Towards Handling Bias in Intelligence Analysis with Twitter

Alexandros Karakikes*
Intelligent Systems Lab,
Dept. of Cultural Technology and Communication
University of the Aegean
Mytilene, Greece
cti22010@aegean.gr

Panagiotis Alexiadis
Intelligent Systems Lab,
Dept. of Cultural Technology and Communication
University of the Aegean
Mytilene, Greece
alexiadisp@gmail.com

Theocharis Theocharopoulos*
Intelligent Systems Lab,
Dept. of Cultural Technology and Communication
University of the Aegean
Mytilene, Greece
theocharopoulos@aegean.gr

Nikolaos Skoulidas
Faculty of Humanities
University of the Aegean
Rhodes, Greece
psed20004@rhodes.aegean.gr

Dimitrios Spiliotopoulos
Department of Management Science and Technology,
University of the Peloponnese
Tripoli, Greece
dspiliot@uop.gr

Konstantinos Kotis*
Intelligent Systems Lab,
Dept. of Cultural Technology and Communication
University of the Aegean
Mytilene, Greece
kotis@aegean.gr

2023 IEEE 10th International Conference on Data Science and Advanced Analytics (DSAA) | 979-8-3503-4503-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/DSAA60987.2023.10302618

***Abstract*—Bias identification and mitigation in the Twitter ecosystem has been lately researched towards achieving a more efficient utilization of the application by different stakeholders and for a wide area of purposes. Among these stakeholders, intelligence services worldwide, collectively called the Intelligence Community (IC), tend to use Twitter, supplementarily to their pre-existent disciplines, for monitoring areas of interest and identifying emerging social, political and security trends/threats. Over time, the IC has identified bias as the major obstacle in information analysis, thus it has developed scientific and empirical methods for bias mitigation, in parallel to those developed by the information and communication technology (ICT) and artificial intelligence (AI) community. As it becomes apparent, it is to both communities' interest to accurately trace bias and ideally eradicate or moderate its effects. In this paper we draw systemic parallels between Intelligence Analysis (IA) and Twitter Analytics (TA), comparatively examine existing bias mitigating methodologies to pinpoint similarities/dissimilarities, and utterly investigate the feasibility of adapting and adjusting methodologies from the first field to the latter. Furthermore, we propose a novel framework for AI-augmented bias mitigation in the IC. Finally, we propose methods and tools, already adapted by the ICT community, for efficiently supporting bias mitigation methodologies adapted by the IC.**

Keywords—Twitter analytics, bias, intelligence analysis

I. INTRODUCTION

Since its inception, Twitter went through several developmental stages, to evolve to its current form, as a microblogging platform used by more than 300 million users on a global scale, posting circa 60.000 tweets per day. Its speed, accessibility to the average user and high volume of omnidirectionally flowing information, elated it from a social media network to a vast pool of information, to be collected, analyzed (predominantly via AI tools - scraping) and exploited for a wide array of purposes. Aforementioned purposes, fluctuate from marketing promotion [1] electoral [2], [3], financial [4] and other events prediction [5], to monitoring conflict zones, emerging and ongoing geopolitical tensions [6], [7].With regards to the latter field, the volume and the promptness of Twitter-derived data, combined with the collection and analysis capabilities that AI provides, has rendered it a highly efficient and exponentially attractive tool for intelligence services worldwide [8].

Bias distorts the validity of the information that are circulating through Twitter, subsequently hindering its utilization from the IC, as one of the available collection methods. Correspondingly, bias has been over the years, one of the main challenges that IC has been facing and mitigating through established practical methodologies, mainly deriving from the cognitive sciences, while employing its traditional disciplines. Presently, significant academic research has been performed towards identifying and mitigating bias in the Twitter ecosystem, predominantly through the usage of AI-powered tools, thus exploiting the processing possibilities they provide, towards examining vast datasets. These approaches address the bias issue either directly, or indirectly. Nonetheless, from an IC perspective, the majority of the aforementioned approaches seem to be fragmentary in nature, not addressing bias in all procedural steps that the IC is following, as part of its structured methodology. In addition, scarce research is devoted to exploring specifically Twitter as a potential intelligence collection and analysis platform.

The aim of the present study is a) to pursue a research study towards understanding bias in Twitter-based intelligence analysis, and b) propose an AI-augmented framework for mitigating bias in IA, based on the capabilities offered by the ICT community. Eventually, the goal is to evaluate the suitability of the proposed framework based on the needs of the IC. The research is performed under the principal postulation that all intelligence gathering activities are to be executed through Twitter, while the proposed AI toolset for supporting the framework is to be comprised of tools suitable to each discernable step of the IA cycle. To accomplish that, a number of academic research papers and studies have been systematically reviewed, to identify trends in bias mitigation methodologies. This research also includes governmental reports and other unclassified, releasable to the public, content pertaining to the IC. The related work is predominantly derived from western countries, due to the high availability stemming from their domestic laws on transparency. In contrast, governmental reports from China, Russia or other countries with similar non-disclosure policies are not included.

Additionally, at this phase of our research no complete implementation of the proposed framework is presented, which is planned for future work. The proposed framework has been partially evaluated with a set of tools and a social dataset, as presented in section V.

The structure of this paper is as follows: in Section II preliminaries on the topic are presented, while in Section III related work is summarized. In Section IV, the proposed framework for bias mitigation in IA is presented, while in Section V early experimentation and results obtained for a different domain is presented. Section VI concludes the paper.

II. Preliminaries

A The Intelligence Cycle

In IC's attempt to effectively provide timely and precise analysis and prediction products for policy/decision makers, it was a paramount priority to systemize its activities. Therefore, a comprehensive procedure of distinctly defined steps was instituted, called the intelligence cycle (called Cycle in the rest of the paper). The Cycle has been adopted, integrally or in differentiations, by the majority of the IC, and interestingly by private sector intelligence/risk assessment corporate entities. Three indicative representations of the Cycle are depicted in Fig. 1 (Fig. 1(a):Federal Bureau of Investigations (left) [9]; US Department of Defense (right) [10] and Fig.1(b): (Director of National Intelligence - US) [11]).

