Pre-print version of Aziani, Jofre, Mancuso (2023) Convergence Between Migrant Smuggling and Trafficking of Goods: Text Analysis of Open-Source Data. Published version available at: https://journals.sagepub.com/doi/10.1177/01979183231200198

Convergence Between Migrant Smuggling and Trafficking of Goods: Text Analysis of Open-Source Data

Alberto Aziani,[1] Maria Jofre,[1] and Marina Mancuso[1]

[1]*Università Cattolica del Sacro Cuore and Transcrime, Milan, Italy*

Abstract

There is debate about the convergence of migrant smuggling and transnational trafficking of goods, the threat it poses to society, and its evolution in response to the COVID-19 pandemic. However, empirical evidence is limited. We tested for convergence using a methodology based on natural language processing. Results show that convergence occurs, but it is rare and episodic. Exploratory results suggest that COVID-19 containment policies had no impact on the simultaneous occurrence of migrant smuggling and transnational trafficking of goods. These findings expand our understanding of migrant smuggling and provide suggestions for policymakers to combat transnational crimes.

Keywords

Smuggling of migrants; Crime convergence; North Africa; COVID-19; Transnational Crime; Open-source data; Text analytics; Natural Language Processing

Introduction

News media, law enforcement agencies, international institutions, and political authorities often claim a convergence—interpreted as the merging of distinct types of

crimes (see Guerette and Aziani 2022)—between smuggling of migrants and other crimes of transit (e.g., trafficking of drugs, wildlife, firearms). These same sources assert that such crime convergence poses a serious threat to the political stability of countries in which migrants and illegal goods transit, to the internal security of countries of destination, and to the life of migrants (e.g., African Union Commission 2018; Europol 2017; Frontex 2021; INTERPOL 2018). For example, the European Migrant Smuggling Centre (EMSC 2018) reported that around 46% of organized criminal groups operating in the Mediterranean area are involved in human trafficking, drug trafficking, money laundering, and organized property crimes in addition to the smuggling of migrants.

If these claims on convergence are confirmed by facts, it means professional smugglers might generate large revenues by engaging in multiple trafficking activities, thus becoming more powerful and influential. Criminal networks might become also more resilient to police actions by diversifying their businesses: whether enforcement pressure targets one of the network's trafficking activities, smugglers might shift to another illegal market. Theoretically, the combination of increased revenues, efficiency, and flexibility would expand the influence and power of smugglers, as well as the magnitude of the illegal flows of people and of illegal goods reaching their destination countries.

In response to the alleged relevance of convergence, in the late 2010s and early 2020s, scholars started to investigate the convergence between migrant smuggling, organized crime, and terrorist activities (e.g., Achilli and Tinti 2019; Campana 2018; Izcara Palacios 2015; Procter 2021; Sanchez 2020; Yates 2021). In line with the bulk of research on migrant smuggling, these studies adopted qualitative methodological approaches to the study of convergence—particularly an ethnographic approach.

According to the results of these studies, there are links between migrant smuggling, organized crime, and terrorist activities. However, in most of the cases, these links do not manifest in the form of criminal groups or terrorist organizations providing illegal migration services, but deriving revenues from the extortion of migrant smugglers, who cross the territories controlled by criminal or terrorist groups.

This form of exploitation of migrant smuggling by organized crime and terrorist organizations does not emerge in every context; on the contrary, it depends on the characteristics of the groups and the contextual factors at play. For instance, on the one hand, Izcara Palacios (2015) observes that migrant-smuggling networks do not transport drugs from Mexico to the United States and criminal groups are not directly involved in migrant smuggling but do extort migrant smugglers who cross their territories. On the other hand, Campana (2018) finds no evidence of Sicilian mafia involvement in smuggling of migrants from North Africa to Sicily.

There is even less information available on the convergence between the smuggling of migrants and the trafficking of illegal goods. In contrast to the widespread claim of a significant and growing convergence, measurements and empirical analyses of this phenomenon are few (Achilli and Sanchez 2021; Andreas 2021). Most of the available studies on the convergence between smuggling of migrants and the trafficking of illegal goods focused on drug trafficking. Relying on 54 interviews with migrant smugglers active in Turkey, Demir, Sever, and Kahya (2017, 381) observe that "in general migrant smugglers only do illegal work in the migrant business and stay away from getting involved in other illicit activities in the underground world, such as dealing in narcotics or the smuggling of goods, and vice versa." In the words of a smuggler: "The migrant smuggler does the migrant job. He does not get involved in another business.

Migrant smugglers who operate on land do not even get involved in a smuggling operation at sea. Everyone has his own limits" (Demir, Sever, and Kahya 2017, 381). Studies analyzing other migration routes reached similar conclusions. Spener (2009) reviewed 197 cases of migrant smuggling across the Mexico-United States border and identified only one case in which drugs were seized. Chin (1999) investigated the convergence between heroin trafficking and smuggling of migrants from China to the United States by interviewing 300 migrants. Only one migrant reported being asked by smugglers to carry drugs.

We contribute to this body of literature by empirically testing whether there is convergence between migrant and goods trafficking whether and how the COVID-19 pandemic affected it. Unlike the available analyses, this study proposes a quantification of the degree of convergence between migrant smuggling and seven types of trafficking: cultural goods, drugs, fuel, medicines, cigarettes and other tobacco products, firearms, and wildlife. The study focuses on the convergence in North Africa and the Mediterranean considering the political relevance of migratory flows from North Africa to Europe.

In response to COVID-19, many countries deployed unprecedented measures to restrict the domestic movement of individuals and to strengthen border controls to reduce the spread of contagions. Policies intended to limit movements of people caused the shrinkage of legal pathways for migrants, and incentivized irregular migration solutions and the recourse to migration facilitators (Bird Ruiz-Benitez de Lugo 2021). As such, the emergence of COVID-19 provides a rare exogenous factor that allows us to explore whether increased border controls impact the convergence of transnational crimes.

To test whether crime convergence actually occurs, detailed information is required on multiple trafficking events that are likely to take place in different

jurisdictions. This requirement makes the use of commonly exploited datasets of criminal statistics produced by national and international institutions unfeasible (e.g., Individual Drug Seizure Cases by the UNODC). Indeed, these datasets usually provide information on only one type of criminal act (e.g., episodes of drug trafficking) without describing the relation with other trafficking activities (e.g., episodes of migrant smuggling).

To overcome these deficiencies in the available data, we present an original data gathering approach suitable for the analysis of complex phenomena such as the convergence of transnational crimes. The proposed methodology permits to collect and process data from newspaper articles available online to obtain information on hidden patterns and trends regarding the simultaneous occurrence of different crimes and phenomena. In doing so, the proposed method allows to collect data that are not available in existing datasets. The worldwide coverage of sources in multiple languages makes it possible to obtain information pertaining to countries that do not provide updated information to international institutions. In turn, this feature allows the identification and analysis of phenomena that cross borders and involve multiple jurisdictions.

This article is structured as follows. The first section presents the setting of the study: convergence between smuggling of migrants and the trafficking of goods in North Africa and the Mediterranean. The second section details the methodology adopted for the assessment of news articles, including all steps of processing and analysis, and discusses their limitations. The third section provides measures of the degree of convergence between migrant smuggling and trafficking activities considered within the study area. The fourth section discusses the results considering the current debate on crime convergence. The conclusions summarize the main messages of the study, suggest

applications of the proposed methodology to other domains, and outline some policy implications of the results of the research.

Background and setting

North Africa has become a hub for migrant smuggling activities for several reasons. On the one hand, the difficult social, economic, and political conditions of many sub-Saharan countries led to migratory flows intended to countries in North Africa and the Global North. The Global North represents the economically developed societies of Europe, North America, Australia, Israel, and South Africa, amongst others (Ekedegwa Odeh 2010). The demographic growth, high unemployment, widespread corruption, lack of political stability, and social conflicts have pushed many people to flee in search of better living conditions (Bel Haj Zerki 2009; Cherti, Balaram, and Szilard 2013; Kari, Malasowe, and Collins 2018). Factors like climate change, deforestation, and other environmental causes have also propelled migration (Borderon et al. 2019; Freeman 2017; Mpandeli et al. 2020). On the other hand, the opportunities to migrate legally in other continents, especially in Europe, have been reduced by the introduction of increasingly reinforced border controls and the selective and limited issuance of visas. Consequently, irregular migration has increased starting from post-1990, which has also led to a growth of migrant smuggling (Moreno-Lax and Lemberg-Pedersen 2019; Sarehane et al. 2009).

The proliferation of migrant smuggling and the presence of other established illicit economic networks in North Africa have stimulated research on the convergence of criminal activities in the area (Shaw 2017). Scholars and institutions have stressed a convergence between smuggling of migrants and other crimes, which are functional to

the smuggling itself, including corruption of officials, document falsification, and money laundering (EMSC 2019; Europol and INTERPOL 2016; Sarehane et al. 2009; UNODC 2019a). Others have pointed out a strong convergence between migrant smuggling and human trafficking since migrants are vulnerable to exploitation and have often been victims of human trafficking before leaving their country of origin (Scherrer 2019). In addition, the same criminals can be involved as offenders in both migrant smuggling and human trafficking (Davy 2017).

The relevant form of convergence for the present study is the one between migrant smuggling and the illegal trafficking of goods. A paper published in 2016 by the Institute for Security Studies highlighted that the strong connections between different economies and business sectors trace back to the involvement of Saharan and Sahel communities and tribes in both the transportation of licit goods (e.g., food staples, cement, fuel, cars, electronics, and mining equipment) and the smuggling of goods (e.g., cigarettes, narcotics, weapons). Also facilitating people movement across borders has been declared as a part of the livelihoods of these communities for centuries (Tinti and Westcott 2016). Despite this, the paper pointed out that migrant smuggling has become a lucrative business for people living in migration hubs when the flow of African people looking for a way to reach Europe passing through North African countries tuned into a growing and constant flow (Tinti and Westcott 2016). The Agadez region in Niger is an example of this dynamic. Indeed, Agadez was a key desert migrant smuggling hub in central Niger since it was used as corridor between Niger and Libya to facilitate the passage of migrants from sub-Saharan Africa to Europe (Tinti and Westcott 2016; Hoffmann, Meester, and Nabara 2017). Here the migration industry was a crucial source of income for many people, who found different job opportunities, for example as drivers or ghetto owners,

at least until the end of 2016, when migratory flows changed their routes due to the implementation of the 2015 Law Against Illicit Smuggling of Migrants (Hoffmann, Meester, and Nabara 2017). One report (2017) also stressed how members of the Tuareg and Toubou active in Agadez engage in trade in legal (e.g., cement, subsidized fuel, and food such as milk and tomato sauce) and illegal goods (e.g., weapons from Libya, and cocaine from Guinea, Niger, and Ghana). They started to be actively involved in the migration industry when some economic sectors declined (e.g., tourism, artisanal handicraft, uranium mining). This process benefited from the expertise that these groups had in smuggling illegal products, their ties to Libya, mostly the last node of the African route before Europe, and the strategic transit geographical position of Agadez (Hoffmann, Meester, and Nabara 2017).

