Unraveling Spatial, Structural, and Social Country-Level Conditions for the Emergence of the Foreign Fighter Phenomenon: An Exploratory Data Mining Approach to the Case Of ISIS

Agustín Pájaro[1]

[1]Director. Academia de Guerra Naval. Viña del Mar, Chile.
E-mail: apajaro@armada.cl

Ignacio J. Duran[2]

[2]Ph.D. Candidate in Management Sciences. Institute for Social Innovation, Univ. Ramon Llull, ESADE Business School. Barcelona, Spain.
E-mail: ignacio.duran@esade.edu

Pablo Rodrigo[3]

[3]Associate Professor. Escuela de Negocios, Universidad Adolfo Ibáñez. Viña del Mar, Chile.
E-mail: prodrigo@uai.cl

Datasets: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/SFT3RT

INTRODUCTION

The *foreign fighter* (FF) phenomenon is not new. FFs are non-nationals who travel to take part in conflicts abroad (Malet, 2010; Mendelsohn, 2011), and scholars indicate they are common in several struggles throughout history. For instance, the presence of FFs has been documented in the Spanish Civil War in the 1930s, the Arab-Israeli War of 1948, the Soviet-Afghan War in the late 1980s, the First Chechen War in the 1990s, and recently in the conflict involving the Islamic State in Iraq and Syria (ISIS) (Chesterman, 2016; Joffé, 2016). The conditions that contribute to the rise of this phenomenon have drawn the attention of political analysts in the case of ISIS (Bakker et al., 2014; Coolsaet, 2016; de Bie et al., 2015; Zafar, 2014; Zammit, 2015).

Lately, research has focused on what could lead non-nationals to participate in cells linked to ISIS (Bousquet, 2012; Byman, 2016) for three main reasons. First, to prevent attacks on the battlefield (Cronin, 2015);

second, because these individuals provide intellectual skills and logistic capabilities that empower the terrorist group across the spectrum of warfare (Bakke, 2014); and third, the return of these jihadists to their home countries poses a risk to national security, since they have obtained military knowledge and absorbed radical doctrines which could compel them to perform terrorist attacks (Vidino, 2014). Aggregate evidence shows that citizens from over 85 nations (e.g. the United States, Australia, Indonesia, Tajikistan) have joined ISIS (Benmelech; Klor, 2020). This has sent FF numbers in Iraq and Syria soaring: over 1,000 in 2011; over 3,500 in 2012, over 8,500 in 2013; over 18,000 in 2014; and more than 25,000 in 2015 (HSC, 2015). Even though efforts from Western countries have reduced these numbers from 2016 on (Speckhard et al., 2017), FFs continue to be a pressing issue because they could commit terrorist attacks in their home countries (Cragin, 2017). The shortcomings of initiatives aimed at stopping FFs may indicate the need for a different approach to better understand the issue.

On the one hand, extant theory has attempted using qualitative and quantitative tools to determine the psychological and/or sociological reasons why people become FFs. However, this tradition has failed to fully comprehend the phenomenon because of outdated methodologies. Also, studies focus on limited settings which do not allow for generalization (Bakker et al., 2014; Coolsaet, 2016; de Bie et al., 2015; Nilsson, 2015; Zafar, 2014; Zammit, 2015; Zelin, 2020), which affects the proposal of effective worldwide policies and strategies to face the problem. On the other hand, some studies have tried to understand several country-level spatial, structural (economic/political), and social variables, but have not accurately pinpointed the way their combination may create the environment required for the emergence of FFs (Benmelech; Klor, 2020; Barrett, 2014; Hegghammer, 2013; Hewitt; Kelley-Moore, 2009; Klausen, 2015; Malet, 2010; Watts, 2008; Weggemans et al., 2014). These studies have only determined general dimensions without elucidating the inherent spatial, structural, and social complexity that underlies the construction of a favorable atmosphere for radical terrorism (Bousquet, 2012; Cronin, 2015)

In the case of ISIS, for example, the FF phenomenon is considered complex due to its geographically dispersed, decentralized, and self-organizing network-like structure, in which agents can acquire knowledge while adapting to contingencies (Bousquet, 2012). From the lens of

complexity theory (Meek, 2010; Mörçol, 2010; Teisman; Klijn, 2008), social science literature has posited that dealing with intricate problems requires efforts to be numerous, intertwined, multidisciplinary, and dynamic to provide valid alternatives (Byrne; Callaghan, 2014; Gilbert; Bullock, 2014). Hence, understanding the complex patterns of country-level variables that underpin the FF problem might allow fighters' countries of origin to focus on designing strategies that include multiple and complementary lines of action, such as regulatory and security programs, border policies and cooperation agreements, educational programs, counterpropaganda activities, etc., so as to reduce risk and provide effective policy options.

In this sense, the purpose of this article is to unveil the underlying conditions that may catalyze the FF phenomenon in a certain country, using ISIS as an example. Our research questions are:

> (1) *Which spatial, structural, and social country-level variables can create an appropriate environment for the increase of ISIS FFs?*
>
> (2) *How do these variables interrelate to form country-level patterns or profiles to further explain this phenomenon?*

To answer them, we use a quantitative exploratory data mining method (Witten et al., 2011) called knowledge discovery in databases (KDD) (Maimon; Rokach, 2010). In order to detect variables and patterns, we access datasets from several domains (Cao et al., 2011; Fayyad et al., 1996) and, after filtering and pre-selecting parameters (Chen et al., 2011; Hall; Holmes, 2003; Maletic; Marcus, 2010), we use the information in decision tree algorithms (de Ville; Neville, 2013; Rokach; Maimon, 2010). Our contribution is twofold: first, we move past traditional social science methods and use an approach associated with complexity theory to provide new academic and practical insight to the quest for ways to face the FF threat. Second, our results may help policy makers improve resource allocation, and thus formulate multifaceted initiatives to extinguish this hazard.

KEY THEORETICAL BACKGROUND

Foreign fighters as a complex phenomenon

Past works have posited two main reasons to uphold that the FF phenomenon is highly complex, especially in the case of ISIS. First, authors such as Barrett (2014) highlight its non-linear nature. Zafar (2014) delves into this issue by arguing that the phenomenon is determined by the interplay of multiple circumstances, and no single variable can be accountable for it. Consistent with complexity theory literature, this implies that any development in an extremist cause is guided by a combination of simultaneous, interwoven forces, in which small variations can have disproportionate consequences (Morçöl, 2010; Teisman; Klijn, 2008). This entails that it is hard to model its antecedents, processes, and outcomes from a public policy lens (Meek, 2010).

Second, ISIS cells are self-organizing and work in a geographically dispersed and decentralized manner. Therefore, we cannot speak of a single and unified terrorist group, but rather of a worldwide network of multiple co-existing clusters which – although sometimes scarcely – interact with each other to wage jihad (Bousquet, 2012; Mendelsohn, 2011). This has been termed a "bottom-up" approach (Watts, 2008), in which transnational insurgents face diverse country scenarios, some of which are more favorable to the development of a terrorist cell – e.g. due to a shared cultural perception of having "no future" (Coolsaet, 2016). These emergent substructures are complemented by the fact that oftentimes these self-organizing factions are linked by religious authorities or veteran FFs, forming a web-like design around the globe (Vidino, 2014). Consequently, even if one particular branch is eradicated, the organization will readapt and keep functioning (Cronin, 2015).

The issues discussed in the two previous paragraphs – chiefly non-linearity, self-organization, geographical dispersion, and emergence – are fundamental macro-level attributes of complex social systems (Byrne; Callaghan, 2014). In these intricate, tough-to-model scenarios we have to discover underlying rules to comprehend their interdependence and their intricate nature (Sull; Eisenhardt, 2015). This explains why literature has started to study the multiple country-level factors which set some conditions that fuel these radical movements. According to our review, parameters in this tradition can be classified in three broad categories: spatial, structural, and social.

Understanding spatial, structural, and social dimensions of the foreign fighter threat

Hewitt and Kelley-Moore (2009) argue that proximity to the conflict – in this case, to ISIS-controlled territories – may influence the emergence of FFs. A shorter distance from Iraq and Syria supposes a more favorable condition for the appearance of the jihadist cause. Two rationales exist: first, and more intuitively, the closer you are, the easier it is to get there (Hegghammer, 2013). For example, European countries are connected to Iraq or Syria by roads, providing a more convenient scenario for the rise of FFs than in faraway countries (Barrett, 2014; Nilsson, 2015). Second, despite the decentralized nature of ISIS, proximity is a condition that might imply closeness to vital communication hubs or nodes that might influence the conformation of FF cells (Bousquet, 2012; Klausen, 2015). In fact, several works portray the importance of distance as an enabling mechanism of other variables (such as access to veteran FFs) which have been considered key in the development of the FF menace (de Bie et al., 2015; Vidino, 2014; Watts, 2008; Zammit, 2015).

Regarding the structural dimension, Benmelech and Klor (2020) evidence the relevance of economic and political variables as country circumstances which may push non-nationals to join ISIS. We categorize these factors as structural, since they describe country-level conditions of the relatively stable arena where potential transnational insurgents are nurtured (i.e., where they grow up, study, and work). Matters such as low political freedom, internet access, unemployment, inequality, meager incomes, and high criminality rates have been conjectured as potential state-level factors leading one to become a FF. According to some authors (Hewitt; Kelley-Moore, 2009; Watts, 2008), the explanation lies in the fact that the complex and unique interrelation of these numerous contextual parameters may lead to a "situational induced motivation" (de Bie et al., 2015), in which deficient living conditions concur with the fact that the internet has curtailed communication restrains (i.e., it has allowed citizens to access ISIS-related information) to potentially create fringe living scenarios that may enhance the development of radical thoughts.

Finally, past studies also consider social parameters as a dimension of this phenomenon. Bousquet (2012) argues that the pattern and nature of interactions in a society, as well as peoples' values and perceptions, are variables that may create an environment conducive to the gener-

ation of FFs. For example, Watts (2008) ascertains that conversations, sermons, print and radio communications, and family members are factors that may lead to the development of radical thoughts and, eventually, terrorism. Also, exposure to religious fundamentalism (i.e., jihad doctrine) in home countries may raise social concern for threats against the Islamic world, creating the conditions for a culture of violence in these communities, which is a basis for the appearance of FFs (Malet, 2010; Weggemans et al., 2014). In some cases, this assessment is complemented by what has been called the "no future" subculture (Coolsaet, 2016). When young people have a shared belief according to which they will not achieve success in life (e.g. due to societal alienation), they search new undertakings (Barrett, 2014). Cronin (2015) evinces that the pursuit of adventure, personal power, sense of self and community, and even sexual desire are social parameters that set some conditions for these radical movements.

Even though our approach is exploratory and follows a data mining (KDD) approach, the number of variables that we could include is infinite. We have thus developed the main dimensions found in extant literature in this section as a starting point to determine which datasets to use. Consistent with our research questions, our work constitutes not only an empirical test of these categories (spatial, structural, and social), but also the proposal of a method to discover country-level patterns which might trigger the FF phenomenon.

METHODOLOGY

Justification of the methodology

This paper addresses research questions which involve the detection of country-level spatial, structural, and social conditions which generate environments favorable to the penetration of ISIS recruitment messages. The discussion above suggests that this phenomenon pertains to the realm of complexity theory.

Political science research dealing with complex phenomena have questioned traditional social science methodologies – both quantitative and qualitative – because they are based on the selection of narrow samples and issues (Meek, 2010; Morçöl, 2010; Teisman; Klijn, 2008), jeopardizing the generalization of the findings (Hair et al., 2010) and

consequently affecting their potential use in policy making. Contrarily, complexity-based social system analysis has the benefit of considering hundreds of parameters to find multiple and dynamic associations among dependent and independent variables without assuming causality *a priori* (Byrne; Callaghan, 2014). This accelerates the search of public policy practitioners for solutions, because it draws researchers closer to a process-oriented – rather than outcome-oriented – study, allowing the development of an intertwined yet understandable pattern which better explains the phenomenon (Trochim; Cabrera, 2005).

