A Novel GWAP System for

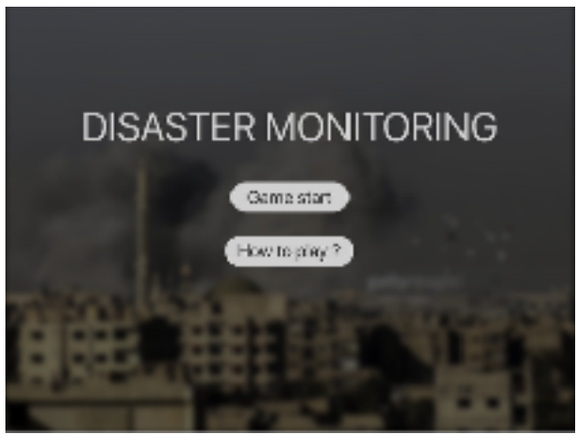
Disaster Monitoring

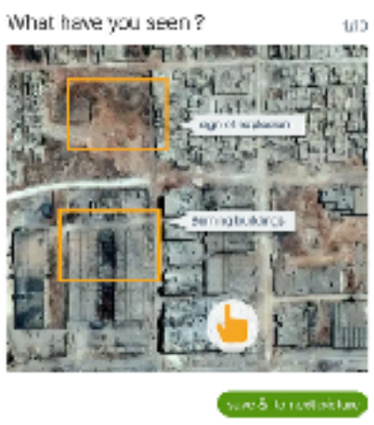
Team Hotpot:
- Changkun Ou
- Yifei Zhan
- Zhe Li

