



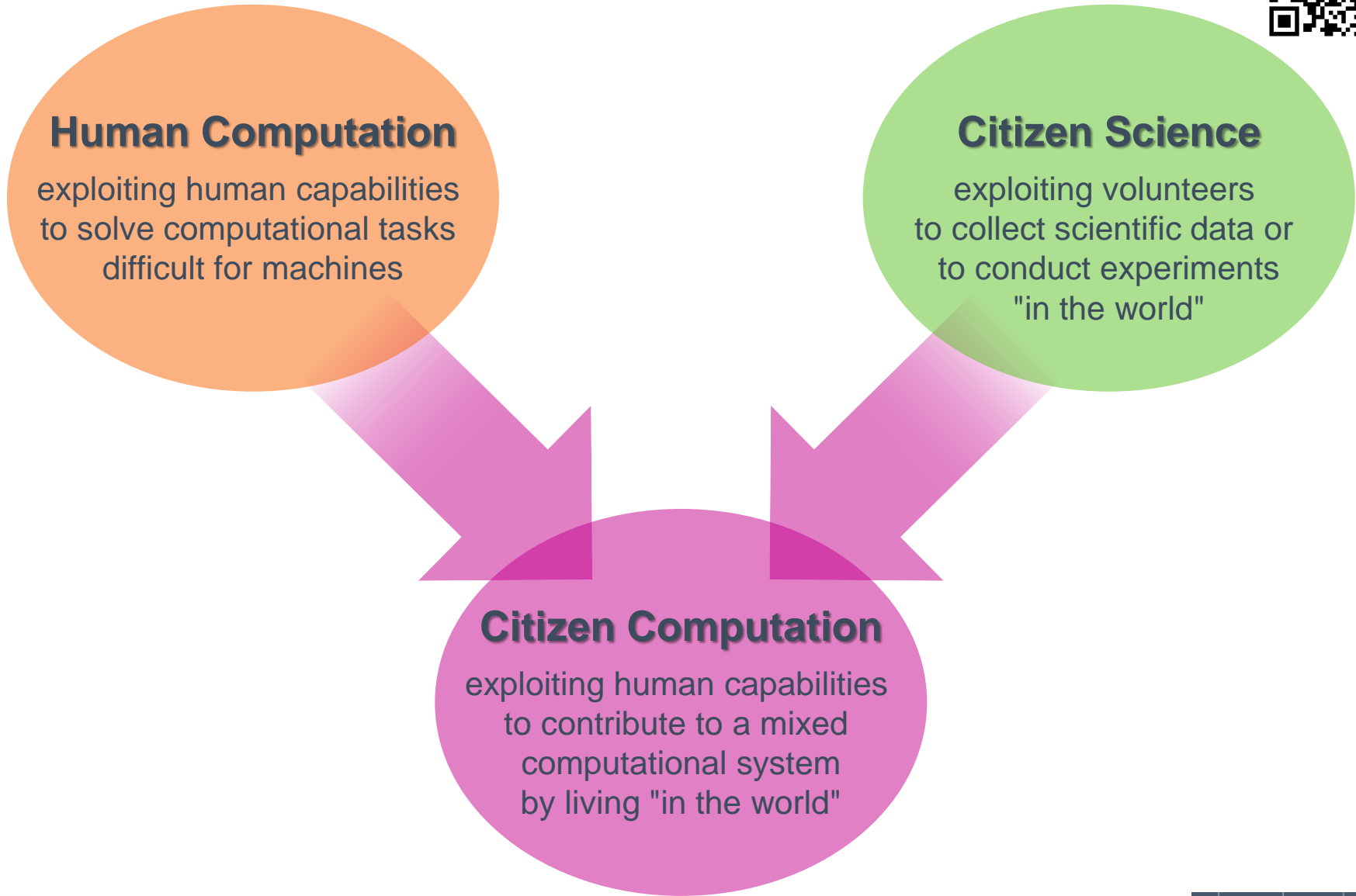
# Linking Smart Cities Datasets with Human Computation: the case of UrbanMatch

Irene Celino, Simone Contessa, Marta Corubolo,  
Daniele Dell'Aglio, Emanuele Della Valle,  
Stefano Fumeo and Thorsten Krüger