More recently, some studies have highlighted the existence of a convergence of offences (i.e., the same offenders involved in more than one crime simultaneously), but without providing measurements and empirical demonstrations. According to one of these studies, criminal networks are expanding their business in different markets (Micallef 2019). They engage in migrant smuggling and other transnational crimes involving different goods, such as drugs, tobacco, weapons, wildlife, cultural goods, medicines, and fuel, due to the high profit margins of such goods. Other contributions pointed out that offender convergence has been enabled by criminal opportunities. For example, it is said smuggled migrants are often exploited for both trafficking and sale of drugs by drug traffickers due to the vulnerable condition of migrants (Sayech 2014). In addition, reports from some research institutes anecdotally highlighted that migrants can be used as “mules” to illicitly transport drugs or other goods (e.g., documents, weapons, works of art, archaeology) hidden in their internal or intimate organs (Mrabet 2018), as

is the case with firearms traffickers in Libya which became a major funder of migrant smuggling after the collapse of the Gadhafi regime in 2011 (Brik 2015).

Some other studies and institutional reports have emphasized a spatial convergence to highlight how criminals use the same routes to perpetrate different transnational crimes. In this regard, for example, drug traffickers in Libya and Algeria seem to exploit the same routes as human smugglers (Brik 2015). In addition, INTERPOL pointed out that criminals involved in the trafficking of stolen cultural goods may have moved such goods along the same routes used by human traffickers and migrant smugglers from West Africa (e.g., Mali) through North Africa (INTERPOL 2018). In this case as well, the existing studies do not provide measurements and empirical evidence, but are mainly anecdotal.

Overall, the economic, social, and political characteristics of the North African countries might enable the successful commission of different crimes and illicit trades, since they ensure a low probability of being detected by reducing the operational costs related to the crime organization (Bel Haj Zerki 2009; Cherti, Balaram, and Szilard 2013; Shaw 2017). Indeed, criminals can realize economies of scale to conduct different crimes, for example, by using the same vulnerable routes, expertise, infrastructure, and well-established criminal contacts. In turn, economies of scale boost their margins, thus increasing the benefits related to the commission of the crimes. Secondly, the historical involvement of tribes in the trade of different goods, both legal and illegal, strengthens the hypothesis (Hoffmann, Meester, and Nabara 2017; Tinti and Westcott 2016). Thirdly, the widespread presence of migrant smuggling and other transnational crimes in North Africa may foment convergence of crimes (African Union Commission 2018; UNODC 2019b). Apart from these possible explanations for the convergence between migrant

smuggling and the trafficking of goods, limited empirical evidence is available in this regard (Achilli and Sanchez 2021; Andreas 2021). Consequently, this paper aims to address this gap by testing whether, and to what extent, there is convergence between migrant smuggling and other transnational crimes. The focus will be on North Africa and the transnational crimes considered include trafficking of drugs, tobacco, weapons, wildlife, cultural goods, medicines, and fuel.

The paper also provides exploratory insights on the impact of the COVID-19 pandemic on this convergence during the years 2020 and 2021. Travel restrictions related to COVID-19 have not stopped migrant smuggling (United Nations 2020; UNODC 2020). These measures have made the commission of crime more difficult, at least in the short term, but the pandemic has exacerbated economic and social problems which, in turn, are likely to lead to increased demand for smuggling services in the absence of legal migration paths. In fact, border closures and other state-imposed mobility restrictions have been redirecting migrants towards more dangerous and expensive landscapes where humanitarian support and rescue are often not available (Bird Ruiz-Benitez de Lugo 2021; Sanchez and Achilli 2020; United Nations 2020; UNODC 2020). Therefore, the pandemic has also exposed migrants to forms of smuggling based on an increasingly widespread use of violence and abuse due to the exploitation of more risky routes and conditions, and to containment measures that forced them to remain for prolonged periods with smugglers without the possibility to receive assistance and support (UNODC 2021a; 2020). Increased controls and enforcement activities have also led to increased operational risks that may push smugglers with a lower risk appetite to abandon crime while also attracting organized crime groups that are more likely to exploit migrants for

ever greater profit (Bird Ruiz-Benitez de Lugo 2021). In this context, COVID-19 could also impact the convergence between migrant smuggling and other transnational crimes.

Methodology

To assess the convergence between migrant smuggling and other transnational crimes in the North Africa countries, we implemented a methodology based on Natural Language Processing (NLP) and text mining techniques using open-source data collected from online media articles. In doing so, we first collected a comprehensive sample of incidents of migrant smuggling covered by newspaper articles in English, French, and Arabic between the years 2011 and 2021. We then estimated the share of the number of episodes of smuggling of migrants in which convergence was detected out of the entire sample of smuggling incidents considered. Based on the information provided by the news themselves, we assessed the convergence at issue, and further aggregated the data collected by year and type of transnational crime.

Information on the intensity of law enforcement activities intended to limit transnational crime is rare (Aziani and Guerette 2022). Previous studies have shown that COVID-19 containment policies led to stricter controls against both legal and undocumented migration (Bird Ruiz-Benitez de Lugo 2021; Sanchez and Achilli 2020). Therefore, we draw on information about the spread of COVID-19 in early 2020 to explore how the intensification of border control impacts the convergence of transnational crimes.

Data

The data considered in this study is collected from news articles available in the Nexis® Metabase (LexisNexis 2022), the largest available repository of international, national, and local news sources, consisting of more than 85,000 sources in over 200 countries. The study focused on articles in English, French, and Arabic, the most widely spoken languages in North African countries and most used in printed media (Central Intelligence Agency 2020; Cengage 2020). We considered articles published during two time periods; first, from January 2011 to December 2019, and then from January 2020 to December 2021, which coincided with the first two years of the COVID-19 pandemic. Several limitations of secondary literature as sources of crime information have been highlighted by previous studies. These limitations are mainly related to the inability to accurately measure such vague and hidden phenomena as human smuggling and transnational crimes (Andreas and Greenhill 2011; Merry 2016). This would be especially true for newspapers that, due to editorial preferences, may be more incline towards more sensationalist or newsworthy events (Aziani et al. 2023; S. Chermak 1995; Ditton and Duffy 1983), and the possible presence of fake news, especially in the era of the COVID-19 pandemic (Gradoń 2020).

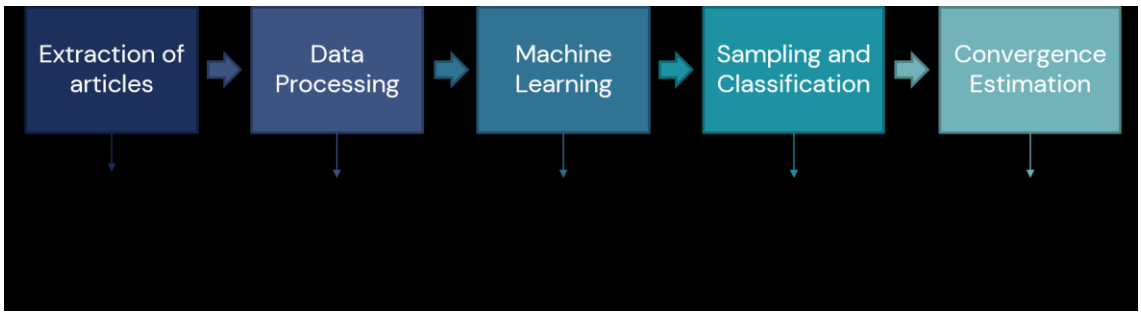
However, several studies focusing on the use of information collected from open-sources to analyze different crimes have highlighted the validity of open-sources for obtaining information regarding criminal events. For instance, scholars have observed that newspapers provide reliable information on homicide victims, perpetrators, and the context of the violent episode (J. M. Feldman et al. 2017; Ozkan, Worrall, and Zettler 2018; Parkin and Gruenewald 2017). In addition, advanced open-source data processing and crawling, along with the use of data and text mining techniques (Ghankutkar et al.

2019), have been shown to expand knowledge about the crimes under study, including human trafficking (Brewster et al. 2014; Brewster, Ingle, and Rankin 2014). NLP technologies have proven to be a popular choice for several different investigative purposes, including the detection of hate speech on social media (Schmidt and Wiegand 2017), investigation of white-collar crime based on financial fraud data (van Banerveld, Kechadi, and Le-Khac 2016), and the detection of criminal organizations from geo-location information (Osorio and Beltran 2020). The utility of NLP and related algorithms has been shown to be multilingual (Asharef et al. 2012).

Process

The process adopted in the present study consisted of the: (a) extraction of media articles from online open-sources based on *ad hoc* keywords; (b) removal of duplicates and articles identified as irrelevant; (c) sampling and manual classification of cases of interests using NLP techniques; (d) machine learning to identify relevant articles; and (e) convergence estimation based on statistical accuracy. The process adopted can be visualized in Figure 1.

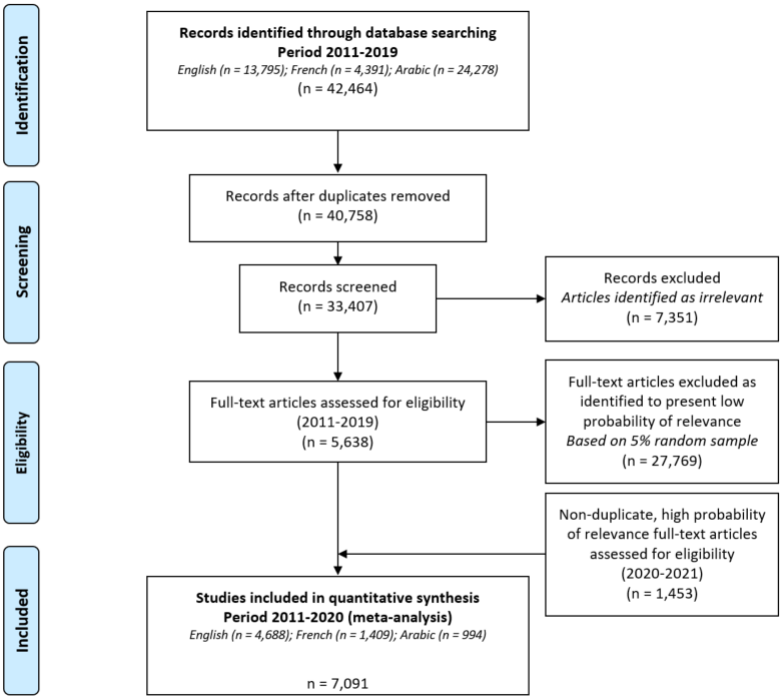
Figure 1: Different steps of the process adopted



The data flow resulting from the process can be depicted in the PRISMA flow diagram

presented in Figure 2 (Moher et al. 2009). The next five subsections describe the different steps in more detail.

