 doi:10.20944/preprints202206.0172.v1

Data Descriptor

MonkeyPox2022Tweets: The First Public Twitter Dataset on the 2022 MonkeyPox Outbreak

Nirmalya Thakur

Department of Electrical Engineering and Computer Science, University of Cincinnati, Cincinnati, OH 45221-0030, U.S.A.; thakurna@mail.uc.edu

Abstract: The world is currently facing an outbreak of the monkeypox virus, and confirmed cases have been reported from 28 countries. Following a recent "emergency meeting", the World Health Organization is considering whether the outbreak should be assessed as a "potential public health emergency of international concern" or PHEIC, as was done for the COVID-19 and Ebola outbreaks in the past. During this time, people from all over the world are using social media platforms, such as Twitter, for information seeking and sharing related to the outbreak, as well as for familiarizing themselves with the guidelines and protocols that are being recommended by various policy-making bodies to reduce the spread of the virus. This is resulting in the generation of tremendous amounts of Big Data related to such paradigms of social media behavior. Mining this Big Data and compiling it in the form of a dataset can serve as a data resource for a wide range of use-cases and applications such as analysis of public opinions, interests, views, perspectives, attitudes, and sentiment towards this outbreak. Therefore, this work presents MonkeyPox2022Tweets, an open-access dataset of Tweets related to the 2022 monkeypox outbreak that were posted on Twitter since the first detected case of this outbreak on May 7, 2022. The dataset is compliant with the privacy policy, developer agreement, and guidelines for content redistribution of Twitter, as well as with the FAIR principles (Findability, Accessibility, Interoperability, and Reusability) principles for scientific data management.

Dataset: https://doi.org/10.5281/zenodo.6635559

Keywords: Monkeypox; monkey pox; Twitter; Dataset; Tweets; Social Media; Big Data; Data Mining; Data Science

1. Background

Monkeypox, caused by the monkeypox virus, which belongs to the Poxviridae family, Chordopoxvirinae subfamily, and Orthopoxvirus genus [1], is a re-emerging zoonotic disease. The monkeypox virus was initially discovered in monkeys in 1958 [2], and the first case of human monkeypox was detected in the Democratic Republic of the Congo (D.R.C.) in a nine-month-old boy in 1970 [3]. The monkeypox virus is closely related to the variola virus (smallpox virus) and results in a smallpox-like disease. The symptoms usually develop within two weeks of exposure. The initial symptoms usually comprise fever, headache, muscle aches and backache, swollen lymph nodes, chills, and exhaustion. Within 1-3 days of these initial symptoms, most infected individuals develop a rash or a sore, which usually appears on the face, followed by rapid spreading to different parts of the body in a centrifugal manner [4]. There are two different clades of this virus: the Congo Basin (central African) clade and the West African clade. These two clades have estimated fatality rates of 10.6% and 3.6%, respectively [5].

The virus had been endemic in the D.R.C. and a few African countries for a very long time, and a few cases outside these geographic regions were recorded only twice –

first in 2003 [6] and then in 2018-19 [7,8]. However, at the time of writing this paper, the world is experiencing a global outbreak of the monkeypox virus, with a total of 780 confirmed cases reported in 28 countries which include – Argentina (2 cases), Canada (58 cases), Mexico (1 case), U.S.A. (19 cases), Morocco (1 case), U.A.E. (8 cases), Austria (1 case), Belgium (12 cases), Czechia (6 cases), Denmark (2 cases), Finland (2 cases), France (33 cases), Germany (57 cases), Hungary (1 case), Ireland (4 cases), Israel (2 cases), Italy (20 cases), Malta (1 case), Netherlands (31 cases), Norway (1 case), Portugal (138 cases), Slovenia (6 cases), Spain (156 cases), Sweden (4 cases), Switzerland (4 cases), United Kingdom (207 cases), Australia (3 cases) [9]. The first case of this 2022 global monkeypox outbreak was confirmed in the United Kingdom on May 7, 2022 [10]. On May 19, 2022, the first draft genome sequence of the monkeypox virus was performed by scientists in Portugal [11]. The genomic data related to this outbreak that has been studied so far indicates that this outbreak is caused by the West African clade [12].

