Discovering identical events between datasets

Welcome to the questionnaire!

The questionnaire is designed and conducted by Manar Attar for his bachleor thesis supervised by Shuai Wang at the Vrije Universiteit Amsterdam.

The goal of this questionnaire is to evaluate the accuracy of an algorithm designed to automatically identify duplicate events included in two different datasets that reports the damage of the Russian invasion of Ukraine since Feburary 2022.

This questionnaire contains 20 questions in total. It is expected to take around 25 mins to complete.

Conditions to participate in the questionnaire
The parcipants should be able to speak both English and Ukrainian at a proficient level. If not, they should quit this questionnaire. Participants should be over 20 years old. There are violent and blood in the content of links of social media post. Please quit this questionnaire if not feeling comfortable with it.

Task of the questionnaire
The participant is expected to identify if two events are identical or not. In this questionnaire, all pairs of events happened in the same city, on the same day. The location, city, postcode, and some description of the events are provided. Furthermore, we provide also information about the distance between the locations of the events. These events are from the datasets of Eyes on Russia and Cilivian Harm.

Participants may take further investigation, the events are backed by a link to the social media post that provides videos and/or photos of the incident. To access some posts on Twitter, you might be asked to verify your age. To access some posts on Telegram, you might be asked to log in when the media is too big. There might be content that requires private access (e.g. a private group on Telegram). Note that the content of some links may not be available.

In case one event is a subevent of the other or with some more general description but the two events are not identical, please choose the option "Unsure".

Otherwise, in case the participant is unsure of the classification of the event, the additional option "Unsure" can be used to answer the question.

In case of any unclearity, the participants can always contact for help or provide some feedback to the authors by sending an email at m.m.attar@student.vu.nl (and cc shuai.wang@vu.nl).

The data is completely anonymous so no personal information is needed.

* Indicates required question

1. **Question 1** *

Do you think the two reported events below are identical? You can use the provided link of social media (e.g. Facebook, Telegram, Twitter) to justify your decision.

---------- Event A ----------

Event A is located in **Mariupol** City with postcode 87512. The exact location can be found on Google Map.

Event A has the following description in the provenence dataset: Image 4 Damage to a shopping mall

Event A was added to the dataset with reference to the link to the social media as evidence.

---------- Event B ----------

Event B is located in **Mariupol** City with postcode 87551. The exact location can be found on Google Map.

Event B has the following description in the provenence dataset: Widespread damage caused by strike.

Event B was added to the dataset with reference to the link to the social media as evidence.

Event A and B are 1.42 km apart from each other.
How likely are these two reported events described above identical?

Mark only one oval.

- ◯ Very likely
- ◯ Likely
- ◯ Unsure
- ◯ Unlikely
- ◯ Very unlikely

2. Question 2 *

Do you think the two reported events below are identical? You can use the provided link of social media (e.g. Facebook, Telegram, Twitter) to justify your decision.

---------- Event A ----------

Event A is located in **Kherson** city with postcode 73040. The exact location can be found on Google Map.

Event A has the following description in the provenence dataset: Victim operated mine on street in Kherson

Event A was added to the dataset with reference to the link to the social media as evidence:

---------- Event B ----------

Event B is located in **Kherson** city

with postcode:73008. The exact location can be found on Google Map.

Event B has the following description in the provenence dataset: Victim operated device on street in Kherson

Event B was added to the dataset with reference to the link to the social media as evidence.

Event A and B are 0.06 km apart from each other

How likely are these two reported events described above identical?

Mark only one oval.

- ◯ Very likely
- ◯ Likely
- ◯ Unsure
- ◯ Unlikely
- ◯ Very unlikely

3. Question 3 *

Do you think the two reported events below are identical? You can use the provided link of social media (e.g. Facebook, Telegram, Twitter) to justify your decision.

---------- Event A ----------

Event A is located in **Mariupol** City with postcode 87512. The exact location can be found on Google Map.

Event A has the following description in the provenence dataset: Incendiary ammunition on Azovstal

Event A was added to the dataset with reference to the link to the social media as evidence.

---------- Event B ----------

Event B is located in **Mariupol** City with postcode 87534. The exact location can be found on Google Map.