Figure 2: PRISMA flow diagram of the data retrieval process



Extraction of online open-source data

The identification of articles was based on the different combinations of keywords related to episodes of migrant smuggling in the North Africa countries during the first period considered (i.e., 2011-2019). Due to the broad and ambiguous terminology often used by the media regarding migrant smuggling, and crime in general, the definition of multiple relevant keywords for the identification of such events required an iterative process. In this process, we manually inspected the articles resulting by the different keywords combinations; then, we selected those keywords that mapped to the most relevant and informative news.

The keywords employed were based on three interconnected elements that accounted for various aspects of the smuggling of migrants in North Africa and the Mediterranean (Table 1). The first element related to the concept of smuggling, that is, the illegal transportation of objects, substances, and/or people across one or more international borders. The second element dealt with the subjects that were transported, in this case, migrant human beings. The third element related to geographical locations referring to the area of interest, namely North African countries. We defined a set of keywords for each of these three elements, first in English, and then translated and adjusted into French and Arabic. We further generated queries composed by different combinations of the selected keywords, and used them to extract news articles that presented one or more of these combinations within their titles and/or contents. More details can be found in Annex 1.

Table 1: Keywords combinations (English)

Pre-print version of Aziani, Jofre, Mancuso (2023) Convergence Between Migrant Smuggling and Trafficking of Goods: Text Analysis of Open-Source Data. Published version available at: https://journals.sagepub.com/doi/10.1177/01979183231200198

Illegal market	*Keyword*	*English*	*French*	*Arabic*
fuel	gasoline	gasoline	gazole OR gasoil	بنزين
fuel	petrol	petrol	essence	نفط
fuel	fuel	fuel	carburant OR combustible	الوقود
fuel	hydrocarbon	hydrocarbon	hydrocarbure	الهيدروكربون
drug	drug	drug	drogue OR stupéfiant	المخدرات
drug	hashish	hashish	haschisch	الحشيش
drug	cannabis	cannabis	cannabis	القنب الهندي
drug	coca	coca	coca OR cocaïne	الكوكايين
drug	amphetamine	amphetamine	amphétamine	الأمفيتامين
drug	methamphetamine	methamphetamine	méthamphétamine	الميثامفيتامين
drug	opiate	opiate	opiacé	أفيون
drug	opium	opium	opium	NULL
drug	opioid	opioid	opioïde	NULL
drug	heroin	heroin	héroïne	الهيروين
tobacco	tobacco	tobacco	tabac	التبغ
tobacco	cigar	cigar	cigare	سيجار
weapon	gun	gun OR handgun OR shotgun	pistolet OR arme à feu OR arme de poing	بندقية
weapon	weapon	weapon	arme	سلاح
weapon	revolver	revolver	revolver	مسدس
weapon	pistol	pistol	pistolet	NULL

weapon	rifle	rifle	fusil	بندقية
weapon	kalashnikov	kalashnikov	kalashnikov	كلاشينكوف
weapon	ammunition	ammunition	munition	الذخيرة
wildlife	wildlife	wildlife	faune	الحيوانات البرية
wildlife	animal	animal	animal	حيوان
wildlife	plant	plant	plante	غرس OR نبات
wildlife	forest	forest	forêt	غابة
wildlife	flora	flora	flore	نباتية
wildlife	fauna	fauna	faune	حيوانات
wildlife	sugar	sugar	sucre	سكر
wildlife	charcoal	charcoal	charbon	فحم
wildlife	diamond	diamond	diamant	ألماس
wildlife	gold	gold	NULL	الذهب
wildlife	mineral	mineral	minéral	غازي OR معدني
wildlife	cobalt	cobalt	cobalt	الكوبالت
wildlife	coltan	coltan	coltan	الكولتان
wildlife	copper	copper	cuivre	نحاس
wildlife	timber	timber	NULL	الأخشاب
wildlife	rosewood	rosewood	bois de rose OR palissandre	خشب الورد
wildlife	ivory	ivory	ivoire	ناب الفيل OR العاج
wildlife	pet	pet	animal de compagnie OR animal domestique	حيوان أليف

wildlife	tusk	tusk	NULL	ناب
wildlife	horn	horn	corne	قرن
wildlife	elephant	elephant	éléphant	الفيل
wildlife	rhino	rhino	rhino	وحيد القرن
wildlife	bird	bird	oiseau OR volatile	عصفور OR طائر
wildlife	parrot	parrot	perroquet	ببغاء
wildlife	lion	lion	lion	أسد
wildlife	cheetah	cheetah	guépard	الفهد
wildlife	leopard	leopard	léopard	NULL
wildlife	crocodile	crocodile	crocodile	التمساح
wildlife	antelope	antelope	antilope	ظبي
wildlife	gazelle	gazelle	gazelle	غزال
wildlife	snake	snake	serpont OR guivre	ثعبان OR أفعى OR حية
wildlife	leather	leather	cuir	جلد OR جلد مدبوغ
wildlife	poach	poach	braconner	الصيد الجائر
cultural artefact	artwork	artwork OR artistic OR work of art	création artistique OR artistique	عمل فني
cultural artefact	antiques	antique	antiquité OR ancienneté	العصور القديمة
cultural artefact	heritage	heritage	patrimoine culturel OR héritage	التراث
cultural artefact	cultural property	cultural property	propriété culturelle OR objet culturel	ملكية ثقافية

cultural artefact	artefact	artefact OR artifact	artefact OR object fabriqué	قطعة أثرية
cultural artefact	archaeological	archaeological	archéologique	أثري
cultural artefact	relic	relic	relique OR vestige	بقايا OR أثر قديم شئ قديم
cultural artefact	vase	vase	vase	مزهرية
cultural artefact	ceramic	ceramic	céramique	خزفي
cultural artefact	sculpture	sculpture	sculpture	النحت
cultural artefact	statue	statue	statue	تمثال
cultural artefact	painting	painting	tableau OR peinture	لوحة فنية
cultural artefact	canvas	canvas	toile OR canevas	لوحة قماشية
cultural artefact	amphora	amphora	amphore	أمفورة
cultural artefact	mask	mask	masque	كمامة OR قناع
cultural artefact	drum	drum	tambour	طبل
cultural artefact	textile	textile	textile	النسيج
cultural artefact	object of worship	object worship	objet de culte OR objet d'adoration	موضوع عبادة
medicine	medicine	medicine	medécine OR médicament	الدواء
medicine	pharmaceutical	pharmaceutical	pharmaceutique	صيدلي OR الأدوية
medicine	antimalarial	antimalarial	antipaludique OR antipaludéen	مضاد للملاريا

Removal of duplicates and irrelevant articles

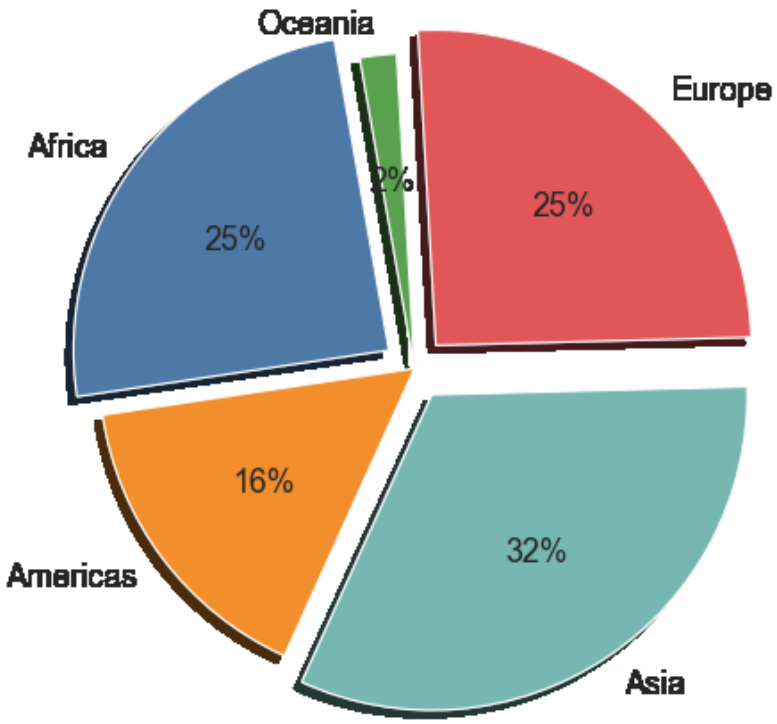
The extracted dataset presented duplicates, that is, more than one article evidencing the same event of migrant smuggling. We considered duplicate articles when they had the same or very similar text bodies, where the content of the text is simply an ordered combination of words. Whenever a keyword was present in an article, regardless of language, a boolean column related to that word was set to one, otherwise zero. Then, a list of binary elements was created by aggregating all boolean variables. To detect duplicate articles, therefore, we transformed the text of each article into a list of binary word-elements, and further generated a variable measuring the similarity between all lists (i.e., all news articles) as a ratio. The higher this ratio, the more similar the articles were, therefore when the ratio was equal to the maximum value of one, the two articles were identical. We set a filtering threshold of one, so all articles presenting 100% similarity to other article(s) were excluded.

We continued with the identification and removal of irrelevant articles, that is, instances that met the keywords criteria but were either unrelated to the issue, or did not provide sufficient information about the incident. These cases included brief summaries of news, highlights of the day, references to books, movies, or series, and reports of stocks and markets, among others. We used specific stop words, selected during the keyword iteration process, to identify and exclude these types of articles in all three languages. There were also cases in which the title of the news met the query criteria, but the content of the article consisted of only a few words, which did not provide minimal information about the incident, and were therefore excluded from the sample.

As result, we were able to identify and extract 33,407 non-duplicate, potentially relevant articles: 12,142 in English, 4,309 in French, and 16,956 in Arabic. Extracted items were retrieved from 1,715 sources mainly from Asia—mostly Saudi Arabia, United

Arab Emirates, and Jordan—Africa—predominantly Egypt—, and Europe—mostly the United Kingdom and France (Figure 3). The data did not distribute evenly over time as more articles were published in the years 2015-2019 (Figure 4 (a)).

