

MiDe22: An Annotated Multi-Event Tweet Dataset for Misinformation Detection

Anonymous Author(s)

A Appendix

In this appendix, we explain tweet crawling process by providing some queries. We describe the annotation tool interface. Finally, we provide the event description and user engagement statistics for each event.

A.1 Tweet Crawling

An example query for crawling tweets from Twitter API for a specific event is given in Table 1. We manually determined the events and query keywords by browsing events in fact checking web pages. The motivation for using a different keyword set for the Other class is that we might not find irrelevant tweets or tweets with no information with the query keywords prepared for the True and False classes. The start and end dates of querying tweets are selected before and after two months of the evidence date, unless restricted by the crawling date (2022-04-06).

We run four consecutive queries for each event. We first collected tweets with media object (image, video, or GIF) and geographic location tags (Query-1). If the number of such tweets was not enough to fulfill the number of target tweets (50 tweets per class, total of 100 tweets for the True and False classes), then we collected tweets without media objects and geographic location tags (Query-2). After running queries for the True and False classes, we collected tweets for the Other class with the same approach by first searching for media objects (Query-3) and then regular tweets (Query-4).

We set the highest number of tweets to be collected for each class (true, false, other) to 50 tweets to provide a balance among classes and limit the total number of tweets to be annotated.

A.2 Annotation Tool

The annotation tool is implemented by using INCEPTION¹. Each annotator had a user account, and could login anytime to continue their annotations from where they have left. After login, the Dashboard screen (Figure 1) provides the main descriptions regarding the annotation task. In the Annotation screen (Figure 2) the corresponding tweets in the dataset are listed for each annotator with a sidebar for labeling the class (True, False, and Other). We did not edit the original tweet contents, and provided any multimedia objects and external links as well.

A.3 Event List

There are four topics in the dataset. The topics are the 2020 Russia-Ukraine War, COVID-19 pandemic, Refugees (Immigration), and Miscellaneous. There are 40 events under four topics for both languages. The list of events along with their topics are given in Table 2 for English, and Table 3 for Turkish.

A.4 User Engagements

The detailed list of tweets and user engagements (like, retweet, reply, and quote) per event are given in Table 4 for both English and Turkish. Note that we provide the average and total numbers of tweets and user engagements at the end of the table.

¹<https://inception-project.github.io/>

Language	English
ID	EN2
Topic	Ukraine
Title	Viral clip shows 'Arma 3' video game not war between Russia and Ukraine
Evidence URL	https://www.usatoday.com/story/news/factcheck/2022/02/21/.../6879521001/
Query keywords	arma 3 russia ukraine
Evidence date	2022-02-21
Query start date	2021-12-21
Query end date	2022-04-06
Other keywords	russia ukraine war video
Sample tweet(s)	1499460925253832707
Query-1	arma 3 russia ukraine lang:en (has:media OR has:geo) -is:retweet 2021-12-21 2022-04-06 100
Query-2	arma 3 russia ukraine lang:en -is:retweet 2021-12-21 2022-04-06 100
Query-3 for Other	russia ukraine war video lang:en (has:media OR has:geo) -is:retweet 2021-12-21 2022-04-06 50
Query-4 for Other	russia ukraine war video lang:en -is:retweet 2021-12-21 2022-04-06 50

Table 1: An example event in the dataset. A part of URL is cropped due to space constraints.