On May 20, 2022, the World Health Organization (WHO) called an "emergency meeting" [13] to discuss the global concerns centered around the rising cases of the monkeypox virus, and currently, the WHO is considering whether the outbreak should be assessed as a "potential public health emergency of international concern" or PHEIC, as was done for the COVID-19 and Ebola outbreaks in the past [14]. On June 6, 2022, the Center for Disease Control (C.D.C.) in the United States raised its monkeypox alert to "Level 2" following the rapid increase in cases [15]. As per C.D.C., "at this time, there are no specific treatments available for monkeypox infection" [16]. However, recently, a vaccine for monkeypox has been approved by the Food and Drug Association (F.D.A.). The vaccine, previously used for smallpox, is called Imvamune or Imvanex or the Jynneos vaccine and was developed by Bavarian Nordic, a Danish biotechnology firm [17]. At the time of writing this paper, the Jynneos vaccine is the only F.D.A.-approved vaccine against monkeypox. Tecovirimat (Tpoxx), an antiviral medication, was approved in the U.S.A., Canada, and Europe for the treatment of human smallpox, and the European Medicines Agency has just approved the same for the treatment of both smallpox and monkeypox [18]. As the cases surge, countries all over the world are taking various forms of preparations, initiatives, and measures to reduce the spread of the virus. These include a lockdown in Belgium [19], the United States ordering 500,000 does of the Jynneos vaccine [20], Canada offering vaccination to high-risk groups [21], health authorities in France and Denmark wanting a vaccine rollout to adults infected by the virus [22], Germany recommending vaccinations for high-risk groups [23], and the United Kingdom advising self-isolation for everyone infected with the virus [24].

The rising cases of monkeypox and the associated recommendations, initiatives, and measures by various countries have led to the general public engaging in conversations for information seeking and sharing related to monkeypox. The Internet of Everything lifestyle of today's living is centered around people engaging in online conversations via the internet, specifically social media platforms, and spending a lot more time on the internet than ever before [25]. As a result, there has been a tremendous increase in the use of social media platforms in the recent past [26]. Conversations on social media include a wide range of topics such as recent issues, global challenges, emerging technologies, news, current events, politics, family, relationships, and career opportunities [27]. Twitter, one such social media platform, is used by people of almost all age groups from all parts of the world [28,29]. At present, there are about 450 million monthly active users on Twitter [30].

Mining of social media conversations, for instance, Tweets, to develop datasets has been of significant interest to the scientific community in the last few years, as can be seen from these recent works where relevant Tweets were mined to develop Twitter datasets on COVID-19 [31,32], 2022 war between Ukraine and Russia [33,34], European Migration Crisis [35], Inflammatory Bowel Disease [36], toxic behavior amongst adolescents [37], music [38], civil unrest [39], drug safety [40], and movies [41]. Such twitter datasets serve as a data resource for a wide range of applications and use-case scenarios related to

 doi:10.20944/preprints202206.0172.v1

studying the associated conversation paradigms as well as for investigating the patterns of the underlying information-seeking and sharing behavior.

The recent outbreak of monkeypox has also led to an increase in research and development in this field. These include - a study on the outbreaks in Europe and North America [41], an analysis of stigmatization of the LGBTQI+ community due to the outbreak [42], studying the increasing cases in England [43], research on monkeypox image data [44], investigating methods of transmission of the virus through sexual contact [45], analyzing public attitude towards monkeypox [46], and predicting the burden and duration of this outbreak [47].

However, none of these works have focused on mining Tweets about the 2022 Monkeypox outbreak. To address these limitations, this work proposes an open-access dataset of about 70,000 Tweet I.D.'s of the same number of Tweets about monkeypox that were posted on Twitter from May 7, 2022, to June 11, 2022. The earliest date was selected as May 7, 2022, as the first case of the 2022 monkeypox outbreak was recorded on this date. June 11, 2022 was the most recent date as per the time of data collection and writing of this paper.

The rest of the paper is organized as follows. Section 2 describes the dataset files. The methodology that was followed for the development of this dataset is presented in Section 3. This section also outlines how the dataset is compliant with the privacy policy, developer agreement, and guidelines for content redistribution of Twitter, as well as with the FAIR principles (Findability, Accessibility, Interoperability, and Reusability) principles for scientific data management. The conclusion and scope for future work are presented in Section 4, which is followed by references.