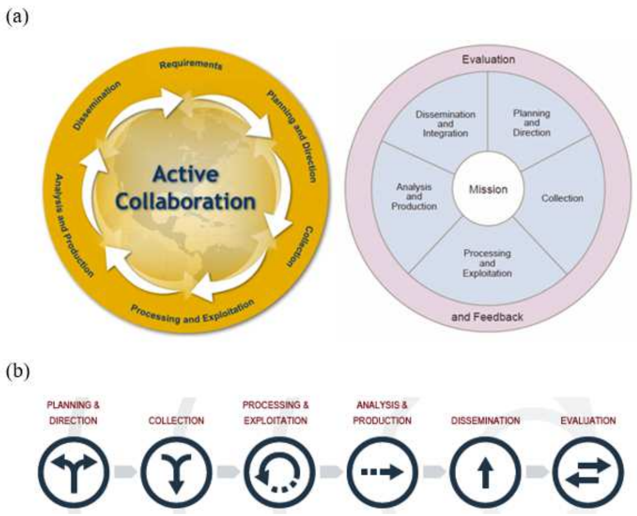

Fig. 1 The Intelligence Cycle

The core phases, that are commonly present in different versions of the Cycle, are: 'Planning & Direction", "Collection", "Processing and Exploitation", "Analysis & Production" and "Dissemination". Thus, for the purpose of the current study, these phases will be further examined for the potential presence of bias phenomena, and for their mitigation through proposed AI tools.

a. Planning & Direction

Being the initial phase of the Cycle, it is decisive for the successful achievement of any intelligence program. At this phase, the decision-making authority, is responsible for stating clearly, to the intelligence team, the information that is required. Accurate direction of the intelligence effort is the task of senior management of the organization, who instructs the intelligence team on the questions required to be answered.

b. Collection

Upon reception of a precisely defined requirement, the intelligence team leverages what relevant information is already available and what are the informational gaps to be filled from available collection resources. Raw information is gathered from sources, which may be electronic, human, open-source media, or other means, leading to the categorization of intelligence in the following disciplines:

- Human Intelligence (HUMINT): information collected from human sources, openly by interviewing sources that willingly collaborate with the IC or conducted through clandestine or covert means (espionage).
- Signals Intelligence (SIGINT): electronic transmissions that can be collected by ground, surface, aerial or space (satellites) platforms. Communications Intelligence (COMINT) is a sub-type of SIGINT and refers to the interception of communications between two parties.
- Imagery Intelligence (IMINT) obtained from the electro-optical and infrared sensors and from synthetic aperture radars (SAR), possibly capable of detecting moving targets.
- Geospatial Intelligence (GEOINT): exploitation of imagery and geospatial information to describe, assess, and visually depict security related activities on the earth. It is produced through an integration of imagery, IMINT, and geospatial information.
- Measurement and Signatures Intelligence (MASINT): scientific and technical information obtained by quantitative and qualitative analysis of data (metric, angle, spatial, wavelength, time dependence, modulation, plasma, and hydromagnetic) derived from specific technical sensors.
- Open-Source Intelligence (OSINT): a broad array of information and sources that are generally available. OSINT sources can be further divided into different sub-categories, including Internet, online publications, blogs, discussion groups, citizen media (i.e. – cell phone videos, and user created content), YouTube, and other social media websites. Social media derived information tends to outperform other traditional disciplines, due to its timeliness and ease of access, a fact that led to the introduction of the term Social Media Intelligence (SMI or SOCMINT). The term is currently semi-officially adopted by the academic community and a few members of the IC, most notably India's Law Enforcement sector [12]. Nonetheless, its importance as a sub-discipline is undisputed and characteristically denoted in "US National Intelligence Strategy 2019" [11].

c. Proccessing and Exploitation

This phase, in combination with the succeeding "Analysis" phase, constitute the conversion of the raw information into completed intelligence products, in a comprehensible format, containing current situation and potential outcomes assessments. The processor screens the data, considering the information's reliability, validity, and relevance. Data are screened and selected according to their relevance, while at the same time a dual evaluation is performed for source reliability and information credibility. This dual evaluation process, has been applied by intelligence practitioners at least since World War II and mainly in the HUMINT discipline, until presently, in the following form:

- Source Reliability: ratings range from “Reliable” (A) to “Unreliable” (E) as shown in Table I [13]. In every instance the rating is based on previous reporting from that source. If there has been no previous reporting, the source must be rated as “F”. An “F” rating does not necessarily mean that the source cannot be trusted, but that there is no reporting history and therefore no basis for making a determination.

TABLE I. EVALUATION OF SOURCE RELIABILITY.

A	Reliable	No doubt of authenticity, trustworthiness, or competency; has a history of complete reliability
B	Usually Reliable	Minor doubt about authenticity, trustworthiness, or competency; has a history of valid information most of the time
C	Fairly Reliable	Doubt of authenticity, trustworthiness, or competency but has provided valid information in the past
D	Not Usually Reliable	Significant doubt about authenticity, trustworthiness, or competency but has provided valid information
E	Unreliable	Lacking in authenticity, trustworthiness, and competency; history of invalid information n in the past
F	Cannot Be Judged	No basis exists for evaluating the reliability of the source

- Information Content Credibility: The highest degree of confidence in reported information is given to that which has been confirmed by external sources. Table II [13] shows evaluation of information content. The degree of confidence decreases if the information is not confirmed, and/or does not seem to make sense. The lowest evaluated rating of “5” means that the information is false. A ranking of “6” does not mean untruthful information, instead specifies that no determination can be made since the information is completely new.

TABLE II. EVALUATION OF INFORMATION CONTENT

1	Confirmed	Confirmed by other independent sources; logical in itself; Consistent with other info on the subject
2	Probably True	Not confirmed; logical in itself; consistent with other information on the subject
3	Possibly True	Not confirmed; reasonably logical in itself; agrees with some other information on the subject
4	Doubtfully True	Not confirmed; possible but not logical; no other information on the subject
5	Improbable	Not confirmed; not logical in itself; contradicted by other information on the subject
6	Cannot Be Judged	No basis exists for evaluating the validity of the information

Combining values of Table I and Table II, for instance, a confirmed information from a reliable source is rated A1, an unknown-validity information from a new source without reputation is F6, an inconsistent illogical information from a known unreliable source is E5, a confirmed information from a moderately doubtful source is C1.

d. Analysis

In this phase, the intelligence analysts employ structured analytic techniques (SATs), to transform processed data and information into a fused, complete intelligence product, contextualized, and in a comprehensible to the decision makers’ format. The IC has identified bias as the major challenge in this phase, thus its mitigation has been standing in the core of the SATs development. The key components of this phase are relevance, accuracy, and completeness in satisfying the original requirement.

e. Dissemination and Feedback

Finally, in this phase, the processed information is collated into reports or other forms of communications and distributed to consumers, which may be either decision or policymakers.

As a closed loop system, the IC ends when the originator of the request provides feedback as to the value of the product. Feedback can be provided via dialogue in a ticketing system, email correspondence, phone call, video conferencing, or an in-person meeting.

