Big Data for Human Security: The case of COVID-19

Pedro Cárdenas[a,*], Ioannis Ivrissimtzis[a], Boguslaw Obara[b], Ibad Kureshi[c],
Georgios Theodoropoulos[d,*]

[a]*Dept. Computer Science, Durham University, United Kingdom*
[b]*School of Computing, Newcastle University, United Kingdom*
[c]*HM Land Registry, Durham, United Kingdom*
[d]*Dept. Computer Science and Engineering, Southern University of Science and Technology (SUSTech), Shenzhen, China*

Abstract

The COVID-19 epidemic has changed the world dramatically since societies are changing their behaviour according to the new normal, which comes along with numerous challenges and uncertainties. These uncertainties have led to instabilities in several facets of society, most notably health, economy and public order. Measures to contain the pandemic by governments have occasionally met with increasing dis-content from societies and have triggered social unrest, imposing serious threats to national security. Big Data Analytics can provide a powerful force multiplier to support policy and decision makers to contain the virus while at the same time dealing with such threats to human security. This paper presents the utilisation of a big data forecasting and analytics framework and its utilisation to deal with COVID-19 triggered social unrest. The paper is an extended version of paper [1] presented at the 2021 International Conference on Computational Science.

Keywords: COVID-19, Epidemics, Big Data, National Security, Data Analytics, Machine Learning

---

[*]Corresponding author

Email addresses: `pedro.cardenas-canto@durham.ac.uk` (Pedro Cárdenas), `ioannis.ivrissimtzis@durham.ac.uk` (Ioannis Ivrissimtzis), `boguslaw.obara@ncl.ac.uk` (Boguslaw Obara), `ibad.kureshi@landregistry.gov.uk` (Ibad Kureshi), `theogeorgios@gmail.com` (Georgios Theodoropoulos)

1. Introduction

Global challenges and emergencies such as climate change, epidemics and natural and man-made calamities present unprecedented governance issues. Going back in time to 2019, the world changed completely because a thitherto virus (SARS-CoV-2 or COVID-19) took people, institutions and governments by surprise. The high contagiousness and rapid spread of the virus and the lack of appropriate treatments resulted to a global pandemic that affected societies not only the healthcare sector, but also in strategic areas that they had not anticipated they would be impacted so seriously, such as the economy and the public order. Governments across the globe took strict decisions aimed at containing the disease and avoiding massive infections, such as curfew, lockdowns or measures such as "stay at home" or clustering domestic territories according to its infection rates [2]. Such containment measures have been accompanied by varying degrees of social discontent and unrest, violent manifestations of which in the forms of demonstrations and riots have been witnessed across the globe [3]. Most countries witnessed severe knock-on effects of pandemic, such as rising levels of unemployment and even sometimes food shortages, which accelerated the internal instability scenarios. The negative social and economic effects of the pandemic also provided a pretext for intensifying pre-existing social discontent and unrest., e.g. the Black Lives Matter movement or the demonstrations in Hong Kong and eventually jeopardised national security.

Policy and decision-makers need to have at their disposal technological tools to act as force multipliers and enable them to gain insights about disasters and unfolding situations and assess the threat to national and international security [4, 5, 6]. Big Data technologies can provide a powerful means in this endeavour [7, 8]. As a result the last decade has witnessed the development of several computational platforms that utilise Big Data analytics to derive insights about disruptive situations that can trigger social unrest [9, 10, 11, 12, 13, 14].

Aspiring to contribute to the global effort towards tools for the prediction and interpretation of disruptive events and national security crises, in earlier

2

work we have presented a framework and associated workflow for the analysis of social media data (Twitter) to derive insights about disruptive events and potential unrest [15, 16, 17, 18, 19]. In the form of intelligence insights, the outcomes facilitate the dissection of complex national security problems that can be used to inform the courses of action by the decision-makers.

In this paper, the framework, which we refer to as *Eunomia*, is utilised to analyse events related to the COVID-19 pandemic. At the time of writing, the pandemic had resulted in more than four million deaths worldwide [20], triggered multiple protests, while its impact on the human components of economic and food security was still unfolding.

The paper is an extended version of [1] and its aim is to further refine and further test the robustness and applicability of our framework in forecasting and analysing global disruptions such as the COVID-19 pandemic. The new material in this paper, compared with [1] is the following: (a) A more complete discussion on National Security components which constitutes the theoretical domain basis for the technical work presented in the paper (section 2); (b) A more detailed description of the Conceptual Framework components (section 3.1); (c) a new use case in a different cultural, political and socio-economic context, namely the COVID-19 related protests in Pakistan in March 2020 (section 4.1).

The rest of the paper is organised as follows. In Section 2 a description of COVID-19 and its impact to national security is addressed. In Section 3 we give a detailed description of our national security analysis framework. In Section 4 we describe the results of the big data analysis of 3 case studies, the protests in Pakistan, Michigan and Texas, and we briefly conclude in Section 5.

2. National Security and Pandemics

National security is a complex concept that entails multiple intertwined components. Nevertheless, there is a widely accepted view that the core part of it lies in maintaining an internal balance amongst its components. As described by [24], the fragile equilibrium between the various national security components

3

tends to be a critical factor of a country's developmental pace. Since there is no development without security, the importance of tracking and understanding the threads of those national security components on which the stability of a state depends is evident.

The work presented in this paper adopts the theoretical framework of the interactions between the various national security components which was first proposed by the United Nations in [22]. Seven components of the national security environment, also called *human security components* were identified in that report, namely: health security, public order, environmental security, communal security, food security, personal security and political security. These components are also inline with the the 2030 Agenda for Sustainable Development adopted by the United Nations in 2015 as a universal call to action to protect the planet, and ensure that by 2030 all people enjoy peace and prosperity [23]. This group of seven components are all related to fundamental human rights, place the individual and the people at the center and facilitate a thorough, and yet still relatively straightforward, way to view, understand and intervene to prevent likely national security threats. To give characteristic examples of each component, assuring a basic income (economic security), providing physically convenient and economically affordable access to food (food security), assuring an environment free of disease and infection (health security), universal access to sanitary water supply and clean air (environmental security), guaranteeing security from physical violence and threats (personal security), protecting people's freedoms and cultural identity (political and communal security) [thesis], are critical factors in ensuring stability and avoiding internal disruptions of the state.

Pandemics such as COVID-19 are recognised as serious National Security threats [25]. Figure 1, provides a schematic description of the impact of COVID-19 on human security components. It illustrates the main thesis of this paper, that the study and analysis of the human security components requires a holistic approach, since the initial severe disruption of one component, here the health component, has knock-on effects on multiple other components, for example,
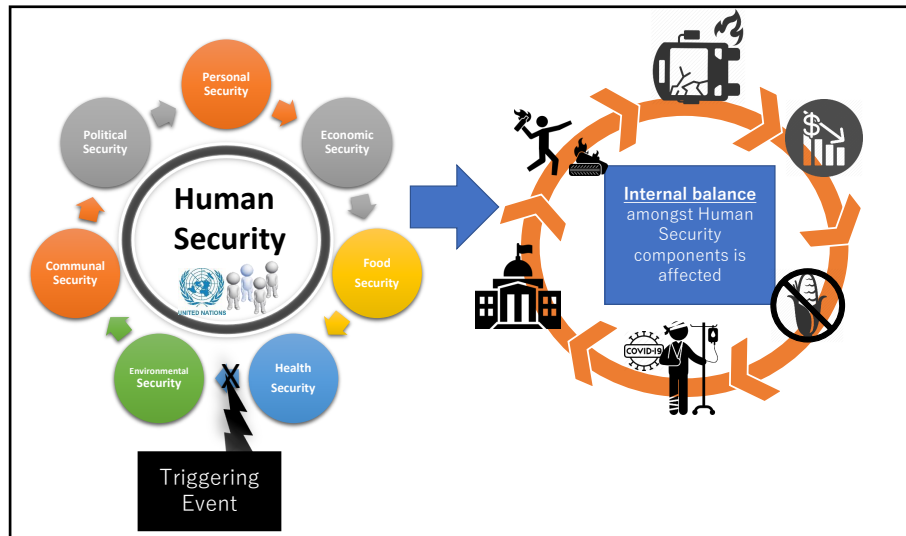
4

Figure 1: Impact of the Human Security Components. The left part of the Figure shows that a triggering event impacted one human security component - health. The right hand side of the Figure illustrates the disruption of the internal balance of the human security elements, with components related to public order, economy, and food security also affected, due to the severity of the impact on people's health.

erosion of trust to public institutions, riots, raising unemployment, financial instability, disruption of the supply chains, and even secondary disruption of the health component from factors such as lack of medicines, or disruption in the operations of the national health systems.