2. Data Description

This section presents the description of the dataset publicly available at https://doi.org/10.5281/zenodo.6635559. The dataset consists of a total of 68,934 tweet I.D.'s of tweets about monkeypox that were posted on Twitter from May 7, 2022, to June 11, 2022. The Tweet I.D.'s are presented in 4 different .txt files based on the timelines of the associated tweets. Table 1 provides the details of these dataset files.

Table 1. Description of all the files present in this dataset.

Filename	No. of Tweet IDs	Date Range of the Tweet IDs
TweetIDs_Part1.txt	19718	June 11, 2022 to June 5, 2022
TweetIDs_Part2.txt	17585	June 5, 2022 to May 27, 2022
TweetIDs_Part3.txt	17705	May 27, 2022 to May 21, 2022
TweetIDs_Part4.txt	13926	May 21, 2022 to May 7, 2022

To comply with the privacy policy, developer agreement, and guidelines for content redistribution of Twitter [48,49], only the Tweet I.D.'s associated with these 68,934 tweets are presented in this dataset. To obtain the detailed information associated with each of these tweets, such as the tweet text, user name, user I.D., timestamp, retweet count, etc., these Tweet I.D.'s need to be hydrated. There are several applications, such as the Hydrator app [50], Social Media Mining Toolkit [51], and Twarc [52], that work by complying with Twitter policies and may be used for hydrating the Tweet I.D.'s in this dataset. A step-by-step process for using one of these applications – the Hydrator app for hydrating the files in this dataset, is presented in Appendix A.

3. Methodology

This section is divided into three parts. Section 3.1 presents the specific steps that were followed for the development of this dataset. Section 3.2 outlines how this dataset is compliant with the privacy policy, developer agreement, and guidelines for content redistribution of Twitter. Section 3.3 states the compliance of this dataset with the FAIR

 doi:10.20944/preprints202206.0172.v1

principles (Findability, Accessibility, Interoperability, and Reusability) principles for scientific data management.

3.1. Steps for Dataset Development

The dataset was developed by searching tweets that comprised the keyword(s) "monkeypox" or "monkey pox", posted from May 7, 2022, to June 11, 2022. This search and the associated mining of tweets were performed as per Twitter API's standard search policies [53]. There are various tools and applications available that comply with these policies and help to search tweets based on one or more keywords. The specific tool that was used for this work is RapidMiner [54]. RapidMiner was used because of its easy-to-use integrated development environment that allows the development of a range of Big Data and Data Mining-based applications using a combination of both built-in and user-defined functionalities. These built-in functionalities are available in the form of "operators" that can be customized as well as integrated for developing a working application on the RapidMiner platform, known as a "process". The platform also allows the user to develop an "operator" from scratch and bundle the same with other built-in or user-defined "operators" to develop a "process".

For this work, RapidMiner studio, version 9.10, was downloaded and installed on a laptop with the Microsoft Windows 10 Home operating system with Intel(R) Pentium(R) Silver N5030 CPU @ 1.10GHz, 1101 Mhz, 4 Core(s), and 4 Logical Processor(s). The specific functionality that was required for this work was searching tweets based on the matching keyword(s) within a date range. This functionality is already available in RapidMiner Studio 9.10 as a built-in "operator" called the Search Twitter "operator" [55] that works by connecting with the Twitter API and by complying with the Twitter API's standard search policies for searching relevant tweets. Here, relevant tweets are defined as those tweets which contain the keyword(s) that are entered as an input to this "operator". So, a "process" was developed in RapidMiner that comprised only the Search Twitter "operator," and it was used to search tweets that contained either "monkeypox" or "monkey pox", posted on Twitter in the date range May 7, 2022, to June 11, 2022. This process was run multiple times in this date range to collect the relevant tweets.