Event B has the following description in the provenence dataset: Extremely clear use of incendiary munition use against Ukrainian positions in Avozstal Plant.

Event B was added to the dataset with reference to the link to the social media as evidence.

Event A and B are 0.22 km apart from each other.

How likely are these two reported events described above identical?

Mark only one oval.

- Very likely
- Likely
- Unsure
- Unlikely
- Very unlikely

4. Question 4 *

Do you think the two reported events below are identical? You can use the provided link of social media (e.g. Facebook, Telegram, Twitter) to justify your decision.

---------- Event A ----------

Event A is located in **Bucha** city with postcode:08101. The exact location can be found on Google Map.

Event A has the following description in the provenence dataset: Bomb remnant landing in pavement in residential area

Event A was added to the dataset with reference to the link to the social media as evidence.

---------- Event B ----------

Event B is located in **Bucha** city with postcode:08293. The exact location can be found on Google Map.

Event B has the following description in the provenence dataset: Rocket motor embedded in pavement in residential area.

Event B was added to the dataset with reference to the link to the social media as evidence.

Event A and B are 0.009 km apart from each other
How likely are these two reported events described above identical?

Mark only one oval.

- ◯ Very likely
- ◯ Likely
- ◯ Unsure
- ◯ Unlikely
- ◯ Very unlikely

5. Question 5 *

Do you think the two reported events below are identical? You can use the provided link of social media (e.g. Facebook, Telegram, Twitter) to justify your decision.

---------- Event A ----------

Event A is located in **Enerhodar** city with postcode 71341. The exact location can be found on Google Map.

Event A has the following description in the provenence dataset: 00:24 - Russian vehicles inside Zaporizhzhya nuclear power plant

Event A was added to the dataset with reference to the link to the social media as evidence.

---------- Event B ----------

Event B is located in **Enerhodar** city

with postcode: 71341. The exact location can be found on Google Map.

Event B has the following description in the provenence dataset: Zaporizhzhia nuclear power plant hit by shelling.

Event B was added to the dataset with reference to the link to the social media as evidence.

Event A and B are 0.81 km apart from each other.

How likely are these two reported events described above identical?

Mark only one oval.

- ◯ Very likely
- ◯ Likely
- ◯ Unsure
- ◯ Unlikely
- ◯ Very unlikely

6. Question 6 *

Do you think the two reported events below are identical? You can use the provided link of social media (e.g. Facebook, Telegram, Twitter) to justify your decision.

---------- Event A ----------

Event A is located in **Bucha** city with postcode 71341. The exact location can be found on Google Map.

Event A has the following description in the provenence dataset: Seven people civillian clothes killed, at least two with hands tied behind their backs. Russian army ration packs & ammunitions container present at the site.

Event A was added to the dataset with reference to the link to the social media as evidence.

---------- Event B ----------

Event B is located in **Bucha** city with postcode 71341. The exact location can be found on Google Map.

Event B has the following description in the provenence dataset: [graphicbodies] seven bodies in civilian clothes killed, at least two with hands tied behind their backs. Russian army ration packs & ammunitions container present at the site.

Event B was added to the dataset with reference to the link to the social media as evidence.

Event A and B are 0.03 km apart from each other

Mark only one oval.

- ◯ Very likely
- ◯ Likely
- ◯ Unsure
- ◯ Unlikely
- ◯ Very unlikely

7. Question 7 *

Do you think the two reported events below are identical? You can use the provided link of social media (e.g. Facebook, Telegram, Twitter) to justify your decision.

---------- Event A ----------

Event A is located in **Kramatorsk** City with postcode 84395. The exact location can be found on Google Map.

Event A has the following description in the provenence dataset: Kramatorsk Psychiatric Hospital damaged

Event A was added to the dataset with reference to the link to the social media as evidence.

---------- Event B ----------

Event B is located in **Kramatorsk** City with postcode 84395. The exact location can be found on Google Map.

Event B has the following description in the provenence dataset: Kramatorsk psychiatric hospital damaged by explosions.

Event B was added to the dataset with reference to the link to the social media as evidence.

Event A and B are 0.12 km apart from each other

Mark only one oval.