CEFRIEL – Politecnico di Milano – SIEMENS



# Citizen Computation



# UrbanMatch: Citizen Computation GWAP



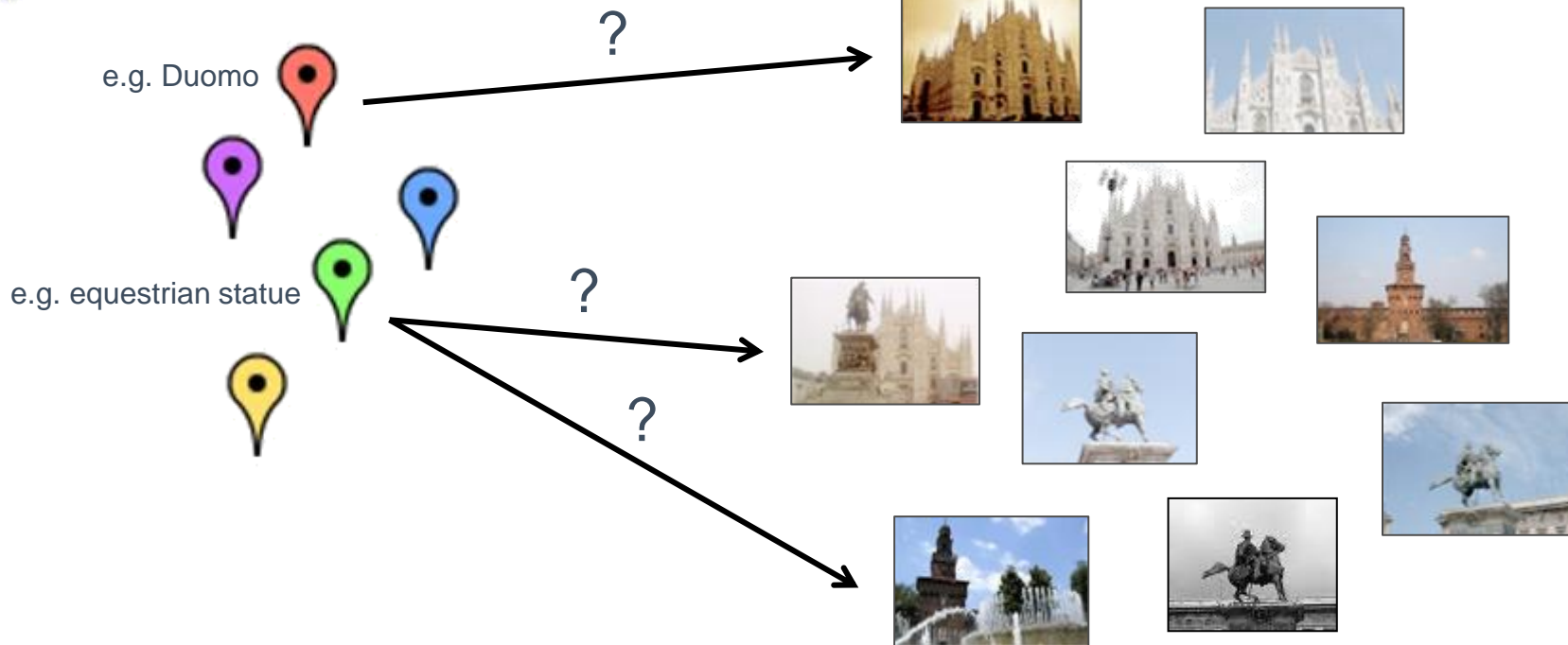
Places & POIs from  
OpenStreetMap

Photos from social media



Wikimedia  
Commons

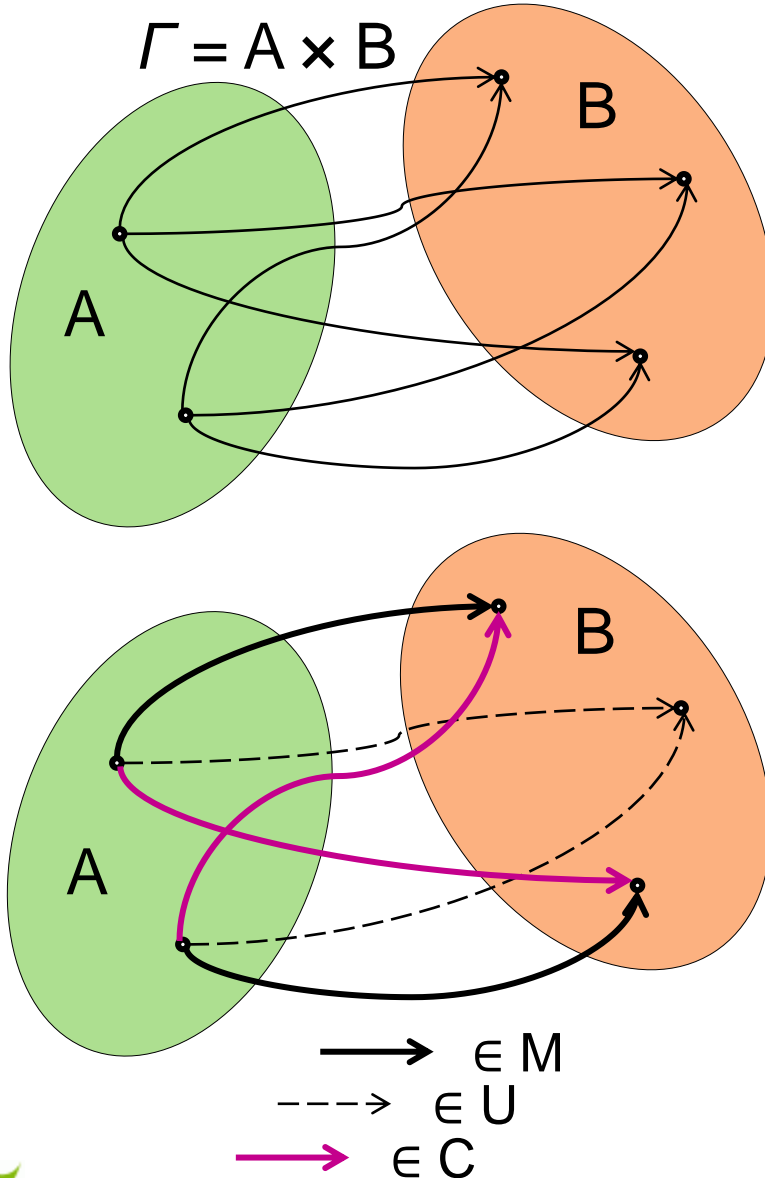
flickr



Which photos *actually* represent the urban POIs?  
Can we *link* POIs to their most representative photos?



# The Record Linkage Problem [Fellegi&Sunter,1969]



- Comparison space
  - $\Gamma = A \times B$
- Purpose is to split  $\Gamma$  in