⁹⁵ Predicting and analysing a crisis is a complex task which gains more relevance and importance when the centre/core of a state (people) can potentially be affected. Digital technologies and Big Data analytics can play a crucial role to the timely spotting of early sparks of societal instability, which later may evolve into national security issues. The rest of the paper discusses such an ¹⁰⁰ approach for the analysis of COVID-19 related societal disruptions. The framework utilises on Machine Learning techniques and Natural Language Processing to analyse data harvested from multiple sources, including Social Networking Sites (SNS), such as Twitter and Instagram, Independent Websites (IW), and
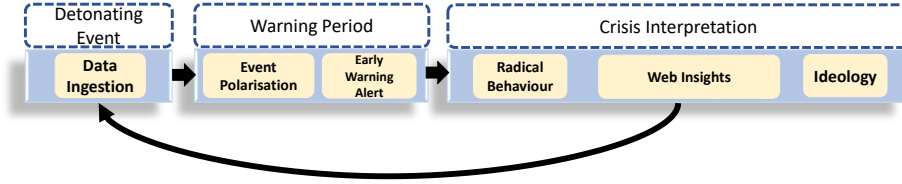
Figure 2: The Eunomia framework, as described in [15, 16, 17, 18, 19].

online data from Non-Profit Organisations.

3. The Eunomia Framework for National Security Analysis

The power and influence of information technologies have changed the world because they link individuals through digital channels such as web sites or social networking services to share ideas and beliefs. This communication cycle works as a valuable tool that can summon big groups of people, and during a catastrophic scenario, it may lead to disruptive episodes that threaten the fragile internal balance of the affected state. A key fact to highlight is that the nature of such events triggers a content demand surge which turns the Internet into a complex and disarranged milieu because the information is conveyed through a variety of flavours such as blogs, websites, video-sharing platforms or social networking sites [26].

During a crisis, data is crucial to augment our understanding of the crisis situation as it unfolds in real-time. Numerous examples can illustrate the way social media has been used during disasters or demonstrations, ranging from earthquakes to social unrests [44, 45, 46].

The Eunomia framework has aspired to bridge well established theories from social and political sciences, and big data analytics to provide a holistic decision support framework and empower competent shareholders in their efforts to predict and manage national security crises [1, 15, 16, 17, 18, 19]. The framework consists of two main stages as illustrated in Figure 2. An initial phase (Warning Period) examines in real-time data from online activity, in a time-continuous

manner to the extent that this is feasible, and spots changes in societal behavioural characteristics. When certain thresholds have been surpassed, it triggers an alert whereupon a subsequent "Crisis Interpretation" phase in activated. This starts to collect and analyse information from a variety of sources, such as web resources and social networking services, and extracts in-depth insights that can help the decision-makers to better understand the unfolding crisis. The following sections illustrate the different elements of this system.

3.1. Analysing a Crisis

A crisis may be defined as an event where the affected group of people witnesses a disruption of routine, social structure, norms and values [26]. The COVID-19 pandemic had all the characteristics of a crisis, albeit of higher magnitudes, both in terms of how deep were the changes in certain behavioural aspects of people's lives, and in terms of its geographical extent. Creating a holistic view of a crisis, encompassing all the affected human security elements, can provide enough insight to explain and interpret the crisis and its impact.

Table 1 summarises the insights that can be provided by the current implementation of the Eunomia framework. The next subsections discuss in more detail all the stages involved, as outlined in Figure 2, and corresponding to the columns of Table 1.

Insights	Stages			
	Early Warning Alert	Radical Behaviour	Ideology	Web Insights
Q1. When do people head towards a situation that evokes that both social stability and national security components can be compromised?	✓	-	-	-
Q2. Which entities are described by people during the crisis?	-	✓	-	-
Q3. What are the radical behavioural traits being conveyed?	-	✓	-	-
Q4. What items are being asked for by individuals in social media?	-	✓	-	-
Q5. Are hostility and authoritarianism traits present during the incident?	-	-	✓	-
Q6. Do embedded web resources in social media texts disclose that the national security components have a horizontal escalation over time?	-	-	-	✓

Table 1: Insights derived from the Analytics Framework described in [15, 16, 17, 18, 19].

3.1.1. Event Polarisation and Early Warning Alert (Q1)

Early warning systems typically encompass disaster risk assessment, hazard forecast, prediction and monitoring, risk communication and emergency preparedness activities [28, 29, 27]. Natural and anthropogenic (human-induced) hazards can lead to disasters [26], which can represent a national security threat

in the worst scenario. In the case of COVID-19 there has been an intense debate to classify it according to the two previous categories. Irrespective of the posture, the core point is identifying when individuals are heading towards a point of no return, which means that human security components have been altered. Frustration, anger, alienation are all factors that may unleash polarised

collective behaviour, such as protests or riots [43]. Timely detection of societal instability sparks represents a core element of dealing with national security problems, as it enables the early prediction of likely major incidents, giving time to the authorities to decide the best course of action to diminish secondary effects. Preparedness for a national security crisis is a crucial step for which

states must be ready for.

The early warning stage of the Eunomia framework is based on an alert mechanism to predict tipping points, which are situations where the unfolding of events acquire their own irreversible dynamic [15]. The Alert Mechanism described in an alert when a certain threshold has been exceeded on three key

components: Global Polarisation (GP), Social Media Connectedness (SMC) and Human Security Impact (HSI) (Figure 3).

Due to the vast nuance of personal views conveyed during a crisis, interpreting messages offers a way to construe subjective information. The first stage (GP) examines individuals' opinions based on a sentiment analysis process. As

part of this stage, a data cleansing procedure has a preponderant role by removing stop words and URLs, replacing contractions and abbreviated words, and correcting errata.

Premised on a previous calibration process, which is supported by the analysis of numerous incidents, a global polarisation index is used as a threshold to
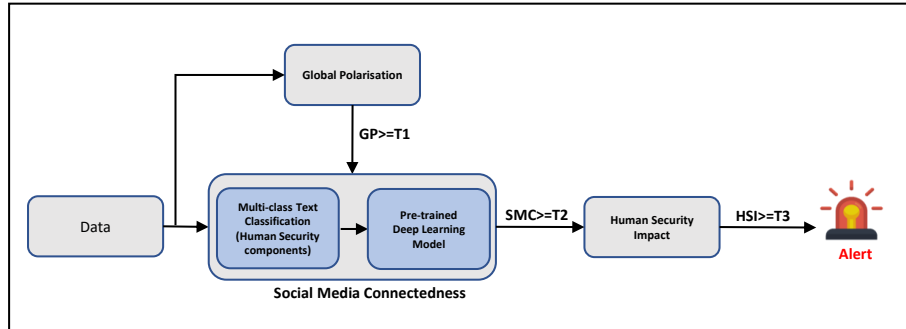
Figure 3: Alert Mechanism Workflow, described in [15].

determine if the studied event has surpassed the pre-established limit, in which case the SMC process is activated.

Then, through the prism of a deep learning model, the SMC stage identifies when individuals felt attracted to the incident by considering that disruptive ideas have remained on trends for more than four continuous hours. However, to see through the deep learning model mentioned earlier, a binary classifier process comes into action. This artificial model analyses language text samples and performs the classification procedure to identify when comments are related to attracting intergroup attention or the call of masses, which is closely linked to the variable "people" of the human security components, as described in [15].

