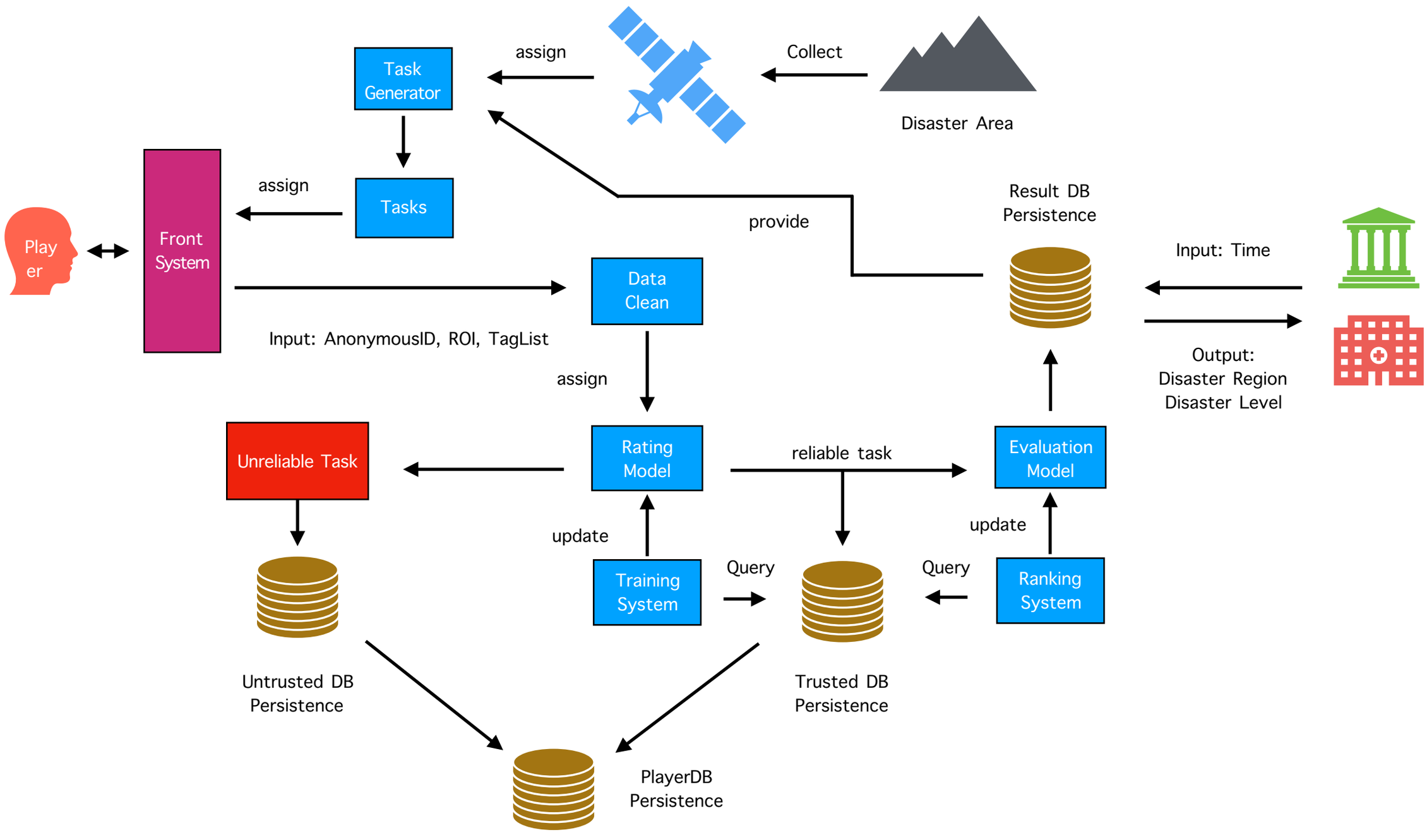
TrustedDB **Fields:**

```
[
  {
    "anonymous_id": number,
    "trust_value": number
    "tasks": [
      {
        "image": image_path,
        "at_time": time,
       "ROI": [
          {
            "latitude": number,
            "longitude": number,
            "tags": [tag1, tag2, ...]
          }, ...
        ]
      }
    ]
  }, ...
]
```

ResultDB **Fields:**

```
[
  {
    “area_id": number,
    "history": [
      {
        “at_time”: time,
        "image": image_path,
       "ROI": [
          {
            "latitude": number,
            “longitude": number,
            "tags": [tag1, tag2, ...]
          }
        ],
       "disaster_level": number
      }, ...
    ]
  }, ...
]
```

System Architecture (*New!*)



Databases (*New!*)

PlayerDB (UntrustedDB & TrustedDB) **Fields:**

```
[
  {
    "anonymous_id": number,
    "reliable": boolean,
    "trust_value": number
    "tasks": [
      {
        "image": image_path,
        "at_time": time,
       "ROI": [
          {
            "latitude": number,
            "longitude": number,
            "tags": [tag1, tag2, ...]
          }, ...
        ]
      }
    ]
  }, ...
]
```

ResultDB **Fields:**

```
[
  {
    “area_id": number,
    "history": [
      {
        “at_time”: time,
        "image": image_path,
        "ROI": [
          {
            "latitude": number,
            “longitude": number,
            "tags": [tag1, tag2, ...]
          }
        ],
        "disaster_level": number
      }, ...
    ]
  }, ...
]
```

Solution - Rethink of the Graph definition (1)

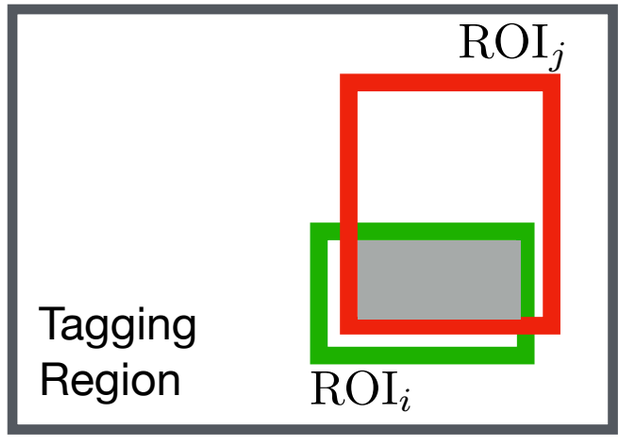
We have multiple player tagging ROIs results and its tags:

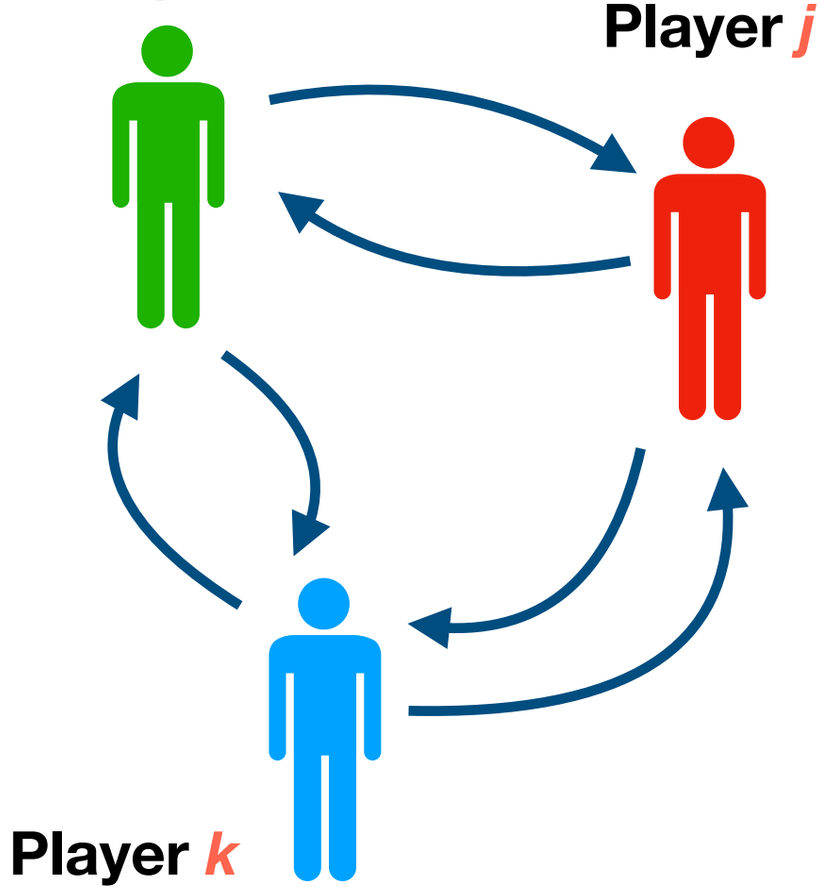
tags = [tag_1, tag_2, ..., tag_n]

For a certain region (or area, or segment image) ***img*** at time ***t*** (ignore due to it's trivial).

Let player ***i*** to rating (->) player ***j***, we have:

$$w_{ij} = \sum_{\mathrm{ROI} \in \mathrm{ROIs}} \frac{\mathrm{ROI}_i \cap \mathrm{ROI}_j}{\mathrm{ROI}_i} \frac{Cov(\mathrm{tags}_i, \mathrm{tags}_j)}{\mathrm{var}(\mathrm{tags}_i)\mathrm{var}(\mathrm{tags}_j)}$$





Solution - Rethink of the Graph definition (1)

We have multiple player tagging ROIs results and its tags:

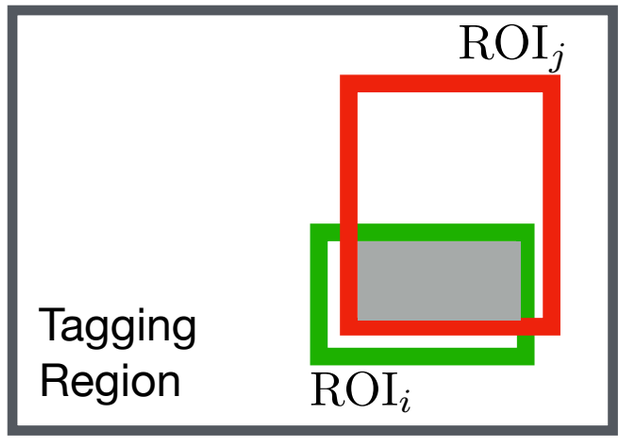
tags = [tag_1, tag_2, ..., tag_n]