  - $M$  (set of matches) and
  - $U$  (set of non-matches)
- Each link  $\gamma$  has a score  $s_\gamma$ 
  - $s_\gamma = P(\gamma \in \Gamma|M)/P(\gamma \in \Gamma|U)$
- Partitioning based on lower/upper thresholds:
  - $s_\gamma > UPPER \Rightarrow \gamma \in M$
  - $s_\gamma < LOWER \Rightarrow \gamma \in U$
  - $LOWER \leq s_\gamma \leq UPPER \Rightarrow \gamma \in C$   
(candidate links to be assessed by experts)

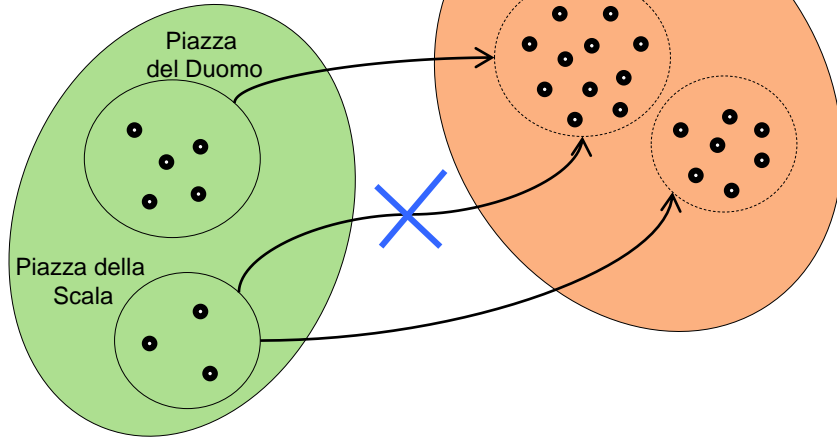
# UrbanMatch Linkage Problem



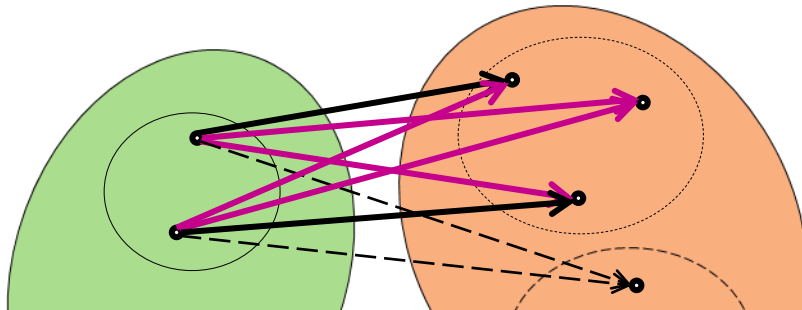
B (photos)