In a similar fashion to the preceding component (GP), a pre-defined threshold has been defined by examining multiple disruptive events. Then, the following stage (HSI) comes to the scene when the threshold, as mentioned earlier, has been exceeded.

The HSI step classifies the data corpus into ten human security aspects (health, public order, transport, economy, people, defence, environment, government, information and life) using unsupervised and supervised learning processes. Finally, using a preconfigured scale it determines if human security components have been compromised.

9

3.1.2. Radical Behaviour (Q2, Q3 and Q4)

A disaster, a crisis or a health crisis (pandemic) are triggers that tend to ramp up the exchange of data and relevant information, which is crucial in the immediate aftermath of an incident [30]. In the middle of these chaotic scenarios, messages containing a vast nuance of ideas, beliefs and thoughts are conveyed. Within this variety of disseminated topics, calls to violence messages are also transmitted. From a national security perspective, the viral propagation of this type of content can represent a severe threat. Because the idea to assassinate a person, destroy critical infrastructure or even exchange ammunition or weapons are activities that can travel from the virtual to the real world and unleash instability scenarios that might affect the human security components.

The analysis and extraction of radical expressions unveil personality traits together with other core societal aspects that facilitate the crisis interpretation process. In line with this idea, and following the workflow illustrated in Figure 4, five different components are interconnected to detect radical behavioural traits: Instability Scenarios, Entity Extraction, Wordlists Creation, Analytics, and Data Interpretation [16].
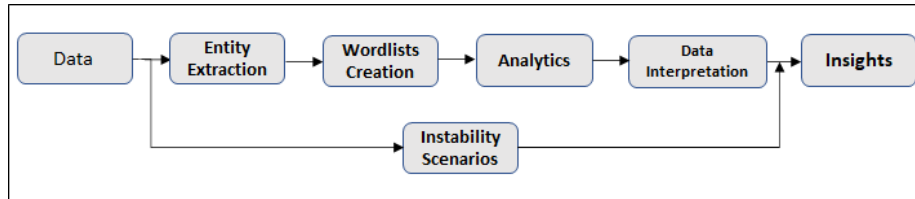

Figure 4: Analysis of Radical Behaviour Workflow, described in [15].

These interlinked stages are supported by dissimilar computational techniques (such as deep learning, natural language processing, supervised and unsupervised learning), and opens up a route to understanding the radicalised environment by providing the following intelligence information:

1. Creation of instability scenarios.

2. Identification of affected entities (people, locations, or facilities).

10

3. Identification of likely affected entities due to their proximity to the inci-
dent.

4. Dissection of the intentions expressed towards an entity.

₂₂₀ 5. Analysis of the dissemination degree of the crisis (widespread or local
incident).

6. Detection of violent expressions.

7. Classification of violent expressions.

8. Necessities shared by individuals amid the crisis.

₂₂₅   It is pertinent to note that the nature of the events can bias the extraction of
meaningful insights, as expressions and vocabulary conveyed during a political
protest, a disaster or a crisis such as a pandemic are diametrically opposite
but rich in specialised words. Word specialisation derived from these dissimilar
scenarios brings to the table some disadvantages. Firstly, a continuous process
₂₃₀ of wordlists update needs to be performed since the detection of necessities
requires two elements, a verb and a noun. The verb indicates the action that
people want to do, and the noun refers to the object/thing that will receive the
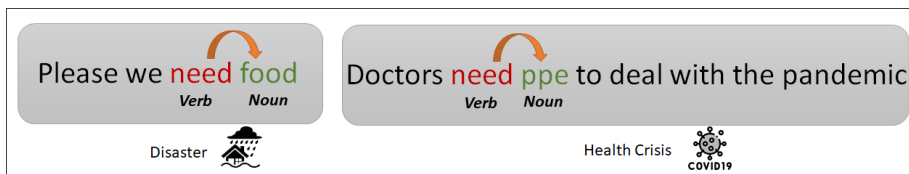action (see Figure 5).

Figure 5: Text language samples conveyed by people during a disaster and the protests due
to COVID-19 restrictions.

  The necessity of a specialised wordlist is displayed in Figure 5, where both
₂₃₅ messages are related to people's needs amidst an incident. On the one hand,
a generic term "food"; but on the other hand, a specialised word that gained
more relevance (popularity) because of the COVID-19 pandemic, "PPE".

  A second limitation is that the detection and classification of violent terms
can be affected similarly to the previous case. By considering the same verb-

11

₂₄₀ noun structure, entities, which in this case are the nouns who received the action, can adopt multiple forms (time-dependent language expressions).

3.1.3. Ideology (Q5)

Information traffic surges in the digital channels during a disaster and a crisis since individuals want to share messages related to a vast nuance of topics such ₂₄₅ as meteorological and seismological information or even political views of an incident [26]. Nevertheless, those posts' ideas, thoughts, and beliefs gain special attention when the disruptive content is directed towards other groups or even authorities. This collective set of emotions/comments can denote appraisals of superiority/inferiority, goal obstruction/injustices, or hostility, as described in ₂₅₀ [31, 32].

Protests and violent events derived from societal problems (e.g. Black Lives Matter) or natural disasters (e.g. pandemics) are two examples of incidents. These disturbing plots have a particular interest as hostility traits may be present and maybe red flags that indicate that the stability of a state may ₂₅₅ have been compromised and a crisis is forthcoming [31, 32].

In addition, such disruptive situations may evolve due to the fact that people do not empathise with decisions or activities performed by those who hold the "proper authority", which can be perceived as a high level of authoritarianism against them, as measured in an aggression, submission and conventionalism ₂₆₀ scale, as described in [35, 36]. For the COVID-19 situation, governments around the globe have introduced various measures to contain the virus. Notwithstanding the differences, the nature of the restrictions have generated traits of hostility and authoritarianism.

The Ideology stage processes emotions to identify such traits of hostility and ₂₆₅ authoritarianism. Then a variational autoencoder model separates ideological traits with unforeseen patterns (anomalies), see Figure 6. Next, the outcomes are compared against precalibrated thresholds that come from the examination of other incidents to determine the presence of the aforementioned ideological features.
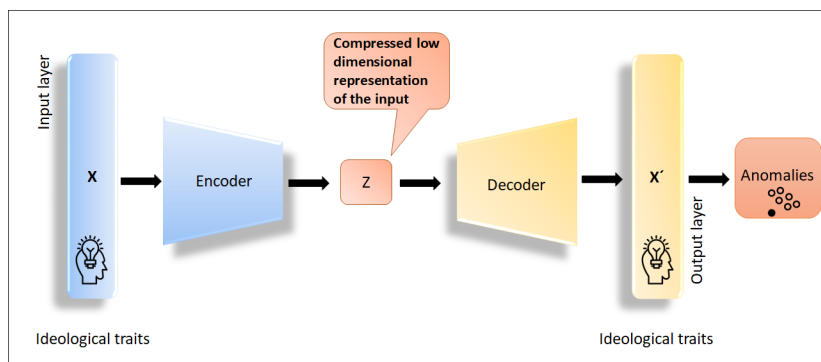
12

Figure 6: Variational Autoencoder model to identify ideological traits anomalies.

3.1.4. Web Insights (Q6)

The analysis of data deluge coming from multiple communication channels is a common problem that many organisations that are in charge of preparing emergency and national security plans are not prepared to deal with. Social media platforms have a preponderant role in this huge communication arena as millions of people have established networks to disseminate information, which outlines the complexity of analysing such a massive volume of data.

Within this heterogeneous mixture of communication channels, the complexity increases when a person posts a message since it may refer to another source (website or social networking service) that contains complementary information that reinforces the very essence of the message. In view of such a complex scenarios, examining the message-embedded informational sources plays a significant role since it helps identify likely national security instability features.