To address these complexity-based issues, we use an exploratory data mining KDD approach (Cao et al., 2011; Fayyad et al., 1996; Maimon; Rokach, 2010). This technique usually aims to gather and process vast amounts of information – stored in large electronic databases – through the use of algorithms in search for patterns. Results then allow researchers to make some generalizations about the problem's complexity, which can lead them to acquire valid knowledge and ultimately provide a basis to explain the behavior of the variables under study (Witten et al., 2011). This information may help future research delve into the multiple associations found in the derived patterns.

The main difference between KDD and other quantitative techniques is its interpretative power. Although researchers also apply algorithms with interpretive purposes in multivariate statistical techniques – for instance, Hair et al. (2010) –, they would have to manage and control all the variables incorporated, which is nearly impossible for human minds dealing with large databases. The number of parameters that could describe a country's structural and social conditions may amount to thousands. Moreover, in traditional quantitative analyses, researchers have to determine the relationships among the variables beforehand – for example, through hypothesizing – and only then test them one by one, precluding the emergence of hidden relationships from the data (Fayyad et al., 1996) and leading to reductionist analyses and conclusions.

In order to find these latent associations, for this study we selected an algorithm called "decision trees", which is based on the idea that some independent variables (attributes) – or a combination thereof – might cause significant changes in one particular dependent variable or target. Decision trees help encode the target behavior in terms of some (not all)

independent variables or attributes. Besides, these algorithms require less data preparation, enabling the use of different types of data, and have robust and fast processing performance (Witten et al., 2011).

Determining the datasets

Based on the three main categories defined in the theoretical background section, we analyzed 950 variables for 211 countries. This constitutes the independent variables set – or attribute database –, with datasets obtained from well-reputed international organizations. The dependent variable – or target variable – is a measure of foreign fighters – for ISIS, in this case –, as in many past studies. To maximize data availability, we only used information from 2014, as other years (2015-2018) presented missing values.

The selection of the 950 variables was based on a two-step method used in many social sciences and related disciplines. First, using the three aforementioned categories as "themes", we undertook a thematic analysis (Braun; Clarke, 2008) of the quantitative FF literature we reviewed, determining the key country-level variables used in past studies – despite their narrower databases. Second, and to complete and/or validate the emerging databases, we used a qualitative "expert panel" technique, submitting the preliminary selection to review and discussion with certain political science scholars and political advisors in order to gain validity (Creswell; Clark, 2006). This step involved experts from an armed forces academic institution in the United States, where one of the researchers spent two years. However, in stark contrast with the traditional quantitative use of expert panels,[1] the experts just gave their opinions regarding the pertinence of including some variables. They did not sort them in order of importance nor did they score them in a quantitative way. This intentional difference enabled the selected algorithm to determine the variables' importance and their interrelations in this large database. Following our thematic review of the literature and expert analysis, the 950 selected parameters were broken down into the following datasets:

Spatial parameter. Our literature review detected that proximity to the ISIS conflict may affect the recruitment of non-nationals. Hence, our first data series ("AA") is distance from a country to Syria in kilometers (TVE, 2015). All distances were measured from capital to capital.[2]

Structural parameters. Economic and political variables are also known to affect the decision to become a FF. Particularly, countries' developmental issues, such as economic growth, income level, education, employment, and institutional quality. Thus, we selected the following datasets:

- Data series "A" – general structural conditions of the countries. World Bank's "World Development Indicators"; attributes A1 to A881 (World Bank, 2014a). A total of 881 parameters per country were obtained here.

- Data series "C" – internet users, growth, and penetration per country; attributes C1 to C8 (Internet Live Stats, 2014).

- Data series "E" – governance variables of each country: voice and accountability, political stability & no violence, government effectiveness, regulatory quality, rule of law, and control of corruption; attributes E1 to E6 (World Bank, 2014b).

- Data series "F" – unemployment per country: total, male, female, youth and adult; attributes F1 to F9 (ILO, 2014).

- Data series "G" – press freedom index; attribute G1 (RSF, 2014).

Social parameters. Social variables that describe human interactions, behaviors, and/or perceptions also might affect the decision to become a transnational insurgent, explaining why we considered the following datasets:

- Data series "B" – estimated Muslim population per country and percentage of total population; attributes B1 and B2 (Pew-Templeton, 2014).

- Data series "D" – age-standardized suicide rates and current alcohol drinkers in the youth population per country; attributes D1 to D8 (WHO, 2014).

- Data series "H" – happiness indexes: life ladder, social support, healthy life expectancy at birth, freedom to make life choices, generosity, the perception of corruption, positive affect, negative affect, confidence in national government, democratic quality, delivery quality, and trust in others; attributes H1 to H13 (SDSN, 2015).[3]

- Data series "I" – Global Peace Index: perceptions of criminality in society, police officers per habitants, incarceration rates, access to small arms and light weapons, the level of internally organized conflict, the level of violent crime, political instability, terrorist acts, and others; attributes I1 to I23 (IEP, 2014).

Dependent or target variable. The dependent or target variable is the number of foreign fighters per country (NFFC). This parameter measures how many non-national combatants have travelled to join ISIS – whether in Syria or Iraq – divided by nation. This information was extracted from The International Centre for the Study of Radicalization and Political Violence (ICSR, 2014). This institution has counted FFs in Syria and Iraq since 2012.

Data mining process

To process all our data, we used the open source software Waikato Environment for Knowledge Analysis (WEKA), version 3.91. Our data mining approach consists of two steps.

The first is "data cleansing", which aims to find the maximum number of independent variables which could impact the outcome (NFFC). This is generally known as "attribute pre-selection" (Fayyad et al., 1996; Maimon; Rokach, 2010), and it is essential in KDD because not all 950 parameters are necessarily associated – both statistically and theoretically – with the dependent variable. This procedure helps avoid spurious relationships (Chen et al., 2011; Hall; Holmes, 2003; Maletic; Marcus, 2010). The procedure for selecting attributes or independent variables followed a mixed approach: first with the use of WEKA, and second based on the authors' interpretation of these preliminary findings. The first phase was conducted with the support of the feature called "select attributes" in WEKA. This software allowed us to use nine different correlation filters to build an indicator of the capacity of each attribute to predict the NFFC.

The filters applied during the "data cleansing" phase were "Chi Squared Attribute Eval", "Correlation Attribute Eval", "CV Attribute Eval", "Filtered Attribute Eval", and "Gain Ratio Attribute Eval", "Info Gain Attribute Eval", "Relief Attribute Eval", "Significance Attribute Eval", and "Symmetrical Uncert Attribute Eval". Using a normalization from 0 to 1 – by dividing values by the maximum number encountered for each variable – and averaging the results of each filter (Witten et al., 2011), it was possible to select all the attributes that had a high predictability score. Following Giudici (2010), we used a 0.40 cut-off criterion. Through this process, we eliminated 925 of the 950 parameters, keeping 25 independent variables which relate to the target. After careful consideration, and based on our knowledge and the literature review, we decided to use only 22 attributes in the decision tree algorithms, as explained in the attribute pre-selection results section below.

Second, we entered the selected attributes into the machine learning scheme. To determine rules, patterns, or find hidden relationships between the depurated independent variable set and the target variable, we used a group of algorithms called "decision trees" (Rokach; Maimon, 2010; Witten et al., 2011). These kinds of algorithm allow a better interpretation when extracting knowledge from data, because they help encode the target variable in terms of several independent variables or attributes, and hence move beyond one-cause/one-effect relationships to discover multiple and simultaneous influences or patterns (de Ville; Neville, 2013). In this sense, although stand-alone independent variables are an important part of explaining the phenomenon, our method suggests that their true importance lies in the specific combination and/or path-dependency – obtained though the selected algorithm – they can have together with other variables (Fayyad et al., 1996).

Besides, as previously mentioned, these algorithms have advantages such as requiring less data preparation or depuration, allowing the use of different kinds of data, being able to handle missing values, and having robust and fast processing performance (Witten et al., 2011). Moreover, decision tree algorithms are also able to "learn" by using examples or instances of attributes and targets, and cutting branches of the tree (known as "pruning") when they have no relationship with the dependent variable. This enhances interpretation by simplifying the resulting model, and therefore makes it easier to determine academic and policy implications.

The instances were configured as the world's nations (211 in total). For this procedure, we deliberately did not include Syria, Iraq, and Afghanistan in the country sample. In the case of the former two, this is due to their being at the heart of the ISIS conflict, being thus considered "receivers" of international combatants, whereas we are trying to understand the conditions of "sender" countries. As for Afghanistan, it was discarded because, according to de Bie et al. (2015) and Malet (2010), its major terrorist menace is still Al-Qaeda. Therefore, as a precautionary measure to avoid bias and confusion between effects of Al-Qaeda and ISIS, we removed it from the sample.

Our decision tree analysis consisted of two distinctive trials following the suggestions made by Witten et al. (2011), Rokach and Maimon (2010), and de Ville and Neville (2013). For Trial #1, we used WEKA's "Experimenter" application, and, for Trial #2, the "Explorer" package. By using a ten-fold cross-validation technique, Trial #1 attempted to determine the best performing algorithms and the best way to codify the target or dependent variable (NFFC). To this end, we tested 26 different algorithms, with a total of 8 different ways of coding the target variable, resulting in 208 different combinations of decision trees. Subsequently, Trial #2 aimed to find the algorithm combination with the best predictive performance – i.e., the one that maximized the number of correctly classified countries in the final sample as senders or non-senders of FFs – and to determine valid patterns of drivers of the FF phenomenon among countries' spatial, structural, and social conditions.

RESULTS

Preliminary results based on attribute pre-selection

The procedure based on Giudici's (2010) predictability score criterion should be carefully used for interpretation purposes, since it only suggests a level of association between just one phenomenon (independent variables) and another (NFFC). Having this caveat in mind, these numbers suggest some preliminary results:

Spatial parameter. Overall, the "AA" series ranked best, with a score of 0.79, providing some insight into the importance of physical proximity to the conflict.

Structural parameters. Only ten attributes of the largest data set ("A") scored higher than the selected cut-off: A6 (0.51), purchasing power parity/gross national income; A185 (0.47), health expenditure per capita; A371/372/379 (0.53, 0.60, 0.61, respectively), carbon dioxide emissions; A529 (0.47) gross domestic product; A621 (0.54) net secondary income; A624 (0.47) total reserves, including gold; A864 (0.57), international migrant stock; and A871 (0.79) international tourists inbound. This indicates that countries' economic conditions do have some relationship with the decision to become a FF.

As far as internet penetration and users are concerned, data series "C" had only two of its six parameters with a high enough predictability score: C1 (0.44), internet users per country; and C3 (0.42), user growth. These precursory results are not surprising, as our literature review indicated social media as vital channels to spread ISIS recruiting messages.

Data series "E", governance variables, also offers relevant insight: attributes E1 (0.51), voice and accountability; and E5 (0.43), the rule of law. This suggests that certain political variables tend to have an association with this phenomenon. It seems that, as countries' governance quality decreases, the more its conditions favor the emergence of FFs. Norway, for example, with a non-normalized E1 value of 1.71, only has 60 NFFC. Finland, with 1.57, has 50-70 NFFC. Denmark, with 1.55, has 100-150 NFFC. In contrast, France, with a non-normalized E1 score of 1.22, has 1,200 NFFC, while the UK, with 1.30, has 500-600 NFFC, and Tunisia, with 0.03, has 1,500-3,000 NFFC.