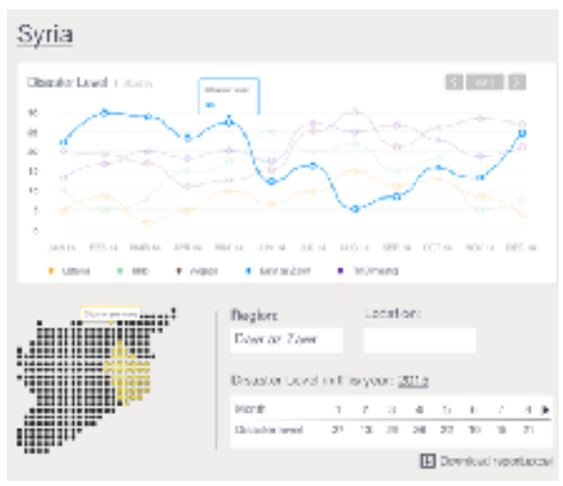
Functionalities

Functionalities: Player & Stakeholder

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



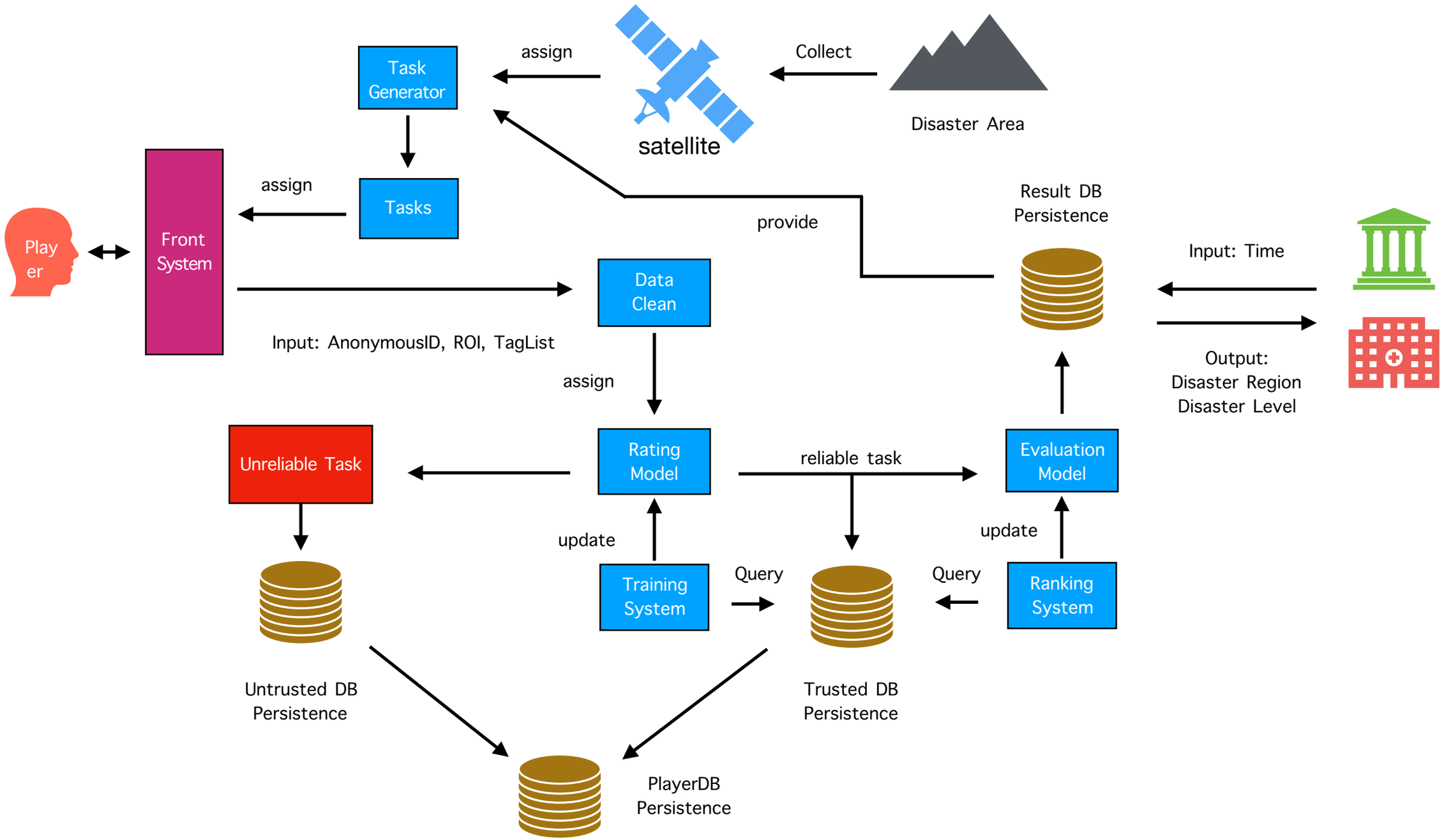




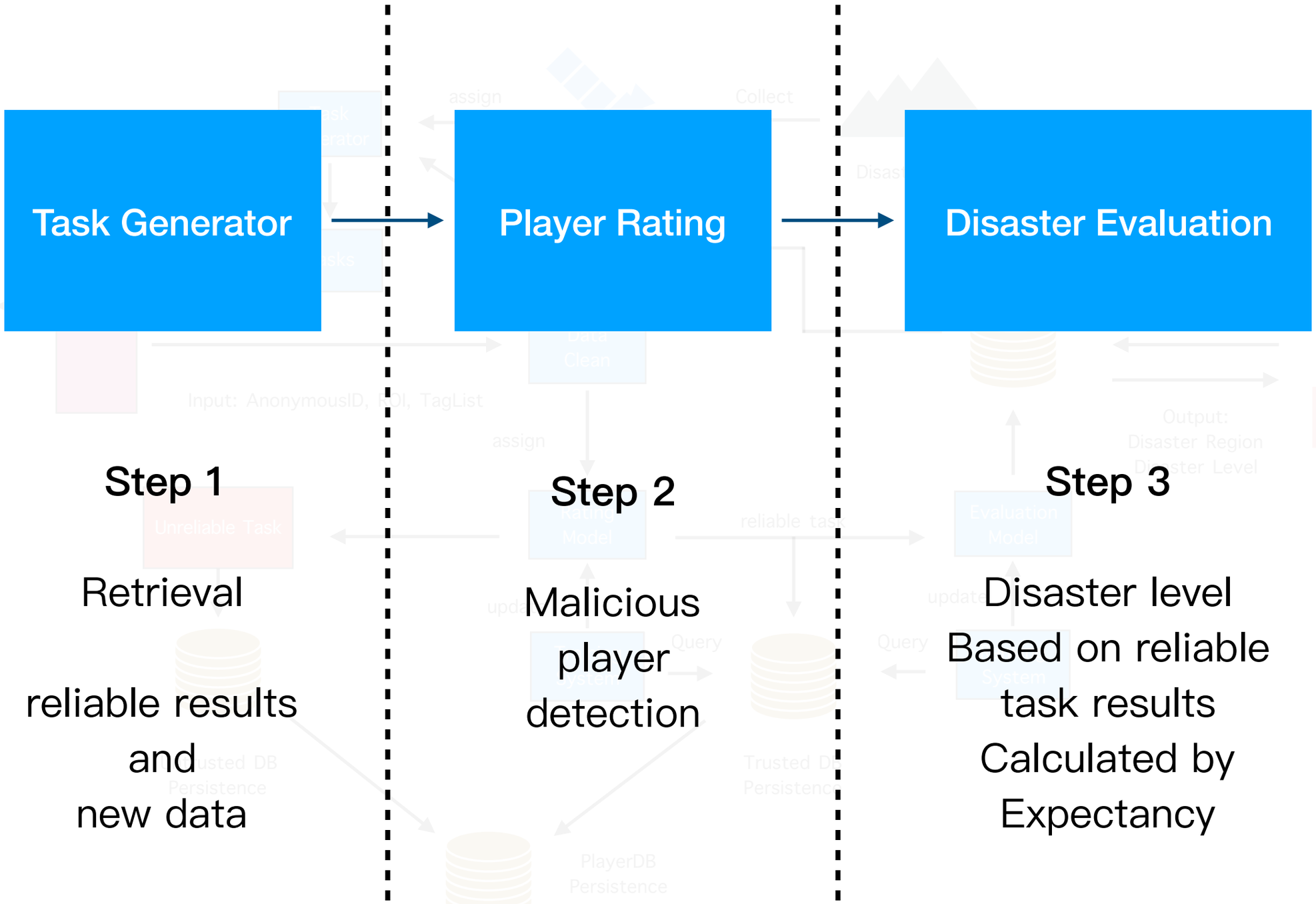
Technology stack suggestion:
Front-end: Polymer, Back-end: Node.js / Python, Database: MongoDB