B. Bias in TA and the Cycle

a. General Bias

Bias emerges as a major challenge to be encountered by the academic community and every other entity which intend to use Twitter as an information collection pool and apply AI methods for analyzing this information. In socio-psychological terms, bias is a disproportionate weight in support of, or in opposition to, an idea or thing, typically in a way that is narrow-minded, prejudicial, or unfair. Its main categories and subcategories comprise of:

- Cultural Biases: age, ethnicity, family roles and connections, education (level and specific discipline), gender, language, nationality, political participation and affiliations, profession, region, religion, social network connections, socio economic position and origins [14].
- Organizational Biases: parochialism, group thinking, excessive compartmentalization [15].
- Idiosyncratic Biases: personal experiences.
- Cognitive Biases: availability bias (beliefs on a topic are formed by whatever information is most easily accessible), framing effects (decisions are influenced by the way information are worded/presented), anchoring bias, confirmation bias (tendency to interpret information in a way that confirms their existent beliefs), bandwagon effects, reinforcement effect, exposure effects, ambiguity effects, less is more effects, decoy effects, Dunning-Kruger effect, priming effects, order effect, peak end rule [16],[17].

b. Bias in TA

Bias in Twitter analysis is a dual challenge: on a data source level (user posting personal opinions or perceptions of events – not necessarily objective) and on an algorithmic level, as it has been observed that algorithms themselves (designed for Twitter analytics) are presenting noticeable indications of bias. This can be explained by the fact that AI soft. engineers are subconsciously instilling their biases to the algorithms, thus perpetuating a problem, which AI is supposed to mitigate. Being admittedly inherent to every human being, bias is unavoidably conveyed to tweets content, Twitter content ranking algorithm [18], as well as the algorithms designed to analyze them, correspondingly. Interestingly, Twitter company itself has admitted the existence of bias in its own algorithms, favoring specific political ideologies over opposing ones [19].

c. Bias mitigation in IA

The IC has already conducted significant research on bias since 1964 [17], recognizing it as the predominant hindrance to objective analysis and the cause for many failures [20]. To this direction, various structured methodologies for mitigating it have been developed and efficiently applied.

Primarily, the Cycle itself was designed to utterly support decision makers with, as impartial as possible, intelligence products, thus it is to be considered a complete bias mitigation architecture. In the overall process, analysis phase bears an increased gravity, and a miscellany of SATs are applied during it (either individually or in combination), towards bias mitigation and objective analysis production, that synoptically are [21]:

- *Diagnostic Techniques:* aiming at exposing assumptions made by the analysts, analytic arguments, or intelligence gaps. Diagnostic Techniques comprise of Key Assumptions Check, Quality of Information Check, Indicators or Signposts of Change and Analysis of Competing Hypotheses (ACH).
- *Contrarian Techniques:* aim to openly pose challenges to prevalent analytic scenarios, potential outcomes and overall current thinking, Subtypes of Contrarian Techniques are Devil's Advocacy, Team A/Team B, High-Impact/Low-Probability and "What If?" Analysis.
- *Imaginative Thinking Techniques:* aim at creating new insights, diversified perspectives and/or suggest alternate outcomes. Brainstorming, Outside-In Thinking, Red Team Analysis and Alternative Futures Analysis, fall under the above-mentioned category of SATs.

III. RELATED WORK

To achieve a more holistic analysis of the proposed research problem, a three-axis thematic review of pertinent literature, is performed. Relevant studies are classified in three categories i.e., "Ontologies and the Cycle", "Indirect Bias Mitigation" and "Direct Bias Mitigation".

A. Ontologies and the Cycle

Studies categorized in the present subsection exhibit the vital role that ontologies can perform in various phases of the Cycle.

In [22], one of the first attempts to introduce the concept of ontologies in various phases of the Cycle is presented, advocating that this would enhance communication among them, especially between the phases of collection and analysis. Additionally, the paper states that an ontology would also assist in consolidating the benefits of emerging technological advances, indirectly referring to AI capabilities.

In [23], the establishment of an ontology-based Integrating Semantic Framework (ISF) is proposed, utilizing the Basic Formal Ontology (BFO) as a basis. BFO is a framework for defining and categorizing entities and concepts that can be applied across different intelligence fields and disciplines. By using BFO, authors contend that intelligence analysts can establish a common understanding and language for discussing intelligence-related concepts, which can improve communication and collaboration among analysts from different agencies and organizations. Additionally, BFO's use of ontological realism as a philosophical foundation can help ensure that the ontology mirrors the structure of the world, which is essential for effective intelligence collection and analysis.

In [24], authors propose the semantic enhancement of data using a formal ontology to improve the accuracy and consistency of intelligence analysis across various domains. To that end, a tri-layered ontology is proposed, consisting of: An individual, small, domain-neutral Upper-level Ontology (ULO), for which in their study BFO is selected; Mid-level Ontologies (MLOs), created by grouping terms pertaining to specific aspects of warfare, or to specific tasks (e.g. inter-agency information sharing) ; and Low-level Ontologies (LLOs) focusing on specific domains, whose main advantage is they might be used as starting points for the development of cross-domain ontologies. Overall, the paper demonstrates the potential of ontology-based approaches to enhance the quality and effectiveness of intelligence analysis.

B. Indirect Bias Mitigation

Studies focusing on establishing evaluation mechanisms for the source of information and the information corpus itself are examined in the current subsection. Ranking the reliability of each source and the information content, by comparing it with already validated sources/information, only the input/output phases of the process are examined, thus bypassing bias in the intermediate process.

In [25], authors propose a generic framework to automatically and in real-time execute credibility analysis of posted messages in social media platforms, including Twitter, which is used as reference. The framework embodies a credibility model based on three aspects: Text credibility (text analysis-based), User credibility (based on attributes about the user's account, such as creation date, verified account), and social credibility (based on attributes that signify social impact, such as followers and following). A fourth aspect, Topic-level credibility, is under consideration by the authors to be added in the future. The proposed future aspect, based on natural language processing (NLP) and sentiment analysis techniques, will measure the level of acceptance of the topic or event referenced in the text.

In [26], authors propose a mainly NLP-based approach, to examine the credibility of Twitter user accounts. To accomplish that, they implement a four-stage linear methodology comprising of: News category analysis stage, Sentiment analysis stage, Source credibility analysis stage and Source credibility visualization stage. At the News category analysis stage, machine learning (ML) algorithms are used, such as Naïve Bayes classifier, Decision Tree and Support Vector Machine. During the Sentiment analysis component, lexicon-based methods and ML-based methods are applied. At the Credibility analysis component, source (Twitter account) credibility evaluation is proposed, by examining 12 features of each account, through the use of K-means clustering method. Finally, the Source credibility visualization component is using a method to depict the impact each Twitter account's post has on its followers (agreement or disagreement). After evaluating their model, authors conclude that it is able to achieve an average accuracy of 68% in assessing a Twitter account's credibility. Interestingly, they compare their study with other related work [27], which is achieving a noticeably higher accuracy, with more inclusion of human factor in the process.

In [27], authors initially use Elasticsearch to access Twitter stream API and gather a dataset of 1.206 tweets on a certain subject. Subsequently, they evaluate the dataset comparatively with two different methods, with the goal to assign to each tweet one of four credibility levels (HC – highly credible, HNC – highly non credible, N – neutral, C – controversial). The first evaluation method employs an ML algorithm in R language achieving an accuracy of 51%-57%, while the second involves two-human agents' team (critics) who applied a set of manually assigned rules and achieved an accuracy of 82%-89%. Finally, by merging both methods, the achieved results were enhanced, raising the precision in the HC and HNC credibility classes by 8%–10% and decreasing low controversial (C) class error by 0,3%.