None of the nine parameters associated with a country's national and youth unemployment levels (data series "F") seemed to set an appropriate environment for the generation of foreign jihadists – maximum score was 0.14. Similarly, freedom of press in a country (attribute G1) scored very low (0.04), indicating that communicative liberty and oppression in a state does not create conditions for the penetration of terrorist recruiting messages.

Social parameters. Both parameters in the "B" series show high levels of correlation with the NFFC: B1 (0.56), the estimated Muslim population per country; and B2 (0.61), the percentage of a state's population

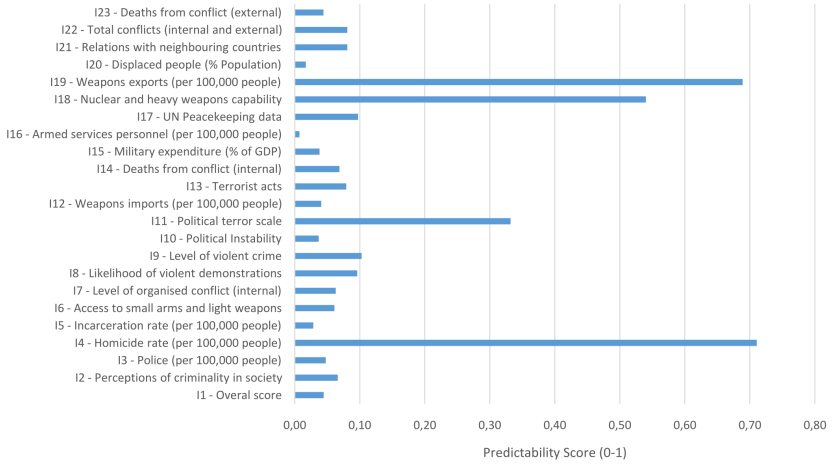
that is Muslim. These preliminary findings are unsurprising, given that extant works report this condition as a relevant factor of the issue under study.

When it comes to social levels of psychological health and alcohol abuse (data series "D"), we detected that only two parameters had a high predictability score: D6 (0.44), percentage of alcohol consumers (both sexes); and D7 (0.52), proportion of alcohol consumers among females. This entails that alcohol consumption could enable an environment conducive to the development of radical ideas and, eventually, to the generation of FFs.

In data series "H" (happiness indexes), less than half the 13 attributes demonstrate a significant enough score of prediction. In broad terms, the freedom to make life choices, trust in others, social support, generosity of the population and confidence in the current national government do not score more than 0.1. Nevertheless, H1 (0.53), the public perception of overall wellness (i.e., life ladder); H7 (0.66), general corruption; and H11 (0.47), democratic quality seem to affect the general sense of well-being, and second, the way people perceive their country's leadership, transparency, and management over time, determining the "social atmosphere" regarding the proneness to join this cause.

The analysis based on each country's Global Peace Index indicators (data series "I") also yielded interesting preliminary results (as an example, see Figure 1). Access to small weapons, the perception of criminality, internal and external conflicts, and the level of political instability in society does not seem to predict the conditions for a higher NFFC. Nevertheless, the number of domestic crimes – homicides rate (I4) – seems to have a high relationship with the radicalization phenomenon (0.72), and, consequently, with the creation of conditions for the penetration of recruiting messages. Besides, attributes I18 (nuclear and massive weapons capability) and I19 (weapons exports) have a high prediction score: 0.54 and 0.68, respectively. Consistent with previous works (Eweiss, 2016; Speckhard, 2006), this suggests that certain perceptions of sender countries' weapons industry can create specific social conditions that trigger radicalism.

Figure 1
Predictability Score: Data series "I"



Source: Elaborated by the authors, using WEKA software.

As explained in the methodology section, following the software's results (the 25 pre-selected parameters), a preliminary assessment of these variables is in order. We discarded the three variables that described countries' CO_2 emissions (A371/372/379), as seemingly they had no conceptual relationship with the NFFC phenomenon whatsoever. To the best of the authors' knowledge – also discussed with other colleagues –, no work has proposed that air pollution variables could favor the development of radical terrorism in a country. Also, we consulted with academic colleagues in other disciplines – such as environmental economics, green politics, and sustainable development – about the extent to which these variables could be boundary conditions for the radicalization of social movements in general – and terrorism in particular – but did not find a clear conceptual link.

The Appendix presents a summary and some background information about the final selection of attributes used.

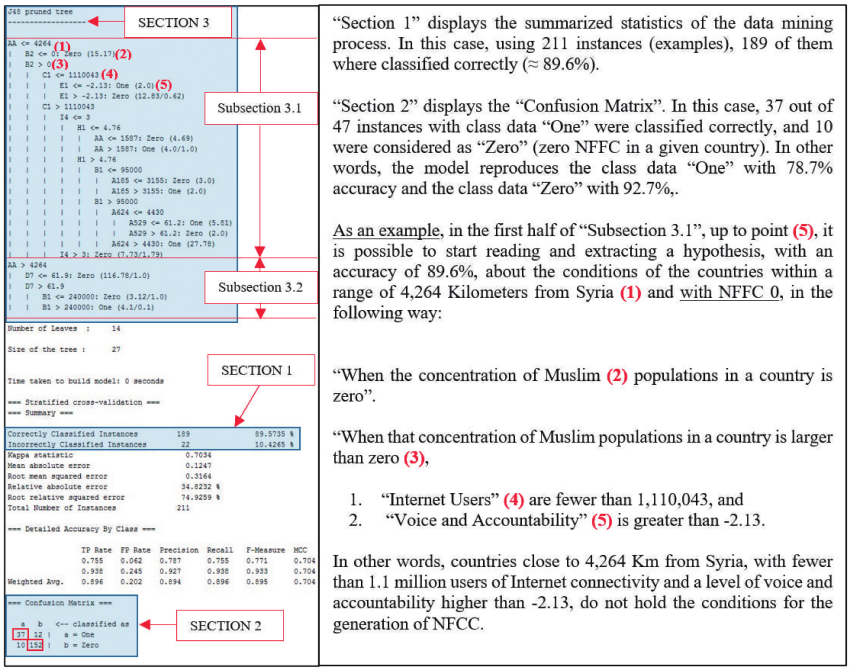
Tree learning algorithms and pattern results

The results of the 208 combinations of Trial #1 determined a total of 13 learning tree algorithms with the highest predictive performance. These procedures were: "Decorate", "Random Committee", "Iterative Classifier Optimizer", "Logit Boost", "J48 Graft", "J48", "Random Sub-

space", "LAD Tree", "Random Tree", "Dagging'", "Rep Free", "CDT", and "PART". This trial also revealed that the best way to code the target or dependent variable was by using two categories: "zero" for countries with no NFFC whatsoever, and "one" for countries with NFFC greater than zero. This categorization aimed to find the prevailing conditions that differentiate states with or without FFs, which is consistent with our goals. These findings reduced the number of possible combinations from 208 to 13, which made it simpler to find the best performing procedure.

Trial #2 consisted in rerunning this final set of algorithms. Out of these 13 combinations, the "J48" decision tree gave the best outcome, with a total of 89.6% countries (instances) correctly classified as to having or not sent foreign fighters to ISIS. Figure 2 presents WEKA's outcome after both trials, and hence, the final result of the J48 tree algorithm with the zero-one coding of the target or dependent variable – including all branches and leaves. Also, Figure 2 has an example of how to interpret each line in the decision tree.

Figure 2
Final outcome of Trials #1 and #2: the best performing decision tree (J48 algorithm with "Zero"/"One" classification) with its branches and leaves.



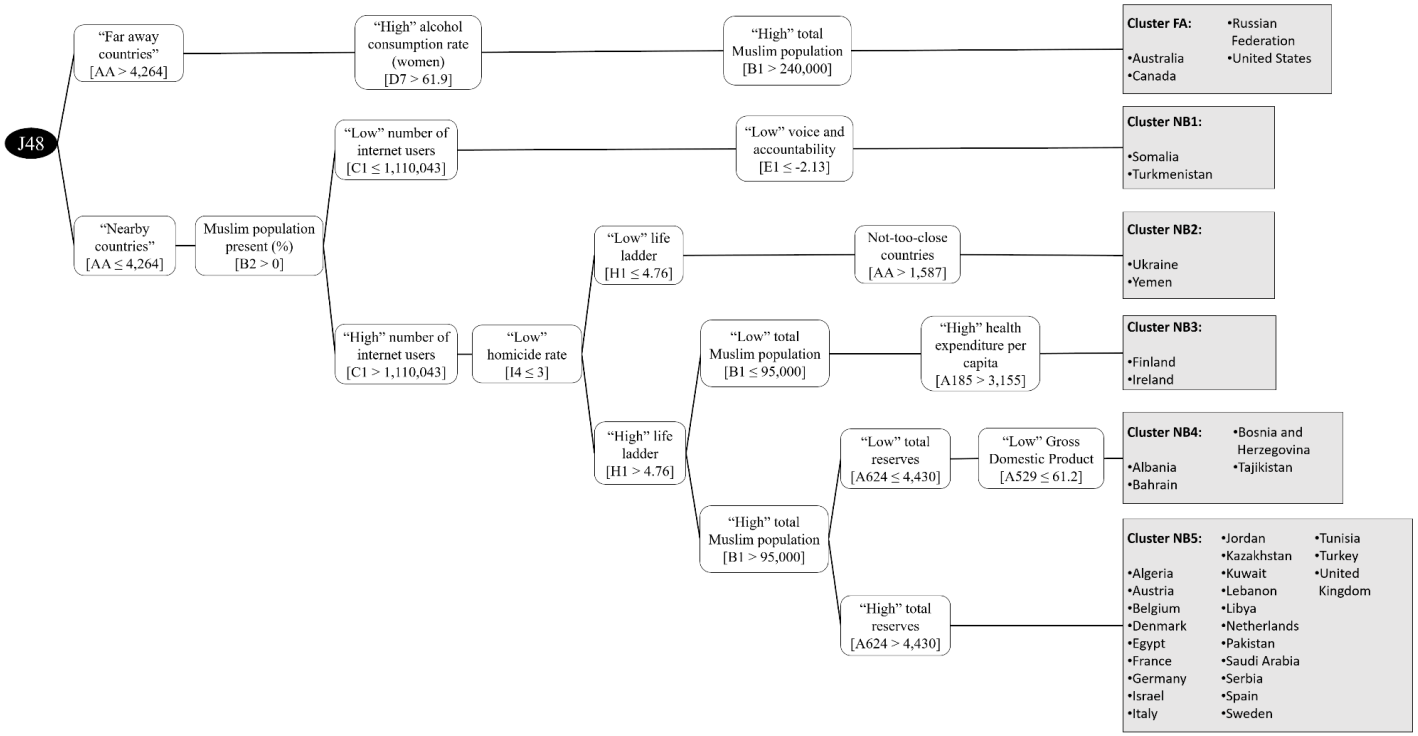
Source: Elaborated by the authors, using WEKA software.

Section 1 in Figure 2 presents the fit statistics of the data mining process. As previously mentioned, this algorithm correctly classified 89.6% – or 189 countries – as senders or non-senders of foreign fighter to ISIS (target variable: = 0/zero and >0/one, respectively). Section 2 in Figure 2 breaks down this number through the "confusion matrix". This reveals that, out of the 47 countries coded as "One", 37 (78.7%) were correctly classified, whereas 152 (92.7%) out of the 164 nations coded as "Zero" were correctly categorized.

Section 3 in Figure 2 presents the actual outcome of the WEKA software, revealing the actual country-level patterns or aggregate conditions that seem to determine the NFFC. In concrete terms, this particular decision tree has two central branches, implying the identification of two main groups of countries that send foreign fighters to ISIS: those that are "nearby" and the ones that are "faraway" (Sections 3.1 and 3.2 in Figure 2, respectively). In this division, the algorithm determined that, to maximize the correct classification of instances, 4,264 kilometers (parameter "AA") was the cut-off distance between both sets. The closer cluster of nations is further subdivided into five subgroups, each one constituting a different sub-branch and leaf of this decision tree. Therefore, our data mining approach discovered a total of six distinct country-level patterns to explain the ISIS foreign fighter phenomenon.