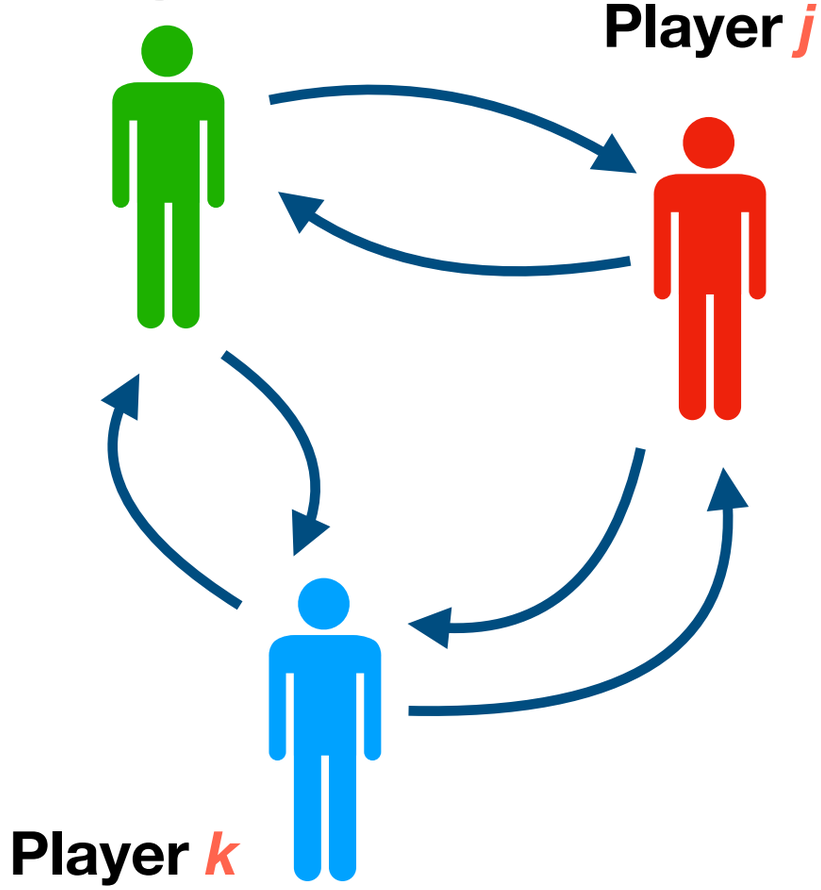
For a certain region (or area, or segment image) ***img*** at time ***t*** (ignore due to it's trivial).

Let player ***i*** to rating (->) player ***j***, we have:

$$w_{ij} = \sum_{\mathrm{ROI} \in \mathrm{ROIs}} \frac{\mathrm{ROI}_i \cap \mathrm{ROI}_j}{\mathrm{ROI}_i} \frac{Cov(\mathrm{tags}_i, \mathrm{tags}_j)}{\mathrm{var}(\mathrm{tags}_i)\mathrm{var}(\mathrm{tags}_j)}$$

Reasonable???

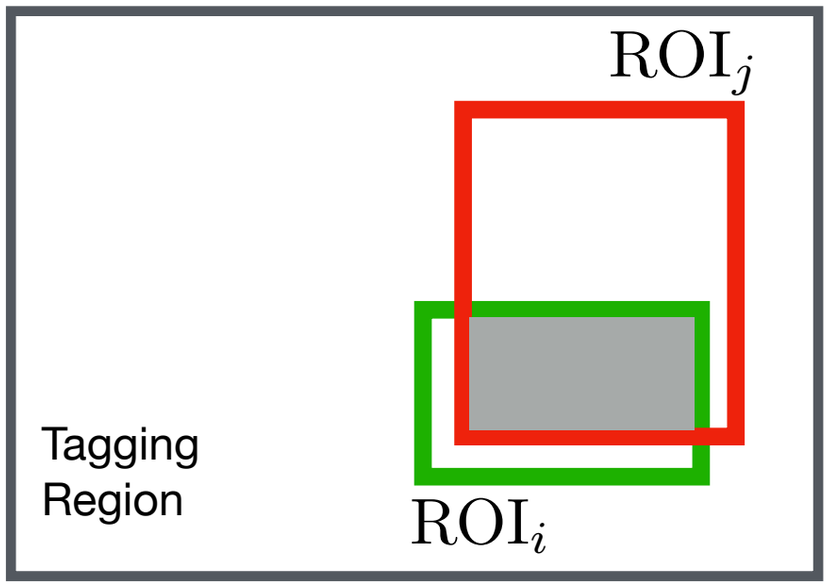


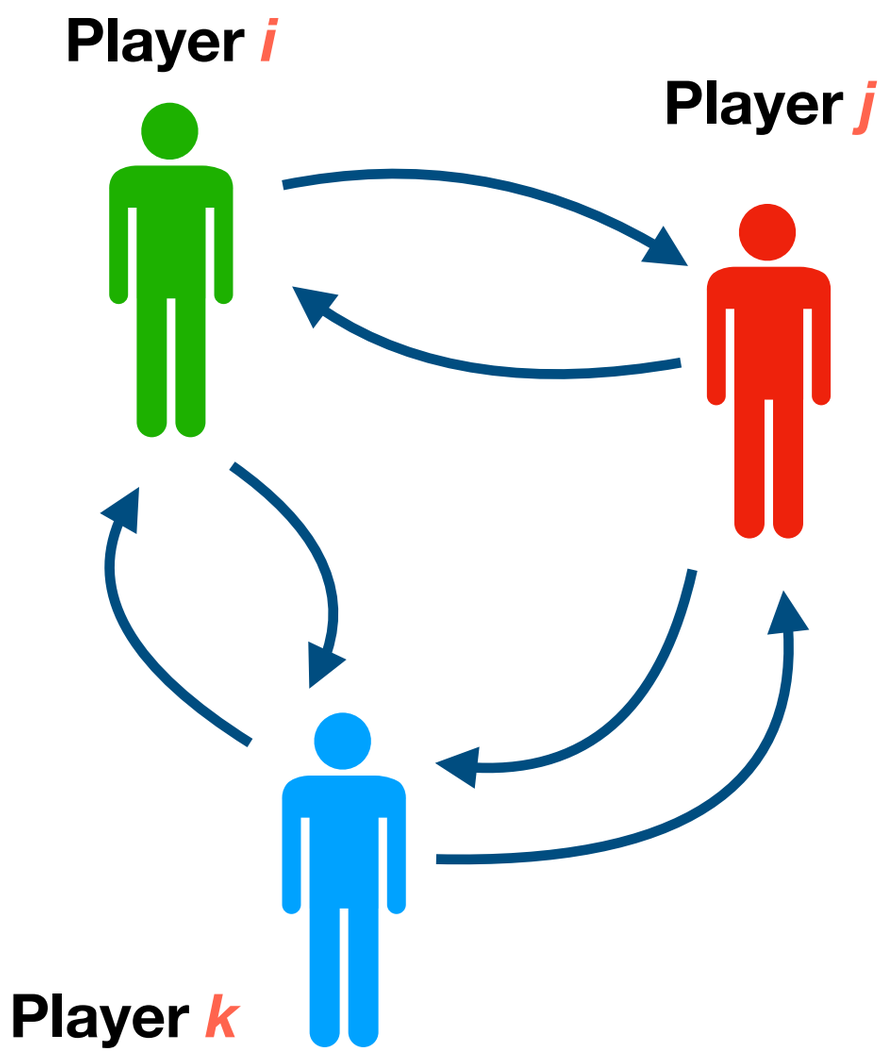


Solution - Rethink of the Graph definition (2)

Which **Matching Area Ratio**?

$$\frac{\mathrm{ROI}_i \cap \mathrm{ROI}_j}{\mathrm{ROI}_i} \qquad \frac{\mathrm{ROI}_i \cap \mathrm{ROI}_j}{\mathrm{ROI}_j} \qquad \frac{\mathrm{ROI}_i \cap \mathrm{ROI}_j}{\mathrm{ROI}_i \cup \mathrm{ROI}_j}$$

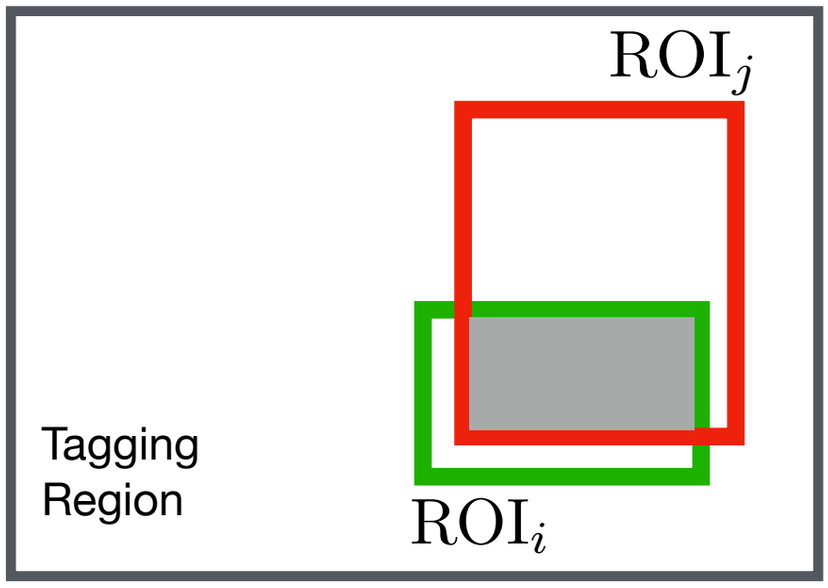


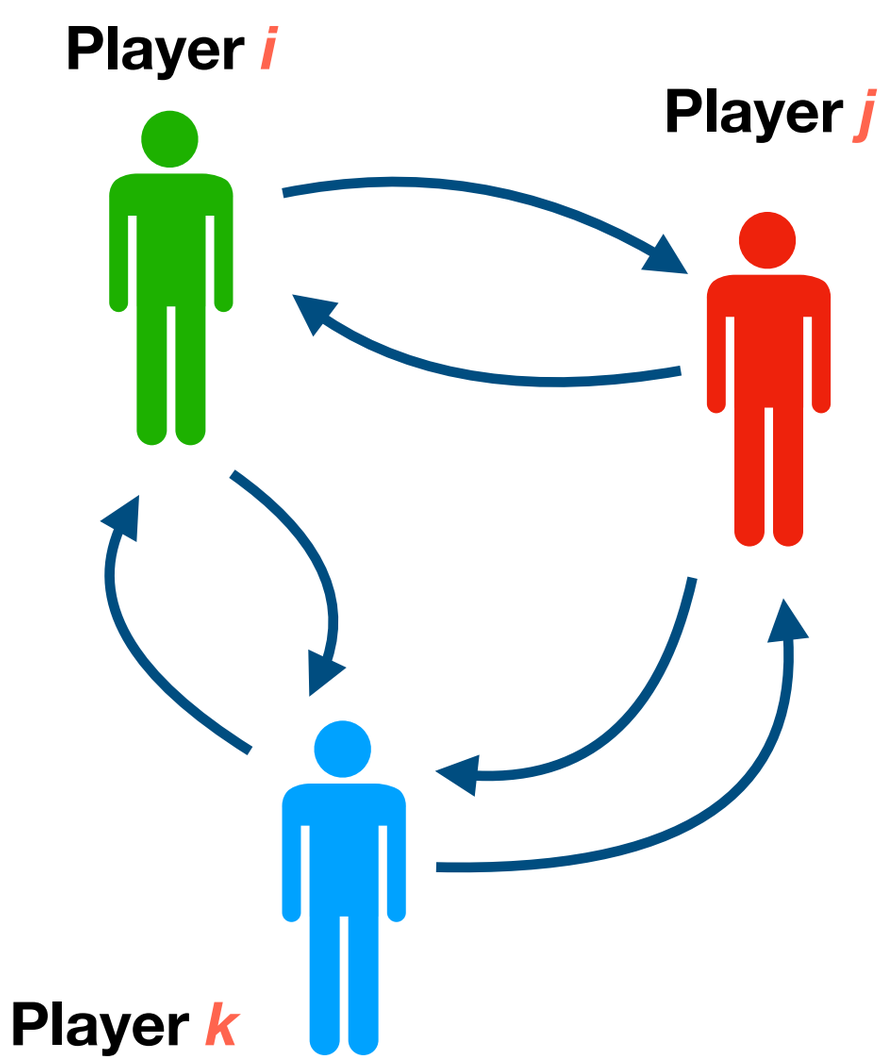


Solution - Rethink of the Graph definition (2)