In [28], authors propose an approach called 'Reliability Index for Twitter', which ascribes every Twitter profile with a numeric value, towards determining a profile's authenticity. This numeric value fluctuates between 0 and 1, where 0 indicates an extremely unreliable and highly likely fake profile while 1 represents a reliable and possibly genuine, human operated account. In this research, 20 factors are taken into account in order to calculate a reliability measure for Twitter accounts. Those parameters have been linked with their corresponding desired values as well as the respective weightage. This study, as summarized in Table III has been exclusively directed in investigating Twitter profile's authenticity, and can be valuable in averting disinformation attempts, carried out either by human operated fake profiles or bots. Supplementary proposals in examining the veracity of the content are to be investigated in other relevant studies.

TABLE III. TWITTER RELIABILITY INDEX FORMULA WITH PARAMETERS

Attributes	Desired Condition	Weight	Scaled Weight
Verified Status	Yes	10	0.03
Followers	>20	9	0.03
Following	>20	5	0.02
Tweets	>30	10	0.03
Media Tweets	>5	16	0.05
Likes	>15	17	0.06
Liked Tweets	>5	17	0.06
Retweets	>25	6	0.02
Retweeted	>10	18	0.06
Account age	>2	7	0.01
Mobile Number Verified	Yes	21	0.07
Email Address Verified	Yes	12	0.03
Ratio of Followers and Followed	>0.8	35	0.12
Hashtags Used	Yes	6	0.02
Bio Added	Yes	12	0.04
Mobile Application attached	Yes	16	0.05
Two Factor Authentication	Yes	21	0.07
Location Tweets	Yes	16	0.05
Contacts Uploaded	Yes	13	0.04
Reported Fake	No	33	0.11
TOTAL		300	1

In [29], authors propose a methodology to automate military intelligence confidence assessments for Twitter messages, predominantly based on the TunkRank algorithm [30]. To that end they initially identify a dataset of Tweets on a specific subject (the reported death of photojournalist Tim Hetherington in Misrata, Libya) by using the Twitter Search API. Then, they classify the content into Tweets expressing that the person is alive, or dead, or neither. Following, they gauge the TunkRank of users and verify the reliability of the user's message (i.e., check for false retweets). Then they determine the independence of sources by observing for explicit and implicit retweets, shared URLs, and distance between users in the Twitter graph.

C. *Direct Bias Mitigation*

Garimella et al. [31], suggest a content-based recommendation approach as alternative to Twitter's built-in recommendation, to mitigate confirmation bias and to assist towards increasing average Twitter user's exposure to conflicting views and beliefs, independently of Twitter user's viewpoint on a specific issue. Their main concept is to merge recommendations from biased groups of Twitter profiles, e.g., pro-Trump and contra-Trump, into a hybrid set that contains tweets from both sides. To implement their approach, they process a dataset of 73.868 tweets from 39.698 profiles and use Apache Solr's MoreLikeThis functionality. For evaluation, they utilized the beyond-accuracy metrics of recommender systems, i.e., diversity and serendipity, concluding that their approach may improve both metrics.

Authors in [32] propose a method for analyzing bias in tweets using NLP techniques, which consists of three steps: data collection, feature extraction, and bias analysis. In the data collection step, authors use the Twitter API to collect tweets related to a particular topic, which are then pre-processed to eliminate any irrelevant information. In the feature extraction step, the authors use NLP to extract features from the tweets, (including sentiment, emotion, and topic), while simultaneously using a sentiment lexicon to identify the sentiment of the tweets. In the bias analysis step, authors use the extracted features to analyze the bias in the tweets and additionally use a supervised ML algorithm to classify the tweets as biased or non-biased. The authors also use a sentiment analysis algorithm to identify the sentiment of the tweets. They test the proposed method using a 2020 US presidential election-related tweets dataset, with the ensuing results showing that the proposed method is able to correctly identify biased tweets. The limitations of their method, according to the authors, include inadequate ability to detect contextual bias, non-textual bias (e.g. video, image), and elusive bias (sarcasm).

In [33], authors present a framework for mitigating biases in ML systems, based on conditional Generative Adversarial Networks (cGANs) which allows the generation of new high-quality synthetic data related to the targeted population groups (minorities). Concurrently, they argue that the main hindrance of the rest mitigation approaches, is that they are model-oriented, thus emphasizing on adjusting the training algorithms to generate fair results, while omitting the fact that the training data can itself be the foremost reason for biased results. The suggested framework enables ML systems engineers to estimate the actual distribution of the original data pertaining to the targeted population groups (population groups that are victims of biases) through establishing a termed "two-player game". The "players" are two models, which are trained in parallel, i.e., the Discriminator (Dis) and the Generator (Gen). Gen is trained to capture the data distribution through trying to maximize the probability of Dis committing a mistake. Dis is trained to maximize the probability that a data sample came from a targeted population group rather than the Gen. Simultaneous training of models is recurrently performed until a generative model that can generate new synthetic data pertaining to the targeted population groups is obtained. The resulting generative model is then used to synthetically produce new data, which are used to augment the training set so as to compensate and overcome the bias problem. In this way, ML algorithms can be trained on these data in order to produce unbiased predictions. Experimental results indicated that the proposed framework efficiently mitigated the biases against targeted population groups, while at the same time enhancing the prediction accuracy of the ML classifiers.

Despite not explicitly stated by the authors, GANs' methods are presently preferred in facial recognition/image analysis related projects. GAN-based methods have been effectively implemented in image synthesis to produce remarkably realistic faces, and a variety of other image types, which are beneficial to training data augmentation and the corresponding recognition projects. [34]. It should be noted though, that within these margins, GANs are also widely selected for satellite imagery related academic researches [35], a fact that may elevate them to a useful tool for GEOINT-IMINT analysis.

IV. The Proposed Framework

Instances of bias can be manifested at each individual phase of the Cycle, either overtly or subconsciously. Thus, it is considered essential to dissect the Cycle, identify the respective manifestations of bias that might be encountered at each individual phase, and investigate the available AI-powered mitigation tools. Upon completing the approach, the abstract design of an integrated bias mitigation framework for IA will be presented.

A. Bias in Planning & Direction

Bias can be introduced by prejudiced framing of research questions, due to decision makers' own biases, which in latter phase may influence the selection of sources or the analysis performed by analysts (confirmation bias).

NLP tools are proposed, for an initial review of an area-of-interest's social media, to recognize prevalent trends, issues, and events pertinent to the intelligence requirement. Subsequently, decision makers are briefed on the findings, to direct their requirements in a more objective way, to all areas of interest (e.g., if an antagonistic country is a traditional military power, decision makers might direct their requirements principally on this sector, thus neglecting the social/political/economic sectors, where potential underlying weaknesses might be missed).