The result of this RapidMiner "process" comprised multiple attributes – "Row no", "Id", "Created-At", "From-User", "From-User-Id", "To-User", "To-User-Id", "Language", "Source Text", "Geo-Location-Latitude", "Geo-Location-Longitude", and "Retweet Count". These refer to the row number of the results, tweet I.D. of the obtained tweet, date and time when the tweet was posted, the username of the Twitter user who posted the tweet, the user I.D. of the Twitter user who posted the tweet, Twitter username of the user whose tweet was replied to (if the tweet was a reply) in the current tweet, Twitter user I.D. of the user whose tweet was replied to (if the tweet was a reply) in the current tweet, the language of the tweet, source of the tweet to determine if the tweet was posted from an Android source, Twitter website, etc., the complete text of the tweet, including embedded URLs, geo-location (latitude) of the user posting the tweet, geo-location (longitude) of the user posting the tweet, and retweet count of the tweet. To comply with the privacy policy, developer agreement, and guidelines for content redistribution of Twitter, multiple data filters were introduced in the RapidMiner "process" to remove all the attributes from the results other than the "Id" attribute. Thereafter the results from multiple runs of this "process" were compiled to develop this dataset.

It is relevant to mention here that Twitter API's standard search does not return an exhaustive list of tweets posted within a date range. Furthermore, Twitter users are allowed to delete a tweet they have posted in the past. For a deleted tweet, there will be no retrievable tweet text and other related information upon hydration of that Tweet I.D. At the time of writing this paper, all the 68,934 Tweet I.D.'s corresponded to Tweets that were not deleted.

3.2. Compliance with Twitter Policies

 doi:10.20944/preprints202206.0172.v1

The privacy policy of Twitter [48] states – *"Twitter is public and Tweets are immediately viewable and searchable by anyone around the world"*. The guidelines for Twitter content redistribution [49] state – "*If you provide Twitter Content to third parties, including downloadable datasets or via an API, you may only distribute Tweet I.D.s, Direct Message I.D.s, and/or User I.D.s (except as described below)*. It also states - *"We also grant special permissions to academic researchers sharing Tweet I.D.s and User I.D.s for non-commercial research purposes. Academic researchers are permitted to distribute an unlimited number of Tweet I.D.s and/or User I.D.s if they are doing so on behalf of an academic institution and for the sole purpose of non-commercial research."* Therefore, it may be concluded that mining relevant tweets from Twitter to develop a dataset (comprising only Tweet I.D.'s) to share the same is in compliance with the privacy policy, developer agreement, and content redistribution guidelines of Twitter.

3.3. Compliance with FAIR

The FAIR principles for scientific data management [56] state that a dataset should have Findability, Accessibility, Interoperability, and Reusability. The dataset is findable as it has a unique and permanent DOI. The dataset is accessible online. It is interoperable due to the use of .txt files for data representation that can be downloaded, read, and analyzed across different computer systems and applications. The dataset is re-usable as the associated tweets and related information such as user I.D., user name, retweet count, etc., for all the Tweet I.D.s can be obtained by the process of hydration in compliance with Twitter policies (Appendix A), for data analysis and interpretation.

4. Conclusion and Future Scope of Work

Since the first case on May 7, 2022, the monkeypox virus has infected 780 people in 28 different countries, and the number of cases is increasing. The World Health Organization is considering whether the outbreak should be assessed as a "potential public health emergency of international concern" or PHEIC, as was done for the COVID-19 and Ebola outbreaks in the past. As a result, several countries are implementing various forms of measures, policies, and guidelines to reduce the spread of the virus. These policies and the increase in cases on a global scale have led to an increase in conversations on social media, specifically Twitter, related to information seeking and sharing about monkeypox. These conversations are leading to the generation of tremendous amounts of Big Data. Mining Twitter conversations on specific topics or viruses or global challenges to develop datasets has been of significant interest to the scientific community in the fields of Big Data, Data Mining, Natural Language Processing, and their related areas in the last few years. While there have been several Twitter datasets developed on different topics in the past, none of those Twitter datasets are a collection of tweets related to the current outbreak of monkeypox. Furthermore, none of the works related to the 2022 monkeypox outbreak have focused on the development of any such datasets thus far. To address this challenge, this work presents MonkeyPox2022Tweets, an open-access dataset of about 70,000 Tweet I.D.'s of the same number of Tweets about monkeypox that were posted on Twitter from May 7, 2022, to June 11, 2022 (the most recent date at the time of writing this paper). Future work on this project would involve updating the dataset with more recent tweets to ensure that the scientific community has access to the recent data in this regard.