Figure 3: Regions of sources of considered news articles

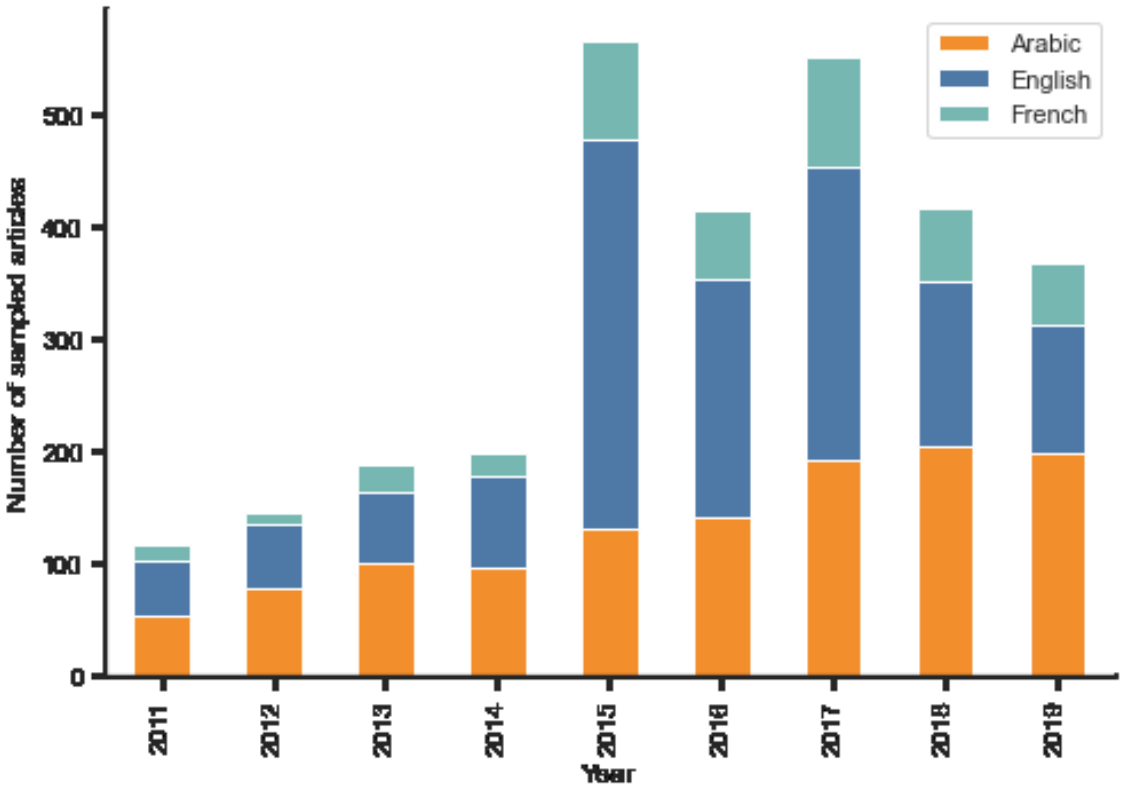


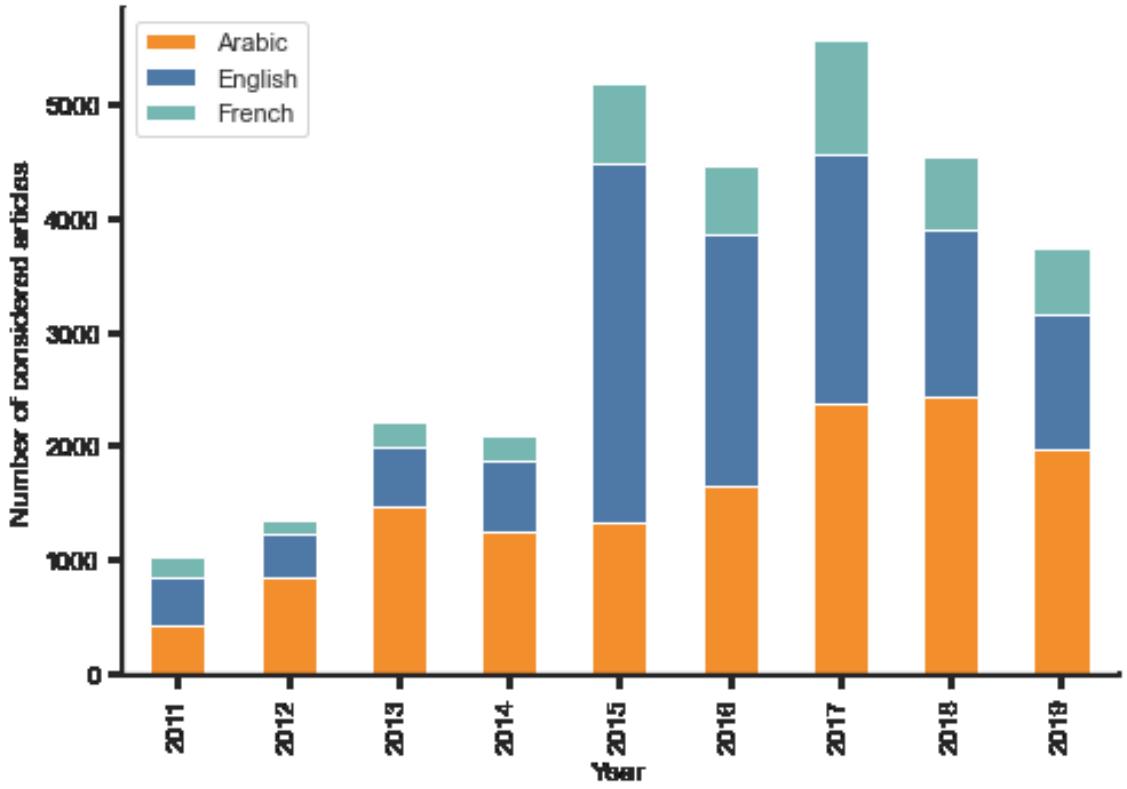
Sampling and classification of cases of interest

A smaller and more workable sample of articles was required to proceed with the manual classification of the news articles considered to verify that the identified articles actually reported information on migrant smuggling incidents in the region of interest. Consequently, a 10% stratified random sample based on the distribution of most relevant keywords was performed for each language, resulting in 1,334 observations in English, 438 in French, and 1,189 in Arabic. The yearly distribution of sampled articles followed a remarkably similar pattern to that of the considered articles (Figure 4 (b)), which is a sign of their temporal representativeness. The sample was read and manually classified as *positive* or *negative*, that is, each article was classified as being a relevant migrant smuggling incident or an irrelevant one. For the case study, articles were classified as

positive if they referred to migrant smuggling incidents (e.g., drowning or rescuing of migrants from shipwrecks, interception of boats carrying irregular migrants), and involved at least one North African country. On the contrary, generic and commentary articles in which migrant smuggling was only broadly mentioned in other contexts and articles in which information was reported by law enforcement agents or other witnesses without reference to any specific episode were labelled as *negative*. For every article classified as positive, we manually confirmed that the public link to the online source actually existed and was not a fake website.

Figure 4: Distribution of articles by year of publication and language





Note: Figure 4 (a) depicts the distribution for the entire population of identified and extracted articles. Figure 4 (b) depicts the distributions for a 10% stratified random sample extracted retrieved from the overall population of articles.

To facilitate the classification task, the sampled articles were further processed to identify different elements relevant to the research at hand. In doing so, we used NLP techniques to generate variables related to different entity types, including geolocations, such as countries, cities, regions, and states, and names of people and criminal groups. NLP algorithms were employed as implemented in the SpaCy library (Honnibal and Montani 2017). As result of the classification process, 425 articles were identified as being *positive*, 59.5% in English, 33.7% in French, and 6.8% in Arabic. Positive articles were ultimately exploited to expand the estimations to a larger sample, and further assess convergence.

Machine learning

We used the classified sample to expand the classification analysis to a larger sample of articles by means of machine learning. This larger sample, hereinafter *likely positive*, included non-duplicate, likely relevant articles from two time periods: (a) same period of the sample (i.e., 2011-2019), and (b) the two subsequent years that coincided with the COVID-19 pandemic (i.e., 2020-2021). The *likely positive* sample was processed and further exploited to predict out-of-sample articles that are more likely to account for migrant smuggling incidents in North Africa. The first step of data transformation was based on standard text processing algorithms, while the second step of prediction was performed by means of Naïve Bayes' modelling.

The first part of text analysis required processing the raw data to convert text bodies into a format that is amenable to analysis (R. Feldman and Sanger 2007). The four steps for text processing included (1) tokenization of articles by splitting the article's body into a list of words, (2) removal of uninformative punctuation, (3) removal of stop words that are not useful for discrimination, and (4) stemming and lemmatization that convert words to a root form, hereinafter referred to as *token*. All algorithms were employed as implemented in the Scikit-Learn library (Pedregosa et al. 2011).

After processing the data, we proceeded with the prediction exercise by identifying most discriminative tokens. A token is 'discriminative' when it helps differentiate between *positive* and *negative* instances, so it can be used to make predictions about new data. For this end, we randomly divided the classified samples of each language into training (80% of the articles) and test (20%) sets. Most relevant tokens were first identified using univariate Naïve Bayes' classifiers, and then ranked in terms of explanatory performance—i.e., proportion of correct classification of both *positive* and

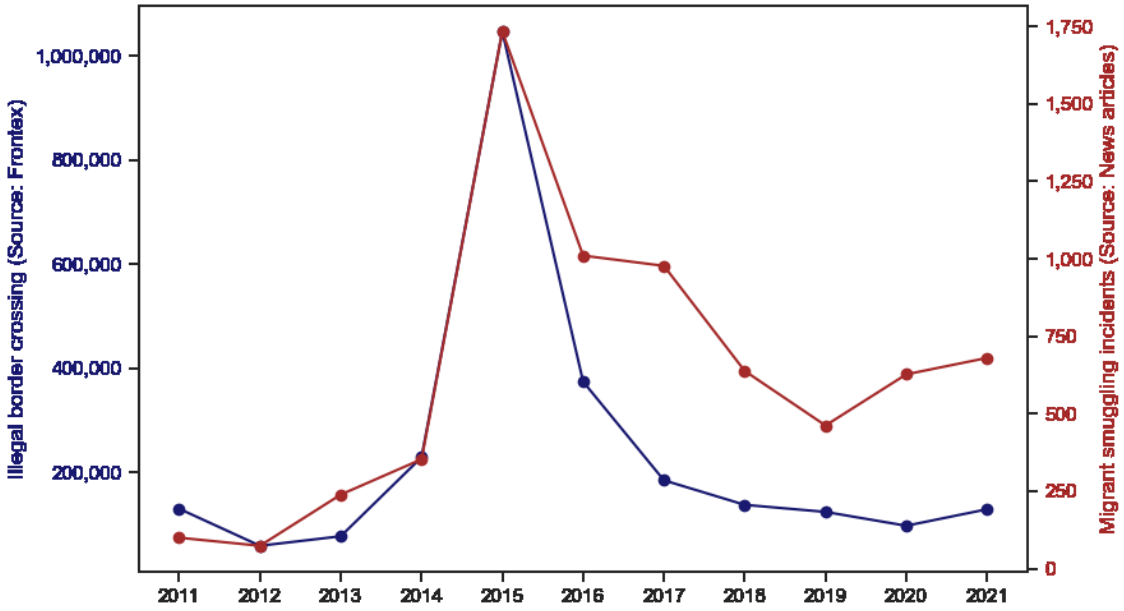
negative instances. We employed a 5-fold cross-validation on the training data to find the optimal number of discriminatory tokens for each of the languages considered based on different accuracy measures, including recall (i.e., proportion of correct *positive* classifications compared to the total number of cases), precision (i.e., proportion of correct *positive* classifications compared to the total number of true positives), and F1-score (i.e., harmonic mean of recall and precision).