- ◯ Very likely
- ◯ Likely
- ◯ Unsure
- ◯ Unlikely
- ◯ Very unlikely

8. Question 8 *

Do you think the two reported events below are identical? You can use the provided link of social media (e.g. Facebook, Telegram, Twitter) to justify your decision.

---------- Event A ----------

Event A is located in **Irpin** city with postcode 08201. The exact location can be found on Google Map.

Event A has the following description in the provenence dataset: Destruction around Irpin

Event A was added to the dataset with reference to the link to the social media as evidence.

---------- Event B ----------

Event B is located in **Irpin** city with postcode 08203. The exact location can be found on Google Map.

Event B has the following description in the provenence dataset: US journalist Brent Renaud killed by small arms fire in Irpin.

Event B was added to the dataset with reference to the link to the social media as evidence.

Event A and B are 0.05 km apart from each other

How likely are these two reported events described above identical?

Mark only one oval.

- ◯ Very likely
- ◯ Likely
- ◯ Unsure
- ◯ Unlikely
- ◯ Very unlikely

9. Question 9 *

Do you think the two reported events below are identical? You can use the provided link of social media (e.g. Facebook, Telegram, Twitter) to justify your decision.

---------- Event A ----------

Event A is located in **Mariupol** city with postcode 87500. The exact location can be found on Google Map.

Event A has the following description in the provenence dataset: Children's hospital in Mariupol heavily damaged by strike. Multiple injuries and fatalities reported. At least one large crater visible.

Event A was added to the dataset with reference to the link to the social media as evidence:.

---------- Event B ----------

Event B is located in **Mariupol** city with postcode 87551. The exact location can be found on Google Map.

Event B has the following description in the provenence dataset: Maternity hospital in Mariupol heavily damaged by strike. Multiple injuries and fatalities reported. At least one large crater visible.

Event B was added to the dataset with reference to the link to the social media as evidence.

Event A and B are 0.02 km apart from each other.

How likely are these two reported events described above identical?

Mark only one oval.

- ◯ Very likely
- ◯ Likely
- ◯ Unsure
- ◯ Unlikely
- ◯ Very unlikely

10. Question 10 *

Do you think the two reported events below are identical? You can use the provided link of social media (e.g. Facebook, Telegram, Twitter) to justify your decision.

---------- Event A ----------

Event A is located in **Kyiv** city with postcode 03028. The exact location can be found on Google Map.

Event A has the following description in the provenence dataset: Civilian vehicles on fire after missile strikes.

Event A was added to the dataset with reference to the link to the social media as evidence.

---------- Event B ----------

Event B is located in **Kyiv** city with postcode 04107. The exact location can be found on Google Map.

Event B has the following description in the provenence dataset: Three civilians reportedly killed, at least 11 injured in a missile strike on Kyiv.

Event B was added to the dataset with reference to the link to the social media as evidence.

Event A and B are 0.04(km) apart from each other

How likely are these two reported events described above identical?

Mark only one oval.

- ◯ Very likely
- ◯ Likely
- ◯ Unsure
- ◯ Unlikely
- ◯ Very unlikely

11. Question 11 *

Do you think the two reported events below are identical? You can use the provided link of social media (e.g. Facebook, Telegram, Twitter) to justify your decision.

---------- Event A ----------

Event A is located in **Mykolaiv** city with postcode 54055. The exact location can be found on Google Map.

Event A has the following description in the provenence dataset: Black Sea Mohyla University partially destroyed in Mykolaiv

Event A was added to the dataset with reference to the link to the social media as evidence.

Event 2:

---------- Event B ----------

Event B is located in **Mykolaiv** city with postcode 54003. The exact location can be found on Google Map.

Event B has the following description in the provenence dataset: Damage to the Black Sea National University in a shelling in Mikolaiv.

Event B was added to the dataset with reference to the link to the social media as evidence:.

Event A and B are 0.10 km apart from each other.

How likely are these two reported events described above identical?

Mark only one oval.

- ◯ Very likely
- ◯ Likely
- ◯ Unsure
- ◯ Unlikely
- ◯ Very unlikely

12. Question 12 *

Do you think the two reported events below are identical? You can use the provided link of social media (e.g. Facebook, Telegram, Twitter) to justify your decision.