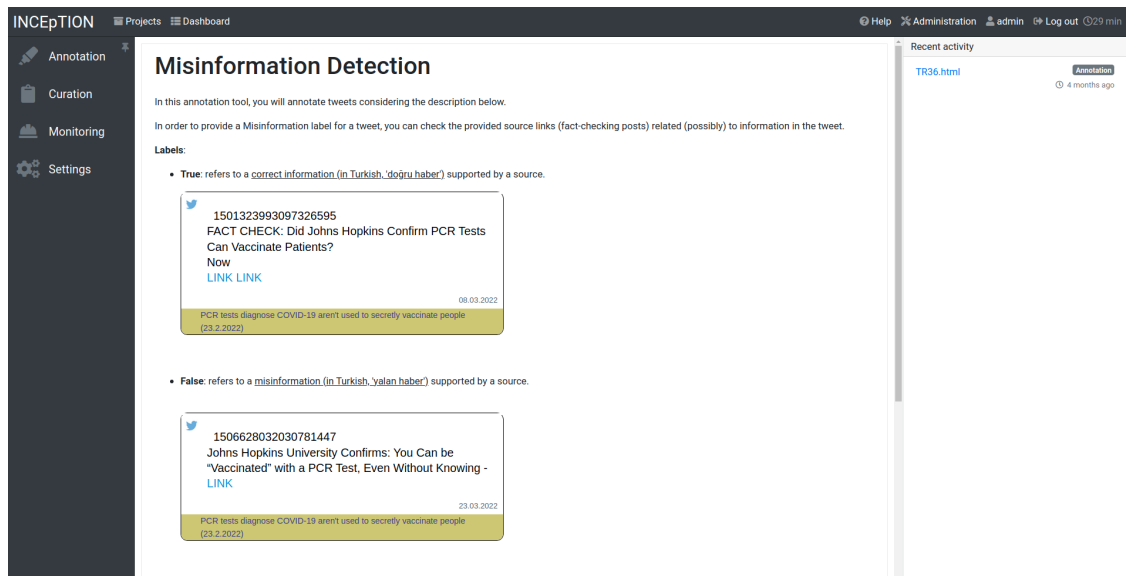


Figure 1: Dashboard screen of the annotation tool.

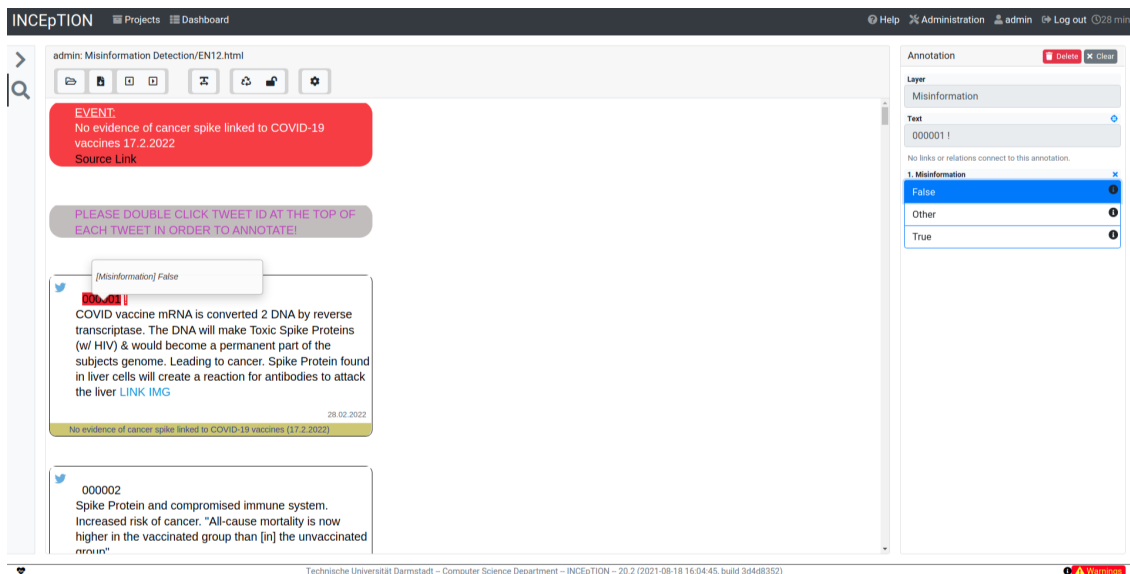


Figure 2: Annotation screen of the annotation tool.