Finished?

$$\frac{\mathrm{ROI}_i \cap \mathrm{ROI}_j}{\mathrm{ROI}_i} \qquad \frac{\mathrm{ROI}_i \cap \mathrm{ROI}_j}{\mathrm{ROI}_j} \qquad \frac{\mathrm{ROI}_i \cap \mathrm{ROI}_j}{\mathrm{ROI}_i \cup \mathrm{ROI}_j}$$



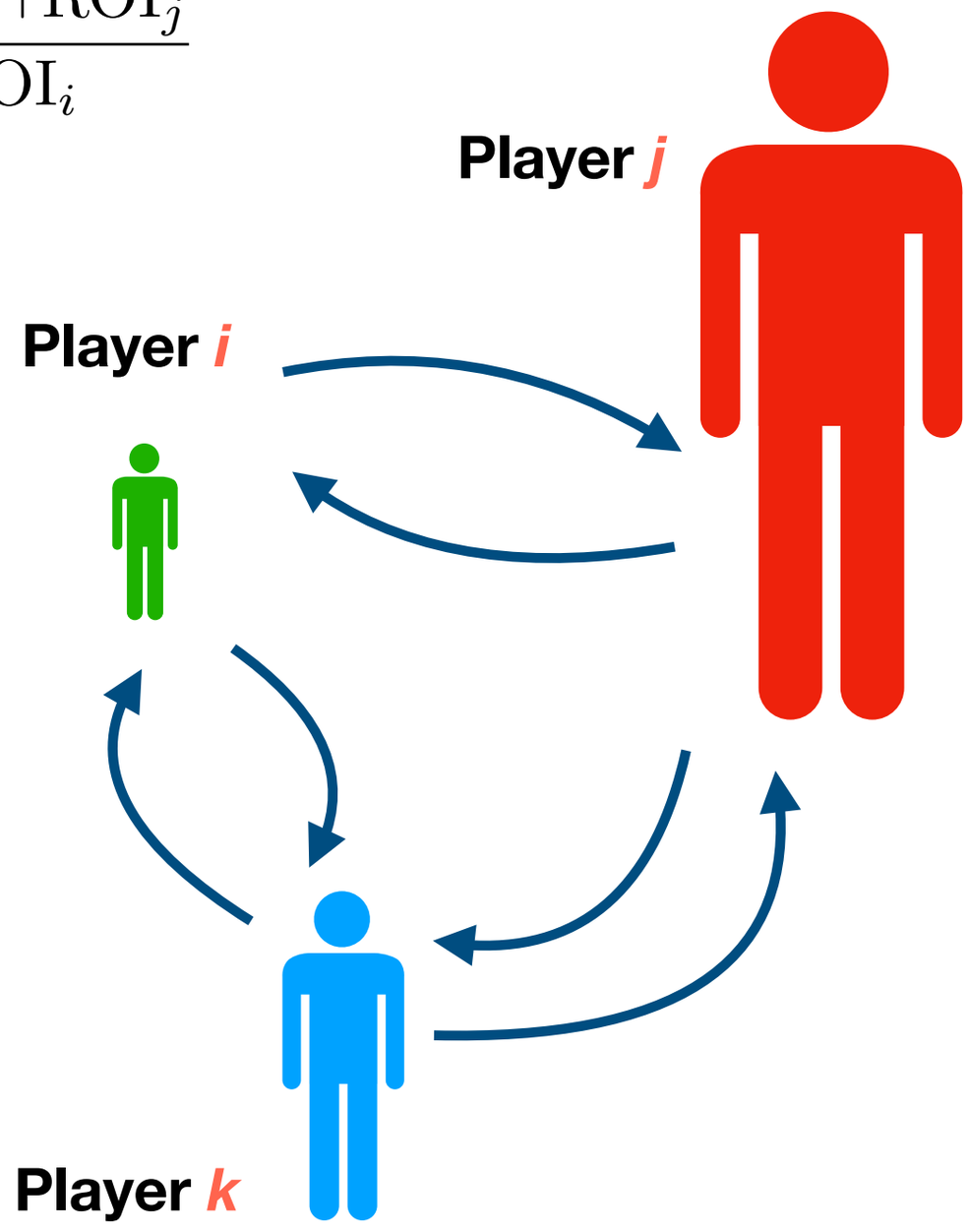


Solution - Rethink of the Graph definition (2)

NO!!!

$$\frac{\mathrm{ROI}_i \cap \mathrm{ROI}_j}{\mathrm{ROI}_i} \longrightarrow \mathrm{TV}_i \times \frac{\mathrm{ROI}_i \cap \mathrm{ROI}_j}{\mathrm{ROI}_i}$$

with ***TV_i*** is the trust value of player ***i***.



Solution - Rethink of the Graph definition (3)

Note that tags are also weighted, assume the tags weight vector is ***v*** (we'll discuss it later), then we use **Weighted Pearson Correlation**:

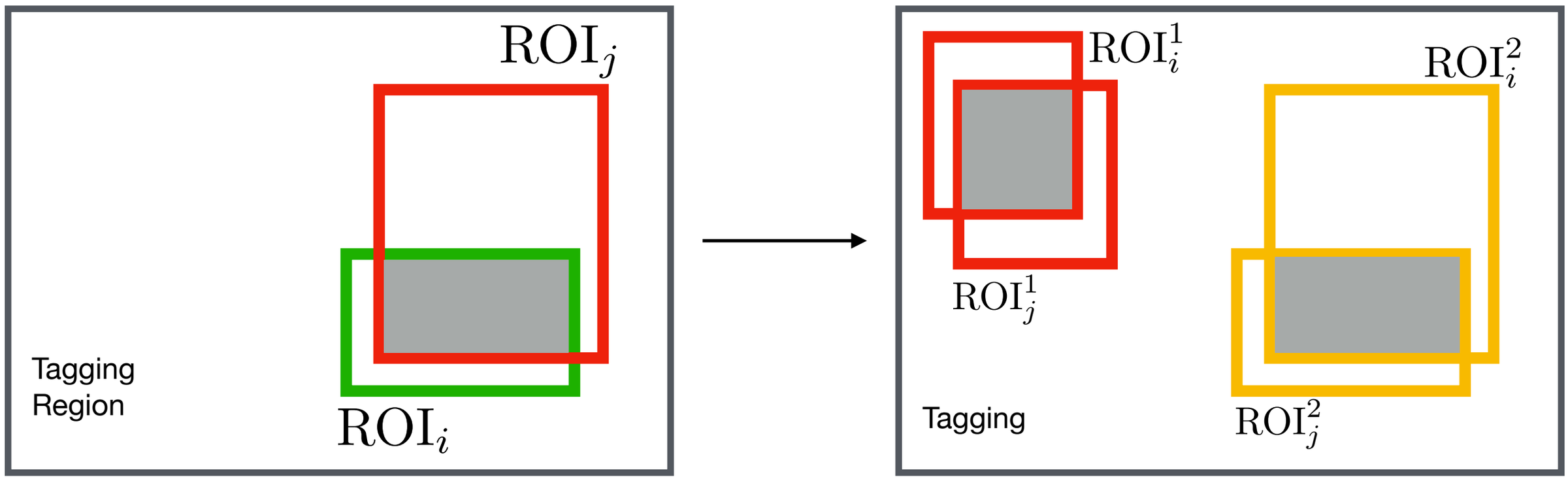
$$\frac{Cov(\text{tags}_i, \text{tags}_j)}{\text{var}(\text{tags}_i)\text{var}(\text{tags}_j)} \longrightarrow \frac{Cov(\text{tags}_i, \text{tags}_j; v)}{Cov(\text{tags}_i, \text{tags}_i; v)Cov(\text{tags}_j, \text{tags}_j; v)}$$

Finally, we have:

$$w_{ij} = \sum_{\text{ROI} \in \text{ROIs}} \text{TV}_i \times \frac{\text{ROI}_i \cap \text{ROI}_j}{\text{ROI}_i} \frac{Cov(\text{tags}_i, \text{tags}_j; v)}{Cov(\text{tags}_i, \text{tags}_i; v)Cov(\text{tags}_j, \text{tags}_j; v)}$$

Solution - Rethink of the Graph definition (4)

What is $\sum_{\mathrm{ROI} \in \mathrm{ROIs}}$??



Solution - Rethink of the Graph definition (4)

What is $\sum_{\mathrm{ROI} \in \mathrm{ROIs}}$??