The horizontal escalation analysis contributes to spot such instability insights by examining the increase in the number of human security components, which works as an indicator of the presence of a national security problem [16], see Figure 7. This method utilises web resources (websites, social networking services, independent websites or information outlets) as the raw material to analyse the national security's components escalation described above, which may have been used as a mouthpiece to propagate the data throughout the
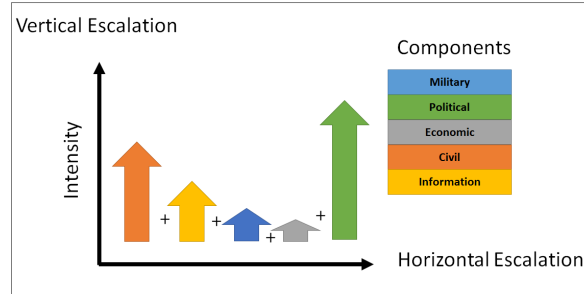
13

crisis timespan.

Figure 7: Horizontal and Vertical Escalation of Instruments of Power, adapted from [38].

The architecture of the web insights procedure begins with the extraction of URLs; then, those digital sources are grouped according to a pre-classified dictionary. Afterwards, a series of machine learning processes analyse the extracted content to perform two core subprocesses. The first one aimed at identifying/classifying violent phrases. The second one determines if the disruptive term relates to an entity of interest as critical infrastructure, people, or a location.

Finally, a multiclass text classification process spots which human security components are being affected and determine an increase in them, which can be construed as a horizontal escalation, see Figure 8.
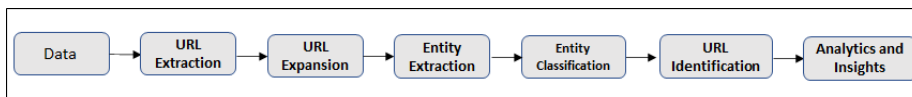

Figure 8: Web Insights Workflow, as described in [16].

4. Deploying Eunomia and Experimental Results

The goal of the Eunomia framework is to monitor the state of the society at any particular moment and in case of an alert, to derive deeper insights about the situation and the threat it may constitute to national security. This section presents the deployment of Eunomia in two different cases of societal disruptions

14

triggered by the COVID-19 pandemic. These cases have been selected to reflect two different cultural and socio-economic contexts: Asia and the USA.

The first case examines the protests in Pakistan in 2020, where the authorities deployed armed forces to enforce countrywide lockdown measures, and as a result, regional protests took place.

The second case investigates protests in the United States of America (Michigan and Texas), when societal disruptions were triggered due to discontent with restriction measures at State level, imposed to control the spread of COVID-19.

The analysis of these cases aim to demonstrate the robustness and applicability of the framework for forecasting and analysing important real-world events : would the framework have been able to provide the competent authorities enough notice and insights to deal with the then forthcoming crisis? The results support the paper's premise that the an insightful examination of imbalances in human security components contributes towards the timely detection of likely national security issues. The analysis also provides interested stakeholders postmortem insights about COVID-19 social crises with the view to contribute to the global effort to tackle this disruptive situation.

4.1. Protests in Pakistan

In early March 2020, small bursts of societal instabilities began to appear in Pakistan due to political views regarding health care sector problems, which were exacerbated by the pandemic. Tensions escalated, and on March 23, 2020, the army was deployed to assist civilian authorities to enforce lockdown measures after all provincial governments ordered a complete or partial lockdown to curb the spread of coronavirus disease (COVID-19) [39].

4.1.1. Data Cleansing

The first stage of the analysis begins with the collection of 78,910 tweets, written in English, between 19th and 31st of March 2020, considering hashtags linked to the incident as shown in Figure 9. Since tweets are unstructured data, the raw text was processed following the steps described below: (1) URLs

were extracted; (2) RT and mention terms were removed; (3) contractions were replaced, for instance, isn't: is not; (4) punctuation marks were removed; (5) emoticons were replaced by suitable terms; (6) duplicated messages were removed; (7) Internet slang was replaced by complete expressions using a preconfigured dictionary, for example, AFAIK: "as far as I know", ASAP: "as soon as possible", or BBL: "be back later".
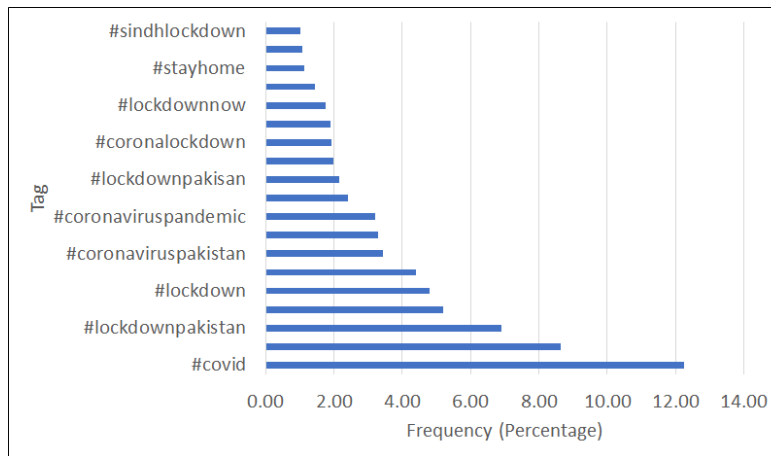

Figure 9: Selected hashtags relevant to the protests in Pakistan in March 2020.

4.1.2. Early Warning Alert (Q1)

Figure 10 depicts the analysis of the collected tweets for the incident in Pakistan, between 19th and 31st of March, 2020. On March 23rd 2020, the local government deployed its army to enforce the lockdown measures imposed to contain the spread of COVID-19. As depicted in figure 10, the Alert Mechanism would raise a red flag (alert) two days before the government deployed its military personnel, suggesting that a tipping point could have been been detected in advance of the escalation in the response of the authorities.

4.1.3. Radical Behaviour (Q2, Q3 and Q4)

The information disseminating during an incident or a crisis, often involves text that includes violent expressions, which embody the feelings and the ideas
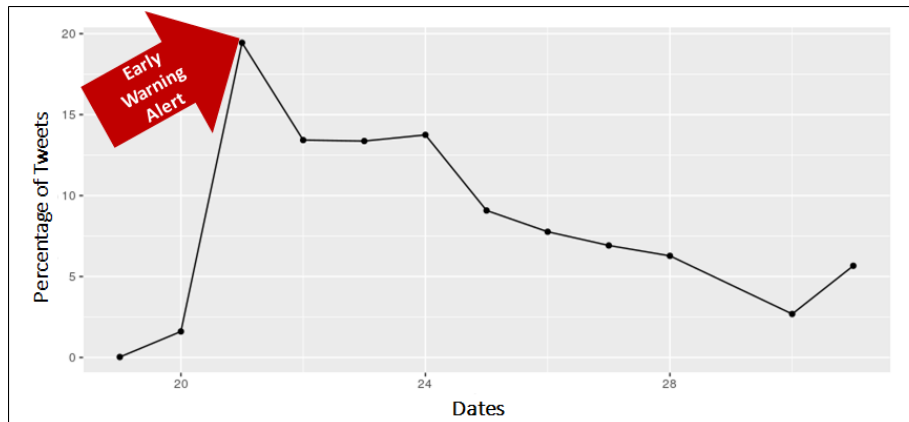
16

Figure 10: Early Warning Alert Detection for the events of Pakistan in March 2020.

of a wider mass of people. Thus, a key element in our approach lies in extracting radical terms, as described in [16].

Q2. As shown in Table 2, a day before the Pakistani government deployed military personnel, people were posting regarding the lockdown in Karachi, and after the deployment of the armed forces, protesters were disseminating messages related to locking the Masjid mosque.