ID	Topic	Event Description
EN01	Ukraine	30mm Russia Embassy in Canada claim that it is not occupying Ukrainian territory
EN02	Ukraine	Viral clip shows 'Arma 3' video game not war between Russia and Ukraine
EN03	Ukraine	Ethnic Russians face "genocide perpetrated by the Kyiv regime."
EN04	Ukraine	Photo shows a Russian tank Ukrainians are selling on eBay
EN05	Ukraine	Where is Zelensky?'Zelensky In Kyiv Has Not Fled To Poland';Ukrainian Parliament Counters Russia's Claim
EN06	Ukraine	Evidence of Russia's war in Ukraine is plentiful and proves it is not 'staged'
EN07	Ukraine	The former Miss Universe Anastasiia Lenna turned in her high heels for combat boots to fight for her country.
EN08	Ukraine	A Time magazine cover shows Vladimir Putin with Hitler's mustache.
EN09	Ukraine	Ukrainian fighter-ace known as the 'Ghost of Kyiv's real name is Samuyil Hyde.
EN10	Ukraine	The claim: Vladimir Putin has banned the Rothschild family from entering Russia
EN11	Covid	COVID-19 vaccines do not contain HIV
EN12	Covid	No evidence of cancer spike linked to COVID-19 vaccines
EN13	Covid	PCR tests diagnose COVID-19 aren't used to secretly vaccinate people
EN14	Covid	Over-vaccination causes faster mutation of the (COVID-19) virus which causes a super virus.
EN15	Covid	Says Bob Saget died from the COVID-19 vaccine.
EN16	Covid	Corona PCR 'test' is implanting a microchip.
EN17	Covid	Pfizer CEO: New Pill Will Have a Microchip That Transmits Info Once You Swallow It!
EN18	Covid	WHO Director-General: The vaccines are being used to kill children.
EN19	Covid	COVID-19 vaccines are gene therapy and a recent Forbes article proves that.
EN20	Covid	COVID-19 vaccines contain luciferase.
EN21	Refugees	Hundreds of 'illegals' were dropped off at Florida hotel.
EN22	Refugees	Biden administration has released over 7000 illegal aliens who were COVID positive just in one Texas city.
EN23	Refugees	There are 22 million illegal aliens living in America and ... voting illegally.
EN24	Refugees	If you are an undocumented immigrant and 55 or older you're eligible for Medicaid.
EN25	Refugees	Report: Biden administration busing 'thousands' of illegals to red states to get permanent status.
EN26	Refugees	The migrant caravan from Tapachula Mexico "is the population of Minneapolis."
EN27	Refugees	Was a 16-year-old Russian beaten to death by a group of Ukrainians in Germany?
EN28	Refugees	Migrants on Polish border were able to enter Belarus thanks to the visa-free regime
EN29	Refugees	A sea of unknown migrants' graves is located at the Polish border
EN30	Refugees	Turkey may introduce ISIS members in the EU for terrorist attacks at Christmas
EN31	Misc	West Coast wildfires are due to "bad forest maintenance" not climate change.
EN32	Misc	Most of the wildfires on the West Coast are "all being started by antifa."
EN33	Misc	Bill Gates hatches 'horribly stupid' plan to block out the sun
EN34	Misc	Five Wisconsin cities received money for the 2020 election from Mark Zuckerberg ... election bribery.
EN35	Misc	Hillary Clinton spied on President Trump.
EN36	Misc	Eating grapefruit "if you're on antidepressants" can "cause a fatal overdose."
EN37	Misc	Obama certificate a fake.
EN38	Misc	Microsoft and Bill Gates created a 1999 video game called "Omikron."
EN39	Misc	The claim: President Joe Biden plans to sell Alaska to Russia
EN40	Misc	The claim: A compass wouldn't work if Earth is spherical

Table 2: **MiDe22-EN Event Titles.** Each event belongs to a topic, namely the Russo-Ukraine War, COVID-19, Refugees, and Miscellaneous.