Considering player ***i*** and player ***j*** with ***two ROIs***:

Then:

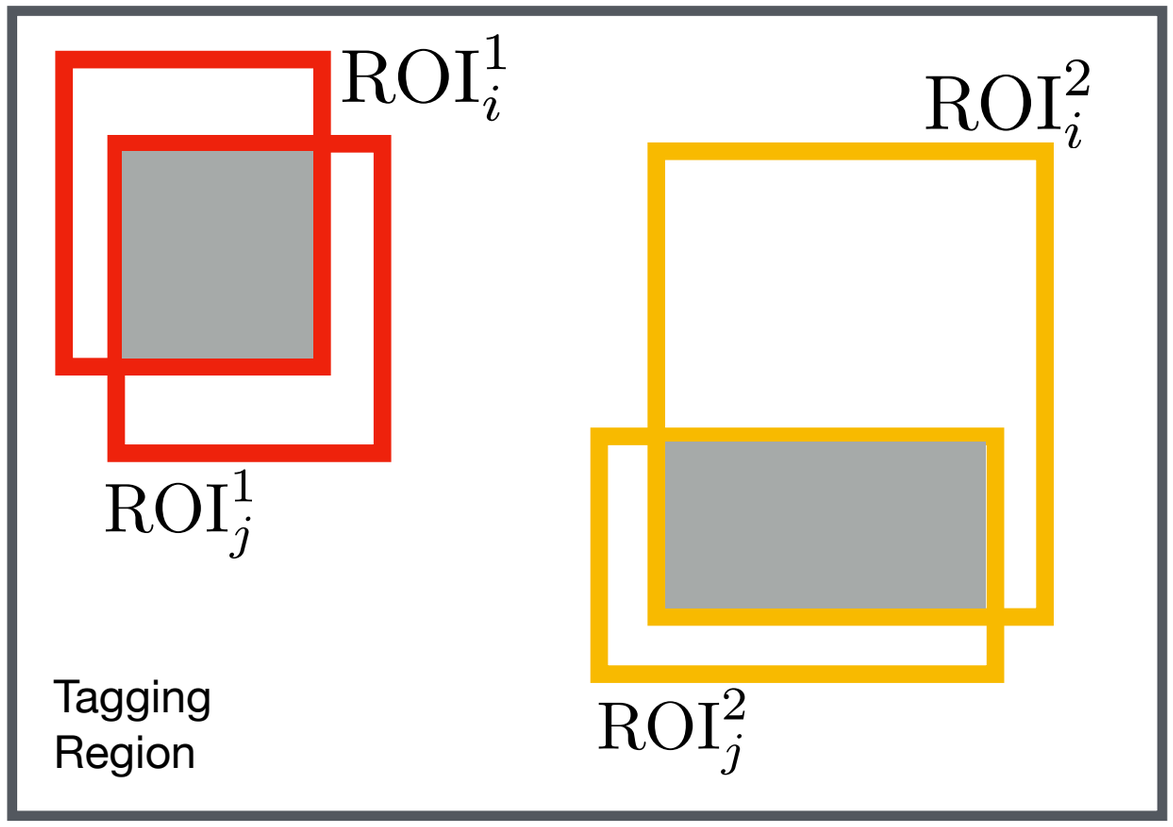
$$w_{ij} = \sum_{\mathrm{ROI} \in \mathrm{ROIs}} \mathrm{TV}_i \times \frac{\mathrm{ROI}_i \cap \mathrm{ROI}_j}{\mathrm{ROI}_i} \ldots$$

$$= \mathrm{TV}_i \times \frac{\mathrm{ROI}_i^1 \cap \mathrm{ROI}_j^1}{\mathrm{ROI}_i^1} \times \ldots$$

$$+ \mathrm{TV}_i \times \frac{\mathrm{ROI}_i^1 \cap \mathrm{ROI}_j^2}{\mathrm{ROI}_i^1} \times \ldots$$

$$+ \mathrm{TV}_i \times \frac{\mathrm{ROI}_i^2 \cap \mathrm{ROI}_j^1}{\mathrm{ROI}_i^2} \times \ldots$$

$$+ \mathrm{TV}_i \times \frac{\mathrm{ROI}_i^2 \cap \mathrm{ROI}_j^1}{\mathrm{ROI}_i^2} \times \ldots$$



Solution - Rethink of the Graph definition (4)

What is $\sum_{\mathrm{ROI} \in \mathrm{ROIs}}$??

Considering player ***i*** and player ***j*** with ***two ROIs***:

Then:

$$w_{ij} = \sum_{\mathrm{ROI} \in \mathrm{ROIs}} \mathrm{TV}_i \times \frac{\mathrm{ROI}_i \cap \mathrm{ROI}_j}{\mathrm{ROI}_i} \ldots$$

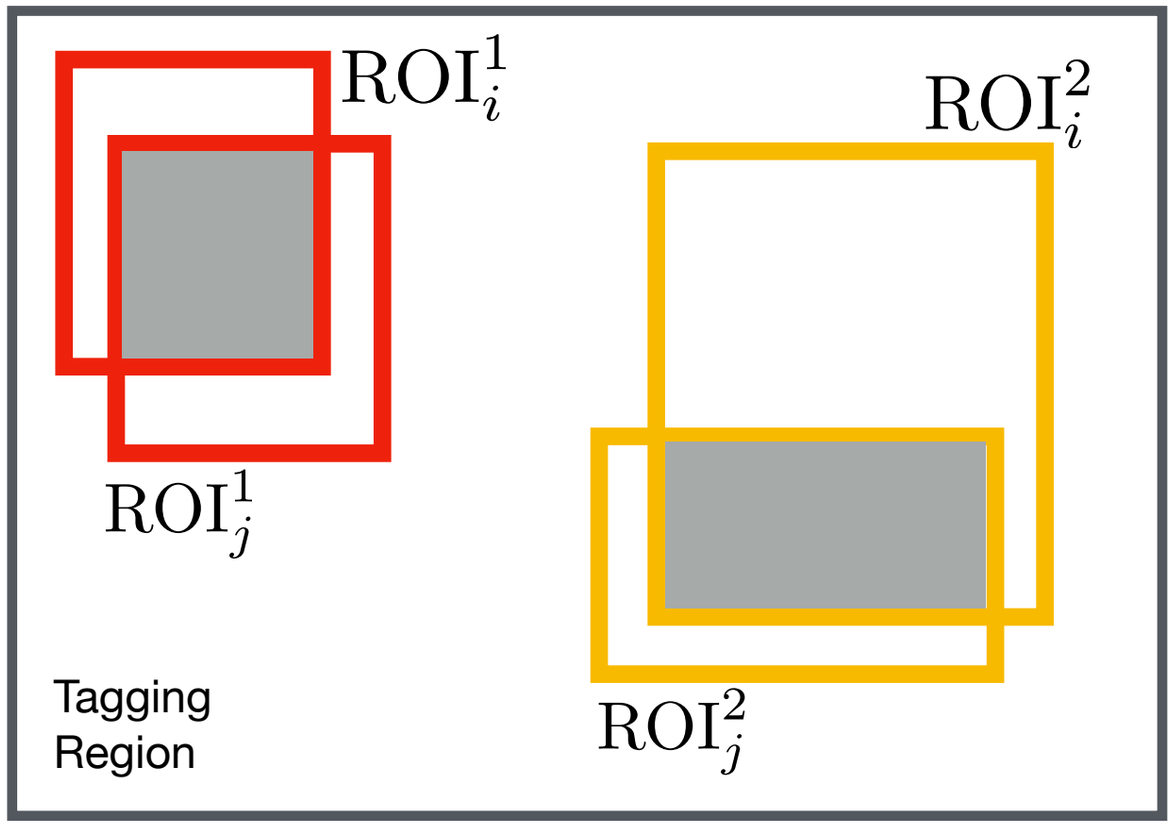
$$= \mathrm{TV}_i \times \frac{\mathrm{ROI}_i^1 \cap \mathrm{ROI}_j^1}{\mathrm{ROI}_i^1} \times \ldots$$

$$+\mathrm{TV}_i \times \frac{\mathrm{ROI}_i^1 \cap \mathrm{ROI}_j^2}{\mathrm{ROI}_i^1} \times \ldots$$

$$+\mathrm{TV}_i \times \frac{\mathrm{ROI}_i^2 \cap \mathrm{ROI}_j^1}{\mathrm{ROI}_i^2} \times \ldots$$

$$+\mathrm{TV}_i \times \frac{\mathrm{ROI}_i^2 \cap \mathrm{ROI}_j^1}{\mathrm{ROI}_i^2} \times \ldots$$

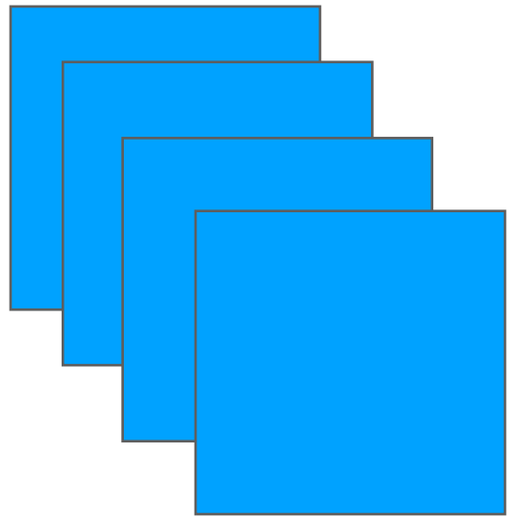
→ **Zero!!**



Solution - Cold Start (1)

Initial Trusted Group

Tagging



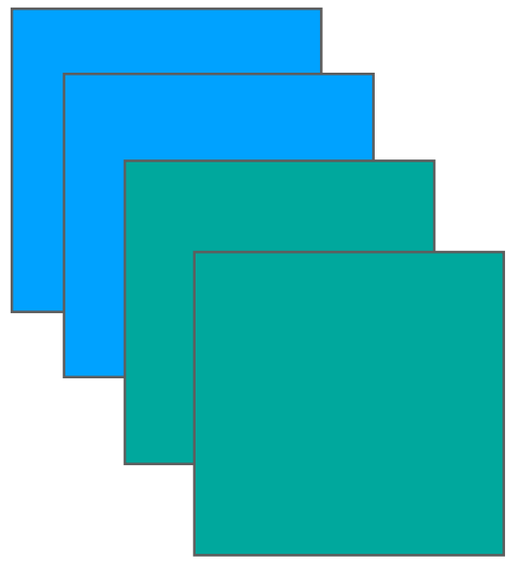
Initial Trusted Results

Untagged

Tagged

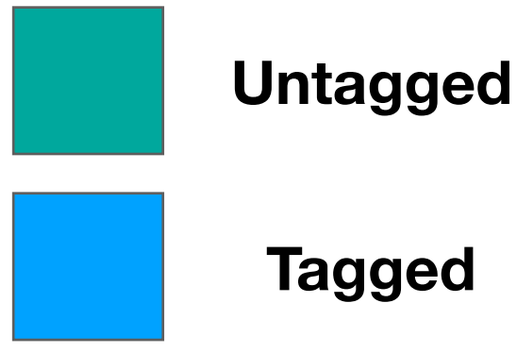
New Player

Tagging

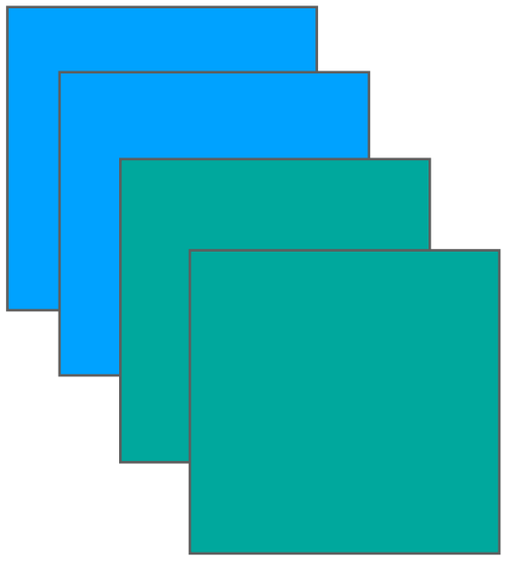


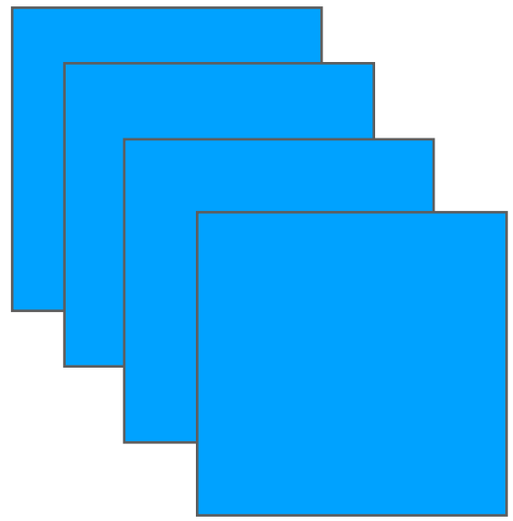
Task

Solution - Cold Start (2)