Q3. To avoid the further spread of COVID-19 among the population, lockdown measures were enforced on March 23rd. When our system triggered a red alert on March 21s, individuals were broadcasting messages associated with the evasion or violation of the lockdown. Their posts contained content indicating refusal, disagreement, defiance, and dissuasion, while the idea of taking to the streets to demand the withdrawal of the measures was also propagated during the following days.

Q4. Finally, as the pandemic affected people's health, various health-related demands were also mentioned during the protests in Pakistan. On March 21st, individuals were asking for ventilators and other specialised equipment, such as (PPE), to help the health services cope with the crisis, see Table 2.

17

Location	Dates (2020)							
	March 19	**March 20**	**March 21**	**March 22**	**March 23**	**March 24**	**March 25**	**March 26**
Pakistan		Aware -> People	Maintain -> Precaution	Impose -> Curfew	Trigger -> Riots	Lockdown -> Karachi	Enforce -> Lockdown	Lift -> Curfew
		Violating -> Lockdown	Provide -> PPE	Disagree -> Curfew	Discover -> Antidote	Take -> Streets	Close -> Mosque	Impose -> Lockdown
		Use -> Sanitizers	Take -> Precaution	Implementing -> Curfew	Declare -> Lockdown	Fight -> Pandemic	Impose -> Lockdown	Protest -> Lockdown
		Avoid -> Lockdown	Violating -> Lockdown	Impose -> Lockdown		Impose -> Lockdown		
		Endure -> Lockdwon	Enforce -> Lockdown	Refuse -> Lockdown				
			Impose -> Lockdown	Need -> Ventilator				
			Need -> Ventilator	Lockdown -> Karachi				
	March 27	**March 28**	**March 29**	**March 30**	**March 31**			
Pakistan	Defying -> Crufew	Remove -> Curfew	Enforce -> Lockdwon	Enforce -> Lockdown	Imposing -> Lockdown			
	Lift -> Curfew			Lockdown -> Masjid	Dissuade -> Lockdown			
	Locking -> Masjid				Enforced -> Lockdown			
	Remove -> Lockdown				Disobey -> Lockdown			
	Tackle -> Menace							

Table 2: Disruptive Expressions extracted using Word Embeddings and Direct Object during the protests in Pakistan in March 2020.

4.1.4. Ideology (Q5)

The public demonstrations in Pakistan against the restrictions imposed by local authorities struggling to contain the coronavirus, are examples of events where individuals shared their beliefs and ideas through digital channels. Within this environment, detecting ideological traits such as hostility and authoritarianism, along with radical expressions, can suggest the onset of a crisis, see Figure 11. Following the framework described in Section 3, hostility and authoritarianism scores were computed and compared against the predefined thresholds proposed in [19].

Figure 11 summarises or measurements on ideological traits. On March 20th, a day before the alert was triggered, traits of hostility and authoritarianism were spotted. A day after, both ideological features increased considerably. These features continued to exceed the predefined thresholds, and on March 23rd, when the armed forces were deployed to enforce the lockdown measures, disgust and aggressiveness, which are components of hostility and authoritarianism,

18

respectively, both witnessed a considerable surge.
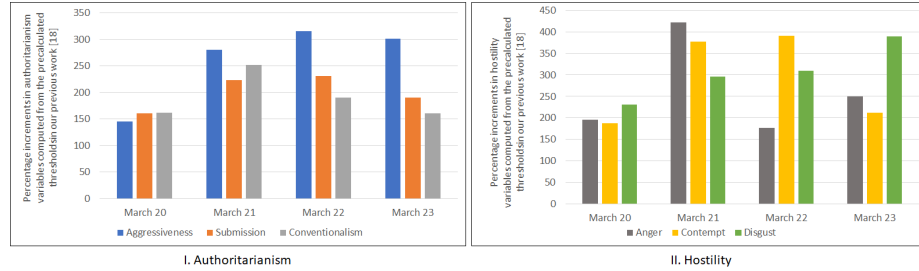

Figure 11: Ideological traits of the protests in Pakistan in March 2020.

4.1.5. Web Insights (Q6)

During an incident, various digital channels are used to propagate ideas and beliefs, ranging from websites to social networking services. During the protests in Pakistan, people used social networking services as the primary communication channel, and as Figure 12 shows, the health component was the most affected, from 19th to 24th of March. Moreover, from a horizontal escalation point of view, from March 20th to March 21st (when the alert was triggered), there was an increase in the number of affected human security components, from five to six. Such an increment suggests national security instability.

4.2. Protests in the United States of America

Two incidents that occurred in Michigan and Texas in April 2020 were studied. In both cases, citizens protested after local governments imposed lockdown rules, notwithstanding other types of the restrictions previously introduced to tackle the pandemic.

In the case of Texas, rallies were organised to show disagreement with the local lockdown measures, and the participants demanded the reopening of the economy [40, 41].

By contrast, in Michigan, a convoy of thousands of motorists drove from all over the state to protest the governor's stay-at-home order extension. The protest, known now as Operation Gridlock, involved clogging with their vehicles
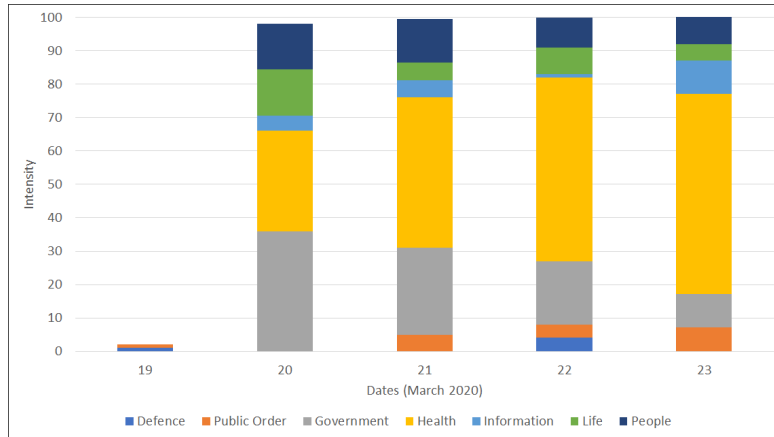
19

Figure 12: Horizontal Escalation of the National Security Components during the protests in Pakistan in March 2020.

the streets surrounding the Michigan State Capitol, including the Capitol Loop, and it drew national attention [42].

4.2.1. Data Cleansing

A data corpus of six million tweets written in English was collected, covering the dates from 10th to 20th April 2020, by considering hashtags such as #covid, #coronavirus, #coronavirusoutbreak and #coronaviruspandemic. Then, two data subsets were extracted from the anterior dataset, each subset containing tweets with a unique combination of specific parameters, such as hashtags that were linked to the studied entities (locations), as depicted in Table 3.

Once these two subsets have been created, tweets appertained to the former clusters were cleansed in a similar fashion to the previous case (protests in Pakistan).

4.3. Early Warning Alert (Q1)

Once an incident is unfolding, the stability of the state can be compromised due to national security components instability, at which point identifying if an event heads toward a significant disruption scenario becomes a primary task. In

20

Dataset 1 (Michigan ,USA)		Dataset 2 (Texas, USA)	
#michigan	#michiganprotest	#texas	#reopentexas
#liberatemichigan	#michiganlockdown	#opentexas	#stayhometexas
#freemichigan	#michiganshutdown	#texasstrong	#texans

Table 3: Popular hashtags posted on April 2020 linked to two locations, namely, Michigan and Texas. The depicted hashtags in the table involve two tokens, the first one associated with a location and the other with a noun/verb. The two types of tokens are shown in different colours - red and black.

light of this idea, the Alert Mechanism used in the preceding case was adopted to identify tipping points.