The probability of an article being *positive* was therefore calculated based on the presence or absence of relevant tokens within its text body. Different probability thresholds were assessed as to identify the optimal value to be used for each language. Multivariate Naïve Bayes' classifiers were fitted this time using the training set, and evaluated on the test data by means of predictive accuracy using same metrics of recall, precision, and F1-score. In comparison to the traditional 50% approach, the optimal probability thresholds were found to be 80% for the English sample, 60% for French, and 70% for Arabic. Consequently, all out-of-sample articles (i.e., extracted articles not considered in the classification exercise from both time periods) presenting a probability of *positive* higher than these optimal thresholds were included in the final dataset of most likely positive articles, along with the articles that were already classified as positives since we know that they were 100% likely to be relevant. The final *likely positive* dataset comprised a total of 7,091 articles, out of which about 66% were in English (4,688), almost 20% in French (1,409), and the remainder 14% in Arabic (994).

To assess the quality and consistency of the news estimates of migrant smuggling incidents, we contrasted the results with official figures related to the latest data provided by Frontex. In particular, we considered data on illegal border crossings detected at the external borders of the European Union Member States (Frontex 2022), which in turn coincided with our area of interest. Human smuggling incidents and illegal border

crossings were found to behave similarly over time, both peaking in 2015 (Figure 5). The discrepancy between the two lines is more pronounced as of 2016, probably because human smuggling in North Africa and the Mediterranean area continued to attract media attention in the years following the Refugee and Migrant Crisis of 2015 (Berry, Garcia-Blanco, and Moore 2016), despite the reduction in the actual number of illegal border crossings. It can be said, however, that despite this difference, the trend of both figures is similar.

Figure 5: Illegal border crossing and human smuggling incidents (2011-2021)



Convergence estimation

For the estimation of convergence between migrant smuggling and crimes of transit, we focused the analysis on the sample of articles that were identified as *likely positive* as to explore evidence of simultaneous occurrence with other transnational crimes, and if so, which illegal markets were involved (i.e., cultural goods, drugs, fuel, medicines, tobacco, weapons, wildlife). In doing so, we assessed, for each crime category, the number of

articles that contained at least one keyword of the crime category at issue (refer to Annex 2 for more details). The hypothesis was that the co-occurrence of keywords related to both migrant smuggling and other transnational crimes within the same article was a proxy of convergence between the two phenomena.

Nevertheless, the co-occurrence of these keywords in a news article may not imply an actual convergence incident. News articles relating to criminal events often contain little information about the incident under consideration, because the information collected by reporters is based primarily on statements by law enforcement officials. These statements may include claims that smugglers are known to be part of organized crime groups or are effectively controlled by such groups. As a result, terms related to trafficking of illegal goods (e.g., “kalashnikov”, “amphetamine”) may appear in incident reports where no concrete evidence of this alleged connection has been found.

To circumvent this issue, a final statistical correction of the convergence estimate values was performed based on the aforementioned precision metric which, in this context, measured the proportion of predicted *positive* articles that referred to an incident involving convergence between migrant smuggling and other transnational crimes. To compute crime-specific precision values, a random 5% sample of the *likely positive* dataset was selected, read, and manually classified. This time, the classification focused on identifying convergence incidents. Specifically, we aimed to distinguish incidents that involved both migrant smuggling and other transnational crimes. The aim was to distinguish these incidents from mere statements made by authorities about migrant

smuggling organizations being involved in other types of smuggling activities.[1] The previously estimated convergence frequencies for each crime category were multiplied by their respective precision rate. As result, it was possible to estimate which transnational crimes were more frequently committed in combination with migrant smuggling in the North Africa area, as well as how the convergence has changed over time.

Limitations and caveats

Several limitations are present in the study. The first is related to the reliability of the data sources. Although several studies have stressed the various limitations of open-source data, in particular news from the media, many others have highlighted the validity of such sources for obtaining information regarding criminal events. However, we cannot ignore the fact that newspapers may be more prone to events of greater popular interest, as well as the possible presence of fake news. To overcome these potential biases, several countermeasures were put in place. Firstly, a vast number of online articles were collected from numerous international, national, and local sources, thereby increasing the likelihood of reaching a diverse portfolio of news. Secondly, different languages were included in the analysis, thus making it possible to consider different political and geographical perspectives on the incident at issue. Finally, the sources of news articles classified as positive were manually verified in an attempt to validate that the related websites were not bogus, thus reducing the likelihood of fake news. The additional contrast between news incidents and illegal border crossing should also serve to support

[1] A sample of 5% provided a sample size that is sufficient to have a confidence level of 95% that the real value is within ±5% of the measured value.

the claim that fake news does not unduly affect our estimates. More can be done to detect and mitigate fakes news and disinformation, including the use of NLP-based solutions and other text mining techniques (Kozik et al. 2022).

Another limitation of the proposed methodology concerns the risk of misclassification given the manual hence human element involved in the classification task. In this regard, we automatized the data processing as much as possible by using NLP technologies as to facilitate the classification of articles. Moreover, the validation of articles predicted as *positive* by the text mining exercise allowed the reduction of biased outcomes that may result from the extensive media coverage granted to specific crimes (e.g., drug trafficking).

The risk of duplicate articles, that is, multiple articles accounting for the same migrant smuggling event, is another inherent limitation of the data. To counter this issue, we first generated a variable indicating the similarity-ratio between the different articles, and further excluded identical copies. We then ordered the remaining articles by date, which significantly helped in the identification of non-identical duplicates during both classification and text mining exercises. As a result, we can confidently state that duplicate articles represented a negligible proportion of the processed dataset.

The final limitation of the proposed methodology concerns the operationalization of the concept of crime convergence. First, and as previously discussed, we assumed that the co-occurrence of keywords related to migrant smuggling and other forms of illegal trade within the same article was representative of the existence of convergence between these crimes. In this respect, it is impossible to fully ensure that all these incidents corresponded to actual cases of convergence. Nonetheless, two measures were implemented to support the validity of this assumption, including the assessment of selected keywords related to the different crime categories during all manual content

exercises together with estimation of crime-specific precision rates for convergence correction. Finally, it should be noted that news articles allow us to identify events that were actually detected, but do not make it possible to verify that an event (e.g., crime convergence) that was not reported in the news did not occur, which might introduce a bias in the estimation of convergence. However, additional evidence of high similarity between news incidents and illegal border crossing should ensure that this bias is not a critical issue, and therefore does not erode the validity of our estimates.

Results

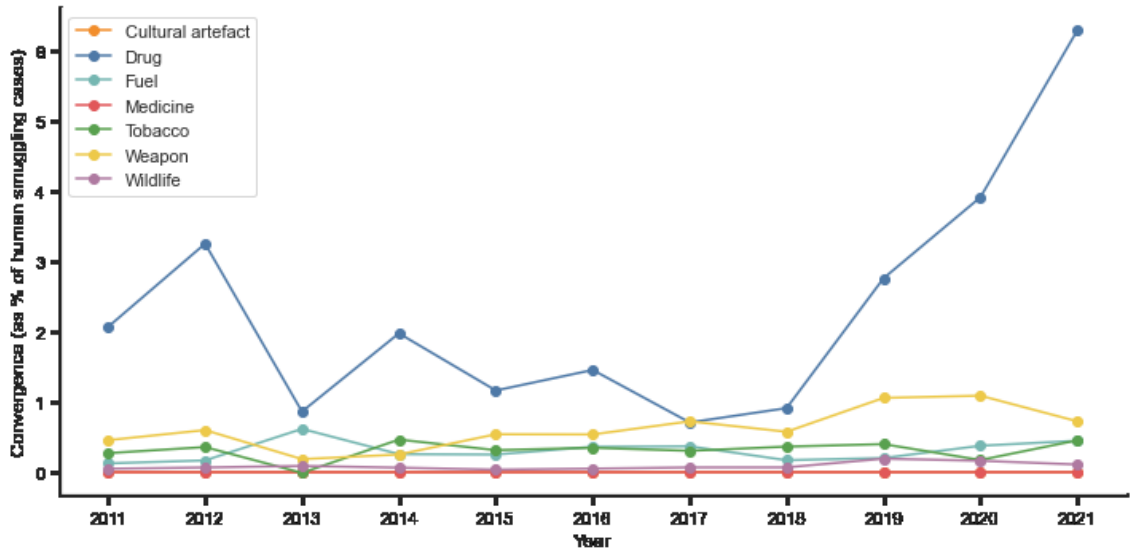
Results showed evidence of a marginal and episodic convergence between migrant smuggling and other transnational crimes given the overall proportion of the identified convergence incidents considering all years (2011–2021), that is, 3.4%. This figure fluctuates over time, presenting the highest value in the last three years, i.e., 4.7% in 2019, 5.8% in 2020, and 8.1% in 2021 respectively (Table 2).

The degree of convergence varies across trafficked goods. The highest estimate of convergence is related to drug trafficking, which emerged in 2% of migrant smuggling cases across the studied period (Table 1 and Figure 6). Convergence with drug trafficking peaked during 2012 and again in 2019; then it continued to increase in 2020 and in 2021. To a lesser extent, convergence with weapons, fuel, and tobacco trafficking can also be identified. The convergence with weapons presented stable figures throughout the period considered, with a small peak during 2019 and 2020. Lastly, no evidence of convergence emerged between the smuggling of migrants and the trafficking of cultural goods and medicines.

Table 2: Convergence as percentage of migrant smuggling cases by category of transnational crime (2011–2021)

Illegal market	*2011*	*2012*	*2013*	*2014*	*2015*	*2016*	*2017*	*2018*	*2019*	*2020*	*2021*	***Total***
Drug	2.1	3.3	0.9	2.0	1.2	1.5	0.7	0.9	2.8	3.9	6.3	**2.0**
Weapon	0.5	0.6	0.2	0.3	0.5	0.5	0.7	0.6	1.1	1.1	0.7	**0.7**
Fuel	0.1	0.2	0.6	0.3	0.3	0.4	0.4	0.2	0.2	0.4	0.5	**0.3**
Tobacco	0.3	0.4	0.0	0.5	0.3	0.4	0.3	0.4	0.4	0.2	0.5	**0.3**
Wildlife	0.1	0.1	0.1	0.1	0.0	0.1	0.1	0.1	0.2	0.2	0.1	**0.1**
Cultural good	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**0.0**
Medicine	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**0.0**
Total	*3.1*	*4.6*	*1.8*	*3.2*	*2.3*	*2.9*	*2.2*	*2.2*	*4.7*	*5.8*	*8.1*	*3.4*

Figure 6: Convergence as percentage of migrant smuggling cases by category of transnational crime (2011–2021)



Discussion

Convergence between smuggling of migrants and crimes of transit related to illegal markets occurred in 3.4% of all considered cases. Therefore, the empirical results indicate

that convergence between smuggling of migrants and other transnational crimes does occur. However, far from being systemic, the convergence between migrant smuggling and trafficking of goods appears to be episodic. Indeed, data show that most of the time migrant smuggling occurs independently from other trafficking activities. Overall, this result expands on the findings of ethnographic studies that suggest a separation between migrant smuggling and trafficking activities perpetuated by organized criminal groups (Achilli and Sanchez 2021; Chin 1999; Demir, Sever, and Kahya 2017; Mancuso and Maldi 2022; Procter 2021; Spener 2009; Yates 2021).