Supplementary Materials: Not applicable.

Funding: This research received no external funding

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are publicly available at https://doi.org/10.5281/zenodo.6635559

Conflicts of Interest: The author declares no conflict of interest

Appendix A

The following is the step-by-step process for using the Hydrator app [50] to hydrate this dataset or, in other words, to obtain the text of the tweet, user I.D., user name, re-tweet count, language, tweet URL, source, and other public information related to all the Tweet I.D.s present in this dataset. The Hydrator app works in compliance with the policies for accessing and calling the Twitter API.

1. Download and install the desktop version of the Hydrator app from https://github.com/DocNow/hydrator/releases.
2. Click on the "Link Twitter Account" button on the Hydrator app to connect the app to an active Twitter account.
3. Click on the "Add" button to upload one of the dataset files (such as Tweet-IDs_Part4.txt). This process adds the dataset file to the Hydrator app.
4. If the file upload is successful, the Hydrator app will show the total number of Tweet I.D.'s present in the file. For instance, for the file - "TweetIDs_Part4.txt", the app would show the Number of Tweet I.D.s as 13926.
5. Provide details for the respective fields: Title, Creator, Publisher, and URL in the app, and click on "Add Dataset" to add this dataset to the app.
6. The app would automatically redirect to the "Datasets" tab. Click on the "Start" button to start hydrating the Tweet I.D.s. During the hydration process, the progress indicator would increase, indicating the number of Tweet I.D.s that have been successfully hydrated and the number of Tweet I.D.s that are pending hydration.
7. After the hydration process ends, a .jsonl file would be generated by the app that the user can choose to save on the local storage.
8. The app would also display a "CSV" button in place of the "Start" button. Clicking on this "CSV" button would generate a .csv file with detailed information about the tweets, which would include the text of the tweet, user I.D., user name, retweet count, language, tweet URL, source, and other public information related to the tweet.
9. Repeat steps 3-8 for hydrating all the files of this dataset.

References

1. McCollum, A.M.; Damon, I.K. Human Monkeypox. *Clin. Infect. Dis.* **2014**, *58*, 260–267, doi:10.1093/cid/cit703.
2. Magnus, P. von; Andersen, E.K.; Petersen, K.B.; Birch-Andersen, A. A Pox-like Disease in Cynomolgus Monkeys. *Acta Pathol. Microbiol. Scand.* **2009**, *46*, 156–176, doi:10.1111/j.1699-0463.1959.tb00328.x.
3. Breman, J.G.; Kalisa-Ruti; Steniowski, M.V.; Zanotto, E.; Gromyko, A.I.; Arita, I. Human Monkeypox, 1970-79. *Bull. World Health Organ.* **1980**, *58*, 165–182.
4. Vaughan, A.; Aarons, E.; Astbury, J.; Brooks, T.; Chand, M.; Flegg, P.; Hardman, A.; Harper, N.; Jarvis, R.; Mawdsley, S.; et al. Human-to-Human Transmission of Monkeypox Virus, United Kingdom, October 2018. *Emerg. Infect. Dis.* **2020**, *26*, 782–785, doi:10.3201/eid2604.191164.
5. Likos, A.M.; Sammons, S.A.; Olson, V.A.; Frace, A.M.; Li, Y.; Olsen-Rasmussen, M.; Davidson, W.; Galloway, R.; Khristova, M.L.; Reynolds, M.G.; et al. A Tale of Two Clades: Monkeypox Viruses. *J. Gen. Virol.* **2005**, *86*, 2661–2672, doi:10.1099/vir.0.81215-0.
6. Centers for Disease Control and Prevention (C.D.C.) Multistate Outbreak of Monkeypox--Illinois, Indiana, and Wisconsin, 2003. *MMWR Morb. Mortal. Wkly. Rep.* **2003**, *52*, 537–540.
7. Public Health England Monkeypox Case Confirmed in England Available online: https://www.gov.uk/government/news/monkeypox-case-confirmed-in-england (accessed on 12 June 2022).
8. Erez, N.; Achdout, H.; Milrot, E.; Schwartz, Y.; Wiener-Well, Y.; Paran, N.; Politi, B.; Tamir, H.; Israely, T.; Weiss, S.; et al. Diagnosis of Imported Monkeypox, Israel, 2018. *Emerg. Infect. Dis.* **2019**, *25*, 980–983, doi:10.3201/eid2505.190076.
9. Multi-Country Monkeypox Outbreak: Situation Update Available online:

https://www.who.int/emergencies/disease-outbreak-news/item/2022-DON390 (accessed on 12 June 2022).
10. Saxena, S.K.; Ansari, S.; Maurya, V.K.; Kumar, S.; Jain, A.; Paweska, J.T.; Tripathi, A.K.; Abdel-Moneim, A.S. Re-Emerging Human Monkeypox: A Major Public-Health Debacle. *J. Med. Virol.* **2022**, doi:10.1002/jmv.27902.
11. First Draft Genome Sequence of Monkeypox Virus Associated with the Suspected Multi-Country Outbreak, May 2022 (Confirmed Case in Portugal) Available online: https://virological.org/t/first-draft-genome-sequence-of-monkeypox-virus-associated-with-the-suspected-multi-country-outbreak-may-2022-confirmed-case-in-portugal/799 (accessed on 12 June 2022).
12. Mauldin, M.R.; McCollum, A.M.; Nakazawa, Y.J.; Mandra, A.; Whitehouse, E.R.; Davidson, W.; Zhao, H.; Gao, J.; Li, Y.; Doty, J.; et al. Exportation of Monkeypox Virus from the African Continent. *J. Infect. Dis.* **2022**, *225*, 1367–1376, doi:10.1093/infdis/jiaa559.
13. WHO Calls Emergency Meeting as Monkeypox Cases Top 100 in Europe Available online: https://www.reuters.com/world/europe/monkeypox-outbreak-europe-largest-ever-region-cases-cross-100-2022-05-20/ (accessed on 12 June 2022).
14. Unlikely Monkeypox Outbreak Will Lead to Pandemic, WHO Says Available online: https://www.reuters.com/business/healthcare-pharmaceuticals/unlikely-monkeypox-outbreak-will-lead-pandemic-says-who-2022-05-30/ (accessed on 12 June 2022).
15. Kelleher, S.R. C.D.C. Raises Monkeypox Travel Alert to Level 2 Available online: https://www.forbes.com/sites/suzannerowankelleher/2022/06/07/cdc-raises-monkeypox-travel-alert-to-level-2/?sh=269eee1e3f93 (accessed on 12 June 2022).
16. Treatment Available online: https://www.cdc.gov/poxvirus/monkeypox/clinicians/treatment.html (accessed on 12 June 2022).
17. Available online: https://assets.publishing.service.gov.uk/government/ (accessed on 12 June 2022).
18. SIGA Receives Approval from the F.D.A. for Intravenous (IV) Formulation of TPOXX® (Tecovirimat) Available online: https://investor.siga.com/news-releases/news-release-details/siga-receives-approval-fda-intravenous-iv-formulation-tpoxxr (accessed on 12 June 2022).
19. Gilchrist, K. Belgium Becomes First Country to Introduce Mandatory Monkeypox Quarantine as Global Cases Rise Available online: https://www.cnbc.com/2022/05/23/belgium-introduces-mandatory-monkeypox-quarantine-as-global-cases-rise.html (accessed on 12 June 2022).
20. Bahl, R. Monkeypox Vaccine: U.s. Orders 500,000 Jynneos Doses as Cases Rise Available online: https://www.healthline.com/health-news/monkeypox-vaccine-existing-vaccines-provide-strong-protection-one-fda-approved (accessed on 12 June 2022).