Additionally, it is essential for the organization to engineer an ontology, which will form the linguistic and semantic basis of all internally exchanged information and knowledge, not only to avoid misinterpretations/ambiguities, but to be able to present information concisely, in a sentimentally neutral language, which will not instigate further bias.

B. Bias in the Collection

In this phase, availability bias may manifest. For instance, algorithmically favored twitter profiles of key figures of an antagonistic country, are more easily available and accessible to domestic IC, leading to overrepresentation of the views they're propagating or the information they're broadcasting in latter stages of the Cycle. Therefore, at this phase it is considered rather advantageous to provide IC the capability of monitoring even algorithmically unfavored, but yet influential profiles, thus collecting, ideally, as-close-as-possible to the entirety of available information, and opposing views, while at the same time overriding limitations set by Twitter API [37]. To this direction, a set of proposed open-source twitter scrapers include Twint [38], Twitter-scraper [39], Twitterscraper [40], TIGMINT [41] (Twitter, Instagram, GeoTagging Media Intelligence, available also in web version [42] supported by Gurugram Police Cyber Security Internship, India), Snscrape [43] (capability to scrape a multitude of social media platforms including Twitter, Instagram, Facebook, Russian VKontakte and Chinese Weibo) and Tweeds [44].

Finally, bots and virtual assistants may be employed to interact with Twitter users, to gather information of interest. Such tools may be modeled similarly to the e-Enemy (e-Voroh) Telegram chatbot, launched on 10th March 2023 by the Ministry of Digital Transformation of Ukraine, which is successfully implemented during the ongoing Russo-Ukrainian conflict. The e-Voroh provides users the possibility to send geolocation, photos, and videos of enemy equipment with the additional possibility to describe in free text what they saw, thus contributing significantly the intelligence collection effort of their intelligence services. Another feature added to e-Voroh is reporting war crimes by occupiers with the opportunity to send photos or videos of war criminals and provide their data. A most recent feature added to the bot is reporting the location of explosive objects thus, the citizens of Ukraine are able to assist friendly forces in the disposal of mines, projectiles, and bombs left by invading forces. Nonetheless, Ukraine had to mitigate disinformation, as verification of intelligence took a long time through the initial version and frequently Russian bots would flood deceptive messages onto the platform. This challenge was effectively solved by requiring users to login and authenticate themselves via the country's e-passport system, Diia. Diia's main function, prior the conflict, was that of a centralized hub for citizens' transactions with their government, allowing user to access their identity cards, fulfill taxational obligations and obtain government services[45].

C. *Bias in Processing*

Bias can be introduced at this stage through the selection of which data to include or exclude, or the interpretation of data.

NLP tools can be used to automatically categorize the assembled data to recognize pertinent information and pre-process data. More specifically, text classification/mining and sentiment analysis tools are proposed. Text classification automatically categorizes text data into specific thematic classes, a valuable task in identifying new, relevant information, sources (influential twitter profiles or ascending twitter profiles), or tracking the activities of specific individuals or groups, especially when the intelligence objective is related to counterterrorism or counter-disinformation field. Text mining techniques such as keyword extraction and named entity recognition can help recognize key concepts and entities that pertain to a certain intelligence objective. It should be emphasized, that findings from this task, should be forwarded back to the collection phase, to be further exploited by the collectors (e.g., exploiting new sources, direct collect efforts to new trends etc.).

NLTK (Natural Language Toolkit) [46], Stanford CoreNLP [47] (supports at the moment 8 languages: Arabic, Chinese, English, French, German, Hungarian, Italian, and Spanish), TweetNLP [48] and GATE NLP [49] (General Architecture for Text Engineering, supports 11 languages), are a few open-source NLP libraries that can also be tested, for suitability in the processing phase.

Additionally, source/information evaluation, open-source frameworks, such as [25] are proposed for this phase. This methodological step contributes to indirect bias mitigation as mentioned in Section II. In addition, all scenarios, even seemingly improbable ones, are retained within the procedural loop, to be evaluated by analysts in a subsequent stage. This feature might prove to be exceptionally useful, particularly considering that IC's history is full of failures stemming from exclusion of "low probability information/scenarios", most tragic of which the 9/11 attacks [50].

D. *Bias in Analysis*

Bias can also be present in the analysis of data, such as through the selection of analytical methods or the interpretation of results. Analysts may subconsciously apply their personal assumptions or prejudices, to the analysis, converging on hypotheses that confirm their predeterminations. To mitigate bias in this phase it is deemed necessary to select which analytic technique is to be implemented in order to opt for the appropriate AI tool. Our research, will focus into the ACH SAT, as it holds a highly prominent position in the IC for its efficiency. ACH was developed for use at the United States Central Intelligence Agency (CIA). By examining the eight distinct steps of ACH as described in [51], the use of supervised ML is proposed for:

- Hypothesis generation: ML algorithms may be utilized to generate multiple, heterogeneous hypotheses established upon available data and preceding knowledge. These algorithms can be trained on historical information (e.g., rival countries' historical bilateral relations) to distinguish patterns and generate new hypotheses accordingly.
- Hypothesis testing: Supervised ML algorithms may additionally be utilized to test competing hypotheses against available evidence. ML algorithms may also be trained on historical information to identify which hypotheses are most consistent with the available evidence and generate corresponding probability estimates.

Open-synthesis [52], is an open-source Python-based framework, which according to its engineers is designed to support the ACH framework. The abovementioned framework is proposed to be tested for supporting IC's objectives in the analysis phase.

Another approach proposed is the use of algorithm adversarial training, for the implementation of SATs that fall within the contrarian techniques or imaginative thinking categories. Contrarian techniques involve challenging assumptions and exploring differentiated interpretations, whereas in a similar manner adversarial training may ensure ML models are not excessively dependent on a single, biased perception. Adversarial training entails producing adversarial examples that are designed to prompt the model to erroneous predictions - and using these examples to train the model to be more effective. This may assist in identifying and rectifying biases in the model.

In addition, adversarial training, may be examined as a potential supportive tool for imaginative thinking technique. By generating adversarial examples (alternative scenarios), models can be trained to manage an array of hypothetical situations, and to produce more precise understanding of their potential outcomes.