Tagging





New Player

Task

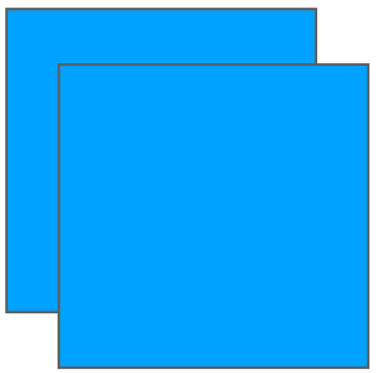
If trustable

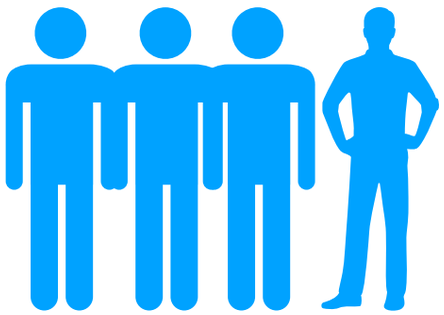
aggregate

untagged images

Initial Trusted Results

+





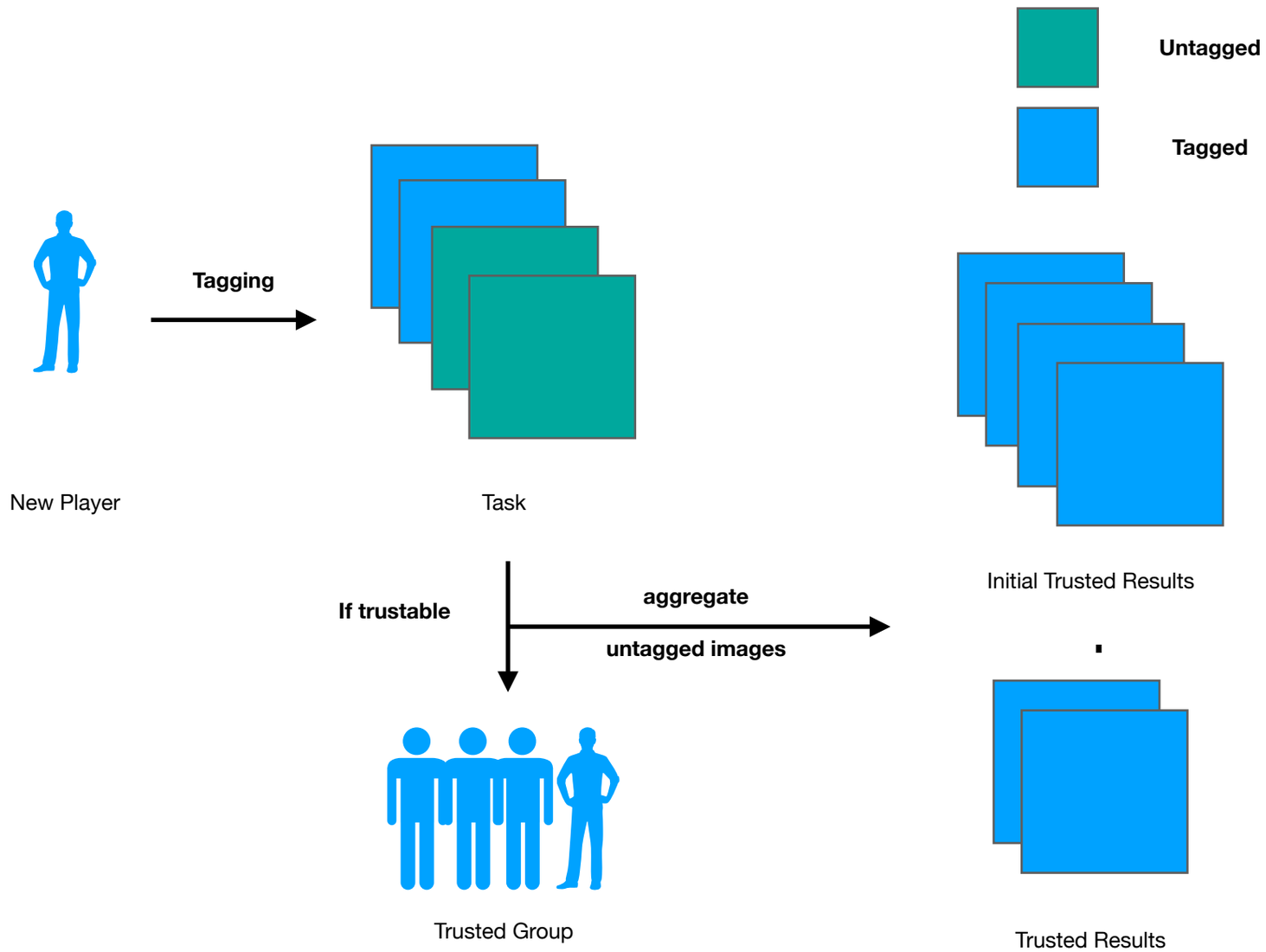
Trusted Group

Trusted Results

Solution - Cold Start (2)

What is a suitable size for Initial Trusted Group?

Only ***Two*** persons!!!



Solution - Cold Start (3)

What about the initial ***TV_i*** ?

$$w_{ij} = \sum_{\mathrm{ROI} \in \mathrm{ROIs}} \mathrm{TV}_i \times \frac{\mathrm{ROI}_i \cap \mathrm{ROI}_j}{\mathrm{ROI}_i} \frac{Cov(\mathrm{tags}_i, \mathrm{tags}_j; v)}{Cov(\mathrm{tags}_i, \mathrm{tags}_i; v) Cov(\mathrm{tags}_j, \mathrm{tags}_j; v)}$$

Solution - Cold Start (3)

What about the initial ***TV_i*** ?

$$w_{ij} = \sum_{\text{ROI} \in \text{ROIs}} \text{TV}_i \times \frac{\text{ROI}_i \cap \text{ROI}_j}{\text{ROI}_i} \frac{Cov(\text{tags}_i, \text{tags}_j; v)}{Cov(\text{tags}_i, \text{tags}_i; v) Cov(\text{tags}_j, \text{tags}_j; v)}$$

Note that ***TV_i*** is in between of 0 and 1, thus:

$$TV_i^{\text{init}} = \frac{1}{|\text{players}^{\text{init}}|}$$

Solution - Handling New Tags

Initial Trusted Group label the initial images with tags, one can calculate the **frequency** of all these initial tags as the **probability (weight) of these tags**:

$$v_i = p(\mathrm{tag}_i) = \frac{|\mathrm{tag}_i|}{\sum_{j \in \mathrm{tags}} \mathrm{tag}_j}$$

a) When a player carries predefined tags: ***Trivial***;

b) When a player carries new tags: ***Directly drop*** —— unreliable, why?

c) When a player carries predefined tags and also new tags:

+ calculate the trust value without new tags;
+ **if** the player is reliable:
    + merge and update all p(tag_i)
+ **else:**
    + drop

Refining

Player Rating Model (refine)

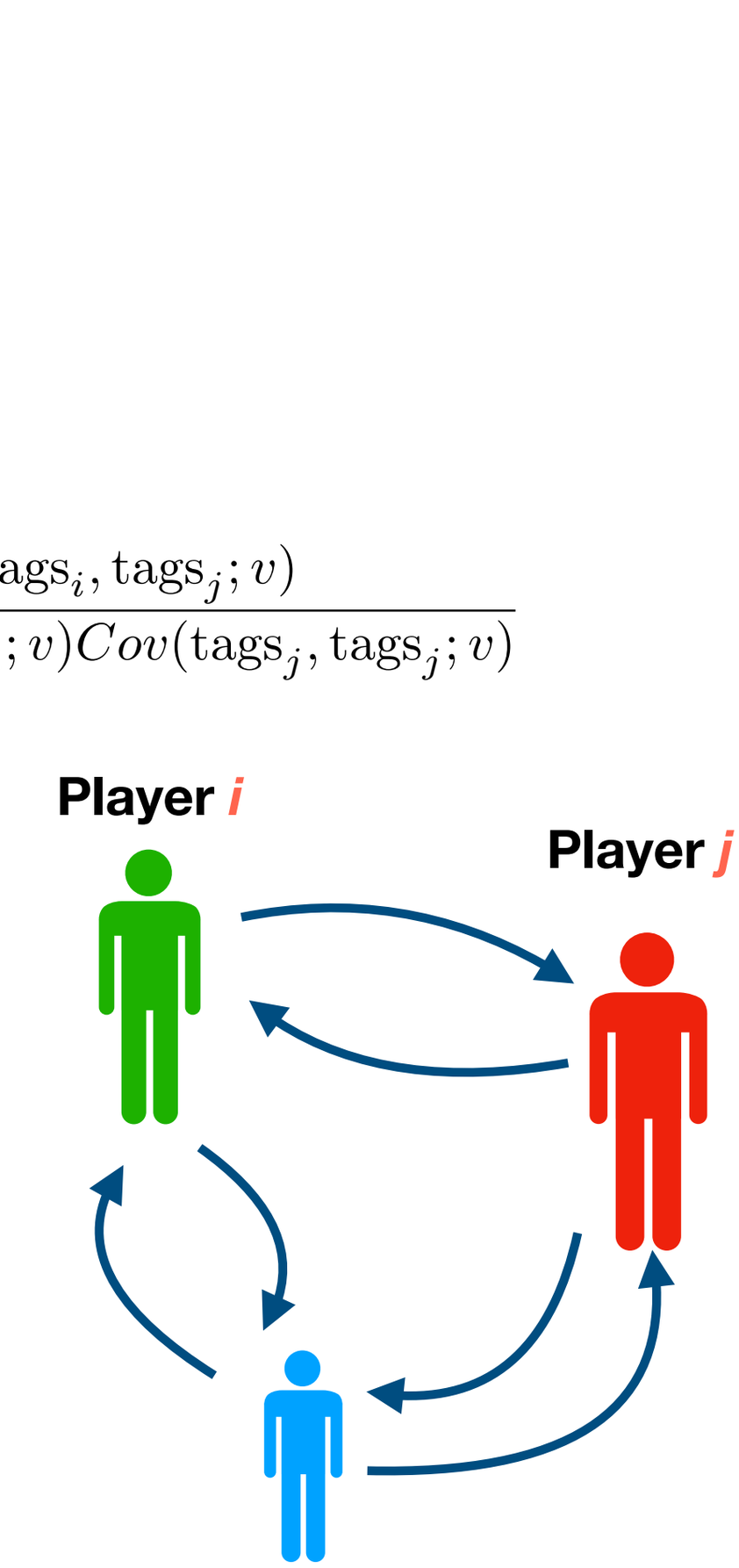
1. **Weight** of Rating Graph:

$$w_{ij} = \sum_{\mathrm{ROI} \in \mathrm{ROIs}} \mathrm{TV}_i \times \frac{\mathrm{ROI}_i \cap \mathrm{ROI}_j}{\mathrm{ROI}_i} \frac{Cov(\mathrm{tags}_i, \mathrm{tags}_j; v)}{Cov(\mathrm{tags}_i, \mathrm{tags}_i; v) Cov(\mathrm{tags}_j, \mathrm{tags}_j; v)}$$

2. **Normalized** Adjacency Matrix

$$A = (\frac{w_{ij}}{\sum_j w_{ij}})$$

3. A is **irreducible, real, non-negative, column-stochastic, and diagonal element being positive***, then eigenvalue of A is the player *i*'s **trust value** ***TV_i***.