---------- Event A ----------

Event A is located in **Kharkiv** City with postcode: 62407. The exact location can be found on Google Map.

Event A has the following description in the provenence dataset: Body on the street, burning vehicle nearby

Event A was added to the dataset with reference to the link to the social media as evidence.

---------- Event B ----------

Event B is located in **Kharkiv** City with postcode: 08464. The exact location can be found on Google Map.

Event B has the following description in the provenence dataset: [graphic: bodies] Two bodies lying in street. What may be a crater is visible.

Event B was added to the dataset with reference to the link to the social media as evidence.

Event A and B are 1.23 km apart from each other
How likely are these two reported events described above identical?

Mark only one oval.

- ◯ Very likely
- ◯ Likely
- ◯ Unsure
- ◯ Unlikely
- ◯ Very unlikely

13. **Question 13** *

Do you think the two reported events below are identical? You can use the provided link of social media (e.g. Facebook, Telegram, Twitter) to justify your decision.

---------- Event A ----------

Event A is located in **Donetska** city with postcode 92911. The exact location can be found on Google Map.

Event A has the following description in the provenence dataset: Artillery fell next to the Tsentr Bus Station in Donetsk on the 14th July resulting in the deaths of several civilians and civilian injuries.

Event A was added to the dataset with reference to the link to the social media as evidence.

---------- Event B ----------

Event B is located in **Donetska** city with postcode 92911. The exact location can be found on Google Map.

Event B has the following description in the provenence dataset: [Graphiccovered bodies, blood] Attack near the Tsentr Bus Station, resulting in at least one death and several reported injuries

Event B was added to the dataset with reference to the link to the social media as evidence.

Event A and B are 0.07 km apart from each other

How likely are these two reported events described above identical?

Mark only one oval.

- ◯ Very likely
- ◯ Likely
- ◯ Unsure
- ◯ Unlikely
- ◯ Very unlikely

14. **Question 14** *

Do you think the two reported events below are identical? You can use the provided link of social media (e.g. Facebook, Telegram, Twitter) to justify your decision.

---------- Event A ----------

Event A is located in **Donetska** city with postcode 86062. The exact location can be found on Google Map

Event A has the following description in the provenence dataset: Fire at Donetsk railway station

Event A was added to the dataset with reference to the link to the social media as evidence

---------- Event B ----------

Event B is located in **Donetska** city with postcode 86062. The exact location can be found on Google Map

Event B has the following description in the provenence dataset: Donetsk railway station on fire.

Event B was added to the dataset with reference to the link to the social media as evidence

Event A and B are 0.02 km apart from each other.

How likely are these two reported events described above identical?

Mark only one oval.

- ◯ Very likely
- ◯ Likely
- ◯ Unsure
- ◯ Unlikely
- ◯ Very unlikely

15. **Question 15** *

Do you think the two reported events below are identical? You can use the provided link of social media (e.g. Facebook, Telegram, Twitter) to justify your decision.

---------- Event A ----------

Event A is located in **Mariupol** city with postcode 87551. The exact location can be found on Google Map

Event A has the following description in the provenence dataset: Medical clinic and residential building damaged

Event A was added to the dataset with reference to the link to the social media as evidence

---------- Event B ----------

Event B is located in **Mariupol** city with postcode 87534. The exact location can be found on Google Map

Event B has the following description in the provenence dataset: Damage to what was reportedly a medical clinic and residential building

Event B was added to the dataset with reference to the link to the social media as evidence

Event A and B are 0.05 (km) apart from each other

Mark only one oval.

- ◯ Very likely
- ◯ Likely
- ◯ Unsure
- ◯ Unlikely
- ◯ Very unlikely

16. **Question 16** *

Do you think the two reported events below are identical? You can use the provided link of social media (e.g. Facebook, Telegram, Twitter) to justify your decision.

---------- Event A ----------

Event A is located in **Kyiv** City with postcode 03146. The exact location can be found on Google Map

Event A has the following description in the provenence dataset: Monument is damaged and dismantled in Kyiv.

Event A was added to the dataset with reference to the link to the social media as evidence

---------- Event B ----------

Event B is located in **Kyiv** City with postcode 02222. The exact location can be found on Google Map

Event B has the following description in the provenence dataset: Cruise missiles strikes on Kyiv.