- A contains POIs
- B contains photos
- $\Gamma$  contains POI-photo links
- "Playable places" group POIs  $\rightarrow$  partitioning of  $\Gamma$  (problem space reduction)

A (POIs)



UrbanMatch links



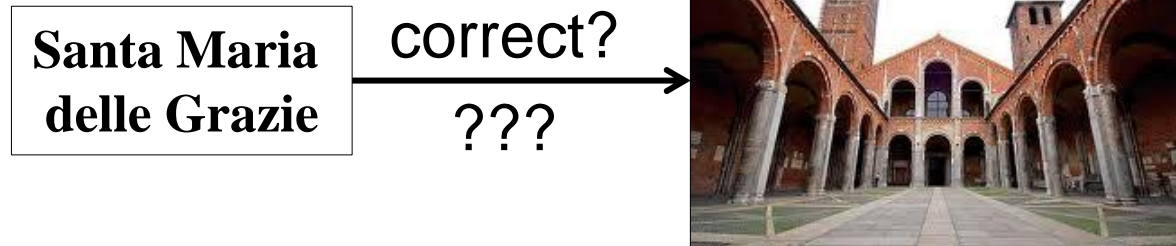
- Input data in UrbanMatch:
  - 34 POIs in 14 playable places
  - 11,089 photos
  - 196 links in M  $\rightarrow$
  - 382 links in U  $---$
  - 37,413 links in C  $\rightarrow$



# Showing links to UrbanMatch players

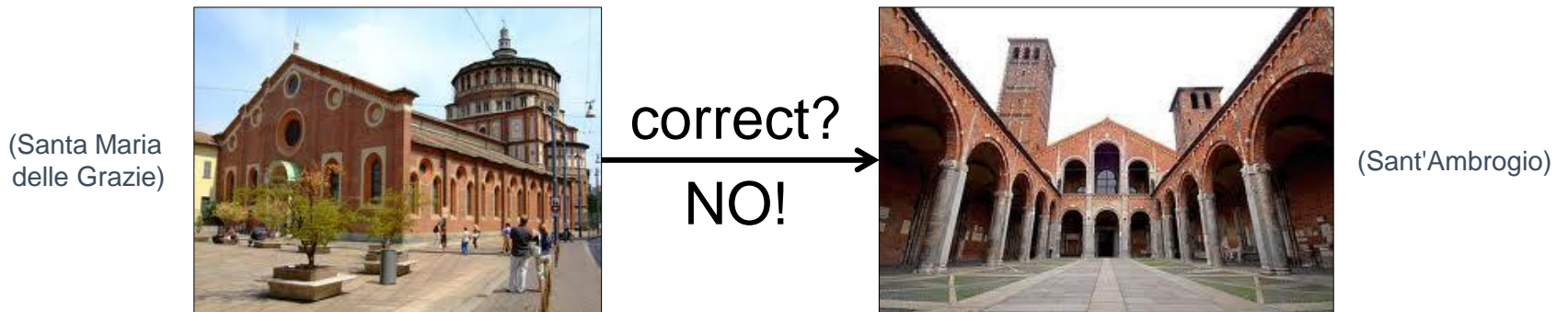


- The POI-photo links to be evaluated could be shown directly to players...



...but the player could not know the name of the POI

- On the other hand, the player can easily "see"



thus we use a "sure" photo as proxy for the POI, thus the UrbanMatch game is designed as a photo-pairing challenge