E. *Bias in Dissemination*

Bias may appear in the dissemination of intelligence products, through the selection of which information to emphasize or the wording of the report. This can affect how decision-makers perceive the intelligence and how they act on it. In addition, organizational bias among various sub-groups of the intelligence service (e.g., collectors vs analysts, planners vs analysts etc.) or IC members (law enforcement, military intelligence, civilian intelligence etc.) can influence the decision to whom the intelligence products will be disseminated. Support in this phase may be offered by AI-powered tools that can be used to automatically disseminate intelligence reports to relevant stakeholders, such as government agencies, military organizations, and law

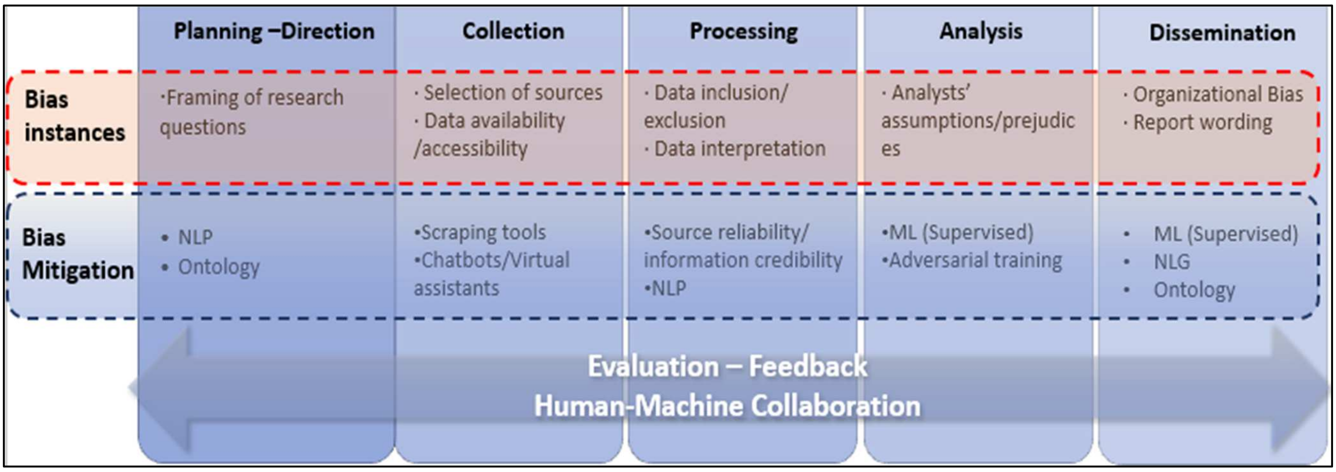

Fig. 2 The proposed AI-augmented bias mitigation framework of Cycle

enforcement agencies. Natural language generation (NLG) tools can be used to generate reports, summaries, and warnings, especially when they pertain to topics of repetitive nature (e.g., Daily Situation Report on a given geographical area, with standard recipients), to keep stakeholders informed about forthcoming threats and trends in a timely manner. To this end, the use and customization of open-source NLG tools like [53], is proposed.

Across all phases, but especially in the Planning & Direction and the Dissemination, it is crucial to maintain an established fair ontology. As stated above, this will be the standardized vocabulary used in all information circulating within the agency, and among collaborating agencies (Intelligence Services of a nation or national Intelligence Services of an Alliance's members), to avoid misinterpretation, information overlapping and time consumption. As emphasized in [54] though, it is of the outmost criticality to engineer a fair ontology, since human or data bias may be encoded in engineered ontologies. Consequently, an unfair ontology may perpetuate bias in IA, instead of mitigating it, thus creating a paradox.

As a conclusive remark it should be stated that, among all phases of the Cycle, a continuous evaluation and feedback has to be performed and monitored, while at the same time the highest possible degree of human-machine collaboration is to be pursued.

The abstract architectural design of the proposed framework is presented in Fig. 2

V. EXPERIMENTATION AND PRELIMINARY RESULTS

In our early experimentation with related technology applied in another domain, we have obtained results from capturing the polarization of Twitter users on recent social issues related to illegal immigration at EU borders (push back of illegal immigrants by authorities) and sexual offence victims (several cases made public).

Towards this aim, two popular hashtags were studied, #metoo and #pushback, which concerned public opinion to a great extent, often extensively covered by the media. The #metoo refers to the phenomenon of harassment, rape and abuse of women both in the workplace and in their social life. The #pushback refers to pushing back immigrants and refugees without examining their individual situation. Through social media, users expressed their thoughts and opinions on these issues, either negatively, neutrally or positively.

For the needs of the research, posts in Greek and English were extracted from Twitter covering the three-year period from 2020 to 2022. Additionally, tweets were selected, that either their creator has declared Greece as their place of residence or the tweets were published from GREECE. In order to be able to mine data produced over such an extended period of time, open-source scraping tools were chosen that make use of Twitter's advanced search endpoint. Usage of aforementioned tools was necessary, since Twitter's API limitations do not allow data set mining with the specific temporal filters. Specifically, Twint [38] and Snscrape [43] were tested, with the latter being preferred due to more consistent and trustworthy results. Before analyzing the sentiment of the posts, word clouds were created that reveal the topics of the discussion.

It became evident very early that in the case of the #metoo hashtag the tweets were focused on the Greek society current affairs, with two dominating well-known cases, as well as numerous complaints from the field of artists. Users expressed support for the complainants calling for strict punishment of the alleged perpetrators. On the other hand, however, there were users who disagreed with the public humiliation of the accused, before final justice verdicts on the accusations, while emphasizing the absence of the #metoo movement in cases of rape where the perpetrators were immigrants, members of minorities (e.g., Roma) or leaning to a specific political ideology.

Similarly, the hashtag #pushback brought together posts of support for people trying to enter Greece illegally, highlighting the need for solidarity and support. Cases of shipwrecks were vividly presented and criticisms of countries that follow the practice of push-backs were published. The opposite pole raised issues of social and national security, characterized the role of NGOs as suspicious and accused those involved of having created a "factory" for personal financial profit, in cooperation with Turkish traffickers. Then an attempt was made to apply sentiment analysis as originally designed with the Textblob[1] tool and alternatively with

[1] https://textblob.readthedocs.io/en/dev/

spaCyTextblob[2] which were unsuccessful with regard to the Greek data set. For this reason, the open-source software Orange3[3], was used. The pre-processing of the data was followed by the sentiment analysis, which for the Greek data set was done with the NLTK corpus and for the English with VADER[4]. Additionally, likes and reposts of the posts were measured as an indication of polarization, since in this case the users show that they agree with the content. Polarization on both hashtags was relatively balanced with negative tweets being on par with positive ones. If neutrals were to be added to the latter, creating the non-negative group, then the polarization would acquire a clearly more positive connotation. The number of likes and retweets per sentiment varied, depending on the case.

Early experimentation and preliminary results obtained (presented in a master thesis of our group, in detail **Error! Reference source not found.Error! Bookmark not defined.**), are important since the tested AI-tools can be examined for applicability in certain phases of the proposed framework (regardless the domain of application). Specifically, tested scraping tools are assessed as suitable for the collection phase, as they will offer the IC community the advantages analyzed in paragraph B of section IV, thus assist to overcome availability bias. Consequently, sentiment analysis may be performed in the processing phase of the framework by using the respective tools used in the experiments, in order to determine polarity of the collected messages, towards identifying potential biases. Determining Twitter messages' polarity through sentiment analysis, can provide useful indicative leads into potential bias: by identifying extreme views on certain topics, thus setting the spectrum within which opinions will fluctuate; identifying echo chambers, where similar-minded, biased opinions are fortified and opposing opinions are censored; and finally revealing specific lingual choices and rhetoric, linked to polarized opinions.