21. Toronto to Offer Monkeypox Vaccine Clinics Targeting High-Risk Communities. *C.B.C. News*.
22. With, A.M. Push for Targeted Monkeypox Vaccine Rollout in France, Denmark Available online: https://www.rfi.fr/en/europe/20220525-push-for-targeted-monkeypox-vaccine-rollout-in-france-denmark (accessed on 12 June 2022).
23. Monkeypox: German Panel Recommends Vaccine for Risk Groups Available online: https://www.dw.com/en/monkeypox-german-panel-recommends-vaccine-for-risk-groups/a-62084728 (accessed on 12 June 2022).
24. U.K. Health Authority Advises Self-Isolation for Monkeypox Infections Available online: https://www.thesundaily.my/world/uk-health-authority-advises-self-isolation-for-monkeypox-infections-NE9314681 (accessed on 12 June 2022).
25. Farias da Costa, V.C.; Oliveira, L.; de Souza, J. Internet of Everything (IoE) Taxonomies: A Survey and a Novel Knowledge-Based Taxonomy. *Sensors (Basel)* **2021**, *21*, 568, doi:10.3390/s21020568.
26. Auxier, B.; Anderson, M. Social Media Use in 2021 Available online: https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2021/04/PI_2021.04.07_Social-Media-Use_FINAL.pdf (accessed on 12 June 2022).
27. Kavada, A. Social Media as Conversation: A Manifesto. *Soc. Media Soc.* **2015**, *1*, 205630511558079, doi:10.1177/2056305115580793.
28. Smith, A.; Brenner, J. Twitter Use 2012 Available online: https://www.pewinternet.org/wp-content/uploads/sites/9/media/Files/Reports/2012/PIP_Twitter_Use_2012.pdf (accessed on 12 June 2022).
29. Morgan-Lopez, A.A.; Kim, A.E.; Chew, R.F.; Ruddle, P. Predicting Age Groups of Twitter Users Based on Language and Metadata Features. *PLoS One* **2017**, *12*, e0183537, doi:10.1371/journal.pone.0183537.
30. Iqbal, M. Twitter Revenue and Usage Statistics (2022) Available online: https://www.businessofapps.com/data/twitter-statistics/ (accessed on 12 June 2022).
31. Lopez, C.E.; Gallemore, C. An Augmented Multilingual Twitter Dataset for Studying the COVID-19 Infodemic. *Soc. Netw. Anal. Min.* **2021**, *11*, 102, doi:10.1007/s13278-021-00825-0.
32. Gupta, R.K.; Vishwanath, A.; Yang, Y. COVID-19 Twitter Dataset with Latent Topics, Sentiments and Emotions Attributes. *arXiv [cs.CL]* **2020**, doi:10.48550/ARXIV.2007.06954.
33. Haq, E.-U.; Tyson, G.; Lee, L.-H.; Braud, T.; Hui, P. Twitter Dataset for 2022 Russo-Ukrainian Crisis. *arXiv [cs.SI]* **2022**, doi:10.48550/ARXIV.2203.02955.
34. Shevtsov, A.; Tzagkarakis, C.; Antonakaki, D.; Pratikakis, P.; Ioannidis, S. Twitter Dataset on the Russo-Ukrainian War. *arXiv*