ID	Topic	Event Description
TR01	Ukraine	Ukraynalıların ele geçirdikleri Rus tanklarını internetten sattıkları iddiası
TR02	Ukraine	CNN Türk ekranında yayınlanan videonun Rusya Ukrayna çatışmasından olduğu iddiası
TR03	Ukraine	Time dergisinin Putin'i Adolf Hitler'e benzeten bir kapak hazırladığı iddiası
TR04	Ukraine	Polonya'ya kaçtı iddiasına jet yanıt
TR05	Ukraine	Toyota'nın Ukrayna'ya 5 bin araç göndererek destek olacağı iddiası
TR06	Ukraine	Ukrayna'nın gönüllü askerlere günde 2 bin dolar vereceği iddiası
TR07	Ukraine	Fotoğrafın Ukrayna güzeli Anastasiia Lenna'nın orduya katıldığını gösterdiği iddiası
TR08	Ukraine	NATO Genel Sekreteri'nin Rusya-Ukrayna çatışması için "Yüzyılın en büyük savaşı olacak" dediği iddiası
TR09	Covid	Üç doz Covid-19 aşısı olanlara HIV bulaştığı iddiası
TR10	Covid	Delmicron adlı yeni bir varyant tespit edildiği iddiası
TR11	Covid	Videonun Koca Turkovac aşısı olurken enjektörün değiştirildiğini gösterdiği iddiası
TR12	Covid	Fatih Erbakan'ın gösterdiği kuyruklu ve tüylü bebek fotoğrafları
TR13	Covid	Bill Gates'in aşılarla dünya nüfusunu azaltmayı amaçladığı iddiası
TR14	Covid	Moderna CEO'sunun bir gecede 400 milyon dolarlık hisse sattığı iddiası
TR15	Covid	Türeci ve Şahin'in fotoğraflarının yeni euro banknotlarda yer alacağı iddiası
TR16	Covid	Covid-19 aşısı olanların bir daha kan veremeyeceği iddiası
TR17	Covid	Fransız ekonomist Jacques Attali'nin 1981'de salgın kaynaklı ötanaziye desteklediğini açıkladığı iddiası
TR18	Covid	Covid-19 aşısı olanların bluetooth ekranında seri numarası gözüktüğü iddiası
TR19	Covid	Dr. Fauci'nin e-postasının koronavirüsün biyolojik silah olduğunu ifşa ettiği iddiası
TR20	Covid	Covid-19 aşısının bazı gençlerde kalp iltihaplanmasına sebep olduğu iddiası
TR21	Refugees	Suriye'de tüm suçları kapsayan genel af çıkarıldığı iddiası
TR22	Refugees	Kızılay'ın kolilerindeki malzemelerin mülteci derneğinin kolilerine aktarıldığı iddiası
TR23	Refugees	MEB'in İzmir'de mültecilere özel dört okul yapacağı iddiası
TR24	Refugees	Videonun Türkiye'nin huduttaki duvarına merdiven dayayan Afgan göçmenleri gösterdiği iddiası
TR25	Refugees	Katarlı gençlerin Türkiye'de sınavsız tıp okuyabileceği iddiası
TR26	Misc	Görüntülerin Alanya Güzelbahçe'de lazerli saldırı sonucu çıkan orman yangınına gösterdiği iddiası
TR27	Misc	A Haber'in referans verdiği raporun Türkiye'nin yangına müdahaledeki başarısını gösterdiği iddiası
TR28	Misc	Time dergisinin Türkiye'deki orman yangınlarını kapaktan haber verdiği iddiası
TR29	Misc	Hırvatistan'dan gelen uçağın THK'nin ateş kuşlarıyla aynı model olduğu iddiası
TR30	Misc	ASELSAN'ın Katar'a satıldığı iddiası
TR31	Misc	ASELSAN'ın Birleşik Arap Emirlikleri'ne satıldığı iddiası
TR32	Misc	A Haber'in zamlardan sonra "Halk her gece ucuz mazot alıyor" alt bandı kullandığı iddiası
TR33	Misc	The Simpsons dizisinin Bill Gates'in 'güneşi karartma planını' bildiği iddiası
TR34	Misc	Türkiye'nin uzay araştırmalarına toplam 20 bin TL ayırdığı iddiası
TR35	Misc	NASA'nın 24 ilahiyatçı istihdam ettiği iddiası
TR36	Misc	Melih Bulu'nun Cengiz Holding'in yeni CEO'su olduğu iddiası
TR37	Misc	THY'nin 2 bin 500 çalışanını işten çıkardığı iddiası
TR38	Misc	The Simpsons'ın Sedat Peker videolarını önceden bildiği iddiası
TR39	Misc	Videonun Bordeaux'daki tramvayları gösterdiği iddiası
TR40	Misc	Udi hindi yağının insan sağlığına olumlu etkileri olduğu iddiası

Table 3: **MiDe22-TR Event Titles.** Each event belongs to a topic, namely the Russian-Ukraine War, COVID-19, Refugees, and Miscellaneous.