In order to ease the interpretation of Figure 2 and to delve into the relevant matters which fulfill our article's purpose, we built a path diagram showing all six country-level patterns found (Figure 3). We labeled the faraway group as "Cluster FA", and the five nearby groups as "Clusters NB" (NB1, NB2, NB3, NB4, and NB5).

Figure 3
Conceptual "J48" tree diagram showing the path to each of the six country-level patterns or clusters.



Source: Elaborated by the authors, using WEKA software.

The first issue that stands out from Figure 3 is that only a combination of 11 out of the 22 attributes originally entered in the J48 tree algorithm (Appendix) was finally used by the procedure, thus discarding the remaining ones for not being part of the derived country-level patterns. This validates the claim according to which inferring something about the target variable starting from just one variable – even if this variable has high predictive power – will somewhat lead to a misinterpretation of the phenomenon (Fayyad et al., 1996). Table 1 presents not only a list of these 11 independent variables used by the decision tree, but also a summary of each cluster's profile. This clarifies the characterization of each group of countries in terms of a set of attributes. We highlight the absence of certain factors previously proposed by extant literature as causes of the foreign fighter phenomenon. For instance, despite the assertions made by Eweiss (2016) and Speckhard (2006), attack capabilities (I18) and the number of exported weapons (I19) did not form part of any cluster.

Table 1
Final 11 attributes or independent variables used by the J48 algorithm, and summary of the cluster profiles or patterns.

Clusters	Independent variables or attributes										
	Spatial	Structural					Social				
	AA	A185	A529	A624	C1	E1	B1	B2	D7	H1	I4
FA	AA > 4,264						> 240,000		> 61.9		
NB1	AA ≤ 4,264				≤ 1,110,043	≤-2,13		> 0			
NB2	1,587 < AA≤ 4,264				> 1,110,043			> 0		≤ 4.76	≤ 3
NB3	AA ≤ 4,264	> 3,155			> 1,110,043		≤ 95,000	> 0		> 4.76	≤ 3
NB4	AA ≤ 4,264		≤ 61.2	≤ 4,430	> 1,110,043		> 95,000	> 0		> 4.76	≤ 3
NB5	AA ≤ 4,264			> 4,430	> 1,110,043		> 95,000	> 0		> 4.76	≤ 3

Source: Elaborated by the authors, using WEKA software.

The second issue from Figure 3 we emphasize is the fact that two prevailing conditions underpin the six country-level patterns: distance from the ISIS conflict (AA; see Table 1, Appendix) and the presence or quantity of Muslim population (B1/B2). All branches of Clusters FA and NB contain at least one of these two components, indicating

that the combination of these spatial and social parameters seem to be paramount factors in the generation of the necessary conditions for the rise of the phenomenon in a country. Besides both conditions, other attributes also further determine each set of countries.

The primary trait of the countries belonging to Cluster FA (Figure 3; Table 1) is their remoteness from the ISIS conflict (AA > 4,264). Moreover, this group possesses a higher alcohol consumption among women (D1 > 61.9) and significant Muslim population in absolute terms (B1 > 240,000). Our analysis determined that four countries present these characteristics: Australia, Canada, the Russian Federation, and the United States.

As for the sets of nearby nations (AA ≤ 4,264; Clusters NB1, NB2, NB3, NB4, and NB5), as previously noted, they also present Muslim populations (B2 > 0). Cluster NB1 (Figure 3; Table 1) is also characterized by a lower number of internet users (C1 ≤ 1,110,043) and a weaker level of institutional quality in terms of voice and accountability (E1 ≤ -2.13). Only two countries fit this description: Somalia and Turkmenistan.

Besides proximity to the conflict and presence of Muslim populations, Clusters NB2, NB3, NB4, and NB5 also share a higher number of internet users (C1 > 1,110,043) and a lower homicide rate (I4 ≤ 3). Besides these attributes, Cluster NB2 (Figure 3; Table 1) is characterized by a lower evaluation of the life ladder (H1 ≤ 4.76) and a medium distance to the ISIS conflict (AA > 1,587). Only two countries belong to this cluster: Ukraine and Yemen.

Although Clusters NB3 (Figure 3; Table 1) and NB2 share their nearness to ISIS territories (AA ≤ 4,264), the presence of Muslim individuals (B2 > 0), a higher number of internet users (C1 > 1,110,043), and a lower homicide rate (I4 ≤ 3), the former is in essence different due to a higher life ladder evaluation (H1 > 4.76), a lower number of Muslim individuals in absolute terms (B1 ≤ 95,000), and a higher health expense per capita (A185 > 3,155). Merely two countries were classified in this set: Finland and Ireland.

Cluster NB4 (Figure 3; Table 1) shares several aforementioned characteristics: nearness (AA ≤ 4,264), a salient presence of Muslim populations (B1 > 95,000; B2 > 0), a higher number of internet users (C1 > 1,110,043), a lower homicide rate (I4 ≤ 3), and a higher life ladder

evaluation (H1 > 4.76). However, it is also defined by a lower degree of international reserves (e.g. monetary gold; A624 ≤ 4,430) and a lower gross domestic product (A529 ≤ 61.2). Four countries represent this cluster: Albania, Bahrain, Bosnia and Herzegovina, and Tajikistan.

Finally, Cluster NB5 (Figure 3; Table 1) differs in its higher number of international reserves (A624 > 4,430). Also, other independent variables distinguish this particular country set: closeness to ISIS (AA ≤ 4,264), an important presence of Muslim people (B1 > 95,000; B2 > 0), a higher number of internet users (C1 > 1,110,043), a lower homicide rate (I4 ≤ 3), and a higher life ladder evaluation (H1 > 4.76). This is by far the largest cluster, with 23 nations in total: Algeria, Austria, Belgium, Denmark, Egypt, France, Germany, Israel, Italy, Jordan, Kazakhstan, Kuwait, Lebanon, Libya, the Netherlands, Pakistan, Saudi Arabia, Serbia, Spain, Sweden, Tunisia, Turkey, and the United Kingdom.

DISCUSSION: COUNTRY PATTERN ANALYSIS

Academic aspects

Before entering the discussion about country patterns or clusters, we highlight that two issues seem to explain a significant part of the foreign fighter phenomenon. First, consistent with the premises of Hewitt and Kelley-Moore (2009), we find that physical proximity (AA, Appendix) to the ISIS conflict is a key aspect accounting for the issue (Clusters NB1-5; Figure 3; Table 1). This corroborates the fact that nations such as Kuwait, Jordan, and Saudi Arabia have sent the greatest numbers of foreign fighters. Our evidence thus supports other authors' assertions concerning the way nearness improves logistic aspects of joining ISIS, for instance, the capability of arriving to Syria (Barrett, 2014; de Bie et al., 2015; Hegghammer, 2013; Nilsson, 2015). Moreover, since veteran foreign fighters who return to their home countries could favor the phenomenon by acting as communicators, closeness to radical terrorism information hubs or nodes remains an important factor for the emergence of FFs (Bousquet, 2012; Klausen, 2015; Vidino, 2014; Weggemans et al., 2014; Zammit, 2015).

Second, besides physical proximity, the number of ISIS foreign fighters seems to be linked to the Muslim religion, as at least one of the parameters (B1/B2, Appendix) has an influence in the NFFC in all six

country-level patterns or clusters (Figure 3; Table 1). A caveat, however: consistent with previous literature (Hewitt; Kelley-Moore, 2009; Klausen, 2015; Malet, 2010; Weggemans et al., 2014), we highlight that it is not the faith itself, but the religious fundamentalist view that some factions have undertaken to violently try to "save" Islam from decaying worldwide. Consequently, and despite exceptions (e.g. Cluster NB3), our interpretation of this finding is that a higher number of Muslims could lead to better conditions for the growth of FF seeds.

Even though the implications discussed in the two previous paragraphs are essential to the phenomenon under study, they are not surprising. It is rather a matter of emphasizing the patterns arising from each set of countries – i.e., the combination of spatial, structural, and social country-level variables. Thus, we warn readers that the following interpretations are aimed at trying to conceptually and abductively explain how some characteristics or conditions in a certain setting (the cluster) emerge, and not to understand individuals' deeper motivations, as this would imply a totally different level of analysis.

Regarding countries in Cluster FA (Figure 3; Table 1), there seems to be an interaction between being distant (AA > 4,264), having a non-Islam majority and a Western culture (Australia, Canada, Russia, and the US), alcohol consumption in women (D7 > 61.9), and large Muslim populations (B2 > 240,000). This is interesting because, although the Islamic religion forbids drinking, evidence shows that individuals in places distant from Muslim countries ingest it nonetheless, and in some cases vastly (Cochrane; Bal, 1990; Neumark et al., 2001). Reasons range from peer pressure to conformity to Western social norms. According to Valentine et al. (2010), the main underlying issue is that alcohol-consuming Muslims are oftentimes stigmatized and alienated in their communities for violating religious dogma. As for Muslim women, this stigmatization could be more severe, since it has been stated that Islamic culture has higher expectations about their role in society and family (e.g. strengthening ties of kinship) (Perešin, 2018).

Hence, we can conjecture that Muslim women in these conditions could become pariahs, because they suffer alienation in two fronts: within Islamic communities due to their alcohol-related habit (Valentine et al., 2010), but also due to xenophobia from other citizens who condemn their religion (Perešin, 2015). The literature emphasizes that this makes them disenchanted with Western values, as they have failed

to reconcile domestic practices (e.g. alcohol consumption) with their traditions (i.e., Sharia Law). The mismatch between religious tradition and (host) country values in this cluster is a (social) factor which might encourage the development and growth of what has been called ISIS's utopian message of camaraderie and sisterhood (Perešin, 2018; Shapiro; Maras, 2019). As a consequence, it may make sense to cross great distances to experiment a change by becoming a FF (e.g. a sort of emancipation from Western oppression) and living according to the Islam lifestyle (Kneip, 2016; Perešin, 2018) while maintaining – according to them – some Western benefits (Perešin; Cervone, 2015; Speckhard, 2016), such as certain foods (like alcohol) or activities (Leong, 2015).

Both Somalia and Turkmenistan (Cluster NB1; Figure 3; Table 1) are societies strongly based on Muslim family clan identity, which makes it the essential social unit (Dyner et al., 2015; Joosse et al., 2015; Karagiannis, 2013). However, their authoritarian governments limit freedom of association (Duyvesteyn; Peeters, 2015; Zogg, 2016). This is consistent with the fact that this country set has a lower degree of voice and accountability (E1 ≤ –2.13), which measures citizens' capability of participating in government, as well as overall freedom (Appendix). According to Dyner et al. (2015) and Duyvesteyn and Peeters (2015), issues such as political repression, marginalization, and poor governance pave the way for the occurrence of the FF phenomenon. Due to these institutional weaknesses, Muslim family clans are not able to freely associate as they wish, and also cannot voice or channel their concerns so governments can improve their day-to-day life.

Consequently, (low) governance conditions in these settings can enhance the search for a place where a community or clan-based life – echoing Islamic principles – is possible. As mentioned above, propaganda is prone to showing images of a good quality of life in ISIS-controlled territories (Leong, 2015; Perešin, 2015; Perešin; Cervone, 2015), which could match their desire to find a venue where they can freely express their culture. Even though this group of countries has a lower internet usage (C1 ≤ 1,110,043), their relative proximity to ISIS suggests closeness to communication hubs or nodes that meet individual needs of information, making access to the worldwide web unnecessary (Dyner et al., 2015; Karagiannis, 2013; Watts, 2008). Therefore, the characteristics of this scenario could be significant drivers or "hooks" for joining this radical cause (Joosse et al., 2015).