Design

Design: Architecture



Design: Data Flow



Step 1

Retrieval

reliable results
and
new data

Step 2

Malicious
player
detection

Step 3

Disaster level
Based on reliable
task results
Calculated by
Expectancy

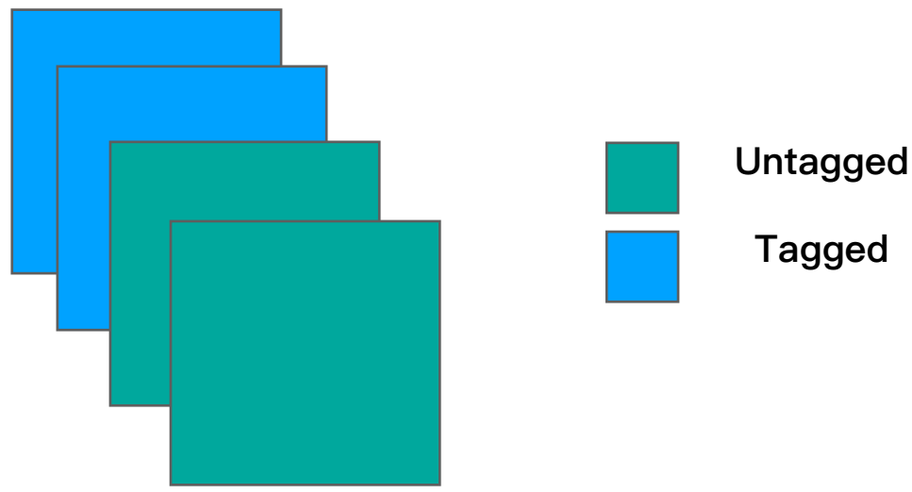
Design: User Task

Task Generator

Step 1

Retrieval

reliable results
and
new data

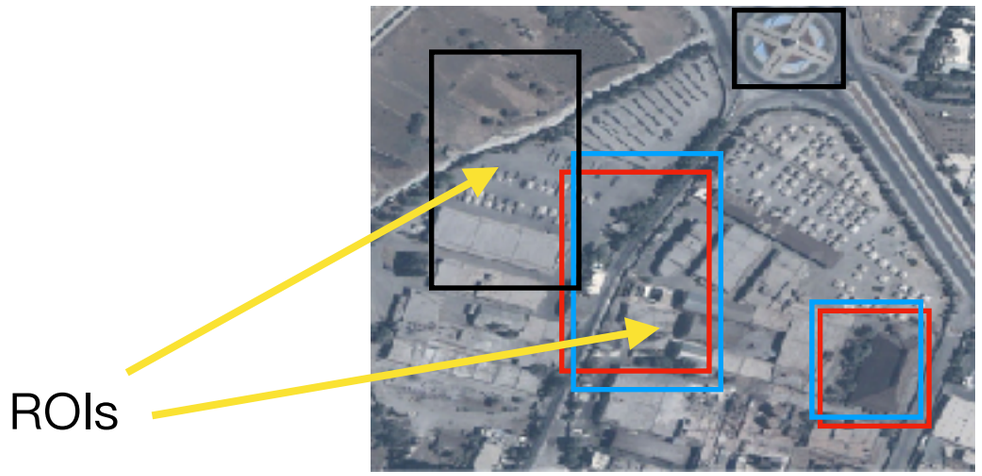


Task

Design: Definition

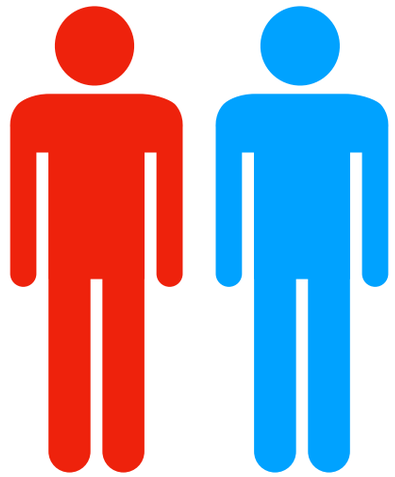
Definition: The **Region of Interests (ROI)** is an indicator that represents the one of the selected two dimensional regions from player. The ***i***-th ROI from player ***p*** in image ***k*** at image create time ***t*** is denoted by $ROI_{p,i,k,t}$.

Considering image ***k*** implies time ***t***, for convenience, we usually simplified $ROI_{p,i,k,t}$ to $ROI_{p,i,k}$.

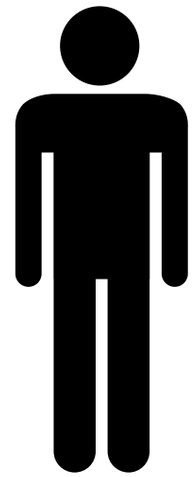


A image from user tasks

Entire monitoring region



Reliable Player



Malicious

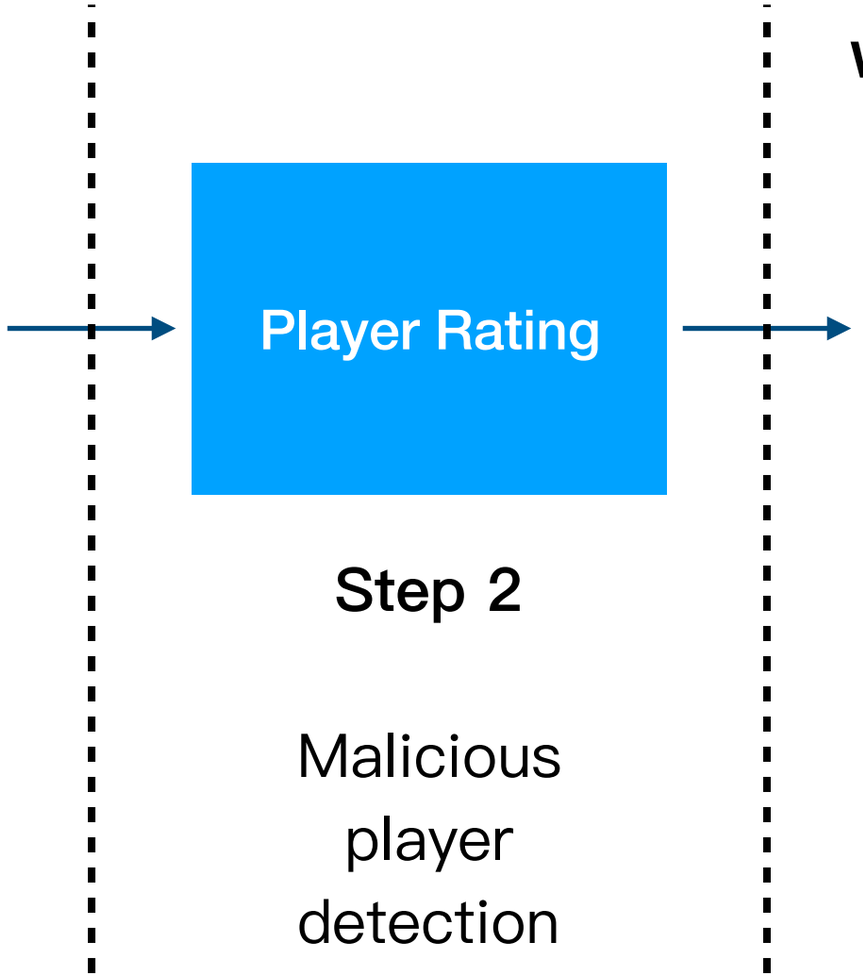
* The unrefined report has notation omits and mistakes

Design: Model Refine (1)

Player Rating Graph (Image Dependent)

For **a certain image** *k*, player *p* rate player *q* by the following:

$$w_{p,q,k} = \sum_{j=1}^{n}\sum_{i=1}^{m}\left(\frac{|ROI_{p,i,k} \cap ROI_{q,j,k}|}{|ROI_{p,i,k}|}\left(2 + \frac{Cov(T_{p,i,k}, T_{q,j,k}; v)}{Cov(T_{p,i,k}, T_{p,i,k}; v)Cov(T_{q,j,k}, T_{q,j,k}; v)}\right)\right)$$



With player *p* selected *m* ROIs, player *q* selected *n* ROIs, where:

$ROI_{p,i,k}$ is the ***i*-th selected** ROI **region** from player ***p***;

$|ROI_{p,i,k}|$ is the **surface** of $ROI_{p,i,k}$

v is the **weight vector** of all tags

$T_{p,i,k}$ is the ***i*-th selected** ROI's **tags vector** from player ***p***

$Cov(x, y; v)$ is the **weighted covariance** of x and y via v;

$A = (a_{p,q,k}) = \left(\frac{w_{p,q,k}}{\sum_q w_{p,q,k}}\right)$ is irreducible, real, non–negative, column–stochastic, and diagonal element being positive.