This result could be explained by the fact that the presence of migrants risks to attract more attention from control authorities, thus increasing the probability that law enforcement will also detects illegal goods (Sanchez and Zhang 2018). Consequently, the combined smuggling of people and goods would expose smugglers to a greater risk of seizure of the goods, and eventually arrest. The increased risk of apprehension due to the simultaneous involvement in migrant smuggling and in the trafficking of goods is relevant for the migration from North Africa to Southern Europe, which mainly occurs by sea. Indeed, every year between 2011 and 2021, hundreds of boats were detected by law enforcement authorities, or were rescued by non-governmental agencies (Cusumano and Villa 2021). This high attention to migrant boats would be clearly negative for the trafficking of goods.

Additionally, migrants, being aware of the traffic of illegal goods, could expose traffickers to further risks acting as witnesses of the trafficking activities. This possibility might reduce the independence and the bargaining power of the traffickers/smugglers towards migrants. The low frequency of actual episodes of convergence suggests that the increase in margins generated by combining multiple activities hardly compensates for the plausible increase in the risk of detection caused by dealing with migrants.

The augmented risk of detection is not the only reason that might explain the rarity of convergence between the considered crimes. The provision of smuggling services requires the development of a relationship between smugglers and migrants, who entrust smugglers with their very lives (Campana and Gelsthorpe 2020; İçduygu 2018). At the same time, managing smuggling operations over long distances is hardly efficient due to the need to control migrants, and deal with police agents in unfamiliar and remote contexts (Campana 2020). These characteristics of migrant smuggling push most smuggling operations to assume a local dimension as migrant smugglers tend to operate in a small area (Aziani 2023; Campana 2018; Majidi 2018; Neske 2006). On the contrary, most of the goods considered in this study are often shipped over long distances from their country of origin to their final markets—e.g., wildlife, opiates, coca derivatives. In the most extreme cases, such as illicit cigarettes, traffickers establish artificially complex routes to make it difficult for authorities to trace the illegal origin of the products (Meneghini, Aziani, and Dugato 2020). Therefore, in this respect, it is the distinct nature of migrant smuggling and the trafficking of goods to limit the intensity of convergence between these activities. Individuals or criminal groups who have the knowhow to operate along relevant smuggling routes (e.g., Amazighs groups along the trans-Saharan trade routes), or who have connections with border authorities and other enforcement corps in relevant trafficking hubs (e.g., Libyan armed militias), might exploit these leverages for multiple traffics without them occurring simultaneously in a same episode.

While the overall convergence between the smuggling of migrants and the trafficking of goods is rare, it recurs more often with some transnational crimes than with others. Convergence emerges with respect to the trafficking of illegal drugs, firearms, tobacco products, fuel, and rarely wildlife, but it does not with respect to cultural goods and medicines. The prominence of the convergence between migrant smuggling and drug

trafficking might partially be due to a higher frequency of the trafficking of drugs in comparison to the trafficking of other goods. Nonetheless, also the distinct nature of the trafficked goods and of the methods used to traffic might contribute to explain these differences in the convergence rate. For instance, while smuggled medicines flow primarily from advanced economies to developing countries, migration mostly goes in the opposite direction. Moreover, the trafficking of most medicines requires the transport and storage of the medicines in refrigerated conditions (Riccardi, Dugato, and Polizzotti 2014). The environmental conditions encountered in the Sahara or the Mediterranean routes, through which migrants move, complicate the preservation of the medicines at the right temperature and humidity. This is not the case for drugs or firearms. Finally, the trafficking of cultural goods often functions alongside—or it is embedded in—apparently legitimate market-operations (Mackenzie et al. 2020). Therefore, the methods through which the transnational trafficking of cultural goods is performed do not align with the most widespread modalities of migrant smuggling.

Lastly, the exploratory analysis of the temporal variations in the incidence of convergence shows no manifest pattern until 2018-2019. Since 2019, the convergence between migrant smuggling and trafficking activities increased. In particular, we observe an increase in the convergence between migrant smuggling and drug trafficking: 2020 and 2021 registered the highest levels of this form of convergence in the 2011-2021 period. In 2020, 3.9% of the episodes of migrant smuggling occurred in parallel with drug trafficking; in 2021, 6.3%. These results are in line with the assertion that COVID-19 has accelerated some drug trafficking patterns, including the increased use of waterway routes (UNODC 2021b). However, they do provide evidence in favour of an increase in converge in reaction to border closure. Indeed, the increase in convergence registered in 2019 with respect to 2018 suggests that the upward trend originated before the

introduction of travel restrictions intended to mitigate the spread of COVID-19. At the same time, the difference between the levels of convergence registered in 2020-2021 compared to previous years is not large enough to support claims about increased migration controls on the impact of convergence. Future studies might decide to analyse convergence using observations aggregated at month or weekly intervals, and produce synthetic counterfactual scenarios to go beyond the exploratory assessment of the present study.

Conclusions

Whether the smuggling of migrants and the transnational trafficking of goods occur simultaneously has relevant policy implications. The fact that different crimes occur together and are managed by the same actors suggests the resilience of the criminals themselves and should shape countering strategies. Therefore, it is not surprising that media, enforcement agencies, and international institutions empathize the occurrence of various forms of convergence. Despite this, actual evidence of the degree to which the smuggling of migrants converges with the trafficking of goods is scant. The anecdotal nature of knowledge on this crime convergence is due to the unavailability of data functional to estimate the actual relevance of the phenomenon. In fact, institutions typically collect and manage separately data on distinct trafficking activities. This makes it challenging to quantitatively investigate the interconnection between different crimes. Assessing crime convergence becomes even harder when adopting a supranational perspective as different countries tend to collect and disseminate data in different formats, levels of detail, and time ranges and intervals.

To mitigate these limits in the study of transnational crimes, we implemented an innovative data gathering approach based on text analytics and open-source data. The

methodology developed in this study allowed the identification and extraction of more than 7,000 new articles reported in three languages (i.e., English, French, and Arabic) and encompassing sources from around the world. Then, the use of NLP and text mining techniques supported the estimation of convergence between smuggling of migrants and the transnational trafficking of seven categories of goods.

The proposed methodology explores new techniques to gather information on criminal phenomena that are often difficult to obtain through traditional and official sources, such as the identification of complex criminal behaviours that involve the simultaneous occurrence of multiple crimes. On the one hand, it shows the potential of open sources and media sources to identifying occurrences of convergent events, despite their limitations. On the other hand, given the open-source nature of the data and the flexibility of the analytical techniques employed, it provides researchers in other domains with a replicable approach to gather massive corpus of text and extract and analyse the most interesting and useful information applying NLP techniques.

Results of our analysis showed the usefulness of the methodology by detecting rare cases of crime convergence in 3.4% of all episodes identified as migrant smuggling events. While the most recurrent convergence is between migrant smuggling and drug trafficking, convergence between trafficking of weapons, tobacco, fuel, wildlife, and migrant smuggling also occurs. To the extent possible, thanks to manual controls, these results have refrained from counting incidents of convergence between migrant-smuggling and other transnational crimes based solely on unverified statements by government officials included in news reports. On the contrary, the results were derived from the analysis of actual smuggling cases.

Furthermore, the convergence between migrant smuggling and drug trafficking largely increased between 2020 and 2021. Nevertheless, the upward trend began in 2018,

before the diffusion of the COVID-19 disease, and the related strengthening of border controls. Therefore, our study provides no evidence to support a significant impact of stricter enforcement on the simultaneous occurrence of migrant smuggling and the trafficking of illegal goods.

The results have policy implications as they expand extant knowledge and provide policymakers with inspirations on how to mitigate the harmful consequences of transnational crimes. Due to the scant crime convergence, the mitigation of the potential harm of migrant smuggling should be carried out by planning specific countermeasures and operations without the purpose to target simultaneously transnational trafficking activities. To be efficient and effective, the strategies to address different crimes and phenomena should be multiple. This aspect is of relevance in the Mediterranean region, due to the increasing responsibilities of Frontex, the European Border and Coast Guard Agency. With the New Frontex Regulation, which entered into force in December 2019, the European Border and Coast Guard Agency is expected to further expand to have 10,000 border and coast guard officers, and extend its mandate from the migration management to the control of any criminal activity that takes place at the external border of the European Union. This may include smuggling of drugs, firearms, tobacco products, mineral oils, alcohol, and wildlife, among other traffics (Frontex 2021). It might be true that "comprehensive operational responses are required" (Frontex 2021, 1) when it comes to areas where Africa and the Schengen Area are separated by a few nautical miles, and where multiple traffic is concentrated (e.g., between Morocco and Spain, or between Tunisia and Italy). However, effective strategies should consider the specificities of each crime, as well as its distinctive dynamics and determinants.

A final implication relates to the need for well-designed legal migration channels that can provide migrants with legitimate ways to leave their country of origin and reach

a destination where they can live safer and in better conditions. Setting up legal migration channels may reduce criminal opportunities by decreasing both the demand for migrant smuggling, and the commission of other ancillary crimes (e.g., corruption, document forgery, money laundering).

References

Achilli, Luigi, and Gabriella E. Sanchez. 2021. 'Introduction to Migration, Smuggling and the Illicit Global Economy'. *Public Anthropologist* 3 (1): 1–7.

Achilli, Luigi, and Alessandro Tinti. 2019. 'Debunking the Smuggler-Terrorist Nexus: Human Smuggling and the Islamic State in the Middle East'. *Studies in Conflict & Terrorism* 45 (5–6): 463–78.

African Union Commission. 2018. 'The Revised Migration Policy Framework for Africa and Plan of Action (2018–2027)'. Addis Ababa, Ethiopia. https://au.int/sites/default/files/newsevents/workingdocuments/32718-wd-english_revised_au_migration_policy_framework_for_africa.pdf.

Andreas, Peter. 2021. 'Criminal Convergence Narratives and the Illicit Global Economy'. *Public Anthropologist* 3 (1): 111–18.

Andreas, Peter, and Kelly M. Greenhill. 2011. *Sex, Drugs, and Body Counts: The Politics of Numbers in Global Crime and Conflict*. Ithaca, NY: Cornell University Press.

Asharef, M, Nazlia Omar, M Albared, Zhou Minhui, Wang Weiming, Zhou Jingjing, and others. 2012. 'Arabic Named Entity Recognition in Crime Documents'. *Journal of Theoretical and Applied Information Technology* 44 (1): 1–6.

Aziani, Alberto. 2023. 'The Heterogeneity of Human Smugglers: A Reflection on the Use of Concepts in Studies on the Smuggling of Migrants'. *Trends in Organized Crime* 26 (1): 80–106.