Contrary to what Barrett (2014) and Zammit (2015) ascertain, we do not find compelling evidence to substantiate the "no culture" hypothesis in faraway countries (Cluster FA). On the contrary, unexpected evidence for this hypothesis was found in Cluster NB2 (Figure 3; Table 1), composed by two nearby countries: Ukraine and Yemen. Although for different reasons, these two countries differ for having been facing internal armed conflicts recently (Byman, 2016; Freedman, 2014). They are namely the Crimea crisis in Ukraine and the Sa'dah revolt and the Southern separatist movement in Yemen (Mantzikos, 2011; Thiele, 2015). According to Mantzikos (2011) and Zhukov (2016), this has had two negative effects: first, from a structural lens, it has reduced job opportunities, as most efforts were redirected to address these potential civil wars. Second, from a social lens, this scenario encourages unemployed Muslim citizens to join a fight to defend certain interests or territories, which has given rise to civilian militias.

Our evidence shows this belligerent scenario is likely to create the necessary conditions for a reduction in the overall perception of well-being and citizens' stand in life (H1 ≤ 4.76) in these countries. Taking up arms seems to be an outlet to this situation. However, these national characteristics suggest that, instead of joining internal conflicts, Muslims (B2 > 0) in these nations channel this external combative incentive by joining jihad-related causes, and particularly ISIS. This is due to the fact that, if they are to engage in combat for the aforementioned contextual causes, they may be more prone to fighting in favor of an ideal aligned with their particular interests – i.e., their religion (Byman, 2016; Thiele, 2015). Therefore, our evidence could indicate that the "no future" hypothesis seems true for "intermediate" countries, such as Ukraine and Yemen (1,587 < AA ≤ 4,264). This is complemented by the fact that, also consistent with our findings (C1 > 1,110,043), internet and social networks seem to be the main means of communication to lure Muslims in this set.

Cluster NB3 (Figure 3; Table 1) is an odd grouping, because it is constituted by two nearby countries, Finland and Ireland, which, according to our results, tend to have high levels of economic and social development (A185 > 3,155), and a small Muslim population (B1 ≤ 95,000). Hence, one might intuitively assume that none of these countries' conditions justify *a priori* the growth of the FF phenomenon, as their citizens in general have a good quality of life and there are no evident religious traditions (and confrontations). Nevertheless, it has been argued that social and

economic prosperity in some Western cultures, such as Finland, may lay the foundations for a kind of "boredom culture" (Haller et al., 2013). This may have an impact on the phenomenon studied, as several political science scholars have claimed that some citizens in developed countries could experiment what has been termed the "Hemingway Effect" (Hegghammer, 2013), which alludes to the search for adventure (Barrett, 2014; Malet, 2010; van Zuijdewijn, 2014; Watts, 2008; Zafar, 2014).

Although Cronin (2015) concurs with the adventurism hypothesis, the author states that it is not the only factor for the emergence of the FF problem in developed countries, such as Finland and Ireland. As people have their basic human needs (e.g. food, safety, health) met, these societies might engender a sense of so-called "humanitarian duty" to try and help other places achieve the same conditions (Mendelsohn, 2011; van Arsdale; Nockerts, 2008; Zafar, 2014). However, following Vaux (2006), these duties are essentially based on biased, distorted, or even naïve ways of understanding and dealing with the ISIS menace, which is possibly due to the low number of Muslims that live in this cluster's countries. In sum, the pattern of characteristics featured by Cluster NB3 apparently creates appropriate conditions to nurture both boredom and humanitarian duty cultures, boosting the ISIS cause in these prosperous settings.

The countries grouped in Cluster NB4 (Figure 3; Table 1) are close to the conflict (AA ≤ 4,264), have a significant number of Muslims (B1 > 0; B2 > 95,000), and, despite having a high life ladder (H1 > 4.76), they are economically underdeveloped in relative terms (A529 ≤ 61.2; A624 ≤ 4,430). The salience of this last issue is consistent with the arguments presented by Dyner et al. (2015) and Shtuni (2015) concerning Albania's, Bosnia and Herzegovina's, and Tajikistan's – three of the four nations embodying this set – economic lag (e.g. GDP, education, competitiveness, urbanization) in comparison with their European and Central Asian counterparts. Consistent with de Bie et al.'s (2015) "situational induced motivation" hypothesis, these authors ascertain that poverty is a strong component which could create an adequate context for migratory processes in the search for better living conditions, such as higher wages and spending in education (Dyner et al., 2015).

This issue, added to the widespread use of the internet in Albania, Bahrain, Bosnia and Herzegovina, and Tajikistan (C1 > 1,110,043) might favor the indoctrination of these poor and uninformed individuals

(Dyner et al., 2015; Shtuni, 2015). This is especially due to the role of ISIS propaganda, which shows they can live an Islamic way of life while trying to spread and save it, as well as improving their personal economic situation (Leong, 2015; Perešin, 2015). Therefore, it is possible to conjecture that the radical message finds fertile ground to expand in nations with these structural characteristics, potentially driving Muslims' decision to be a part of ISIS.

Finally, characterizing Cluster NB5 was cumbersome, due to the heterogeneity of its constituting countries. Despite the emergence of a common pattern (Figure 3; Table 1), we have nations from very different backgrounds: Africa (Algeria, Egypt, Libya, Tunisia), Europe (Austria, Belgium, Denmark, France, Germany, Italy, The Netherlands, Serbia, Spain, Sweden, and the United Kingdom), and Asia (Israel, Jordan, Kazakhstan, Kuwait, Lebanon, Pakistan, Saudi Arabia, and Turkey). In general terms, these countries are close to ISIS-controlled territories (AA $\leq$ 4,264), have a significant Muslim population (B1 > 0; B2 > 95,000), and an overall high life ladder, which fits the sound financial position of their Central Banks or corresponding monetary authority (A624 > 4,430). In fact, concerning this last point, according to the United Nations, none of the nations in Cluster NB5 are among the world's least developed countries. They are either developed (e.g. Austria, Denmark, and the United Kingdom), in transition (Kazakhstan and Serbia), or developing (e.g. Algeria, Egypt, Pakistan) (UN, 2014).

Hence, although higher economic development in these terms does not necessarily translate into social betterment, the pattern these countries follow seems to suggest that they do not adhere to the "situational induced motivation" hypothesis (de Bie et al., 2015), as none of the abovementioned negative contextual conditions (e.g. poverty or poor education) seem to necessarily contribute to the occurrence of the FF phenomenon in these settings. Although we analyze the patterns that describe this cluster of countries, it is important to note that our aggregate quantitative evidence seemingly contradicts some recent studies, at least regarding specific nations in this group. Such is the case of Tunisia, where Zelin's (2020) historical ethnography suggests the occurrence of a "low life ladder", as the expectations of young people regarding employment – especially after the fall of president Ben Ali in 2011 – have not been met, not living the life they expected, and consequently looking for alternatives (e.g. joining ISIS).

Beyond this caution, a distinctive characteristic emphasized by extant literature is that these somewhat financially developed countries share a vaster access to the internet (C1 > 1,110,043) (Bakker et al., 2014; Hegghammer, 2013; Heinke; Raudszus, 2015; Nilsson, 2015; Vidino, 2014; Watts, 2008; Weggemans et al., 2014; Zafar, 2014). As we have indicated, this is one of the main channels through which recruiters reach foreigners (Leong, 2015; Perešin, 2015; Perešin; Cervone, 2015; Speckhard, 2016). This should come as no surprise, as the effect of social media on people has been extensively researched, and it seems that the ISIS FF phenomenon is not an exception. Hence, our interpretation is that Muslims living in this diverse group of nations are likely to constitute a significant portion of the internet user base. In this scenario, and consistent with several scholars' position, access to the worldwide web seems to be the main enabling structural condition for Islam followers to receive ISIS propaganda, to search for more information about this terrorist cell's operations on their own, and to contact key information nodes or hubs to gain knowledge on how to travel and join the jihadist cause (Malet, 2010; Mendelsohn, 2011; van Zuijdewijn, 2014).

General practical recommendations

From some of our findings, we can propose practitioner-oriented recommendations for governments and policy makers worldwide, to aid them in adequately facing the FF problem at country level and, hopefully, winning the battle against ISIS. Our suggestions are specifically aimed at operationalizing our findings. Hence, from a practical standpoint, an essential contribution of this study is showing that understanding countries as clusters or sets enables the proposition of distinct and specific solutions for each of them. Thus, state officials can improve resource allocation and build personalized undertakings for each group of countries.

With reference to the farthest away cluster (FA), we find that the "no future" subculture is not necessarily an issue as suggested by previous works (Barrett, 2014; Zammit, 2015). Contrarily, a specific social aspect – high alcohol consumption among women –, added to religion, seems to be a triggering factor, because it alienates Muslim women. Therefore, to stop the rise of foreign fighters, countries such as Australia, Canada, Russia, and the United States should focus their social program efforts on two different fronts: first, social integration between locals

and Muslims; and second, reducing xenophobia among non-Muslims. Consistent with what Valentine et al. (2010) mention, it is recommended that these governments manage the consumption of alcohol in Muslim-majority neighborhoods and eliminate discrimination to avoid creating feelings of social marginalization.

Our tree algorithm in Figure 3 (Table 1) reveals five specific groups of nearby countries which send foreign fighters (NB1-5). Governments fighting ISIS should prioritize investments that stop the effect of communication hubs and nodes, as previous works have implied (Malet, 2010; Mendelsohn, 2011). In concrete terms, veteran FFs constitute a significant factor enhancing the proclivity of enrollment (Bakker et al., 2014; Nilsson, 2015; van Zuijdewijn, 2014; Watts, 2008). Thus, governments' initiatives are to be targeted at either blocking the entrance of individuals returning from conflict areas in Syria and/or Iraq, or permanently keeping an eye on them once returned.

The public policy recommendation for Cluster NB1 is clear: the governments of Somalia and Turkmenistan must improve their institutional quality. This is especially true in the case of these nations' voice and accountability (e.g. overall freedom), as their authoritarian governments hinder a family clan lifestyle and the possibility of improving governance aspects (Duyvesteyn; Peeters, 2015; Zogg, 2016). Because low levels of this variable could encourage people to seek a sense of community elsewhere – and particularly within ISIS as FFs –, it is essential to guarantee freedom of association and the ability to participate in law-making (Dyner et al., 2015; Joosse et al., 2015). All this could be considered obvious or simple; however, it is not an easy endeavor to achieve, because it implies vast amounts of time. A first step to start a long-term institutional process to improve democracy in these countries could be to empower civil society organizations (e.g. non-governmental organizations – NGOs), as their characteristics have been proposed to make them key means to effectively spread the democratic message (Edwards et al., 1999; Kamat, 2003). In this sense, and following Barber's (1998) ideas, NGOs can use technology – like ISIS does – to educate the population about democracy. Yet, as these countries have low internet usage rates, our recommendation would be that NGOs must directly engage in alliances with local authorities (e.g. municipalities) or local independent organizations to build the foundations for participatory democratic governance (de Witt, 2000; Edwards et al., 1999).

In the specific case of Cluster NB2, policy should be targeted at improving society's perception of the life ladder. This seems tough in the short term, due to the ongoing armed conflicts in Ukraine and Yemen (Byman, 2016; Freedman, 2014). However, expansive macroeconomic policies may ameliorate overwhelming unemployment in these nations. This way, people's economic well-being can improve – at least for some time –, hopefully deterring Muslim citizens' decision to join civilian militias or ISIS.