[cs.SI] **2022**, doi:10.48550/ARXIV.2204.08530.

35. Urchs, S.; Wendlinger, L.; Mitrovic, J.; Granitzer, M. MMoveT15: A Twitter Dataset for Extracting and Analysing Migration-Movement Data of the European Migration Crisis 2015. In Proceedings of the 2019 IEEE 28th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE); IEEE, 2019; pp. 146–149.
36. Stemmer, M.; Parmet, Y.; Ravid, G. What Are IBD Patients Talking about on Twitter? In *I.C.T. for Health, Accessibility and Wellbeing*; Springer International Publishing: Cham, 2021; pp. 206–220 ISBN 9783030942083.
37. Wijesiriwardene, T.; Inan, H.; Kursuncu, U.; Gaur, M.; Shalin, V.L.; Thirunarayan, K.; Sheth, A.; Arpinar, I.B. ALONE: A Dataset for Toxic Behavior among Adolescents on Twitter. In *Lecture Notes in Computer Science*; Springer International Publishing: Cham, 2020; pp. 427–439 ISBN 9783030609740.
38. Zangerle, E.; Pichl, M.; Gassler, W.; Specht, G. #nowplaying Music Dataset: Extracting Listening Behavior from Twitter. In Proceedings of the Proceedings of the First International Workshop on Internet-Scale Multimedia Management - WISMM '14; A.C.M. Press: New York, New York, U.S.A., 2014.
39. Sech, J.; DeLucia, A.; Buczak, A.L.; Dredze, M. Civil Unrest on Twitter (CUT): A Dataset of Tweets to Support Research on Civil Unrest. In Proceedings of the Proceedings of the Sixth Workshop on Noisy User-generated Text (W-NUT 2020); Association for Computational Linguistics: Stroudsburg, PA, U.S.A., 2020; pp. 215–221.
40. Tekumalla, R.; Banda, J.M. A Large-Scale Twitter Dataset for Drug Safety Applications Mined from Publicly Existing Resources. *arXiv [cs. I.R.]* **2020**, doi:10.48550/ARXIV.2003.13900.
41. Mahase, E. Monkeypox: What Do We Know about the Outbreaks in Europe and North America? *BMJ* **2022**, *377*, o1274, doi:10.1136/bmj.o1274.
42. Bragazzi, N.L.; Khamisy-Farah, R.; Tsigalou, C.; Mahroum, N.; Converti, M. Attaching a Stigma to the LGBTQI+ Community Should Be Avoided during the Monkeypox Epidemic. *J. Med. Virol.* **2022**, doi:10.1002/jmv.27913.
43. Mahase, E. Seven Monkeypox Cases Are Confirmed in England. *BMJ* **2022**, *377*, o1239, doi:10.1136/bmj.o1239.
44. Ahsan, M.M.; Uddin, M.R.; Luna, S.A. Monkeypox Image Data Collection. *arXiv [eess.IV]* **2022**, doi:10.48550/ARXIV.2206.01774.
45. Heskin, J.; Belfield, A.; Milne, C.; Brown, N.; Walters, Y.; Scott, C.; Bracchi, M.; Moore, L.S.; Mughal, N.; Rampling, T.; et al. Transmission of Monkeypox Virus through Sexual Contact - A Novel Route of Infection. *J. Infect.* **2022**, doi:10.1016/j.jinf.2022.05.028.
46. Yoo, J.-H. Once Bitten, Twice Shy: Our Attitude towards Monkeypox. *J. Korean Med. Sci.* **2022**, *37*, e188, doi:10.3346/jkms.2022.37.e188.
47. Bisanzio, D.; Reithinger, R. Keep Calm and Carry on: Projected Case Burden and Duration of the 2022 Monkeypox Outbreak in Non-Endemic Countries. *bioRxiv* 2022.
48. Privacy Policy Available online: https://twitter.com/en/privacy/previous/version_15 (accessed on 12 June 2022).
49. Developer Agreement and Policy Available online: https://developer.twitter.com/en/developer-terms/agreement-and-policy (accessed on 12 June 2022).
50. Hydrator: Turn Tweet I.D.s into Twitter JSON & CSV from Your Desktop! Available online: https://github.com/DocNow/hydrator (accessed on 12 June 2022).
51. Tekumalla, R.; Banda, J.M. Social Media Mining Toolkit (SMMT). *Genomics Inform.* **2020**, *18*, e16, doi:10.5808/GI.2020.18.2.e16.
52. Twarc: A Command Line Tool (and Python Library) for Archiving Twitter JSON Available online: https://github.com/DocNow/twarc (accessed on 12 June 2022).
53. Standard Search API Available online: https://developer.twitter.com/en/docs/twitter-api/v1/tweets/search/api-reference/get-search-tweets (accessed on 12 June 2022).
54. Mierswa, I.; Wurst, M.; Klinkenberg, R.; Scholz, M.; Euler, T. YALE: Rapid Prototyping for Complex Data Mining Tasks. In Proceedings of the Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '06; A.C.M. Press: New York, New York, U.S.A., 2006.
55. RapidMiner GmbH Search Twitter - RapidMiner Documentation Available online: https://docs.rapidminer.com/latest/studio/operators/data_access/applications/twitter/search_twitter.html (accessed on 12 June 2022).
56. Wilkinson, M.D.; Dumontier, M.; Aalbersberg, I.J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.-W.; da Silva Santos, L.B.; Bourne, P.E.; et al. The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Sci. Data* **2016**, *3*, 160018, doi:10.1038/sdata.2016.18.