4.3.1. Michigan

Figure 13.I shows that the system would generate an alert on 14th April 2020, a day before protests began because the governor's "stay at home" order was declared, and five days before protests escalated (19th April 2020). The triggered alarm suggests that the internal cohesion amongst national security components has been disrupted.
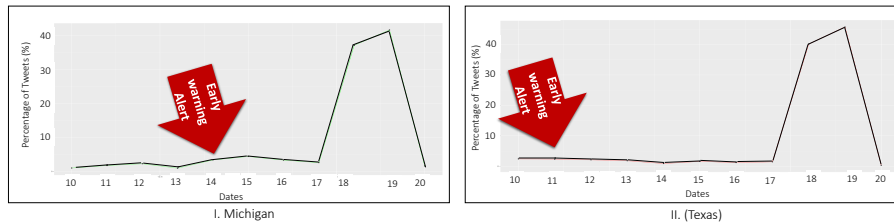

Figure 13: Early Warning Alert Detection in the events of Michigan and Texas in April 2020.

4.3.2. Texas

As depicted in Figure 13.II on 11th April 2020, an alert was triggered by the early warning process, eight days before protests against Coronavirus policies intensified (19th April 2020).

4.4. Radical Behaviour (Q2, Q3 and Q4)

The analysis of radical behavioural traits can lead to critical and actionable insights. Table 1 demonstrates how addressing Q2, Q3 and Q4 enables the identification of entities, behavioural traits and required objects; justifying the use of the radical behaviour analysis methodology proposed in our previous work [17] to enrich this part of the analysis.

4.4.1. Michigan

Q2. and Q3. In order to facilitate the narrative, Q2 and Q3 will be presented together. It can be seen in Table 4.I that on April 15th 2020 protesters were conveying messages about violating the lockdown as well as expressing disagreement against the measure. In contrast, two days later messages that expressed an intention to take to the streets were disseminated, coupled with messages that urged people to wear masks while protesting; moreover, messages suggesting the location of the protest, namely Michigan, were conveyed too.

Location	Dates (2020)							
	April 10	April 11	April 12	April 15	April 16	April 17	April 18	April 19
I. Michigan				Violating -> Lockdown		Stop -> Insanity	Cancel -> Lockdown	Avoid -> Quarantine
				Disagree -> Curfew		Take -> Streets	Protest -> Lockdown	Violating -> Distancing
				Protest -> Rally		Break -> Demand	Take -> Lockdown	Michigan -> Edict
				Protest -> Virus		Protest -> Michigan	Protest -> Distancing	Need -> Lawmaker
						Wear -> Masks	Demand -> Reopening	Break -> Curfew
						Shut -> Now	Liberate -> Lockdown	Want -> Cure
							Block -> Roads	Want -> PPE
							Rally-> Arizona	Take -> Streets
II. Texas		Allow -> Business	Halt -> Covid			Reopen -> Government	Spreading -> Frustation	Hoarding -> PPE
		Avoid -> Corona	Develop -> Diarrhea			Need -> Michigan	See -> Outrage	Open -> Quarantine
		Lifting -> Quarantine	Open -> Employment			Close -> Schools	Wear -> PPE	Observe -> Distancing
		Reopen -> Texas	Help -> Employees				Wear -> Facemask	Protesting -> Coronavirus
		Care -> Lives	Puts -> Halt				Authorizing -> Reopen	Violate -> Lockdown
		Rise -> Deaths						Protest -> Lockdown
		Want -> Nurses						Support -> Boycott
								Make -> Masks
								Rally -> Whattsapp
								Rally -> Austin Texas

Table 4: Disruptive Expressions extracted using Word Embeddings and Direct Object (Texas and Michigan).

On 18th and 19th April 2020, demands related to the lift of the lockdown

and messages urging to continue protesting against the imposed measures were spread. In addition, some other ideas were present, such as demands to reopen, lift the quarantine, and liberate from the lockdown. Incitement to actions affecting various public thoroughfares, such as blocking roads or taking to the streets, were also present.

Q4. Lastly, messages were individuals conveyed their personal needs for PPE (personal protective equipment), or urge for action towards a cure for COVID-19 were shared likewise, see Table 4.I.

4.4.2. Texas

Q2. Radical behavioural traits revealed that individuals expressed ideas linked to reopening a specific location, namely, Texas, see Table 4.II. According to the Levin's classification[33], the verb "need" expresses that a person desires something. Following Levin's analysis, on April 17th 2020, messages were posted conveying the desire that a different location, Michigan, would join the incident. As argued in [34], this mention of different locations, suggests that the state is dealing with a widespread event.

Q3. Social media users (Twitter) expressed concepts connected to the demand of allowing business in the city, lifting the quarantine, and contempt towards Coronavirus, as described in Table 4.II. On the other hand, in the following days (17th, 18th and 19th April 2020), messages that instigate violations of the lockdown, urge protest and boycott, close schools, wear PPE, and spread the frustration, were shared.

Q4. Concerns about health were also transmitted, related for example to the need for more nurses, and the rise of deaths.

4.5. Ideology (Q5)

The ideological traits of authoritarianism and hostility reveal important social characteristics. Authoritarianism denotes that individuals do not empathise with decisions or activities performed by those who hold the "proper authority"

23

[35, 36]. Hostility enables the identification of collective emotions which are seen whilst disruptive events take place [32, 37].

In order to begin the dissection of ideology in the COVID-19 datasets, a sentiment analysis process was performed, then tweets with negative polarisa-tion were selected accordingly. In both cases, negative sentiments played the predominant role; Michigan had the highest percentage with 51%, while Texas had 35%. Then authoritarianism and hostility traits were computed using the methodology and thresholds proposed in our previous study [19]. Consistently with our previous approach, when the calculated variables of authoritarianism (aggressiveness, submission and conventionalism) and hostility (anger, contempt and disgust) were above the predefined thresholds, the results were deemed to suggest that the aforementioned ideological traits were present.
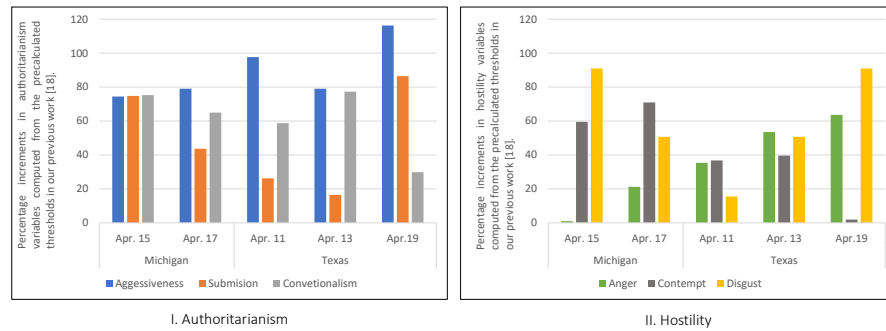

Figure 14: Ideological traits (Michigan and Texas).

4.5.1. Michigan

A day after the early warning alert was triggered (April 14th 2020), signs of authoritarianism and hostility were detected (April 15th 2020), the same date that the local government imposed the lockdown, see Figure 14.

4.5.2. Texas

On April 11th 2020, ideological traits were detected, the same date that the early warning alert was triggered, see Figure 14. This specific point turns into a modular axis, since people were concerned about competing aspects such as

24

the COVID-19 death toll, and lifting the quarantine, see the radical behaviour analysis in subsection 4.3 and Table 4.II.

Regarding authoritarianism, it should be noted that in both of the studied cases, aggressiveness was above 60% of the precalculated threshold, while the other two variables showed irregular increments. The consistent increase in aggressiveness suggests that people were conveying messages indicating prejudice/intolerance against a specific topic [36], here a lockdown, a curfew, or a quarantine, see Table 4.

4.6. Web Insights (Q6)

During an incident or a health crisis such as COVID-19, individuals and organisations use digital channels to disseminate information and data such as breaking news, messages or pictures, the analysis of which can help understand whether a crisis is escalating over time. Hence, as described in Section III, the web insights methodology proposed in our previous work [16] enables the study of the horizontal escalation of national security components. Following the methodology there, URLs were classified according to a comprehensive list of entities created over the Wikidata knowledge base. Then, a web scrapping process was conducted to retrieve the content of such web resources.