Concerning Cluster NB3, which is composed by Finland and Ireland, state officials should attempt to improve the lifestyle of citizens who are considering joining ISIS (Barrett, 2014; Coolsaet, 2016; Malet, 2010; van Zuijdewijn, 2014; Watts, 2008; Zafar, 2014). Specific social programs, for example, may provide adventurous alternatives for young people aged 20-35. Governments can offer planned social activities in favor of other parts of the world where there is a lesser risk of becoming a FF (e.g. Western Africa or Latin America), which can nonetheless satisfy citizens' need to help a humanitarian cause. The goal is to redirect any tendency to enroll in a terrorist cell toward a truly altruistic initiative, and hence avoid the decision to help while undertaking risky activities.

Cluster NB4 reveals that economic conditions could be a triggering factor when added to a large Muslim population and pervasive internet access (Dyner et al., 2015; Shtuni, 2015). These countries are relatively poorer – in average – than those in other country sets. Therefore, the governments of Albania, Bahrain, Bosnia and Herzegovina, and Tajikistan should focus on improving structural conditions, particularly the amount of money spent on individuals' education. Following our discussion above, this could have the side-effect of discouraging ideological indoctrination through the internet by showing the truth about ISIS.

Finally, Cluster NB5 points to the importance of internet access in spreading jihad around the world. Despite having a considerable Muslim population with good overall social and economic development, they still have incentives to join ISIS as FFs. According to our findings and discussion above, access to the internet seems to be a vital factor for the number of foreign fighters, as it constitutes a key source of extremist information (Bakker et al., 2014; Hegghammer, 2013; Heinke; Raudszus, 2015; Nilsson, 2015; Watts, 2008; Zafar, 2014). The nations in this set could perhaps use internet monitoring systems that would allow the detection of people searching for terrorist content. Our evidence suggests that the

number of users is critical in the decision to become an ISIS non-national combatant. These policies, however, must be implemented in specific cases and in strict agreement with democratic principles – i.e., only with the endorsement of an independent judicial system –, so they do not become authoritarian surveillance tools.

CONCLUSION

The main contribution of this work is moving beyond traditional social science methods towards a more sophisticated exploratory procedure based on data mining (KDD) to better understand the country-level conditions which enable the rise of the FF phenomenon. To this end, we apply this procedure to the case of ISIS, discovering attributes and patterns which could explain NFFC rates. After identifying attributes from over 950 different parameters and inquiring into their relationship with the dependent variable, six different country profiles were constructed to better understand and address this threat. Although we have focused on ISIS in this article, the methodological approach undertaken can be used to find potential factors of social environments favorable to the escalation of armed conflict or terrorist initiatives – for instance, to examine the ongoing violent issues in Chechnya and Somalia (Duyvesteyn; Peeters, 2015).

Although the six country sets we identified improve our understanding of a complex issue such as the necessary country-level characteristics that could enhance FF movements, no work is free from limitations. The methodology used leads to the necessary consideration of at least three potential limitations. First, the possibility of being in the presence of an "ecological fallacy" problem (Wakefield; Lyons, 2010), understood as making inferences about individuals' behavior based on macro-level data (King et al., 2004). Although we are linking country-level (macro) variables to the aggregate number of FFs (NFFC), we have been circumspect regarding the research objective and questions, and the discussions as well, in order not to confuse readers. In fact, we have tried to be clear in stating that we are neither inferring nor discussing people's deeper motivations to join ISIS. Instead, we are establishing general spatial, structural, and social conditions which, to a certain degree, could create an "environment" which generally fosters the emergence of the FF phenomenon in specific sender coun-

tries. In this sense, our NFFC variable is also a country-level variable being affected by an intricate relationship of national variables, none of which can explain the phenomenon on its own.

Second, one must consider the information used to configure the target data (NFFC), because its numbers were based on open-source evidence obtained from the International Centre for the Study of Radicalization and Political Violence. This institution has kept count of the number of foreign fighters in Syria and Iraq since 2012, but the data is not always totally precise, given the difficulty of measurement, which can lead to some loss of richness and/or precision in the analysis. Therefore, future work is to replicate this study with new, larger, and more precise and depurated databases.

Third, since the method applied dropped some variables – both in attribute pre-selection and data mining processes –, and although all indicators of goodness of fit show a robust model (see, Figure 2), we have to deal with the loss of some important variables which could partially explain the phenomenon: in the search for parsimony in our findings, we lose a certain degree of internal validity. Moreover, as a study based on a quantitative technique, we cannot compete with the explanatory depth of qualitative studies, such as ethnographic ones (e.g. Van San, 2018). In fact, these qualitative studies can grasp the phenomenon in multiple directions and levels, offering alternative explanations for it. We, instead, are limited to interpreting the results by strictly following the variables considered important by the algorithm, which could be considered a narrow view for these social phenomena. However, our main strength is the gain of external validity and generalization power due to the technique and database used (Shadish et al., 2002).

Actually, the main virtue of using data mining techniques to research complex social systems is their ability to find which elements in vast amounts of data might be the precursory foundations of a phenomenon. Based on our results, future investigations should now deploy other social science methodologies (both quantitative and qualitative) to deepen our findings. For example, from a quantitative lens, scholars can first use confirmatory factor analysis to further validate the existence of the six country-level clusters, to substantiate the attributes, and subsequently undertake cluster analysis and see if sets coincide.

From a qualitative perspective, researchers can endorse our findings internally by doing ethnographic studies which may explain the phenomenon further and deeper.

There is no denying that the initiatives of governments worldwide to deal with ISIS should be multifaceted. A key front requiring the attention of policy makers is knowing how to reduce the number of foreigners that travel to join this jihadist cause. We hope that our findings lead state officials to start thinking in a more comprehensive way – considering several variables and country-level patterns – to reallocate resources. This way they can design and implement better customized strategies to deal with this menace, depending on the specific country cluster identified by our article.

(Received on December 2, 2019)
(Resubmitted on December 15, 2020)
(Accepted on March 22, 2021)

NOTES

1. It is important to avoid confusing qualitative and quantitative uses of this technique. In the latter, the objective goes beyond the validation and right selection of variables, but rather aims to determine the predictive power of some variables. See, for example, Giannarou and Zervas (2014) and Marquardt and Pemstein (2018).
2. The exception to this criterion was the Russian Federation, due to its size. In this case, we used the distance between Damascus and a midpoint of the Russian Federation.
3. The World Happiness Report is a survey of the state of global happiness based on different perspectives, such as health, government statistics, polls, economy indexes, and psychology.

REFERENCES

BAKKE, Kristin. (2014), "Help wanted? The mixed record of foreign fighters in domestic insurgencies". *International Security*, v. 38, n. 4, pp. 150-187.

BAKKER, Edwin; PAULUSEEN, Christoph; ENTENMANN, Eva. (2014), "Returning jihadists foreign fighters". *Security and Human Rights*, v. 25, n. 1, pp. 11-32.

BARBER, Benjamin. (1998), *A Place for Us: How to Make Society Civil and Democracy Strong*. New York: Hill and Wang.

BARRETT, Richard. (2014), *Foreign Fighters in Syria*. New York: The Soufan Group.

BENMELECH, Efraim; KLOR, Esteban. (2020), "What explains the flow of foreign fighters to ISIS?". *Terrorism and Political Violence*, v. 32, n. 7, pp. 1458-1481.

BOUSQUET, Antoine. (2012), "Complexity theory and the war on terror: Understanding the self-organizing dynamics of leaderless jihad". *Journal of International Relations and Development*, v. 15, n. 3, pp. 345-369.

BRAUN, Virginia; CLARKE, Victoria. (2006), "Using thematic analysis in psychology". *Qualitative Research in Psychology*, v. 3, n. 2, pp. 77-101.

BYMAN, Daniel. (2016), "ISIS goes global: Fight the Islamic State by targeting its affiliates". *Foreign Affairs*, v. 95, n. 2, pp. 76-85.

BYRNE, David; CALLAGHAN, Gill. (2014), *Complexity Theory and the Social Sciences: The State of the Art*. Routledge.

CAO, Longbing; ZHANG, Huaifeng; ZHAO, Yanchang; LUO, Dan; ZHANG, Chengqi. (2011), "Combined mining: Discovering informative knowledge in complex data". *IEEE Transactions on Systems, Man, and Cybernetics*, v. 41, n. 3, pp. 699-712.

CHEN, Xiaolin; WU, Jia; CAI, Zhihua. (2011), "Learning the attribute selection measures for decision tree". *Energy Procedia*, v. 13, pp. 3333-3341.

CHESTERMAN, Simon. (2016), "Dogs of war or jackals of terror? Foreign fighters and mercenaries in international law". *International Community Law Review*, v. 18, n. 5, pp. 389-399.

COCHRANE, Raymond; BAL, Sukhwant. (1990), "The drinking habits of Sikh, Hindu, Muslim and white men in the West Midlands: A community survey". *British Journal of Addiction*, v. 85, n. 6, pp. 759-769.

COOLSAET, Rik. (2016), *Facing the fourth foreign fighters wave*. London: Egmont Institute.

CRAGIN, Kim. (2017), "The challenge of foreign fighter returnees". *Journal of Contemporary Criminal Justice*, v. 33, n. 3, pp. 292-312.

CRESWELL, John; CLARK, Vicki. (2006), *Designing and Conducting Mixed Method Research*. London: Sage.

CRONIN, Audrey. (2015), "ISIS is not a terrorist group: Why counterterrorism won't stop the latest jihadist threat". *Foreign Affairs*, v. 94, n. 2, pp. 87-98.

DE BIE, Jasper; DE POOT, Christianne; VAN DER LEUN, Joanne. (2015), "Shifting modus operandi of jihadist foreign fighters from the Netherlands between 2000 and 2013". *Terrorism and Political Violence*, vol. 27, n. 3, pp. 416-440.

DE VILLE, Barry; NEVILLE, Padraic. (2013), *Decision Trees for Analytics using SAS Enterprise Miner*. Cary: SAS Institute.

DE WIT, Joop. (2000), "Towards good governance at the local level: The role of grassroots institutions". *ISS Working Papers – General Series*, v. 325, pp. 1-40.

DUYVESTEYN, Isabelle; PEETERS, Bram. (2015), *Fickle foreign fighters? A cross-case analysis of seven Muslim foreign fighter mobilizations (1980-2015)*. International Centre for Counter--Terrorism.

DYNER, Anna; LEGIE , Arkadiusz; R KAWEK, Kacper. (2015), *Ready to go? ISIS and its presumed expansion into Central Asia*. Warsaw: The Polish Institute of International Affairs.

EDWARDS, M.; HULME, D.; WALLACE, T. (1999). "NGOs in a global future: Marrying local delivery to worldwide leverage". *Public Administration and Development*, v. 19, n. 2, pp. 117-136.

EWEISS, Nada. (2016), *Non-state actors & WMD: Does ISIS have a pathway to a nuclear weapon?*. London: British American Security Information Council.

FAYYAD, Usama; PIATETSKY-SHAPIRO, Gregory; SMYTH, Padhraic. (1996), "From data mining to knowledge discovery in databases". *AI Magazine*, v. 17, n. 3, pp. 37-54.

FREEDMAN, Lawrence. (2014), "Ukraine and the art of limited war". *Survival*, v. 56, n. 6, pp. 7-38.

GIANNAROU, Lefkothea; ZERVAS, Efthimios. (2014), "Using Delphi technique to build consensus in practice". *International Journal of Business Science & Applied Management*, v. 9, n. 2, pp. 65-82.

GILBERT, Nigel; BULLOCK, Seth. (2014), "Complexity at the social science interface". *Complexity*, v. 19, n. 6, pp. 1-4.

GIUDICI, Paolo. (2010), "Data mining model comparison", *in* O. Maimon and L. Rokach (eds.), *Data Mining and Knowledge Discovery Handbook*. New York, Springer, pp. 641-654.

HAIR, Joseph.; BLACK, William; BABIN, Barry; ANDERSON, Rolph. (2010), *Multivariate Data Analysis*. Upper Saddle River: Prentice Hall.