MiDe22-EN						MiDe22-TR					
ID	Tweet	Retweet	Reply	Quote	Like	ID	Tweet	Retweet	Reply	Quote	Like
EN01	147	203	129	61	458	TR01	74	77	35	8	446
EN02	123	1,913	885	447	5,861	TR02	149	1,083	326	179	4,683
EN03	141	2,112	478	1,257	8,085	TR03	73	149	46	44	935
EN04	143	312	114	35	2,443	TR04	122	214	135	34	904
EN05	149	52	25	19	206	TR05	108	115	92	19	1,309
EN06	147	540	242	69	1,602	TR06	137	228	75	39	1,256
EN07	113	253	168	41	1,016	TR07	140	529	135	49	4,009
EN08	142	227	31	20	1,833	TR08	141	670	141	57	3,161
EN09	79	132	29	6	1,891	TR09	145	501	104	117	2,977
EN10	147	92	33	12	302	TR10	145	99	39	10	918
EN11	131	142	37	19	431	TR11	147	309	183	162	1,038
EN12	141	100	52	33	456	TR12	115	366	459	167	1,946
EN13	141	216	116	58	1,460	TR13	148	344	73	44	1,310
EN14	133	36	8	7	260	TR14	64	212	60	40	842
EN15	144	104	230	17	526	TR15	93	280	70	73	1,460
EN16	103	90	110	12	452	TR16	148	180	76	70	1,111
EN17	150	149	63	46	317	TR17	140	1,134	237	178	2,457
EN18	146	322	168	57	880	TR18	145	62	173	85	747
EN19	148	201	69	56	678	TR19	147	899	117	121	3,053
EN20	107	172	33	51	1,334	TR20	127	150	32	45	757
EN21	140	230	75	41	461	TR21	150	270	109	49	1,576
EN22	129	350	58	11	1,478	TR22	83	658	133	76	2,812
EN23	115	519	157	84	1,696	TR23	74	89	150	48	562
EN24	84	79	28	16	300	TR24	115	370	111	34	1,235
EN25	150	291	54	30	520	TR25	139	183	62	537	1,325
EN26	131	1,471	493	228	4,839	TR26	143	801	187	72	2,521
EN27	63	778	37	71	1,629	TR27	146	804	1,031	99	3,286
EN28	120	534	129	128	1,954	TR28	97	162	97	12	1,229
EN29	146	484	129	60	1,362	TR29	91	893	131	214	3,011
EN30	140	441	194	160	776	TR30	144	112	17	3	655
EN31	146	125	21	15	559	TR31	141	526	147	29	2,722
EN32	145	188	85	14	637	TR32	146	122	92	88	1,048
EN33	140	194	105	26	573	TR33	146	204	77	20	878
EN34	150	1,068	126	64	2,190	TR34	149	266	35	15	2,200
EN35	122	33	15	3	125	TR35	149	153	161	37	1,106
EN36	133	242	118	96	945	TR36	149	578	110	61	2,328
EN37	124	173	2	9	899	TR37	149	298	50	20	1,327
EN38	146	139	125	22	449	TR38	92	119	26	79	1,530
EN39	147	250	45	29	631	TR39	103	221	175	82	3,053
EN40	138	115	193	33	821	TR40	150	400	12	73	1,393
Avg.	132.10	376.80	130.23	86.58	1,333.38	Avg.	126.60	370.75	138.03	79.73	1,777.90
Std. dev.	± 20.63	± 473.41	± 162.53	± 204.96	± 1588.82	Std. dev.	± 27.27	± 291.28	± 167.73	± 90.54	± 1036.01
Total	5,284	15,072	5,209	3,463	53,335	Total	5,064	14,830	5,521	3,189	71,116

Table 4: Event-based user engagement statistics (the number of observations) in MiDe-22.