4.6.1. Michigan

It can be seen in Figures 15.I and 15.II, that only two media resources were embedded in people's messages while posting a tweet, namely Independent Websites (IW) and Social Networking Services (SNS).

On April 14th 2020, when the early warning alert was triggered, SNS (Instagram and Twitter) were used to convey that one national security component was being affected, in this case, health. One day later, messages posted on those social media sites showed that four national security components were unbalanced (information, defence, business and health). Such increment in the number of affected components (from one to four), demonstrates a horizontal

escalation, which according to [38] may represent a disruptive situation, see

₅₂₅ Figure 15.I.

It should be noted that both business and government components had the highest intensities, which may complement the behavioural traits previously extracted that referred to violating the lockdown and the disagreement towards that measure (see Table 4.I).

₅₃₀ On the other hand, IW showed on April 15th 2020, that three national security components were affected, namely defence, information and government, with government having the highest intensity figure. The result suggests that those web resources were providing a more detailed description of the government's activity (see Figure 15.II).
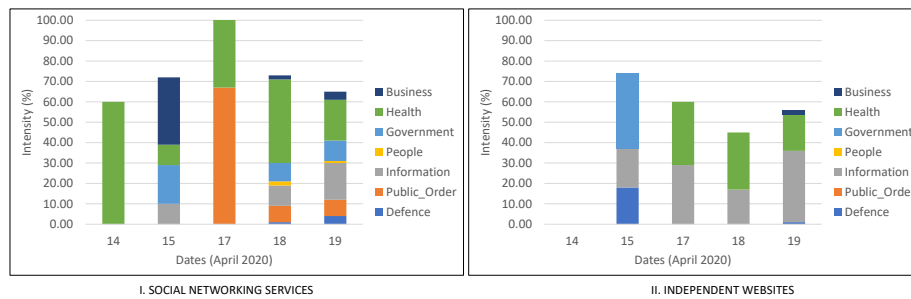

Figure 15: Horizontal Escalation of the National Security Components during the protests in Michigan (April 2020)

₅₃₅ On the following days (17th, 18th and 19th April 2020), both IW and SNS published content affecting the health and information components. The result is relevant since, on April 19th, COVID-19 cases began to spike [42]. By contrast, only SNS revealed information about two other components (people and public order), as displayed in Figure 15.I.

₅₄₀ *4.6.2. Texas*

Figure 16.I, 16.II and 16.III show that three digital web resources were used by people to disseminate information amidst the protests, namely, IW, SNS and Non-Profit Organisations.
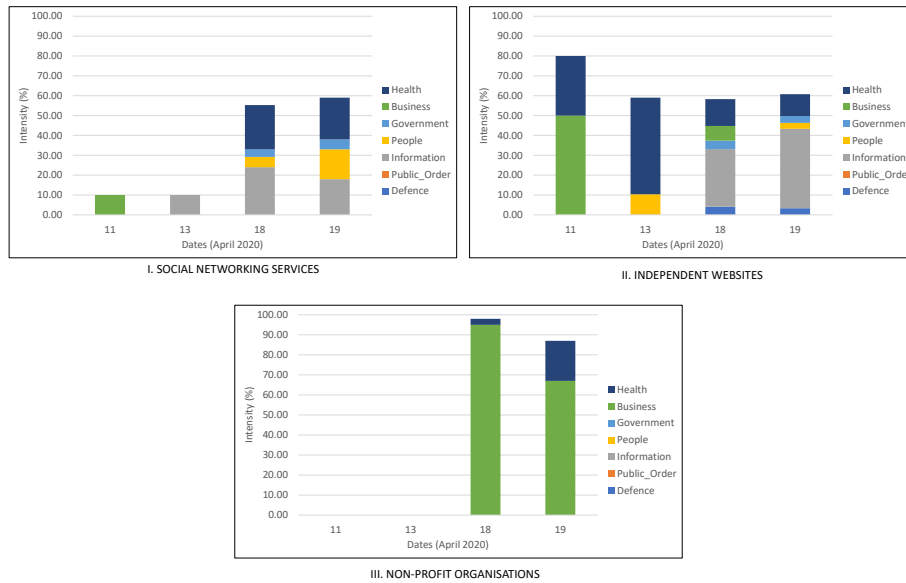
26

Figure 16: Horizontal Escalation of the National Security Components during the protests in Texas (April 2020).

As mentioned earlier, the early warning alert and the emergence of ideolog-
ical traits happened on the same date (11th April 2020). Unlike the previous
case, the Independent Websites were used more intensively and they unveil that
two national security components were disrupted, business and health; while
SNS showed that only the business component was affected, with 80% less in-
tensity than in IW.

Visible changes were displayed between 13th and 18th April 2020, as the IW
and SNS showed an increased number of affected components, which exposed a
horizontal escalation across the national security factors, which went from two
affected factors to five for the IW, and from one to four for SNS.

In addition, Non-profit organisations played a crucial role on the 18th and
19th April 2020, because topics in business and health were affected by them.
Moreover, intensity health levels had a considerable increase of 85%, while,
by contrast, health levels in SNS and IW showed little change, around 7% on
average. Such a difference indicates that Non-profit organisations were stressing

27

issues linked to health.

Finally, it should be noted that on April 19th, when the highest burst of online activity took place (see Figure 13.II), SNS were used to convey more messages linked to people, as indicated by an increase of 70%; whereas IW were focused on disseminating data related to the information component, which had an increment close to 28%.

5. Conclusions

As COVID-19 has so clearly demonstrated, pandemics constitute a severe National Security threat, and they can become breeding grounds for societal instabilities, which can affect multiple human security components. Given the complexity of analysing and detecting national security threats, and motivated by the importance of providing relevant information to facilitate the decision-making process, this paper has presented a holistic analytics framework for analysing national security aspects.

In the era of digital media, social networks and other communication channels use the Internet to disseminate information faster than ever before. They act as a multiplier of ideas, thoughts, beliefs, and, therefore, they disseminate ideologies that can turn into violent actions or societal disruptions. Quite often, such disruptions feed back into their initial trigger, as it is the case with the COVID-19 pandemic, where the infection rate can increase as a result of the aforementioned societal instabilities. In turn that leads to other human security elements been impacted in parallel, such as economic and food security.

This paper has discussed the utilisation of the framework for analysing national security aspects in the context of COVID-19. Two real-world cases were considered where authorities' measures to contain the disease via lockdowns resulted in protests and social unrest. The cases represent different cultural, political and socio-economic contexts. The results have demonstrated the ability of the system to provide early warning well in advance of the manifestions of social unrest. It also demonstrated its capacity to provide insights enabling

28

the better understanding and interpretation of the crisis.

Future work will fully integrate and automate the framework, utilising additional data analytics approaches, and making the various thresholds involved adaptive to the cultural and socio-economic context of the algorithm's deployment.

Acknowledgements

References

[1] Cárdenas P., Obara B., Ivrissimtzis I., Kureshi I., Theodoropoulos G.: Big Data for National Security in the Era of COVID-19. Computational Science – ICCS 2021. (2021).

[2] Kharroubi S, Saleh F. Are Lockdown Measures Effective Against COVID-19?. Front Public Health. 2020;8:549692. Published 2020 Oct 22. doi:10.3389/fpubh.2020.549692.

[3] Peretti-Watel, P., Seror, V., Cortaredona, S., Launay, O., Raude, J., Verger, P., et al. Attitudes about COVID-19 Lockdown among General Population, France, March 2020. Emerg Infect Dis. 2021;27(1):301-303. https://dx.doi.org/10.3201/eid2701.201377. (2021).

[4] Kureshi,I., Theodoropoulos, G., Magina,E., O'Hare, G., Roche,J., Towards An Info-Symbiotic Framework For Disaster Risk Management, the 19th IEEE/ACM International Symposium on Distributed Simulation and Real Time Applications (DS-RT 2015), Xinhua International Hotel, Chengdu, China, October 14 - 16. (2015).