HALL, Mark; HOLMES, Geoffrey. (2003), "Benchmarking attribute selection techniques for discrete class data mining". *IEEE Transactions on Knowledge and Data Engineering*, v. 15, n. 6, pp. 1437-1447.

HALLER, Max; HADLER, Markus; KAUP, Gerd. (2013), "Leisure time in modern societies: A new source of boredom and stress?", *Social Indicators Research*, v. 111, n. 2, pp. 403-434.

HEGGHAMMER, Thomas. (2013), "Should I stay or should I go? Explaining variation in Western jihadists' choice between domestic and foreign fighting". *American Political Science Review*, v. 107, n. 1, pp. 1-15.

HEINKE, Daniel; RAUDSZUS, Jan. (2015), "German foreign fighters in Syria and Iraq". *CTC Sentinel*, v. 8, n. 1, pp. 18-21.

HEWITT, Christopher; KELLEY-MOORE, Jessica. (2009), "Foreign fighters in Iraq: A cross--national analysis of jihadism". *Terrorism and Political Violence*, v. 21, n. 2, pp. 211-220.

HOMELAND SECURITY COMMITTEE (HSC). (2015), *Combating terrorist and foreign fighter travel*. Washington: US House of Representatives.

INSTITUTE FOR ECONOMICS & PEACE (IEP). *Global peace index* [online]. 2014 [15-09-2015]. Retrieved from http://economicsandpeace.org/wp-content/uploads/2015/06/2014--Global-Peace-Index-REPORT_0-1.pdf.

INTERNATIONAL CENTRE FOR THE STUDY OF RADICALIZATION AND POLITICAL VIOLENCE (ICSR). *Foreign fighter total in Syria/Iraq now exceeds 20,000* [online]. 2014 [15-09-2015]. Retrieved from http://icsr.info/2015/01/foreign-fighter-total-syriairaq--now-exceeds-20000-surpasses-afghanistan-conflict-1980s/.

INTERNATIONAL LABOR ORGANIZATION (ILO). *Global Employment Trends 2014: Supporting Data Sets* [online]. 2014 [02-09-2015]. Retrieved from www.ilo.org/global/research/global-reports/global-employment-trends/2014/WCMS_234879/lang--en/index.htm.

INTERNET LIVE STATS. *Internet users by country* [online]. 2014 [30-08-2015]. Retrieved from www.internetlivestats.com/internet-users-by-country/.

JOFFÉ, George. (2016), "Global jihad and foreign fighters". *Small Wars & Insurgencies*, v. 27, n. 5, pp. 800-816.

JOOSSE, Paul; BUCERIUS, Sandra; THOMPSON, Sara. (2015), "Narratives and counter-narratives: Somali-Canadians on recruitment as foreign fighters to Al-Shabaab". *British Journal of Criminology*, v. 55, n. 4, pp. 811-832.

KAMAT, Sangeeta. (2003), "NGOs and the new democracy". *Harvard International Review*, v. 25, n. 1, pp. 65-69.

KARAGIANNIS, Emmanuel. (2013), "Transnational Islamist networks: Western fighters in Afghanistan, Somalia and Syria". *International Spectator*, v. 48, n. 4, pp. 119-134.

KING, Gary.; ROSEN, Ori; TANNER, Martin. (2004), "Information in ecological inference: An introduction", *in* G. King, O. Rosen, and M. Tanner (eds.), *Ecological Inference: New Methodological Strategies*. Cambridge, Cambridge University Press, pp. 1-12.

KLAUSEN, Jytte. (2015), "Tweeting the *jihad*: Social media networks of Western foreign fighters in Syria and Iraq". *Studies in Conflict & Terrorism*, v. 38, n. 1, pp. 1-22.

KNEIP, Katharina. (2016), "Female jihad – Women in ISIS". *IAPSS Political Science Journal*, v. 29, pp. 88-106.

LEONG, Dymples. (2015), *Why ISIS appeals to Muslim women in Western countries*. Singapore: Rajaratnam School of International Studies.

MAIMON, Oded; ROKACH, Lior. (2010), "Introduction to knowledge discovery and data mining", *in* O. Maimon and L. Rokach (eds.), *Data Mining and Knowledge Discovery Handbook*. New York, Springer, pp. 1-15.

MALET, David. (2010), "Why foreign fighters? Historical perspectives and solutions". *Orbis*, v. 54, n. 1, pp. 97-114.

MALETIC, Jonathan; MARCUS, Andrian. (2010), "Data cleansing: A prelude to knowledge discovery", *in* O. Maimon and L. Rokach (eds.), *Data Mining and Knowledge Discovery Handbook*. New York, Springer, pp. 19-32.

MANTZIKOS, Ioannis. (2011), "Somalia and Yemen: the links between terrorism and state failure". *Digest of Middle East Studies*, v. 20, n. 2, pp. 242-260.

MARQUARDT, Kyle; PEMSTEIN, Daniel. (2018), "IRT models for expert-coded panel data". *Political Analysis*, v. 26, n. 4, pp. 431-456.

MEEK, Jack (2010), "Complexity theory for public administration and policy". *Emergence: Complexity & Organization*, v. 12, n. 1, pp. 1-14.

MENDELSOHN, Barak. (2011), "Foreign fighters: Recent trends". *Orbis*, v. 55, n. 2, pp. 189-202.

MORÇÖL, Göktu . (2010), "Issues in reconceptualizing public policy from the perspective of complexity theory". *Emergence: Complexity & Organization*, v. 12, n. 1, pp. 52-60.

NEUMARK, Yehuda; RAHAV, Giora; TEICHMAN, Meir; HASIN, Deborah. (2001), "Alcohol drinking patterns among Jewish and Arab men and women in Israel". *Journal of Studies on Alcohol*, v. 62, n. 4, pp. 443-447.

NILSSON, Marco. (2015), "Foreign fighters and the radicalization of local Jihad: Interview evidence from Swedish jihadists". *Studies in Conflict & Terrorism*, v. 38, n. 5, pp. 343-358.

PEREŠIN, Anita. (2015), "Fatal attraction: Western Muslimas and ISIS". *Perspectives on Terrorism*, v. 9, n. 3, pp. 21-38.

_____. (2018), "Why women from the West are joining ISIS". *International Annals of Criminology*, v. 56, n. 1-2, pp. 32-42.

PEREŠIN, Anita; CERVONE, Alberto. (2015), "The Western *Muhajirat* of ISIS". *Studies in Conflict & Terrorism*, v. 38, n. 7, pp. 495-509.

PEW-TEMPLETON (GLOBAL RELIGIOUS FUTURES PROJECT). *Data Explorer: Mapping the Global Muslim Population* [online]. 2014 [01-09-2015]. Retrieved from www.pewforum.org/2009/10/07/mapping-the-global-muslim-population.

REPORTERS WITHOUT BORDERS (RSF). *World Press Freedom Index* [online]. 2014 [10-09-2015]. Retrieved from http://index.rsf.org/#!/index-details.

ROKACH, Lior; MAIMON, Oded. (2010), "Classification trees", *in* O. Maimon and L. Rokach (eds.), *Data Mining and Knowledge Discovery Handbook*. New York, Springer, pp. 149-174.

SHADISH, William; COOK, Thomas; CAMPBELL, Donald. (2002), *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin.

SHAPIRO, Lauren; MARAS, Marie-Helen. (2019), "Women's radicalization to religious terrorism: An examination of ISIS cases in the United States". *Studies in Conflict & Terrorism*, v. 21, n. 1-2, pp. 88-119.

SHTUNI, Adrian. (2015), "Ethnic Albanian foreign fighters in Iraq and Syria". *CTC Sentinel*, v. 8, n. 4, pp. 11-14.

SPECKHARD, Anne. (2006), "The new global jihad, 9-11 and the use of weapons of mass destruction: Changes in mindset and modus operandi". *Democracy and Security*, v. 2, n. 2, pp. 287-298.

_____. (2016), "Talking to terrorists: what drives young people to become foreign fighters for ISIS and other terrorist groups and what can be done in response". *Freedom From Fear*, v. 11, n. 2, pp. 24-28.

SPECKHARD, Anne; SHAJKOVCI, Ardian; YAYLA, Ahmet. (2017), "What to expect following a military defeat of ISIS in Syria and Iraq?". *Journal of Terrorism Research*, v. 8, n. 1, pp. 81-89.

SULL, Donald; EISENHARDT, Kathleen. (2015), *Simple Rules: How to Thrive in a Complex World*. Boston: Houghton Mifflin Harcourt.

SUSTAINABLE DEVELOPMENT SOLUTION NETWORK (SDSN). *World Happiness Report 2015* [online]. 2015 [04-09-2015]. Retrieved from http://worldhappiness.report/download/.

TEISMAN, Geert; KLIJN, Erik-Hans. (2008), "Complexity theory and public management". *Public Administration Review*, v. 10, n. 3, pp. 287-297.

THIELE, Ralph. (2015), "Crisis in Ukraine – The emergence of hybrid warfare". *ISPSW Strategy Series: Focus on Defense and National Security*, v. 347, pp. 1-13.

TOPOGRAFISCH VERBOND ELBRUZ (TVE). *Capital to Capital Distance* [online]. 2015 [30-08-2015]. Retrieved from www.elbruz.org/the-world/capital-to-capital-distance/.

TROCHIM, William; CABRERA, Derek. (2005), "The complexity of concept mapping for policy analysis". *Emergence: Complexity & Organization*, v. 7, n. 1, pp. 11-22.

UNITED NATIONS (UN). *Country Classification* [online]. 2014 [26-06-2017]. Retrieved from www.un.org/en/development/desa/policy/wesp/wesp_current/2014wesp_country_classification.pdf [accessed 26 June 2017].

VALENTINE, Gill; HOLLOWAY, Sarah; JAYNE, Mark. (2010), "Contemporary cultures of abstinence and the nighttime economy: Muslim attitudes towards alcohol and the implications for social cohesion". *Environment and Planning A*, v. 42, n. 1, pp. 8-22.

VAN ARSDALE, Peter; NOCKERTS, Regina. (2008), "A theory of obligation". *The Journal of Humanitarian Assistance*.

VAN SAN, Marion. (2018), "Belgian and Dutch young men and women who joined ISIS: Ethnographic research among the families they left behind". *Studies in Conflict & Terrorism*, v. 41, n. 1, pp. 39-58.

VAN ZUIJDEWIJN, Jeanine. (2014), "The foreign fighters' threat: what history can (not) tell us". *Perspectives on Terrorism*, v. 8, n. 5, pp. 59-73.

VAUX, Tony. (2006). "Humanitarian trends and dilemmas". *Development in Practice*, v. 16, n. 3-4, pp. 240-254.

VIDINO, Lorenzo. (2014), "European foreign fighters in Syria: Dynamics and responses". *European View*, v. 13, n. 2, pp. 217-224.

WAKEFIELD, Jonathan; LYONS, Hillary. "Spatial aggregation and the ecological fallacy, *in* A. Gelfand, P. Diggle, M. Fuentes and P. Guttorp (eds.), *Handbook of Spatial Statistics* [online]. 2010 [02-12-2020]. Retrieved from http://carlosreynoso.com.ar/archivos/territorio/Gelfand-Handbook-of-Spatial-Statistics.pdf.

WATTS, Clinton. (2008), "Foreign fighters: How are they being recruited? Two imperfect recruitment models". *Small Wars Journal*, v. 22.

WEGGEMANS, Daan; BAKKER, Edwin; GROL, Peter. (2014), "Who are they and why do they go? The radicalization and preparatory processes of Dutch jihadist foreign fighters". *Perspectives on Terrorism*, v. 8, n. 4, pp. 100-110.