Aziani, Alberto, Gianluca A. Bertoni, Maria Jofre, and Michele Riccardi. 2023. 'COVID-19 and Organized Crime: Strategies Employed by Criminal Groups to Increase Their Profits and Power in the First Months of the Pandemic'. *Trends in Organized Crime* 26: 114–35.

Aziani, Alberto, and Rob T. Guerette. 2022. 'Future Directions in the Study of Displacement and Convergence of Transnational Crimes'. In *The Evolution of Illicit Flows: Displacement and Convergence among Transnational Crime*, edited by Ernesto U. Savona, Rob T. Guerette, and Alberto Aziani. Cham, Switzerland: Springer.

Banerveld, Maarten van, Mohand-Tahar Kechadi, and Nhien-An Le-Khac. 2016. 'A Natural Language Processing Tool for White Collar Crime Investigation'. In *Transactions on Large-Scale Data- and Knowledge-Centered Systems XXIII: Selected Papers from FDSE 2014*, edited by Abdelkader Hameurlain, Josef Küng, Roland Wagner, Tran Khanh Dang, and Nam Thoai, 1–22. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-662-49175-1_1.

Bel Haj Zerki, A. 2009. 'La Migration de Transit En Tunisie: État Des Lieux et Impacts et Avancement de La Recherche Sur La Question [Transit Migration in Tunisia:

State of Play and Impacts and Progress of Research on the Issue]'. CARIM Analytic and Synthetic Notes 16. San Domenico di Fiesole (FI): Robert Schuman Centre for Advanced Studies, European University Institute. https://cadmus.eui.eu/handle/1814/11297.

Berry, M., I. Garcia-Blanco, and K. Moore. 2016. 'Press Coverage of the Refugee and Migrant Crisis in the EU: A Content Analysis of Five European Countries'. Project Report. Geneva: United Nations High Commissioner for Refugees. http://www.unhcr.org/56bb369c9.html.

Bird Ruiz-Benitez de Lugo, Lucia. 2021. 'Human Smuggling in the Time of COVID-19. Lessons from a Pandemic'. In *The Routledge Handbook of Smuggling*, edited by Max Gallien and Florian Weigand, 341–52. London: Routledge.

Borderon, Marion, Patrick Sakdapolrak, Raya Muttarak, Endale Kebede, Raffaella Pagogna, and Eva Sporer. 2019. 'Migration Influenced by Environmental Change in Africa: A Systematic Review of Empirical Evidence'. *Demographic Research* 41: 491–544.

Brewster, Ben, Timothy Ingle, and Glynn Rankin. 2014. 'Crawling Open-Source Data for Indicators of Human Trafficking'. In *2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing*, 714–19. https://doi.org/10.1109/UCC.2014.116.

Brewster, Ben, Simon Polovina, Glynn Rankin, and Simon Andrews. 2014. 'Knowledge Management and Human Trafficking: Using Conceptual Knowledge Representation, Text Analytics and Open-Source Data to Combat Organized Crime'. In *Graph-Based Representation and Reasoning*, edited by Nathalie Hernandez, Robert Jäschke, and Madalina Croitoru, 104–17. Lecture Notes in Computer Science. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-08389-6_10.

Brik, Taher. 2015. '، الجهود التشريعية لدول شمال إفريقيا في مجال مكافحة الهجرة غير الشرعية' [The Legislative Efforts of North African Countries in the Field of Combating Illegal Immigration]'. *Knowledge Journal for Research and Historical Studies* 13: 114–51.

Campana, Paolo. 2018. 'Out of Africa: The Organization of Migrant Smuggling across the Mediterranean'. *European Journal of Criminology* 15 (4): 481–502.

———. 2020. 'Human Smuggling: Structure and Mechanisms'. *Crime and Justice*. https://doi.org/10.1086/708663.

Campana, Paolo, and Loraine Gelsthorpe. 2020. 'Choosing a Smuggler: Decision-Making Amongst Migrants Smuggled to Europe'. *European Journal on Criminal Policy and Research*. https://doi.org/10.1007/s10610-020-09459-y.

Cengage. 2020. 'Newspapers And Print Media: Arab Countries'. Encyclopedia.Com. 12 November 2020. https://www.encyclopedia.com/humanities/encyclopedias-almanacs-transcripts-and-maps/newspapers-and-print-media-arab-countries.

Central Intelligence Agency. 2020. 'The World Factbook'. Central Intelligence Agency. 2020. https://www.cia.gov/library/publications/resources/the-world-factbook/.

Chermak, Steven. 1995. 'Image Control: How Police Affect the Presentation of Crime News'. *American Journal of Police* 14 (2): 21–43.

Cherti, Myriam, Brhmie Balaram, and Miklos Szilard. 2013. 'Welcome Back Everyone: The Return and Reintegration of Irregular Migrants from Morocco'. Institute for Public Policy Research. http://www.ippr.org/files/images/media/files/publication/2013/07/welcome-back-morocco_ENG_June2013_11052.pdf?noredirect=1.

Chin, Ko-lin. 1999. *Smuggled Chinese: Clandestine Immigration to the United States*. Philadelphia: Temple University Press.
Cusumano, Eugenio, and Matteo Villa. 2021. 'From "Angels" to "Vice Smugglers": The Criminalization of Sea Rescue NGOs in Italy'. *European Journal on Criminal Policy and Research* 27 (1): 23–40.
Davy, D. 2017. 'Unpacking the Myths: Human Smuggling from and within the Horn of Africa'. RMMS Briefing Paper 6. Nairobi, Kenya: Regional Mixed Migration Secretariat (RMMS). https://reliefweb.int/sites/reliefweb.int/files/resources/RMMS%20BriefingPaper6%20-%20Unpacking%20the%20Myths.pdf.
Demir, Oguzhan Omer, Murat Sever, and Yavuz Kahya. 2017. 'The Social Organisation of Migrant Smugglers in Turkey: Roles and Functions'. *European Journal on Criminal Policy and Research* 23 (3): 371–91.
Ditton, Jason, and James Duffy. 1983. 'Bias in the Newspaper Reporting of Crime News'. *British Journal of Criminology* 23 (2): 159–72.
Ekedegwa Odeh, Lemuel. 2010. 'A Comparative Analysis of Global North and Global South Economies'. *Journal of Sustainable Development in Africa* 12 (3): 338–48.
EMSC. 2018. 'Two Years of EMSC: Activity Report Jan 2017-Jan 2018'. Europol. https://www.europol.europa.eu/sites/default/files/documents/two_years_of_emsc_report.pdf.
———. 2019. 'European Migrant Smuggling Centre: 3rd Annual Activity Report - 2018'. Europol. https://www.europol.europa.eu/sites/default/files/documents/emsc_report_final_2019_2final.pdf.
Europol. 2017. *European Union Serious and Organised Crime Threat Assessment 2017: Crime in the Age of Technology*. 2021 revised version. The Hague: Europol. https://www.europol.europa.eu/activities-services/main-reports/european-union-serious-and-organised-crime-threat-assessment-2017.
Europol, and INTERPOL. 2016. 'Migrant Smuggling Networks: Joint Europol-INTERPOL Report'. Executive Summary. https://www.europol.europa.eu/publications-documents/europol-interpol-report-migrant-smuggling-networks.
Feldman, Justin M., Sofia Gruskin, Brent A. Coull, and Nancy Krieger. 2017. 'Quantifying Underreporting of Law-Enforcement-Related Deaths in United States Vital Statistics and News-Media-Based Data Sources: A Capture–Recapture Analysis'. *PLoS Medicine* 14 (10): e1002399.
Feldman, Ronen, and James Sanger. 2007. *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge University Press.
Freeman, Laura. 2017. 'Environmental Change, Migration, and Conflict in Africa: A Critical Examination of the Interconnections'. *The Journal of Environment & Development* 26 (4): 351–74.
Frontex. 2021. 'Cross-Border Crime'. Frontex. 2021. https://frontex.europa.eu/we-support/cross-border-crime/.
———. 2022. 'Migratory Map'. 2022. https://frontex.europa.eu/we-know/migratory-map/.
Ghankutkar, Surabhi, Neelabh Sarkar, Pooja Gajbhiye, Sanyukta Yadav, Dhananjay Kalbande, and Nida Bakereywala. 2019. 'Modelling Machine Learning For Analysing Crime News'. In *2019 International Conference on Advances in*

Computing, Communication and Control (ICAC3), 1–5. https://doi.org/10.1109/ICAC347590.2019.9036769.
Gradoń, Kacper. 2020. 'Crime In The Time Of The Plague: Fake News Pandemic And The Challenges To Law-Enforcement And Intelligence Community'. *Society Register* 4 (2): 133–48.
Guerette, Rob T., and Alberto Aziani. 2022. 'The Displacement and Convergence of Transnational Crime Flows'. In *The Evolution of Illicit Flows: Displacement and Convergence among Transnational Crime*, edited by Ernesto U. Savona, Rob T. Guerette, and Alberto Aziani, 9–25. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-95301-0_2.
Hoffmann, Anette, Jos Meester, and Hamidou Manou Nabara. 2017. 'Migration and Markets in Agadez: Economic Alternatives to the Migration Industry'. CRU Report. Netherlands Institute of International Relations Clingendael. https://www.clingendael.org/sites/default/files/2017-10/Migration_and_Markets_Agadez.pdf.
İçduygu, Ahmet. 2018. 'Middle East'. In *Migrant Smuggling Data and Research: A Global Review of the Emerging Evidence Base*, edited by Anna Triandafyllidou and Marie McAuliffe, 2:19–44. Geneva: IOM. https://doi.org/10.18356/5347a578-en.
INTERPOL. 2018. 'Overview of Serious and Organized Crime in North Africa'. Analytical Report. https://www.interpol.int/en/content/download/10231/file/North%20Africa.pdf.
Izcara Palacios, Simón Pedro. 2015. 'Coyotaje and Drugs: Two Different Businesses'. *Bulletin of Latin American Research* 34 (3): 324–39.
Kari, Umar, Godwin Malasowe, and Ogbu Collins. 2018. 'A Study of Illegal Migration Trends and The Pull and Push Factors in Nigeria, 2011 - 2017'. *World Journal of Innovative Research* 5 (6): 53–60.
Kozik, Rafał, Sebastian Kula, Michał Choraś, and Michał Woźniak. 2022. 'Technical Solution to Counter Potential Crime: Text Analysis to Detect Fake News and Disinformation'. *Journal of Computational Science* 60 (April): 101576. https://doi.org/10.1016/j.jocs.2022.101576.
LexisNexis. 2022. 'Nexis® Metabase'. https://www.lexisnexis.com/en-us/professional/data-as-a-service/daas-metabase-product.page.
Mackenzie, Simon, Neil Brodie, Donna Yates, and Christos Tsirogiannis. 2020. *Trafficking Culture: New Directions in Researching the Global Market in Illicit Antiquities*. Abingdon, Oxon and New York: Routledge.
Majidi, Nassim. 2018. 'Community Dimensions of Smuggling: The Case of Afghanistan and Somalia'. *The ANNALS of the American Academy of Political and Social Science* 676 (1): 97–113.
Mancuso, Marina, and Francesca Maldi. 2022. 'Does Human Smuggling Converge with Other Transnational Crimes in North Africa and the Mediterranean Area?' In *The Evolution of Illicit Flows: Displacement and Convergence among Transnational Crime*, edited by Ernesto U. Savona, Rob T. Guerette, and Alberto Aziani. Cham, Switzerland: Springer.
Meneghini, Cecilia, Alberto Aziani, and Marco Dugato. 2020. 'Modeling the Structure and Dynamics of Transnational Illicit Networks: An Application to Cigarette Trafficking'. *Applied Network Science* 5 (1): 1–27. https://doi.org/10.1007/s41109-020-00265-3.
Merry, Sally Engle. 2016. *The Seductions of Quantification: Measuring Human Rights, Gender Violence, and Sex Trafficking*. Chicago: University of Chicago Press.