* The unrefined report has notation omits and mistakes

Design: Model Refine (2)

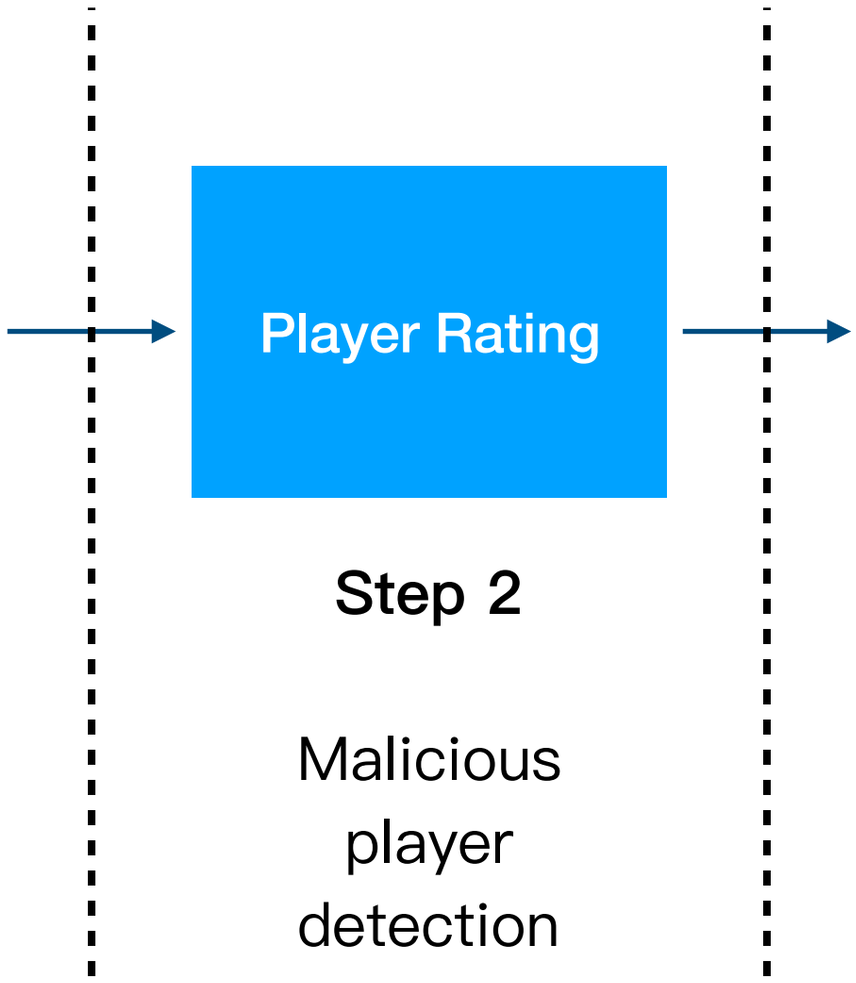
FAQ on Player Rating Graph (Image Dependent)

What is the "2"?

– It is a translation from [–1, 1] to [1, 3]

For **a certain image** *k*, player *p* —> (rate) player *q* by the following:

$$w_{p,q,k} = \sum_{j=1}^{n}\sum_{i=1}^{m}\left(\frac{|ROI_{p,i,k} \cap ROI_{q,j,k}|}{|ROI_{p,i,k}|}\left(2 + \frac{Cov(T_{p,i,k}, T_{q,j,k}; v)}{Cov(T_{p,i,k}, T_{p,i,k}; v)Cov(T_{q,j,k}, T_{q,j,k}; v)}\right)\right)$$



Why we need Pearson Correlation?

Because user inputs have two parts:

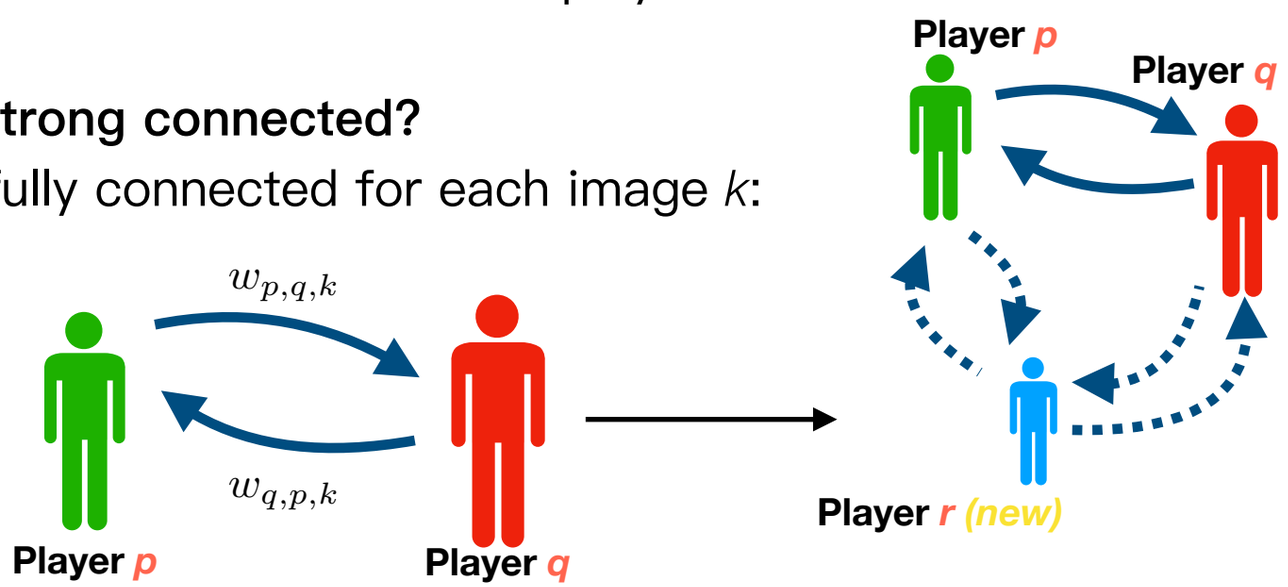
– ROI —> handle by ROI intersection over rating person ROI

– Tags —> handle by Pearson Correlation

It is a metric measures the tags relation as well as the ROI Intersection division between two players

Is graph strong connected?