Despite offering strong indications of prejudiced opinions, it is important to point that polarity determination does not constitute undoubtful evidence of bias, and thus might not suffice to entirely identify bias in a comprehensive manner, due to a number of limitations: insufficient contextual comprehension, as analyzing polarization requires understanding the broader context of the social issue, the history, and the motivations of the users involved; algorithmic biases, since, as above stated, algorithms can have their own biases, influencing the accuracy of the analysis; dissimilar interpretations, as polarization doesn't automatically equate to bias. Users might express strong opinions, as a result of sincere disagreements or genuine concerns.

VI. CONCLUSION

Bias mitigation in TA has attracted significant interest from the academia. Given the position Twitter holds in our daily communication, but also at more theme-specific opinion expressions (politics, economy, social issues etc.), bias is going to retain its place as an open issue of concern. Additionally, bias in the AI domain raises not only ethical but practical issues as well, since tasks performed by humans, including decision making, are constantly assigned to AI. In this paper we propose a framework for AI-augmented bias mitigation in the IC. Also, we propose a toolset, already adapted by the ICT community, for supporting this framework of bias mitigation methodologies adapted by the IC.

Depending on their design and how they are employed and incorporated into the proposed framework, the use of AI-powered tools can cause not only positive, but also negative impact on the trustworthiness of the output intelligence products. Subsequently, inefficient exploitation of AI tools may challenge the overall implementation of the framework and utterly pose threats to the validity of our work. Negative impact may include: manipulation of AI tools, as malicious/adversarial actors may intentionally feed them with misinformation or fake sources; complexity of source (especially human) evaluation: AI tools might not factor in crucial aspects of human sources, such as the credibility of an informant or their incentives and personal agendas; unconditional trust on automation: depending entirely on AI tools in the Cycle, might lead to excessive reassurance among human analysts, who could then overlook critical nuances that only human judgment and intuition may perceive.

Our future plans are oriented towards evaluating the proposed tool-supported framework, detect potential weaknesses, examine the need for accordingly customizing its AI tools and exploring other AI tools that may enhance the effectiveness of the Cycle towards mitigating bias. Emphasis will be given on assessing Twitter scraping tools' role in the Collection phase and in the Processing phase correspondingly, the role of sentiment analysis and source reliability/information credibility frameworks. Our ongoing work intends to capitalize on the meticulous, evaluation-supported research, already conducted by our team, in assessing Twitter profiles scraping, through experimentation with a number of currently available open-source and free tools, as it was already performed in early experiments, presented in Chapter V.

REFERENCES

[1] "Twitter Marketing." https://marketing.twitter.com/en (accessed Feb. 02, 2023).

[2] E. Tunggawan and Y. E. Soelistio, "And the winner is …: Bayesian Twitter-based prediction on 2016 U.S. presidential election," in 2016 International Conference on Computer, Control, Informatics and its Applications (IC3INA), Oct. 2016, pp. 33–37. doi: 10.1109/IC3INA.2016.7863019.

[3] D. A. Kristiyanti, Normah, and A. H. Umam, "Prediction of Indonesia Presidential Election Results for the 2019-2024 Period Using Twitter Sentiment Analysis," in 2019 5th International Conference on New Media Studies (CONMEDIA), Oct. 2019, pp. 36–42. doi: 10.1109/CONMEDIA46929.2019.8981823.

[4] X. Guo and J. Li, "A Novel Twitter Sentiment Analysis Model with Baseline Correlation for Financial Market Prediction with Improved Efficiency," in 2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS), Oct. 2019, pp. 472–477. doi: 10.1109/SNAMS.2019.8931720.

[5] G. A. Ruz, P. A. Henríquez, and A. Mascareño, "Sentiment analysis of Twitter data during critical events through Bayesian networks classifiers," Future Generation Computer Systems, vol. 106, pp. 92–104, May 2020, doi: 10.1016/j.future.2020.01.005.

[6] J. S. Pohl, M. V. Seiler, D. Assenmacher, and C. Grimme, "A Twitter Streaming Dataset collected before and after the Onset of the War between Russia and Ukraine in 2022," SSRN Electronic Journal, 2022, doi: 10.2139/ssrn.4066543.

[7] S. Sazzed, "The Dynamics of Ukraine-Russian Conflict through the Lens of Demographically Diverse Twitter Data," in 2022 IEEE

[2] https://spacy.io/universe/project/spacy-textblob

[3] https://orangedatamining.com/

[4] https://github.com/cjhutto/vaderSentiment

International Conference on Big Data (Big Data), Dec. 2022, pp. 6018–6024. doi: 10.1109/BigData55660.2022.10020274.

[8] J. Rovner, "Intelligence in the Twitter Age," International Journal of Intelligence and CounterIntelligence, vol. 26, no. 2, pp. 260–271, Jun. 2013, doi: 10.1080/08850607.2013.757996.

[9] "Intelligence Cycle Graphic," *United States of America, Federal Bureau of Investigations*. https://www.fbi.gov/image-repository/intelligence-cycle-graphic.jpg/view (accessed Feb. 01, 2023).

[10] "Joint Intelligence Joint Publication 2-0," *USA, Chairman of the Joint Chiefs of Staff*. 2013.

[11] United States of America, Office of the Director of National Intelligence, "U.S. National Intelligence : An Overview" 2013.

[12] Government of India, Bureau of Police Research and Development, Ministry of Home Affairs, "Manual on Social Media Intelligence (SOCMINT) for Law Enforcement Agencies," 2021.

[13] "Intelligence Analysis-Army Techniques Publication (ATP) 2-33.4," Headquarters, Department of the Army (United States of America). Washington, DC, 2020. [Online]. Available: https://atiam.train.army.mil/catalog/dashboard

[14] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a Feather: Homophily in Social Networks," Annu Rev Sociol, vol. 27, pp. 415–444, 2001, [Online]. Available: http://www.jstor.org/stable/2678628

[15] M. I. Handel, "Intelligence and the problem of strategic surprise∗," Journal of Strategic Studies, vol. 7, no. 3, pp. 229–281, Sep. 1984, doi: 10.1080/01402398408437190.

[16] L. Azzopardi, "Cognitive Biases in Search," in Proceedings of the 2021 Conference on Human Information Interaction and Retrieval, Mar. 2021, pp. 27–37. doi: 10.1145/3406522.3446023.

[17] R. J. Heuer, "Strategic Deception and Counterdeception: A Cognitive Process Approach," International Studies Quarterly, vol. 25, no. 2, pp. 294–327, 1981, doi: 10.2307/2600359.

[18] F. Huszár, S. I. Ktena, C. O'Brien, L. Belli, A. Schlaikjer, and M. Hardt, "Algorithmic amplification of politics on Twitter," Proceedings of the National Academy of Sciences, vol. 119, no. 1, Jan. 2022, doi: 10.1073/pnas.2025334119.