Event B was added to the dataset with reference to the link to the social media as evidence

Event A and B are 3.21 (km) apart from each other

How likely are these two reported events described above identical?

Mark only one oval.

- ◯ Very likely
- ◯ Likely
- ◯ Unsure
- ◯ Unlikely
- ◯ Very unlikely

17. **Question 17** *

Do you think the two reported events below are identical? You can use the provided link of social media (e.g. Facebook, Telegram, Twitter) to justify your decision.

---------- Event A ----------

Event A is located in **Mariupol** City with postcode 87512. The exact location can be found on Google Map

Event A has the following description in the provenence dataset: Alleged use of incendiary munition against Ukranian positions in Avozstal Plant.

Event A was added to the dataset with reference to the link to the social media as evidence

---------- Event B ----------

Event B is located in **Mariupol** City with postcode 87534. The exact location can be found on Google Map

Event B has the following description in the provenence dataset: Extremely clear use of incendiary munition use against Ukrainian positions in Avozstal Plant.

Event B was added to the dataset with reference to the link to the social media as evidence

Event A and B are 0.04 (km) apart from each other
How likely are these two reported events described above identical?

Mark only one oval.

- ◯ Very likely
- ◯ Likely
- ◯ Unsure
- ◯ Unlikely
- ◯ Very unlikely

18. **Question 18** *

Do you think the two reported events below are identical? You can use the provided link of social media (e.g. Facebook, Telegram, Twitter) to justify your decision.

---------- Event A ----------

Event A is located in **Kharkiv** City with postcode 64566. The exact location can be found on Google Map

Event A has the following description in the provenence dataset: Bomb in front of store, injured person and foot seen near store

Event A was added to the dataset with reference to the link to the social media as evidence

---------- Event B ----------

Event B is located in **Kharkiv** City with postcode 64566. The exact location can be found on Google Map

Event B has the following description in the provenence dataset: [Graphic: Video shows at least two dead, bloodied bodies lying on the street] Explosions kill at least two people and damage civilian infrastructure.

Event B was added to the dataset with reference to the link to the social media as evidence

Event A and B are 2.37 (km) apart from each other

How likely are these two reported events described above identical?

Mark only one oval.

- ◯ Very likely
- ◯ Likely
- ◯ Unsure
- ◯ Unlikely
- ◯ Very unlikely

19. **Question 19** *

Do you think the two reported events below are identical? You can use the provided link of social media (e.g. Facebook, Telegram, Twitter) to justify your decision.

---------- Event A ----------

Event A is located in **Kharkiv** City with postcode 61128. The exact location can be found on Google Map

Event A has the following description in the provenence dataset: Shelling in Kharkiv, reported casualties

Event A was added to the dataset with reference to the link to the social media as evidence

---------- Event B ----------

Event B is located in **Kharkiv** City with postcode 64566. The exact location can be found on Google Map

Event B has the following description in the provenence dataset: Damage to Sarzhyn Yar public park. Reportedly one women died

Event B was added to the dataset with reference to the link to the social media as evidence

Event A and B are 2.46 (km) apart from each other

How likely are these two reported events described above identical?

Mark only one oval.

- ◯ Very likely
- ◯ Likely
- ◯ Unsure
- ◯ Unlikely
- ◯ Very unlikely

20. **Question 20** *

Do you think the two reported events below are identical? You can use the provided link of social media (e.g. Facebook, Telegram, Twitter) to justify your decision.

---------- Event A ----------

Event A is located in **Kharkiv** City with postcode 61057. The exact location can be found on Google Map

Event A has the following description in the provenence dataset: Damage to a large retail store in centrak Kharkiv

Event A was added to the dataset with reference to the link to the social media as evidence

---------- Event B ----------

Event B is located in **Kharkiv** City with postcode 61046. The exact location can be found on Google Map

Event B has the following description in the provenence dataset: Munition remnants embedded in pavement

Event B was added to the dataset with reference to the link to the social media as evidence

Event A and B are 1.72 (km) apart from each other

How likely are these two reported events described above identical?

Mark only one oval.

- ◯ Very likely
- ◯ Likely
- ◯ Unsure
- ◯ Unlikely
- ◯ Very unlikely