Yes, it is fully connected for each image *k*:



Design: Model Refine (3)

FAQ on Player Rating Graph (Image Independent) ?

Can we give a global image independent Player Rating Graph?

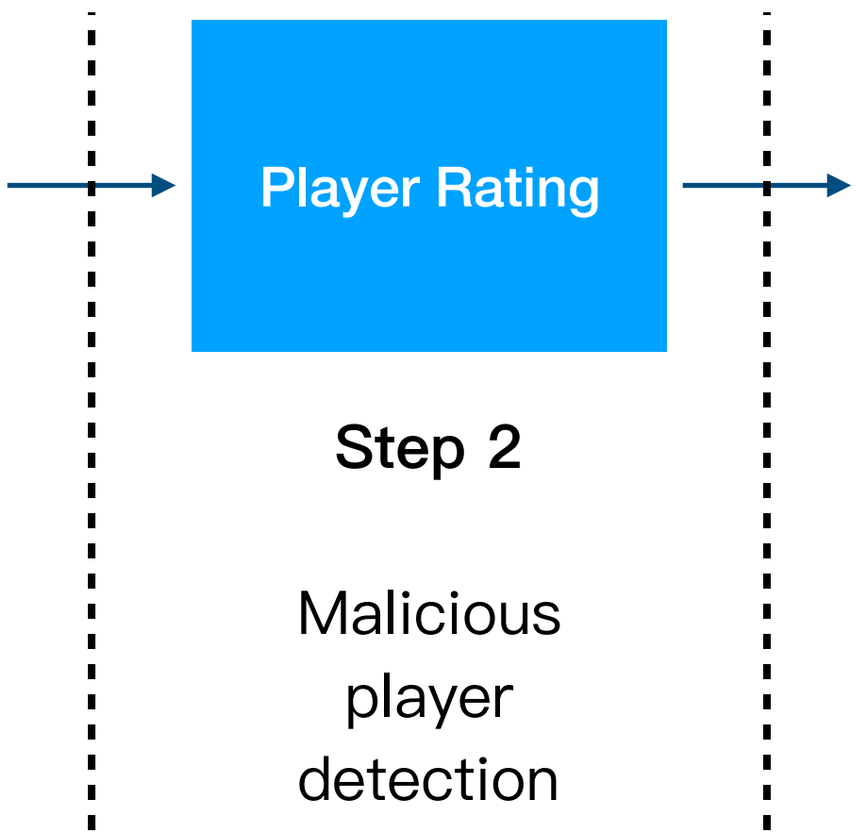
Might not possible and **not** make sense. **Counter example:** Considering:

Player A: ImageA, ImageB, ImageC

Player B: ImageB, ImageC, ImageD

Player C: ImageA, ImageE, ImageF

A directly sum for all image's ROI intersection does not make sense. From A to C, only one image sum; but from A to b sums two images



Assume we fixed this issue by a hard thinking,

Can we use agent from trusted people by tagging all images to ensure the image independent model's graph globally strong connected?

Two major issues for this solution:

1. Cost large labor (requires trusted people)
2. If the labor finished tags, we have no reason to use this HC system.

The fact is in the early stage of this project, I considered it first but abandoned (found few consideration on the manuscript).

We don't know solution yet, deep thinking on this issue is needed, But beyond the achievements we achieved in this semester.

* The unrefined report has notation omits and mistakes

Design: Model Refine (4)

Disaster Evaluation

Step 3

Disaster level
Based on reliable
task results

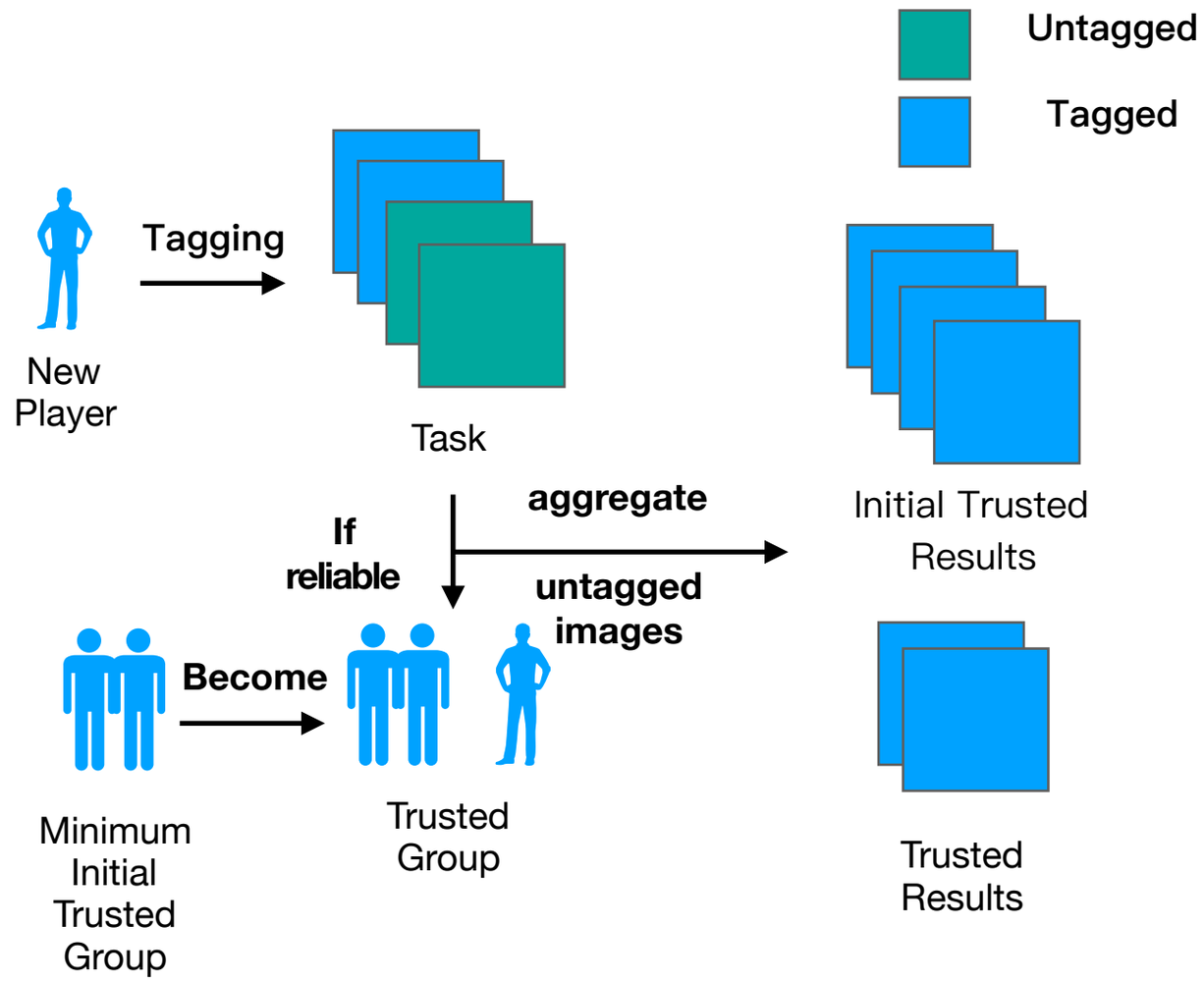
After the Step2, player result is determinated.
Finally, we can calculate the disaster level.