[19] L. Belli, "Examining algorithmic amplification of political content on Twitter," Oct. 21, 2021. https://blog.twitter.com/en_us/topics/company/2021/rml-politicalcontent (accessed Feb. 01, 2023).

[20] K. Knorr, "Failures in National Intelligence Estimates: The Case of the Cuban Missiles," World Polit, vol. 16, no. 3, pp. 455–467, Apr. 1964, doi: 10.2307/2009582.

[21] "A Tradecraft Primer: Structured Analytic Techniques for Improving Intelligence Analysis," Center for the Study of Intelligence, Central Intelligence Agency (CIA), USA. 2009.

[22] R. Desimone and D. Charles, "Towards an Ontology for Intelligence Analysis and Collection Management," 2002.

[23] B. Mandrick and B. Smith, "Philosophical foundations of intelligence collection and analysis: a defense of ontological realism," Intelligence and National Security, vol. 37, no. 6, pp. 809–819, Sep. 2022, doi: 10.1080/02684527.2022.2076330.

[24] "Ontology for the intelligence analyst," 2012. [Online]. Available: https://www.researchgate.net/publication/290966313

[25] Y. Cardinale, I. Dongo, G. Robayo, D. Cabeza, A. Aguilera, and S. Medina, "T-CREo: A Twitter Credibility Analysis Framework," IEEE Access, vol. 9, pp. 32498–32516, 2021, doi: 10.1109/ACCESS.2021.3060623.

[26] M. Wijesekara and G. U. Ganegoda, "Source credibility analysis on Twitter users," in Proceedings - International Research Conference on Smart Computing and Systems Engineering, SCSE 2020, Sep. 2020, pp. 96–102. doi: 10.1109/SCSE49731.2020.9313064.

[27] L. Krzysztof, S.-W. Jacek, J.-L. Michal, and G. Amit, "Automated Credibility Assessment on Twitter," Computer Science, vol. 16, no. 2, p. 157, 2015, doi: 10.7494/csci.2015.16.2.157.

[28] K. Sharma, "Reliability Index for Twitter – Twitter Handles' Credibility Assessment," HELIX, vol. 8, no. 5, pp. 4094–4099, Aug. 2018, doi: 10.29042/2018-4094-4099.

[29] M. M. Kokar and B. Ulicny, "Automating Military Intelligence Confidence Assessments for Twitter Messages," 2014. [Online]. Available: http://blogs.aljazeera.net/twitter-dashboard

[30] Daniel Tunkelang, "A Twitter Analogy to PageRank." https://thenoisychannel.com/2009/01/13/a-twitter-analog-to-pagerank/ (accessed Mar. 17, 2023).

[31] K. Garimella, G. De Francisci Morales, A. Gionis, and M. Mathioudakis, "Mitigating Confirmation Bias on Twitter by Recommending Opposing Views," in WSDM 2017 - Proceedings of the 10th ACM International Conference on Web Search and Data Mining, Feb. 2017, pp. 81–90. doi: 10.1145/3018661.3018703.

[32] E. Tankard, C. Flowers, J. Li, and D. B. Rawat, "Toward Bias Analysis Using Tweets and Natural Language Processing," in 2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC), Jan. 2021, pp. 1–3. doi: 10.1109/CCNC49032.2021.9369461.

[33] A. Abusitta, E. Aïmeur, and O. A. Wahab, "Generative Adversarial Networks for Mitigating Biases in Machine Learning Systems," May 2019.

[34] S. Li and W. Deng, "Deep Facial Expression Recognition: A Survey," IEEE Trans Affect Comput, vol. 13, no. 3, pp. 1195–1215, Jul. 2022, doi: 10.1109/TAFFC.2020.2981446.

[35] H. Du, H. Shi, D. Zeng, X.-P. Zhang, and T. Mei, "The Elements of End-to-end Deep Face Recognition: A Survey of Recent Advances," Sep. 2020.

[36] H. Mansourifar and S. J. Simske, "GAN-Based Object Removal in High-Resolution Satellite Images," Jan. 2023.

[37] "Twitter API Rate limits" https://developer.twitter.com/en/docs/twitter-api/rate-limits.

[38] "Twint." https://github.com/twintproject/twint (accessed Mar. 16, 2023).

[39] "Twitter-scraper." https://github.com/bisguzar/twitter-scraper (accessed Mar. 16, 2023).

[40] "Twitterscraper." https://github.com/taspinar/twitterscraper (accessed Mar. 16, 2023).

[41] "TIGMINT." https://github.com/TIGMINT/TIGMINT (accessed Mar. 16, 2023).

[42] "TIGMINT web version," Supported by Gurugram Police Cyber Security Internship, India. https://sakshi-25.github.ioTIGMINT-web (accessed Mar. 16, 2023).

[43] "Snscrape." https://github.com/JustAnotherArchivist/snscrape (accessed Mar. 16, 2023).

[44] "Tweeds ." https://github.com/achyuthjoism/tweeds (accessed Mar. 16, 2023).

[45] Druziuk, Y. (2022, April 18). A Citizen-like chatbot allows Ukrainians to report to the government when they spot Russian troops — here's how it works. https://www.businessinsider.com/ukraine-military-e-enemy-telegram-app-2022-4

[46] "Natural Language Toolkit (NLTK)." https://www.nltk.org/ (accessed Mar. 20, 2023).

[47] "Stanford CoreNLP." Stanford CoreNLP (accessed Mar. 20, 2023).

[48] "TweetNLP." https://github.com/cardiffnlp/tweetnlp (accessed Mar. 20, 2023).

[49] "GATE." https://gate.ac.uk/ (accessed Mar. 20, 2023).

[50] United States Senate Intelligence Committee, "H. Rept. 107-792 - JOINT INQUIRY INTO INTELLIGENCE COMMUNITY ACTIVITIES BEFORE AND AFTER THE TERRORIST ATTACKS OF SEPTEMBER 11, 2001," Washington DC, Dec. 2002.

[51] Heuer RJ, Psychology of Intelligence Analysis. Washington D.C: CIA, Center for the Study of Intelligence, 1999.

[52] "Open Synthesis." https://github.com/twschiller/open-synthesis (accessed Mar. 21, 2023).

[53] "Natural Language Summary Generation from Structured Data." https://github.com/akanimax/natural-language-summary-generation-from-structured-data (accessed Mar. 16, 2023).

[54] E. Paparidis and K. Kotis, "Towards Engineering Fair Ontologies: Unbiasing a Surveillance Ontology," in 2021 IEEE International Conference on Progress in Informatics and Computing (PIC), Dec. 2021, pp. 226–231. doi: 10.1109/PIC53636.2021.9687030.

[55] P. Alexiadis, " Capturing the polarization of Twitter users on key social issues using Intelligent Web method," Master thesis, Intelligent Systems Lab, Dept. of Cultural Technology and Communication, University of the Aegean, Feb 2023. [Online]. Available: https://hellanicus.lib.aegean.gr/handle/11610/25654.