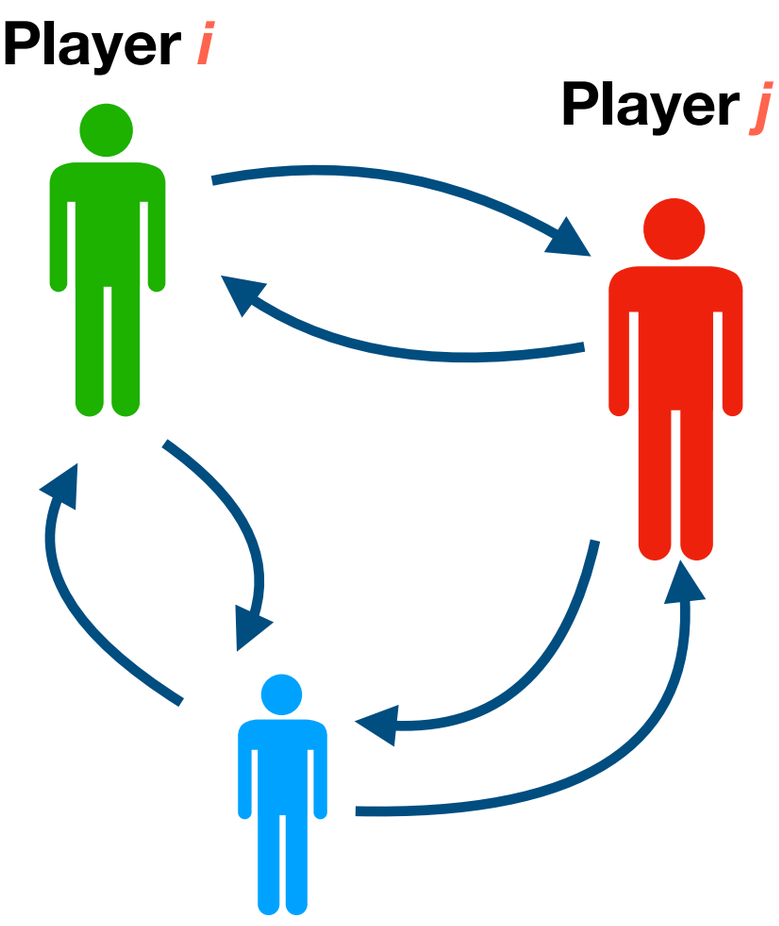
*** Proof in the report, but trivial**

Player Rating Criteria (recap)

Let *TV_new* is the trust value of new player, then a player is reliable if:

$$TV_{new} \geq \frac{1}{|\text{players}|} \sum_{i \in \text{players}} TV_i = \frac{1}{|\text{players}|} tr(A)$$

Otherwise drop and mark it as unreliable.



Disaster Evaluation Model (restatement)

Every image has its tagging history. For a certain image img at time t, it has disaster tags from different players, it has its own disaster level:

$$DL_{img} = \sum_{i \in \text{tags}} v_i \times \#(i)$$

One can define ***Disaster Level*** for a global monitoring area as follows:

$$DL = \sum_{img \in \text{area}} DL_{img}$$

System Evaluation

Criteria - Tech Evaluation

Note that malicious player detection is a classification problem.

One can generate **random data** and test the Rating Model through ***accuracy*** and ***recall***, even ***ROC*** curve.

For generate ROI:

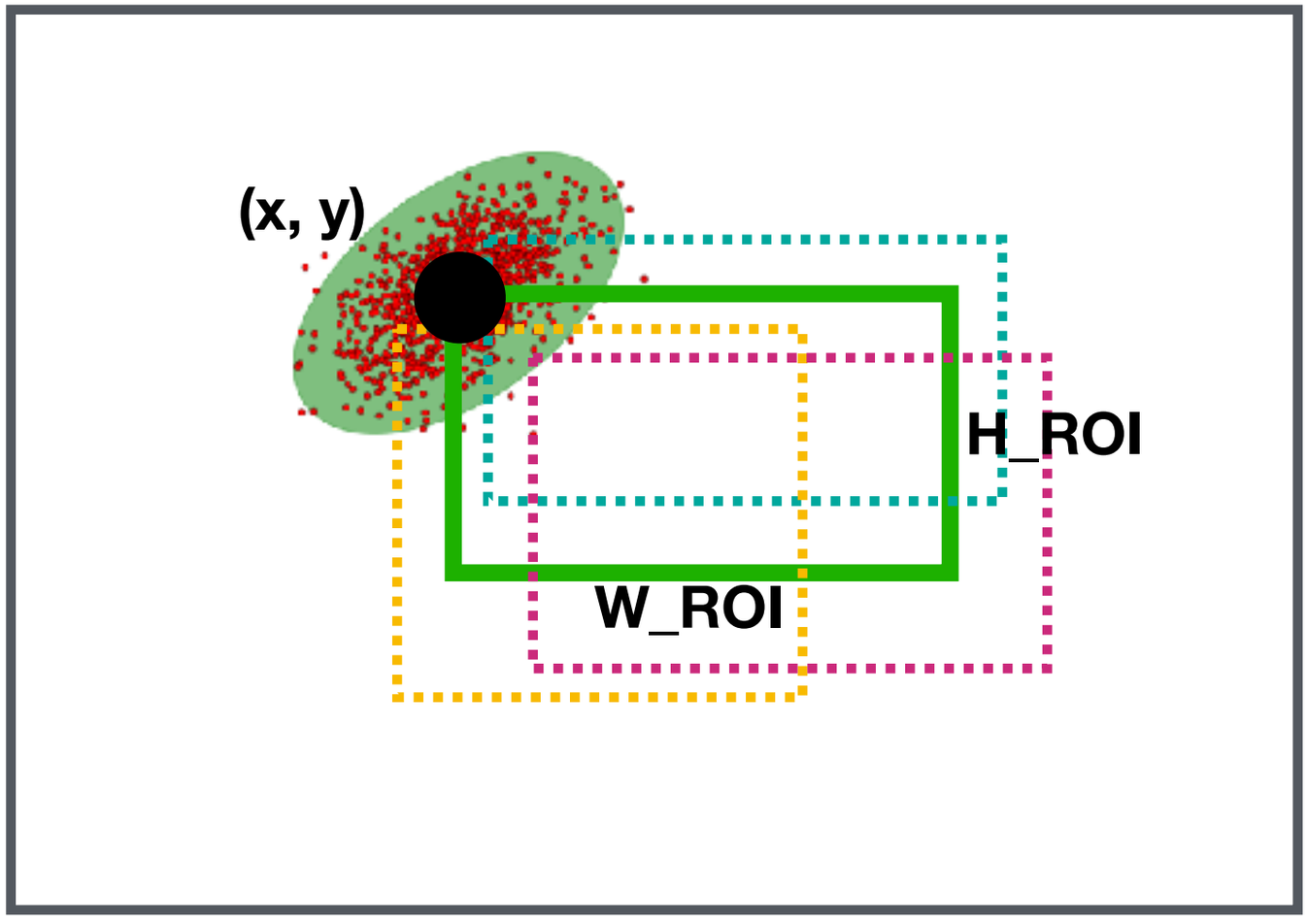
$$(x + N(0, \delta), y + N(0, \delta))$$

$$H_{\mathrm{ROI}} + N(0, \delta)$$

$$W_{\mathrm{ROI}} + N(0, \delta)$$

For generate tags:

- randomly pick random number of tags



Criteria - Success Criteria

1. The number of players
    A. more users: more tags (higher accuracy of our level of disaster)
    B. more users: trustworthy (higher trusted value which can filter malicious groups)

2. The time interval of satellite photos of the same region
    A. smaller time interval. It’s important to obtain the latest info.

Criteria - Social & Ethics

1. Data Security,Privacy at Risk
    A. Leaked to ordinary users
    B. Leaked to terrorists

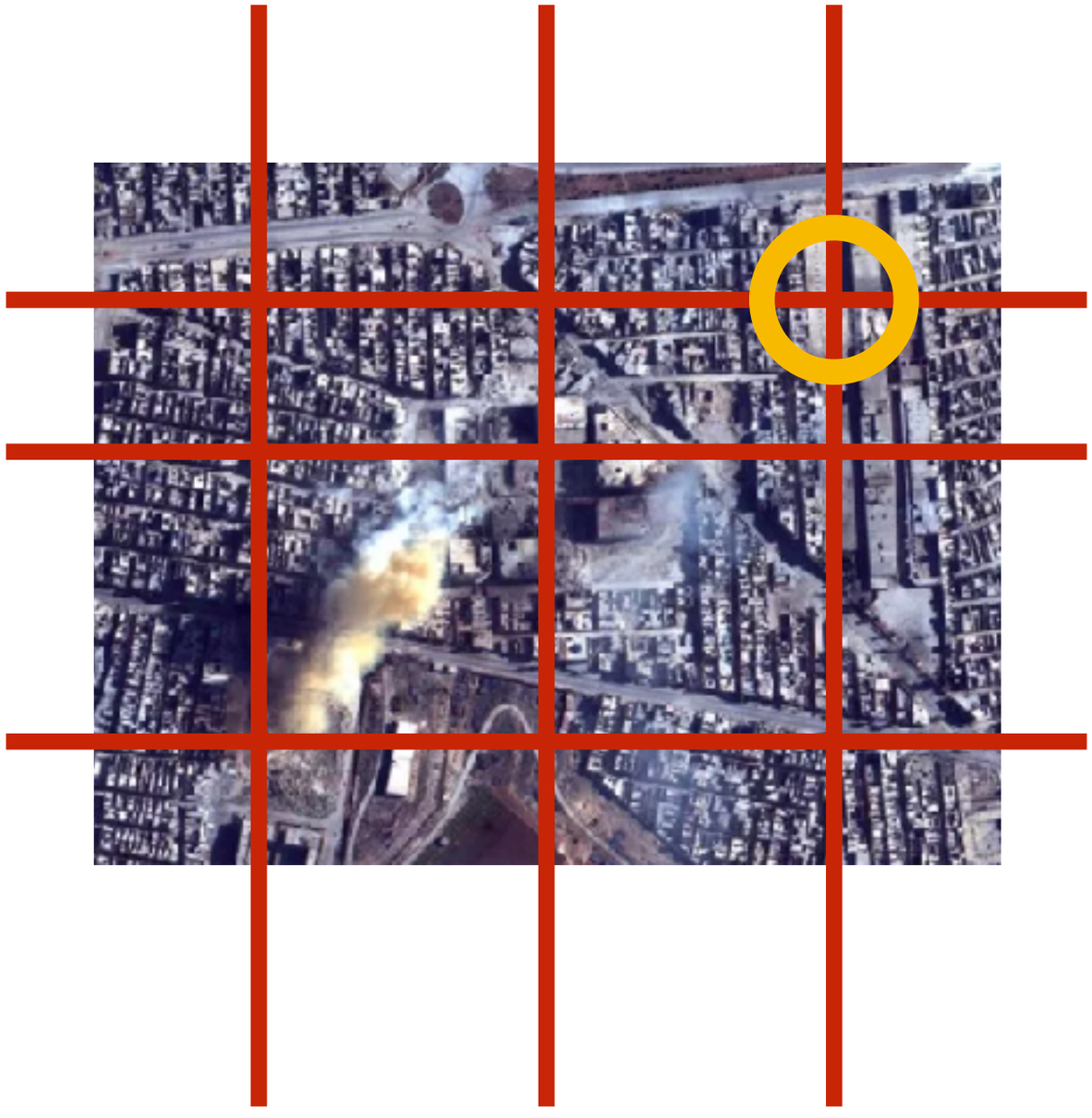
2. IPR (Intellectual Property Right)