Definitions:
– **Tags vector** $T_{p,i,k}$: player ***p***, ***i***–th ROI on image ***k***;
– **Weight vector** ***v***: globally frequency for all tags as a vector;

Disaster level (DL):

$$DL = \sum_{k} \left(v \cdot \left(\sum_{i} T_{p,i,k} \right) \right)$$

Design: Initialization



Is enough player required for initial the system?

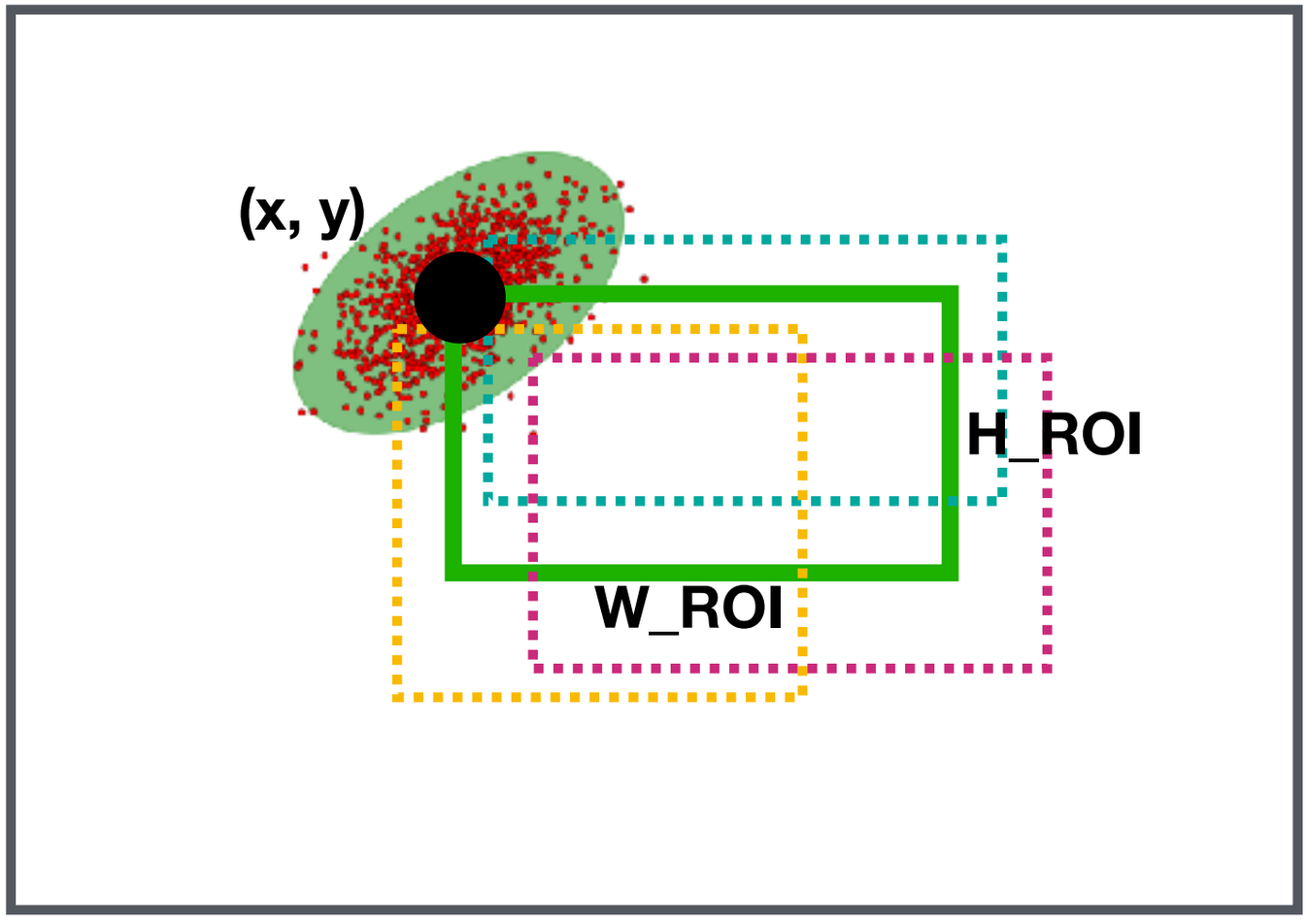
No, it only needs two pre-trusted persons.

Evaluations

Evaluations: Tech Criteria by Simulation

Generate **random data** and test the Rating Model through **accuracy** and **recall**, even **ROC** curve. (This is the common indicators for classification problem)

$ROI_{p,i,k}$ has four parameters: start coordinates (x,y), height and width.



FAQ:

Does the start point influence this simulation idea (Generate random datasets)?
No, it doesn't. Direction is not Important in this case.

Evaluations: Social & Ethics Aspects (1)

– Leakage of data

– Besides leak to ordinary users, the employees of UNICEF should have no right to access the entire database.