https://www.degruyter.com/document/doi/10.7208/9780226261317/html?lang=de.
Micallef, Mark. 2019. 'Shifting Sands — Libya's Changing Drug Trafficking Dynamics on the Coastal and Desert Borders'. Switzerland: Global Initiative Against Transnational Organized Crime. https://globalinitiative.net/wp-content/uploads/2019/11/EDMR2019_BackgroundReport_Libya.pdf.
Moher, David, Alessandro Liberati, Jennifer Tatzlaff, Douglas G. Altman, and The PRISMA Group. 2009. 'Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement'. *PLoS Medicine* 6 (7): 1–6.
Moreno-Lax, Violeta, and Martin Lemberg-Pedersen. 2019. 'Border-Induced Displacement: The Ethical and Legal Implications of Distance-Creation through Externalization'. *Questions of International Law* 56: 5–33.
Mpandeli, Sylvester, Luxon Nhamo, Sithabile Hlahla, Dhesigen Naidoo, Stanley Liphadzi, Albert Thembinkosi Modi, and Tafadzwanashe Mabhaudhi. 2020. 'Migration under Climate Change in Southern Africa: A Nexus Planning Perspective'. *Sustainability* 12 (11). https://doi.org/10.3390/su12114722.
Mrabet, E.A. 2018. 'La Migration Africaine [The African Migration]'. Morocco: Royal Institute for Strategic Studies. https://ires.ma/images/Publications/ETUDE-IRES_MIGRATION-AFRICAINE.pdf.
Neske, Matthias. 2006. 'Human Smuggling to and through Germany'. *International Migration* 44 (4): 121–63. https://doi.org/10.1111/j.1468-2435.2006.00382.x.
Osorio, Javier, and Alejandro Beltran. 2020. 'Enhancing the Detection of Criminal Organizations in Mexico Using ML and NLP'. In *2020 International Joint Conference on Neural Networks (IJCNN)*, 1–7. https://doi.org/10.1109/IJCNN48605.2020.9207039.
Ozkan, Turgut, John L. Worrall, and Haley Zettler. 2018. 'Validating Media-Driven and Crowdsourced Police Shooting Data: A Research Note'. *Journal of Crime and Justice* 41 (3): 334–45.
Parkin, William S., and Jeff Gruenewald. 2017. 'Open-Source Data and the Study of Homicide'. *Journal of Interpersonal Violence* 32 (18): 2693–2723.
Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, et al. 2011. 'Scikit-Learn: Machine Learning in Python'. *Journal of Machine Learning Research* 12 (85): 2825–30.
Procter, Caitlin. 2021. 'Coordinated Mobility: Disrupting Narratives of Convergence in the Irregular Migration of Youth from the Gaza Strip'. *Public Anthropologist* 3 (1): 93–110.
Riccardi, Michele, Marco Dugato, and Marcello Polizzotti. 2014. *The Theft of Medicines from Italian Hospitals*. Transcrime Research in Brief 2. Trento-Milano: Transcrime – Joint Research Centre on Transnational Crime.
Sanchez, Gabriella E. 2020. *Migrant Smuggling in the Libyan Context: Re-Examining the Evidence*. International Organization for Migration. https://cadmus.eui.eu//handle/1814/68575.
Sanchez, Gabriella E., and Luigi Achilli. 2020. 'Stranded: The Impacts of COVID-19 on Irregular Migration and Migrant Smuggling'. 2020/20. Policy Brief. European University Institute - Migration Policy Centre. https://cadmus.eui.eu/bitstream/handle/1814/67069/PB_2020_20_MPC.pdf?sequence=1&isAllowed=y.
Sanchez, Gabriella E., and Sheldon Zhang. 2018. 'Rumors, Encounters, Collaborations, and Survival: The Migrant Smuggling–Drug Trafficking Nexus in the U.S.

Southwest'. *The ANNALS of the American Academy of Political and Social Science* 676 (1): 135–51.
Sarehane, F., N. Baba, A. Ezzine, and C. Lautier. 2009. 'Traite Internationale Des Personnes: Etat Des Lieux et Analyse Des Réponses Au Maroc [Transnational Human Trafficking: State of Play and Analysis of Responses in Morocco]'. Royaume du Maroc Ministère de la Justice and International Organsation for Migration. https://www.unhcr.org/50a4c7579.pdf?fbclid=IwAR1UigBcSuPAwZo4G_s3qn-rQ2eUrXv0xYcR9VakdX41FtcwCl1WE6TPTG0.
Sayech, A. 2014. '، مكافحة تهريب المهاجرين السريين [Combating the Smuggling of Illegal Migrants]'. Algeria: University Mouloud Mammeri, Faculty of law and political sciences. https://dl.ummto.dz/bitstream/handle/ummto/74/%D8%B5%D8%A7%D9%8A%D8%B4%20%D8%B9%D8%A8%D8%AF%20%D8%A7%D9%84%D9%85%D8%A7%D9%84%D9%83.pdf?sequence=1&isAllowed=y.
Scherrer, A. 2019. 'Detecting and Protecting Victims of Trafficking in Hotspots: Ex-Post Evaluation'. Brussels: European Parliament Research Service (EPRS). https://www.europarl.europa.eu/RegData/etudes/STUD/2019/631757/EPRS_STU(2019)631757_EN.pdf.
Schmidt, Anna, and Michael Wiegand. 2017. 'A Survey on Hate Speech Detection Using Natural Language Processing'. In *Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media*, 1–10. Valencia, Spain: Association for Computational Linguistics. https://doi.org/10.18653/v1/W17-1101.
Shaw, Mark. 2017. 'Africa's Changing Place in the Global Criminal Economy'. Continental report 01. ENACT. https://journals.co.za/doi/abs/10.10520/EJC-b181ee16e.
Spener, David. 2009. *Clandestine Crossings: Migrants and Coyotes on the Texas-Mexico Border*. 1st ed. New York: Cornell University Press.
Tinti, Peter, and Tom Westcott. 2016. 'The Niger-Libya Corridor: Smugglers' Perspectives'. ISS Paper 299. Institute for Security Studies. https://sahelresearch.africa.ufl.edu/files/ISS-Niger-Libya-migrant-corridor.pdf.
United Nations. 2020. 'Impact of Natural Disasters, Conflicts and Crises, Such as the COVID-19 Pandemic, on Trends in Organized Criminal Groups and on Routes for the Smuggling of Migrants, as Well as Good Practices to Support Effective Law Enforcement Cooperation during Such Crises to Detect, Investigate and Prosecute Such Cases'. In *Working Group on the Smuggling of Migrants*. Vienna: United Nations. https://undocs.org/ctoc/cop/wg.7/2020/2.
UNODC. 2019a. 'Facilitators of Smuggling of Migrants in Southeast Asia: Fraudulent Documents, Money Laundering, and Corruption'. Bangkok: United Nations Office on Drugs and Crime. https://www.unodc.org/documents/southeastasiaandpacific/Publications/2019/UNODC_Facilitators_of_Smuggling_of_Migrants_in_Southeast_Asia.pdf.
———. 2019b. 'Migrant Smuggling FAQs'. United Nations Office on Drugs and Crime. 2019. https://www.unodc.org/unodc/en/human-trafficking/faqs-migrant-smuggling.html.
———. 2020. 'How COVID-19 Restrictions and the Economic Consequences Are Likely to Impact Migrant Smuggling and Cross-Border Trafficking in Persons to Europe and North America'. Vienna: United Nations Office on Drugs and Crime.

https://www.unodc.org/documents/islamicrepublicofiran//2020/05/Covid-related-impact-on-SoM-TiP.PDF.

———. 2021a. ‘COVID-19 and the Smuggling of Migrants. A Call for Safeguarding the Rights of Smuggled Migrants Facing Increased Risks and Vulnerabilities’. Vienna: United Nations Office on Drugs and Crime. https://www.unodc.org/documents/human-trafficking/SOM_and_COVID-19_Publication_final_EN_final.pdf.

———. 2021b. ‘World Drug Report 2021’. United Nations Office on Drugs and Crime. www.unodc.org/unodc/en/data-and-analysis/wdr2021.html.

Yates, Caitlyn. 2021. ‘Fusion Points: The Perceived, Performed, and Passive Merging of Criminality and Mobility in Mexico’. *Public Anthropologist* 3 (1): 56–72.

Annexes

Annex 1: Keywords and queries used for extraction

English query:

((morocc* OR algeria* OR tunisia* OR libya* OR egypt* OR mauritania* OR ((western) PRE/1 (sahara))) AND ((smuggl*) NEAR/2 (human* OR people OR person OR migra* OR refugee* OR clandestine*))) AND feedClass:(News Print) AND language: English[2]

French query:

((maroc* OR algérie* OR tunisie* OR libye* OR égypt* OR mauritanie* OR ((sahara) PRE/1 (occidental))) AND (((contreband* OR trafi*) NEAR/2 (humain* OR personne OR migra* OR immigr* OR clandestin* OR réfugié*)) OR ((passage*) NEAR/2 (migra* OR immigr* OR clandestin* OR réfugié*)))) AND feedClass:(News Print) AND language: French

Arabic query:

((المغرب OR مغربي OR الجزائر OR جزائري OR ليبيا OR ليبي OR مصر OR مصري OR موريتانيا OR موريتاني OR ((الغربية) NEAR/1 (الصحراء))) AND ((المهرب OR مهرب OR التهريب OR تهريب) NEAR/1 (البشر OR البشري OR الإنساني OR الأشخاص OR الأفراد OR الناس OR شخص OR فرد OR إنسان OR مهاجر OR لاجئ OR سري))) AND feedClass:(News Print) AND language: Arab

[2] The asterisk (*) is a wildcard symbol that broadens a search by finding words that start with the same letters. For example, "smuggl*" allowed for the inclusion of search words that refer to smuggling activities (e.g., smuggle, smuggler, smuggling and smuggled).