# UrbanMatch level definition



## Set A of POIs

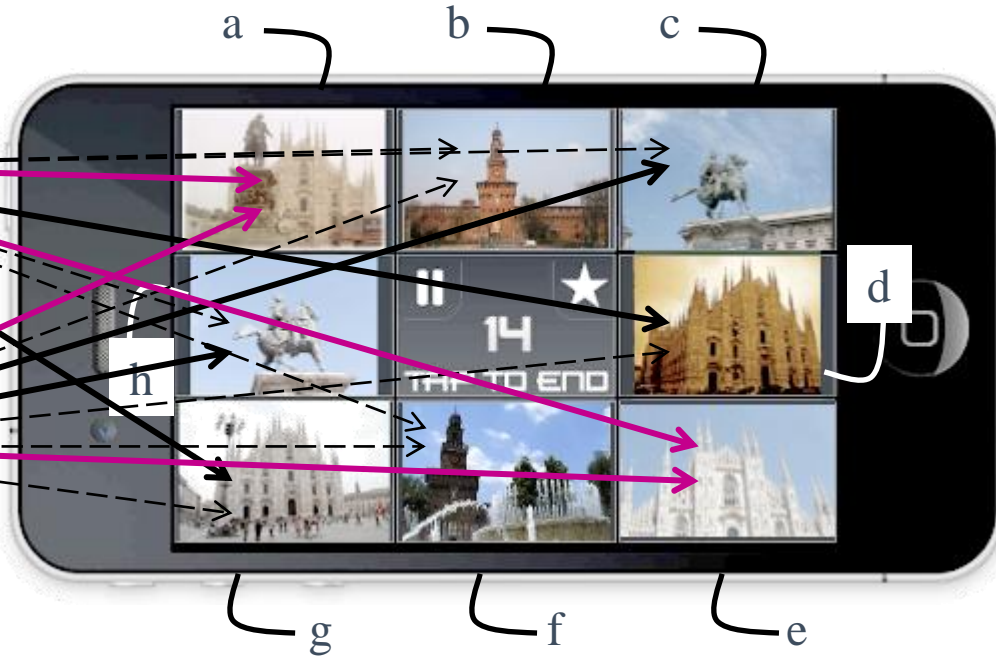


POI 1  
(Duomo)



POI 2  
(statue)

## Set B of photos



- Photos *d* and *g* depict *POI 1* (and do not depict *POI 2*)
- Photos *c* and *h* depict *POI 2* (and do not depict *POI 1*)
- Photos *b* and *f* do not depict *POI 1* nor *POI 2* (distractors)
- Photos *a* and *e* maybe depict *POI 1* and/or *POI 2*



# Feedbacks to UrbanMatch players



If the selected pair is correct (two sure photos) or plausible, UrbanMatch gives a positive feedback



If the selected pair is wrong (at least one selected photo is a "distractor" option), UrbanMatch gives a negative feedback (and counts the player's mistake)





# UrbanMatch: gameplay



Video at: <http://bit.ly/um-video>



# Link score in UrbanMatch



- Effect of players' evidences on the score  $s_\gamma$ 
  - If there is a *positive evidence* (an UrbanMatch player says that the link  $\gamma$  is correct),  $s_\gamma$  is increased at most by  $K_{pos}$
  - If there is a *negative evidence* (an UrbanMatch player says that the link  $\gamma$  is incorrect),  $s_\gamma$  is decreased at most by  $K_{neg}$
  - Each evidence is weighted with a factor that represents UrbanMatch *player's reliability*

$$r_p = e^{-\varepsilon_p/2}$$

where  $\varepsilon_p$  is the number of errors made by the player, thus  $r_p \in [0..1]$  (reliability is 1 if the player makes no mistake, 0.6 if he makes 1 error, almost zero if he makes 6 mistakes or more)

- For each link  $\gamma$  in  $C$ , its score varies as follows
  - Positive evidence:  $s'_\gamma = s_\gamma + K_{pos} * r_p$
  - Negative evidence:  $s'_\gamma = s_\gamma - K_{neg} * r_p$



# Evaluating UrbanMatch precision



## ■ Accuracy definition

(metrics specific to Data Quality)

game ability to *correctly*  
assess candidate links

candidate links *correctly* assessed to  
be either trustable or incorrect

$$\text{Accuracy} = \frac{(CM - FP) + (CU - FN)}{CM + CU}$$

candidate links assessed to be  
either trustable or incorrect

- CM = right links, i.e. candidate links moved from C to M (because  $s_\gamma > UPPER$ )
- CU = wrong links, i.e. candidate links moved from C to U (because  $s_\gamma < LOWER$ )
- FP = false positives (candidate links assessed as trustable but actually incorrect)
- FN = false negatives (candidate links assessed as incorrect but actually trustable)

***accuracy requires ground truth → computed only on a manually analyzed set of links for evaluation's sake !***



# Evaluating UrbanMatch as a GWAP



- Throughput and ALP definition  
(metrics specific to Games with a Purpose)

game ability to assess  
candidate links

$$\textit{Throughput} = \frac{\textit{SolvedProblems}}{\textit{PlayedTime}}$$

game ability to  
engage players

$$\textit{ALP} = \frac{\textit{PlayedTime}}{\textit{ActivePlayers}}$$

where  $\textit{SolvedProblems} = \textit{CM} + \textit{CU}$  in which:

- $\textit{CM} = \textit{right links}$ , i.e. candidate links moved from C to M  
(because  $s_\gamma > \textit{UPPER}$ )
- $\textit{CU} = \textit{wrong links}$ , i.e. candidate links moved from C to U  
(because  $s_\gamma < \textit{LOWER}$ )



# Setting the thresholds for link score $s_\gamma$



- Accuracy and Throughput as a function of *LOWER*

<b>LOWER</b>	<b>0.05</b>	<b>0.10</b>	<b>0.15</b>	<b>0.20</b>
CU	321	348	1152	<b>1216</b>
FN	4	5	7	<b>8</b>
Accuracy	98.75 %	98.56 %	99.39 %	<b>99.34 %</b>
Throughput	108.08	117.17	387.87	<b>409.42</b>

- Accuracy and Throughput as a function of *UPPER*

<b>UPPER</b>	<b>0.60</b>	<b>0.65</b>	<b>0.70</b>	<b>0.75</b>	<b>0.80</b>	<b>0.85</b>	<b>0.90</b>
CM	227	225	<b>68</b>	65	61	60	49
FP	48	47	<b>4</b>	4	3	3	1
Accuracy	78.85 %	79.11 %	<b>94.11 %</b>	95.38 %	95.08 %	95.00 %	97.95 %
Throughput	76.43	75.75	<b>22.89</b>	21.88	20.53	20.20	16.49



# UrbanMatch evaluation results



## Accuracy, Throughput and ALP results

- Evaluation data
  - 54 unique players, 290 games (781 game levels)
  - 2006 input links, 1284 assessed (trustable/incorrect)
  - upper threshold 0.70, lower threshold 0.20
- **Accuracy** (on the manually-checked subset)
  - **99.06%** → very good result 😊
- **Throughput and ALP** (on all processed links)
  - **485 links/h** → very good result 😊
  - **3m 17s** → one-time game, needs improvement 😐



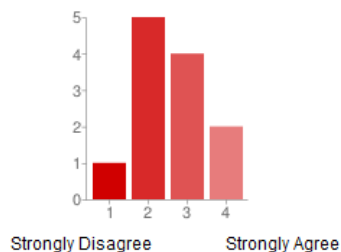
# Evaluation UrbanMatch as a whole



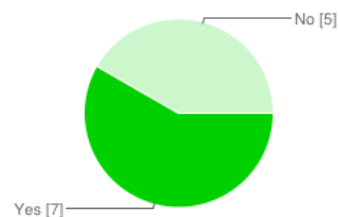
User-based evaluation: actual "engagement"

- Based on game evaluation literature and integrated with our research-specific questions
- Questionnaire at <http://bit.ly/um-survey>
- Findings:

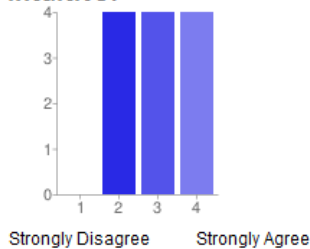
**Are the game rules simple, clear and intuitive?**



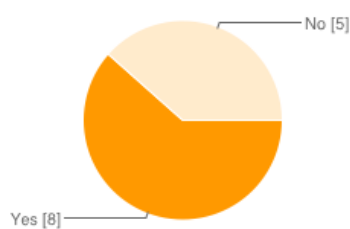
**Is it useful playing in the urban space?**



**Was the game interface clear and intuitive?**



**Did you suggest your friends to play?**



# Conclusions



- Citizen Computation games are effective to leverage **linking capabilities of "humans"**
- Game is to be designed to keep in consideration users' capabilities and *usage context*:
  - (lack of) **background knowledge**
  - small **mobile-phone** screen
- **Different tasks** to be solved need different game "missions"
  - *Linking* task → *pairing* mission
  - *Verifying* task → *rating/selecting* mission
  - *Collecting* task → *selecting/editing* mission



- 😊 *Advertisement* 😊 try out **Urbanopoly** game here in Boston, to collect and assess **geo-spatial linked data** participant to Semantic Web Challenge <http://bit.ly/urbanopoly>





# Thanks for your attention! Any question?



Irene Celino – CEFRIEL, ICT Institute Politecnico di Milano  
email: [Irene.Celino@cefriel.it](mailto:Irene.Celino@cefriel.it) – web: <http://swa.cefriel.it>