Evaluations: Social & Ethics Aspects (2)

The number of players

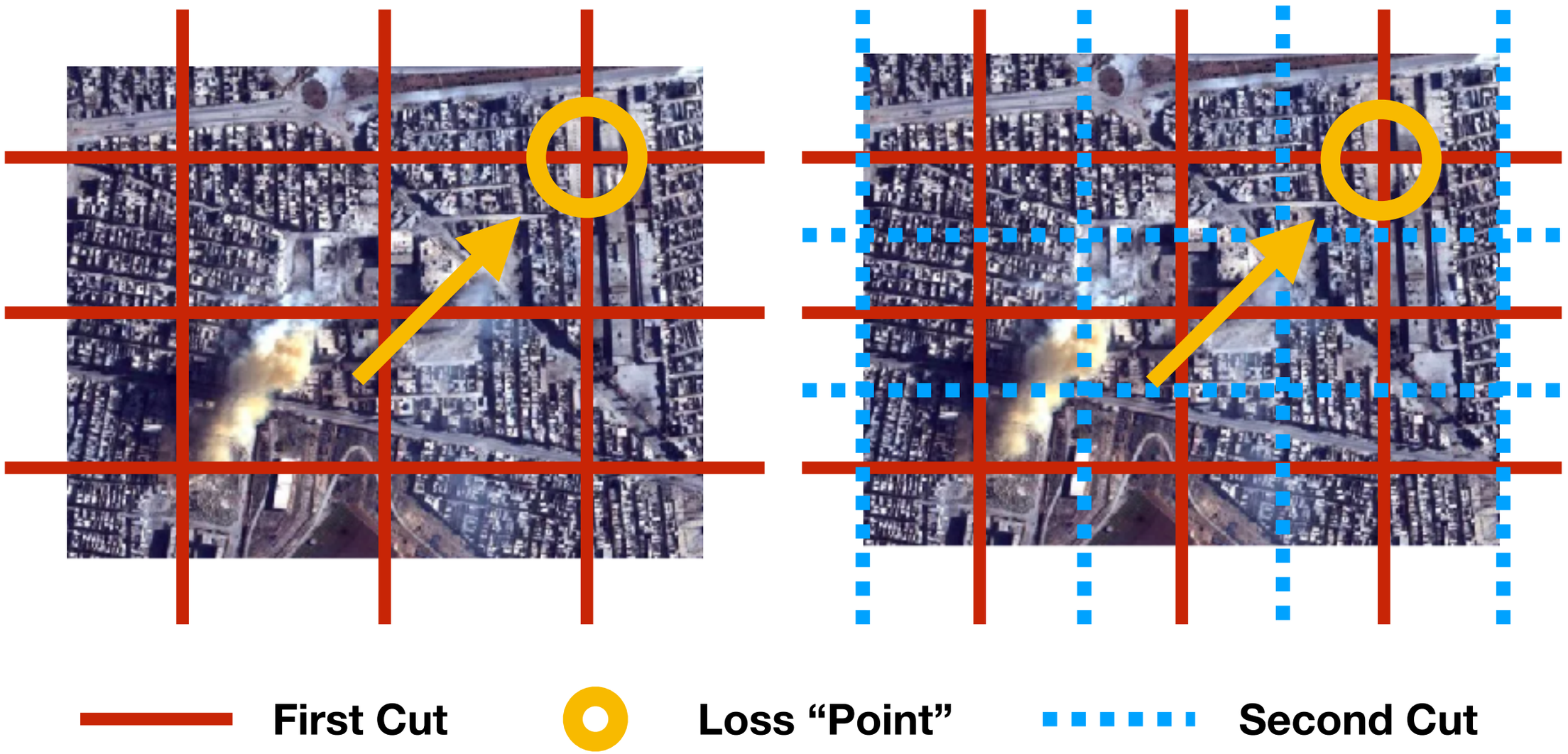
A. more users: more tags (higher accuracy of our level of disaster)
B. more users: trustworthy (higher trusted value which can filter malicious groups)

IPR (Intellectual Property Rights)

Evaluations: Social & Ethics Aspects (3)

Potential problem for leakage solution: Information loss

Solution: “Half Shifting” Cut



Limitations

Limitations (1): Evaluation Outdated

Each evaluation get invalid if region image is outdated

Limitations (2): Game Play and Playability

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

– E.g. Lake, forest, mountain, desert, some city area, nothing to tag.

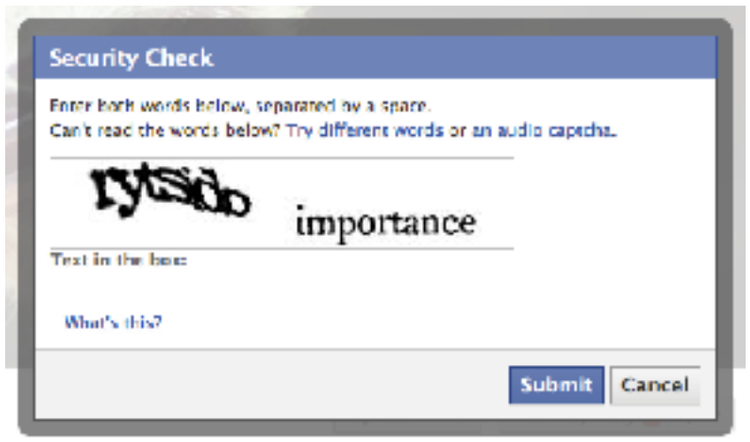
Future Works

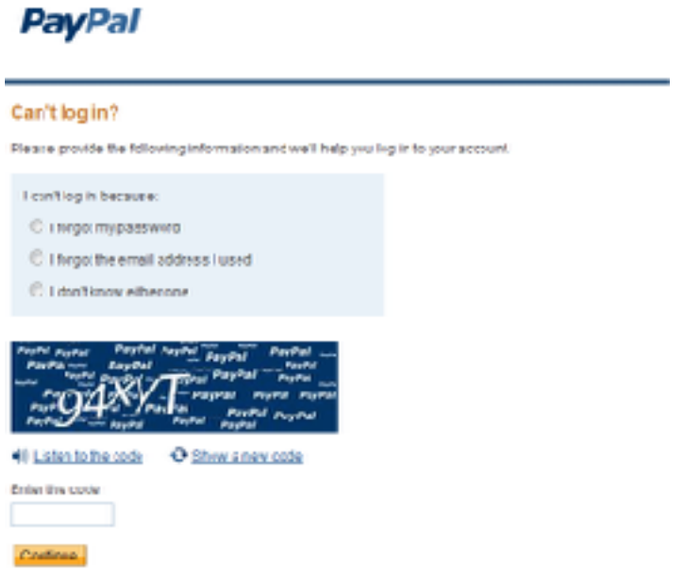
Future Extensions & Interactions

1. General System with Replaceable Rating Model: malicious detection algorithm can be replaced by any other machine learning algorithm with enough data.