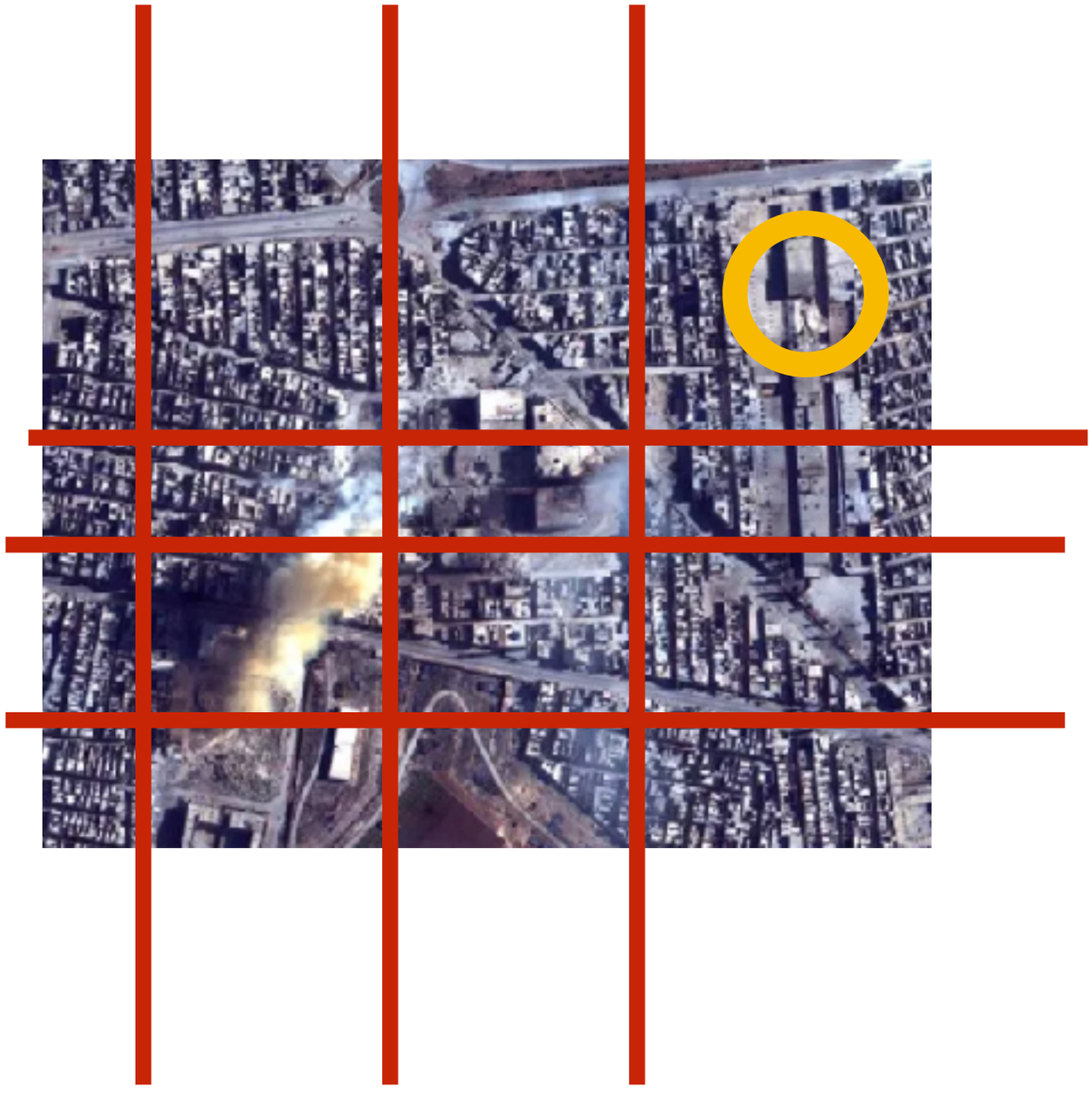
3. Uses Considered Misuses
    A. used by terrorists or some Hostile countries.
    B. used for profitability.

Limitations

Limitations - Information Loss

- Potential problem : Information lost



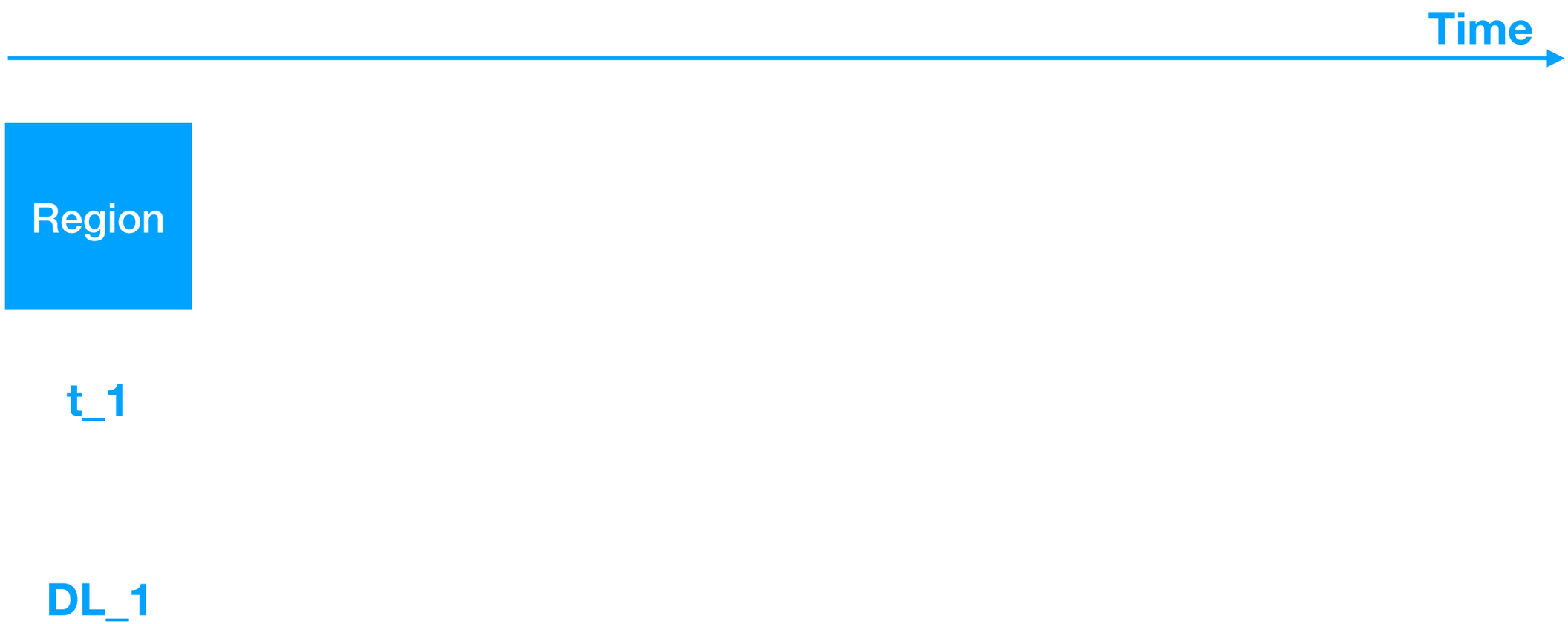


Limitations - Outdate of Evaluation

Each evaluation get invalid if region image outdate.

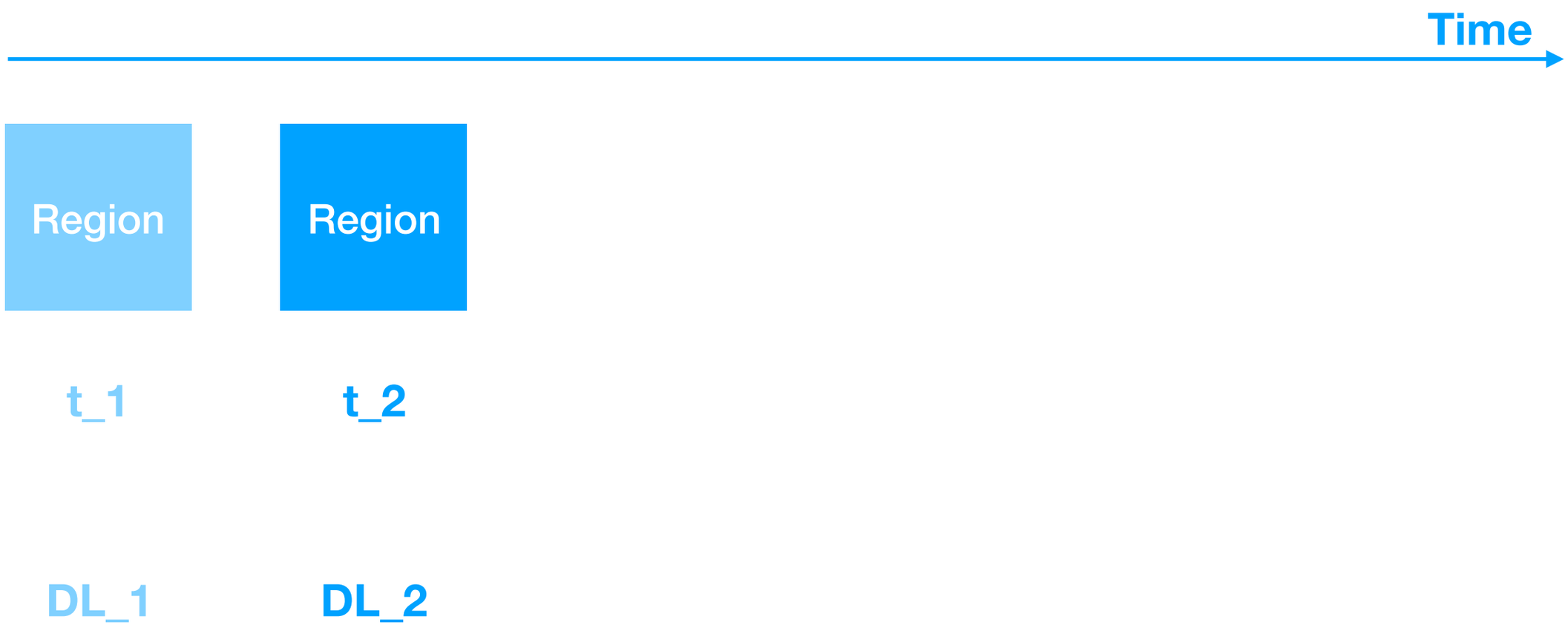
Limitations - Outdate of Evaluation

Each evaluation get invalid if region image outdate.