[5] Toth T., Theodoropoulos G., Boland S., Kureshi I., Ghandar A.: Global Challenge Governance: Time for Big Modelling?, 2019 IEEE 18th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC), Milan, Italy, pp. 244-253. (2019).

[6] Peter J. Haas and Georgios Theodoropoulos. 2020. Introduction to the Special Issue for Towards an Ecosystem of Simulation Models and Data. ACM Trans. Model. Comput. Simul. 30, 4, Article 20 (December 2020), 3 pages. DOI:https://doi.org/10.1145/3425907

[7] Roff, H., Uncomfortable Ground Truths: Predictive Analytics and National Security. Brookings National Security Report. (2020).

[8] Akhgar, B., Saathoff, G., Arabnia, H., Hill, R., Staniforth, A., Bayerl, P., Application of Big Data for National Security: A Practitioner's Guide to Emerging Technologies. Butterworth-Heinemann, Paperback ISBN: 9780128019672, eBook ISBN: 9780128019733. (2015).

[9] Dolyle, A., et al. The EMBERS architecture for streaming predictive analytics, IEEE International Conference on Big Data, USA. (2014).

[10] ICEWS Homepage, `https://lmt.co/3owTj8L`. Last accessed 10 Jan. 2021.

[11] Muthiah S., Butler P., Khandpur R. P., Saraf P., Self N., Rozovskaya A., Zhao L., Cadena J., Lu C.-T., Vullikanti A., Marathe A. , Summers K., Katz G., Doyle A., Arredondo J. , Gupta D. K., Mares D., and Ramakrishnan N.: EMBERS at 4 years: Experiences operating an open source indicators forecasting system, Proceedings of the 22nd ACM SIGKDD. New York, NY, USA. (2016).

[12] Parang S. and Naren R.: EMBERS AutoGSR: Automated Coding of Civil Unrest Events. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16). Association for Computing Machinery, New York, NY, USA. (2016).

[13] Zhang X., Li J., Zhang X., Fu J., and Wang D.: Construction of a Geopolitical Environment Simulation and Prediction Platform Coupling Multi-Source Geopolitical Environmental Factors, Science and Technology Review 36, no, 3. (2018).

[14] Van Puyvelde, D. , Coulthart, S. and Hossain, M. S.: Beyond the buzzword: big data and national security decision-making. International Affairs, 93(6). (2017).

[15] Cárdenas P., Theodoropoulos G., Obara B. and Kureshi I.: Defining an alert mechanism for detecting likely threats to National Security. IEEE International Conference on Big Data. USA. (2018).

[16] Cárdenas P., Theodoropoulos G. and Obara B.: Web Insights for National Security: Analysing Participative Online Activity to Interpret Crises, IEEE International Conference on Cognitive Informatics and Cognitive Computing, Italy. (2019).

[17] Cárdenas P., Obara B., Theodoropoulos G. and Kureshi I.: Analysing Social Media as a Hybrid Tool to Detect and Interpret likely Radical Behavioural Traits for National Security, IEEE International Conference on Big Data. USA. (2019).

[18] Cárdenas P., Theodoropoulos G., Obara B. and Kureshi I.: A Conceptual Framework for Social Movements Analytics for National Security. The International Conference on Computational Science. (2018).

[19] Cárdenas P., Theodoropoulos G., Obara B. and Kureshi I.: Unveiling Ideological Features Through Data Analytics to Construe National Security Instabilities, IEEE International Conference on Big Data. USA. (2020).

[20] World Health Organization. WHO Coronavirus (COVID-19) Dashboard. World Health Organization Homepage `https://bit.ly/3pGWvCa`. Last accessed 28 Oct. 2021.

[21] Wilson, J., US lockdown protests may have spread virus widely, cellphone data suggests, The Guardian, 18 May. (2020).

[22] Human Development Report, United Nations Development Program: New York and Oxford: Oxford University Press, 22-33 (1994).

[23] Transforming our world: The 2030 Agenda for Sustainable Development, United Nationas Report A/RES/70/1, 2015.

[24] Thiago C. J.: Seguridad Nacional, Poder Nacional y Desarrollo. México: CISEN, (1991).

[25] Klarevas, L., Clarke, C., COVID-19 Is a Threat to National Security. Let's Start Treating It as Such, Just Security. (2020).

[26] Castillo C.: Big Crisis Data: Social Media in Disasters and Time-Critical Situations. Cambridge University Press. (2016).

[27] Marchezini V, Horita FEA, Matsuo PM, Trajber R, Trejo-Rangel MA and Olivato D (2018) A Review of Studies on Participatory Early Warning Systems (P-EWS): Pathways to Support Citizen Science Initiatives. Front. Earth Sci. 6:184. doi: 10.3389/feart.2018.00184

[28] Villagrán de León, J. C. (2012). "Early warning principles and practices," in Handbook of Hazards and Disaster Risk Reduction and Management, eds B. Wisner, J. C. Gaillard, and I. Kelman (Abingdon: Routledge),481–492.

[29] Villagrán de León, J. C., Bogardi, J., Dannenmann, S., and Basher, R. (2006). Early warning systems in the context of disaster risk management Entwickl. Ländlicher Raum. 2, 23–25.

[30] Bjerge B, Clark N, Fisker P, Raju E.:Technology and Information Sharing in Disaster Relief. PLOS ONE 11(9).(2016).

[31] Maslow, A. H. Authoritarian character structure. Soc. Psychol., (1943).

[32] Matsumoto D. et al.: The Role of Intergroup Emotions in Political Violence, Current Directions in Psychological Science, (2015).

[33] Levin B. English Verb Classes and Alternations: A Preliminary Investigation. University of Chicago Press, Chicago. (1993).

[34] Reed T., Belvin K., Hickman M., Kokolus M., Mendoza F., Phelps A. and Reisler K.: Open Source Indicators Program Handbook. The MITRE Corporation, USA. (2017).

[35] Saunders B.A., Ngo J.: The Right-Wing Authoritarianism Scale. In: Zeigler-Hill V., Shackelford T. (eds) Encyclopedia of Personality and Individual Differences. Springer, Cham.(2017).

[36] Dunwoody, P., Funke, F.. The Aggression-Submission-Conventionalism Scale: Testing a New Three Factor Measure of Authoritarianism. Journal of Social and Political Psychology, North America. (2016).

[37] Troost D., Emotions of Protest, Emotions in Politics. Palgrave Studies in Political Psychology series. Palgrave Macmillan, London. (2013).

[38] Patrick J. Cullen and Erik Reichborn-Kjennerud: MCDC Countering Hybrid Warfare Project:Understanding Hybrid Warfare. Multinational Capability Development Campaign. (2017).

[39] Gul, A., Pakistan Deploys Army to Deal With Coronavirus Outbreak, The Voice of America. 23 March. (2020)

[40] Jeffery, A., Scenes of protests across the country demanding states reopen the economy amid coronavirus pandemic, CNBC. 18 April. (2020).

[41] Hundreds protest COVID-19 orders at Texas Capitol, FOX4News Homepage `https://bit.ly/38IA1rl`. Last accessed 14 Jan. 2021.

[42] Hutchinson, B., Operation Gridlock': Convoy in Michigan's capital protests stay-at-home orders, ABCNews. April 16. (2020).

[43] Jasper, J.M.: The Emotions of Protest: Affective and Reactive Emotions In and Around Social Movements. Sociological Forum 13, (1998).

[44] H. N. Nguyen and T. Dang.:EQSA: Earthquake Situational Analytics from Social Media. IEEE Conference on Visual Analytics Science and Technology (VAST), (2019).

[45] Tan L., Ponnam S., Gillham P., Edwards [v1] B. and Johnson E..:Analyzing the impact of social media on social movements: A computational study on Twitter and the Occupy Wall Street movement. IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013), (2013).

[46] Statista Homepage. Most popular social networks worldwide as of January 2021, ranked by number of active users.https://bit.ly/3s5uYZy. Last accessed March 19 2021.