2. Collaborative Computing for Playability: Collaborative computer vision computing improves playability.

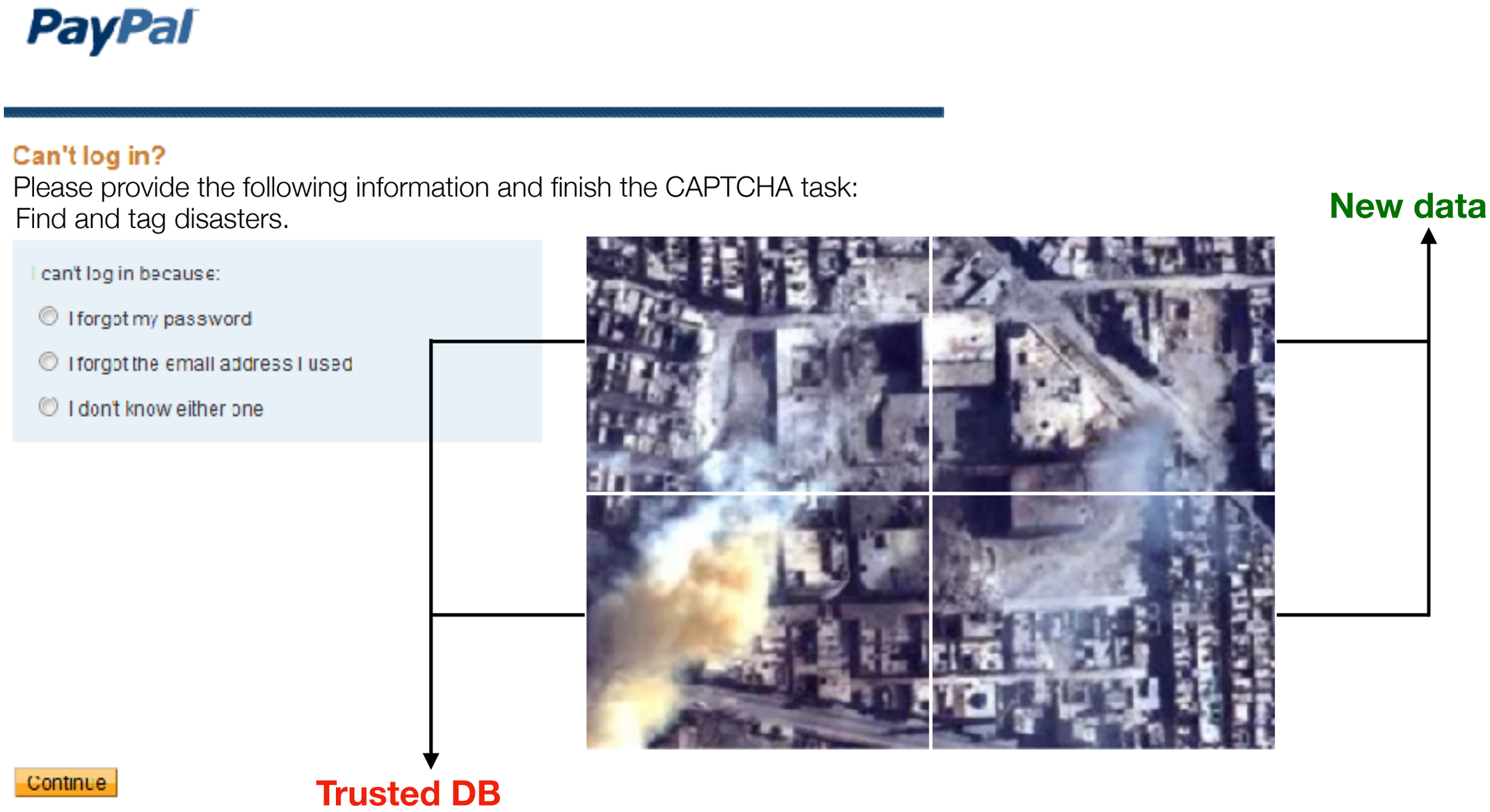
3. CAPTCHA Integrated System: the grate wastes of the computing power of humankind.





Future Extensions & Interactions

4. CAPTCHA Integrated System:



Conclusions

Pros Cons

1. GWAP, Image Tagging HC System (**Image Dependent**)
   + Only requires **two initial persons**
   + Participants play game **individually without registration**
   + Web–based, **cross all platforms**
   + Theoretical **proved**, with theoretical criterial

2. Social & Ethics
   + Incentivization: **Altruism**
   + Leakage & Privacy: **Fragments**
   + Malicious Player: **Model algorithm**

3. Limitations
   + Evaluation **outdated**
   + Playability for **senseless area**

4. Further Extensions & Interactions
   + The “general” system
   + Machine vision improvements
   + CAPTCHA integration

Achievements Summary

– Efforts:
  + Changkun Ou: System Design
  + Yifei Zhan: Mockup
  + Zhe Li: Issues discovery

– Front end **Mockup**, back end system **design (clear architecture with two math models)**, thoughts on possible **issues** yield by this design

– **Beamers**: lab1, lab2, lab3, final exam

– **Report**:
  + Version Control: https://github.com/changkun/hc-ss17-disaster-monitoring

Unrefined version		
Changkun Ou	**Yifei Zhan**	**Zhe Li**
* Typesetting * Abstract * Ch3: Design * Ch4 Intro * Ch4.1.1: Model Evaluation * Ch4.2.1: Evaluation Outdated * Ch5.1: Conclusions	* Chapter 2: Functionalities * Ch5 Intro * Chapter 5.2: System Extensions	* Ch1: Introduction * Ch4.1.2: Social&Ethical Issues * Ch4.2.2: Information Loss * Ch4.2.3: Playability * Ch5.3: Other Interactions