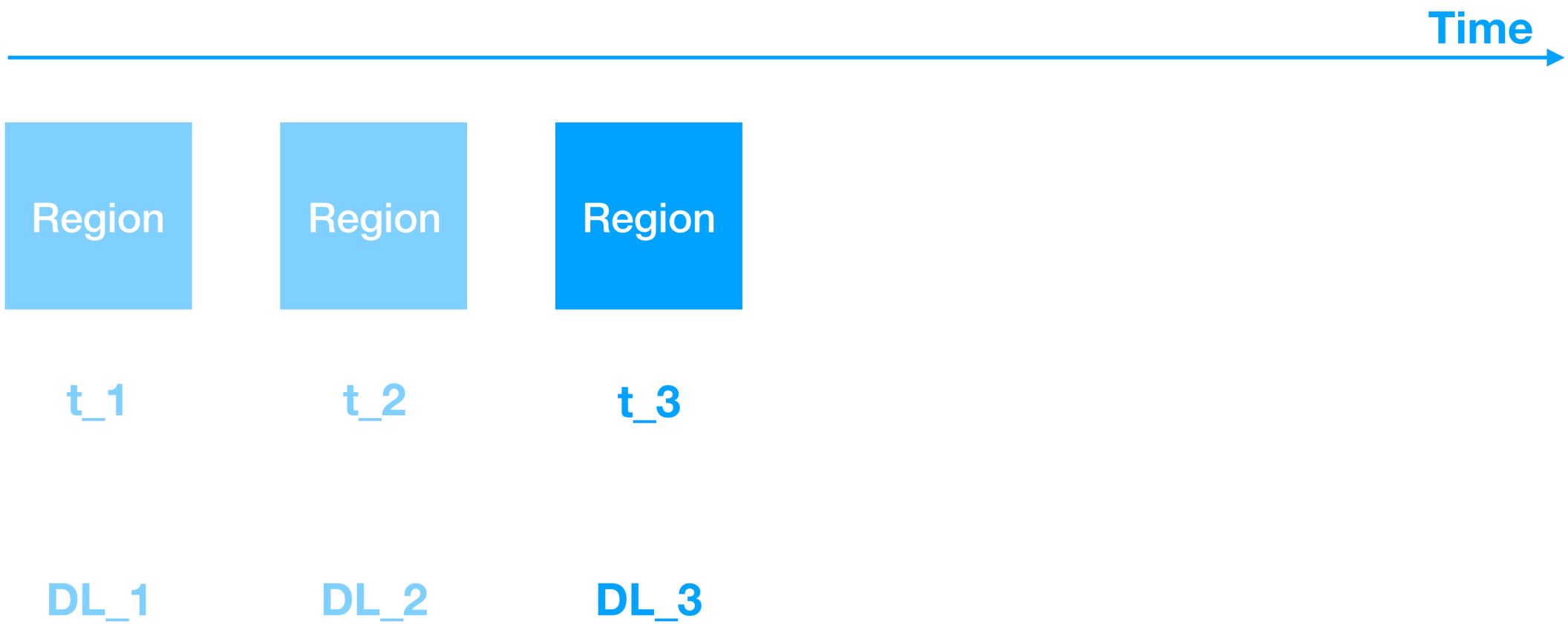
Limitations - Outdate of Evaluation

Each evaluation get invalid if region image outdate.



Limitations - Outdate of Evaluation

Each evaluation get invalid if region image outdate.



Limitations - Outdate of Evaluation

Each evaluation get invalid if region image outdate.

Limitations - Outdate of Evaluation

Each evaluation get invalid if region image outdate.

Limitations - Gameplay and Playability

- Users may meet the situation that there is no available ROI in several continuous rounds.

Limitations - Gameplay and Playability

- Users may meet the situation that there is no available ROI in several continuous rounds.

Lake, forest, mountain, desert…

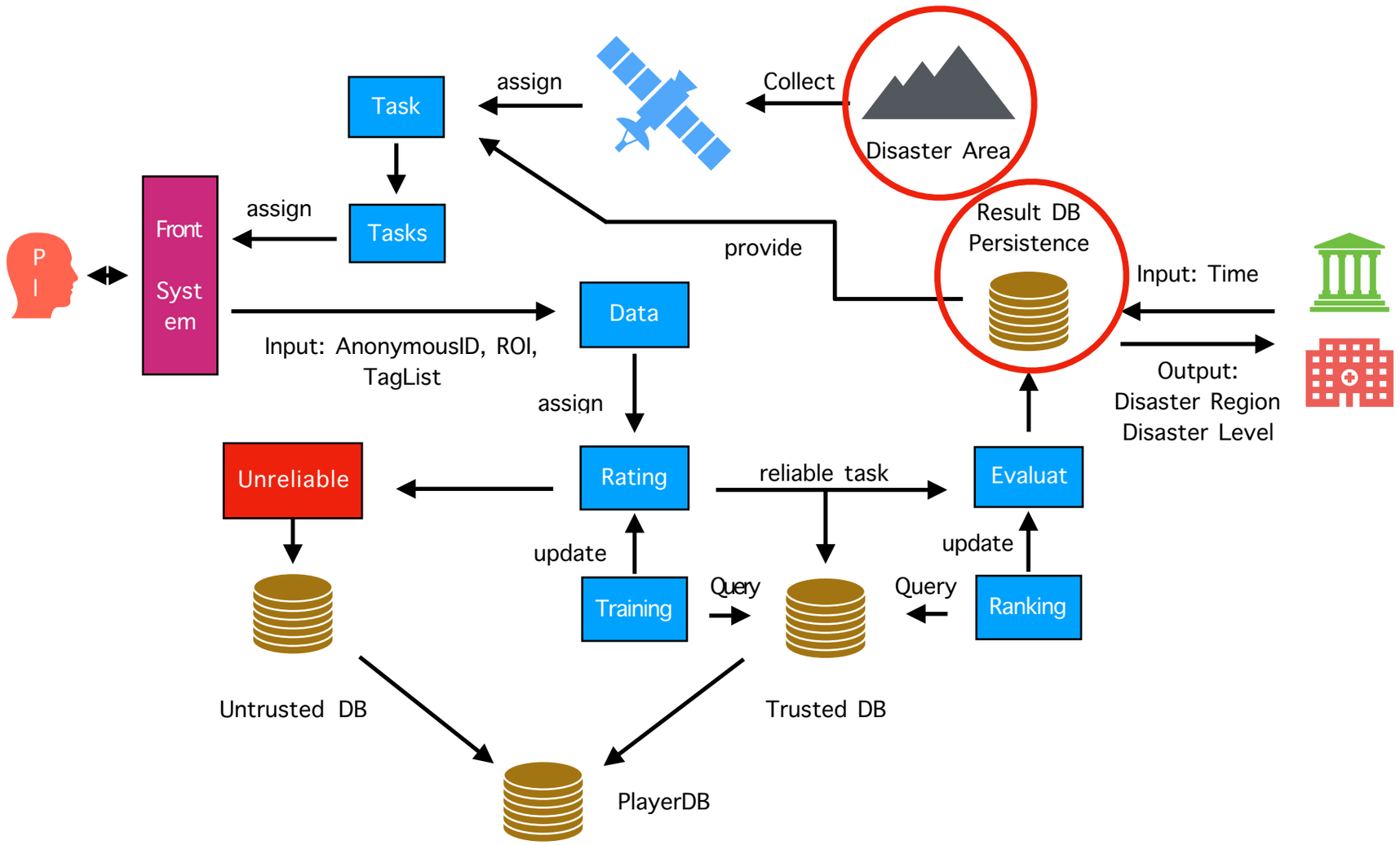
Limitations - Gameplay and Playability

- Users may meet the situation that there is no available ROI in several continuous rounds.

city area, but no disaster

Limitations - Gameplay and Playability

- Users may meet the situation that there is no available ROI in several continuous rounds.
- Solution



Future Works

Future Works

1. **General System**: Theoretically, any of HC problems with tagging ROI tasks fit our system.

2. **Replaceable Rating Model**: In general, malicious player filtering is a classification problem. Once we have stored enough data, then the entire rating model can be replaced by any other machine learning algorithm.

3. **Collaborative Computer Factors**: what if we apply computer vision method for disaster detection that against human inputs?

4. **Gameplay and Playability**: Using computer vision to shield interference maps.

FAQ?

References

[1] L. Page, S. Brin, R. Motwani, and T. Winograd, “The page rank citation ranking: Bringing order to the web.,” tech. rep., Stanford InfoLab, 1999.

[2] P. Bonacich and P. Lloyd, “Eigenvector-like measures of centrality for asymmetric relations,” *Social networks*, vol. 23, no. 3, pp. 191–201, 2001.

[3] R. RealandJ. M. Vargas,“The probabilistic basis of jaccard’s index of similarity,”*System atic biology*, vol. 45, no. 3, pp. 380–385, 1996.

[4] L. Von Ahn and L. Dabbish, “Labeling images with a computer game,” in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 319–326, ACM, 2004.

[5] C. Wieser, F. Bry, A. Bérard, and R. Lagrange, “Artigo: building an artwork search engine with games and higher-order latent semantic analysis,” in *First AAAI Conference on Human Computation and Crowdsourcing*, 2013.

[6] J. A. Hanley and B. J. McNeil, “The mean in gand use of the area under a receiver operating characteristic (roc) curve.,” *Radiology*, vol. 143, no. 1, pp. 29–36, 1982.

[7] X. Bi, Y. Li, and S. Zhai, “FFitts law: modeling finger touch with Fitts’ law,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1363–1372, ACM, 2013.