WORLD HEALTH ORGANIZATION. "Substance use and mental health". *Global Health Observatory* [online]. 2014 [01-09-2015]. Retrieved from http://apps.who.int/gho/data/?theme=home.

WITTEN, Ian; FRANK, Eibe; HALL, Mark. (2011), *Data Mining: Practical Machine Learning Tools and Techniques*. Burlington: Morgan Kaufmann.

WORLD BANK. *World Development Indicators* [online]. 2014a [30-08-2015]. Retrieved from http://wdi.worldbank.org/tables.

WORLD BANK. *Worldwide Governance Indicators* [online]. 2014b [02-09-2015]. Retrieved from http://info.worldbank.org/governance/wgi/index.aspx#home.

ZAFAR, Shaarik. (2014), *Western Foreign Fighters in Syria*. Washington: Washington Institute for Near East Policy.

ZAMMIT, Andrew. (2015), *Australian Foreign Fighters*. Sidney: Lowy Institute for International Policy.

ZELIN, Aaron. (2020), *Your Sons Are at Your Service. Tunisia's Missionaries of Jihad*. New York: Columbia University Press.

ZHUKOV, Yuri. (2016), "Trading hard hats for combat helmets: The economics of rebellion in eastern Ukraine". *Journal of Comparative Economics*, v. 44, n. 1, pp. 1-15.

ZOGG, Benno. (2016), *Putin's Next Steppe: Central Asia and Geopolitics*. Zurich: Center for Security Studies.

APPENDIX

Basic information of the 22 pre-selected attributes or independent variables which may be associated to NFFC.

Dimension	Variable code	Description	Details	Predictability score
Spatial	AA	Distance from a particular country to Syria in kilometers	Measured from capital to capital.	0.79
Structural	A6	Purchasing power parity gross national income (US$ billions)	PPP GNI is gross national income converted to international dollars using purchasing power parity rates. Gross national income is the sum of value added by all resident producers plus any product taxes (minus subsidies) not included in the valuation of output plus net receipts of primary income (compensation of employees and property income) from abroad. Data series is in current international dollars.	0.51
	A185	Health expenditure per capita (current US$)	Total health expenditure is the sum of public and private health expenditures as a ratio of total population. It covers the provision of health services (preventive and curative), family planning activities, nutrition activities, and emergency aid. Data are in current US dollars.	0.47
	A529	Gross domestic product (current US$ billions)	GDP at purchaser's prices is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. Data are in current U.S. dollars	0.47

Basic information of the 22 pre-selected attributes or independent variables which may be associated to NFFC (cont.).

Dimension	Variable code	Description	Details	Predictability score
Structural	A621	Net secondary income (US$ millions) - 2014	Secondary income refers to transfers recorded in the balance of payments whenever an economy provides or receives goods, services, income, or financial items without a *quid pro quo*. All transfers not considered to be capital are current. Data are in current U.S. dollars.	0.54
	A624	Total reserves (includes gold, US$ millions)	Total reserves comprise holdings of monetary gold, special drawing rights, reserves of IMF members held by the IMF, and holdings of foreign exchange under the control of monetary authorities. Data are in current U.S. dollars.	0.47
	A864	International migrant stock (thousands)	International migrant stock is the number of people born in a country other than that in which they live (including refugees). The data used to estimate the international migrant stock at a particular time are obtained mainly from population censuses.	0.57
	A871	International tourist inbound (thousands)	International inbound tourists are the number of tourists who travel to a country other than that in which they have their usual residence, but outside their usual environment, for a period not exceeding 12 months and whose main purpose in visiting is other than remunerated activity within the country visited.	0.79

Basic information of the 22 pre-selected attributes or independent variables which may be associated to NFFC (cont.).

Dimension	Variable code	Description	Details	Predictability score
Structural	C1	Internet users per country	The live counters display the current estimate based on statistical analysis of data provided by the most reliable sources available. Data is gathered and retrieved from official sources, then processed and elaborated by the source in order to obtain the best possible estimate.	0.44
	C3	Internet user growth in 1 year	Data generated by us by means of subtraction.	0.42
	E1	Voice and accountability	Reflects the extent to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media.	0.51
	E5	Rule of law	Reflects the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, police, courts, as well as the likelihood of crime and violence.	0.43

Basic information of the 22 pre-selected attributes or independent variables which may be associated to NFFC (cont.).

Dimension	Variable code	Description	Details	Predictability score
Social	B1	Estimated Muslim population	Projections of these variables are generated from models commonly used by demographers around the world to forecast changes in population size and composition. The models follow what is known as the cohort-component method, which starts with a baseline population (in this case, the current number of Muslims in each country) divided into groups, by age and sex. Each cohort is projected into the future by adding likely gains – new births and immigrants – and subtracting likely losses – deaths and emigrants.	0.56
	B2	Percentage of population that is Muslim		0.61
	D6	Alcohol consumers (%), past 12 months – both sexes	Total number of individuals in a population who have consumed alcohol in the past 12 months divided by the total population.	0.44
	D7	Alcohol consumers (%), past 12 months – female		0.52
	H1	Life ladder	The English wording of the question is "Please imagine a ladder, with steps numbered from 0 at the bottom to 10 at the top. The top of the ladder represents the best possible life for you and the bottom of the ladder represents the worst possible life for you. On which step of the ladder would you say you personally feel you stand at this time?" This measure is also referred to as Cantril life ladder.	0.53

Basic information of the 22 pre-selected attributes or independent variables which may be associated to NFFC (cont.).

Dimension	Variable code	Description	Details	Predictability score
Social	H7	Perception of corruption	Perceptions of corruption are the average of binary answers to two questions: "Is corruption widespread throughout the government or not?" and "Is corruption widespread within businesses or not?".	0.66
	H11	Democratic quality	Individuals' perceptions of voice and accountability, political stability and absence of violence, government effectiveness, regulatory quality, rule of law, and control of corruption.	0.47
	I4	Number of homicides per 100,000 people	Intentional homicide refers to death deliberately inflicted on a person by another person, including infanticide. The figures refer to the total number of penal code offences or their equivalent, brought to the attention of the police or other law enforcement agencies and recorded by one of those agencies.	0.72
	I18	Nuclear and heavy weapons capabilities	This indicator concerns the perception of peace based on a categorized system for rating the destructive capability of a country's stock of heavy weapons. Holdings are those of government forces and do not include holdings of armed opposition groups.	0.54

Basic information of the 22 pre-selected attributes or independent variables which may be associated to NFFC (cont.).

Dimension	Variable code	Description	Details	Predictability score
Social	I19	Volume of transfers of major conventional weapons as supplier (exports) per 100,000 people	Measures the total volume of major conventional weapons exported by a country between 2008 and 2012 divided by the average population during this time period. It covers all international sales and gifts of major conventional weapons and the technology necessary for their production. Major conventional weapons include: aircraft, armored vehicles, artillery, radar systems, missiles, ships, and engines.	0.68

RESUMO

Desvendando as Condições Espaciais, Estruturais e Sociais de Países para o Surgimento do Fenômeno do Combatente Estrangeiro: Uma Abordagem Exploratória do Caso ISIS pela Mineração de Dados

Na atualidade, um dos maiores desafios para os governos é lidar com organizações terroristas radicais e, em particular, a forma como elas recrutam combatentes estrangeiros. Porém, pesquisadores até agora ignoraram que este fenômeno pertence ao reino da teoria da complexidade, falhando em determinar a combinação de variáveis de países que podem catalisar o problema. Esta é uma questão importante para nações que querem elaborar estratégias sociopolíticas eficazes para diminuir a adesão de combatentes estrangeiros a grupos extremistas ou, pelo menos, para reduzir a penetração de suas mensagens radicais. Portanto, para estudar este fenômeno, fizemos um estudo através da prática de mineração de dados para descobrir padrões a nível nacional que poderiam influenciar a inclusão desses jovens em grupos terroristas, usando como exemplo o caso do chamado Estado Islâmico. Após um procedimento de pré-seleção, as 950 variáveis inicialmente selecionadas foram reduzidas para 22, e subsequentemente utilizadas em árvores de decisão. As descobertas revelam a existência de seis grupos específicos de países, caracterizados por variáveis espaciais, estruturais (econômicas e políticas) e sociais que criam um ambiente favorável para o surgimento deste fenômeno. Posteriormente, discutimos recomendações acadêmicas e práticas.

Palavras-chave: teoria da complexidade; mineração de dados; combatentes estrangeiros; ISIS; descoberta de conhecimento em bancos de dados.

ABSTRACT

Unraveling Spatial, Structural, and Social Country-Level Conditions for the Emergence of the Foreign Fighter Phenomenon: An Exploratory Data Mining Approach to the Case Of ISIS

Governments face a tough and timeless challenge: dealing with the capability of radical terrorist organizations to recruit foreign fighters. However, scholars so far have ignored that this phenomenon pertains to the realm of complexity theory, failing to determine the combination of country-level variables able to catalyze this issue. This is an important concern if countries want to design effective socio-political strategies aimed at decreasing terrorist groups' capability to enroll foreign fighters or, at least, to curtail the penetration of their radical message. Thus, to address this issue we undertake an exploratory data mining approach (knowledge discovery in databases) to discover country-level patterns which might engender conditions that induce people to join an extremist organization, based on the case of ISIS. After a pre-selection procedure, the 950 variables initially selected were

reduced to 22, and subsequently used in decision tree algorithms. Findings reveal the existence of six specific country clusters, which are characterized by some spatial, structural (economic and political), and social variables that create favorable conditions for the emergence of the phenomenon. Academic and practical recommendations are then discussed.

Keywords: complexity theory; data mining; foreign fighters; ISIS; knowledge discovery in databases.

RÉSUMÉ

Découvrir les Conditions Spatiales, Structurelles et Sociales des Pays pour L'Émergence du Phénomène des Combattants Étrangers: Une Approche Exploratoire du Cas ISIS par le Data Mining

L'un des plus grands défis pour les gouvernements aujourd'hui est de traiter avec les organisations terroristes radicales et, en particulier, la façon dont elles recrutent des combattants étrangers. Cependant, les chercheurs ont jusqu'à présent ignoré que ce phénomène appartenait au domaine de la théorie de la complexité, n'ayant pas réussi à déterminer la combinaison de variables nationales pouvant catalyser le problème. Il s'agit d'un enjeu important pour les nations qui souhaitent concevoir des stratégies sociopolitiques efficaces pour diminuer l'adhésion des combattants étrangers aux groupes extrémistes ou, à tout le moins, réduire la pénétration de leurs messages radicaux. Par conséquent, pour étudier ce phénomène, nous avons mené une étude à travers la pratique de l'exploration de données pour découvrir des modèles au niveau national qui pourraient influencer l'inclusion de ces jeunes dans des groupes terroristes, en prenant comme exemple le cas du soi-disant État Islamique. Après une procédure de présélection, les 950 variables initialement sélectionnées ont été réduites à 22, puis utilisées dans les arbres de décision. Les résultats révèlent l'existence de six groupes spécifiques de pays, caractérisés par des variables spatiales, structurelles (économiques et politiques) et sociales qui créent un environnement favorable à l'émergence de ce phénomène. Nous discutons ensuite des recommandations académiques et pratiques.

Mots-clés: théorie de la complexité; exploration de données; combattants étrangers; ISIS; Découverte des Connaissances dans les Bases de Données.

RESUMEN

Descubriendo las Condiciones Espaciales, Estructurales Y Sociales De Países para la Aparición del Fenómeno del Combatiente Extranjero: Un Enfoque Exploratorio del Caso ISIS por la Minería De Datos

Hoy en día, uno de los mayores desafíos para los gobiernos es tratar con organizaciones terroristas radicales y, en particular, cómo estas reclutan a combatientes extranjeros. Sin embargo, hasta ahora los investigadores han ignorado que este fenómeno pertenece al ámbito de la teoría de la complejidad, fallando en determinar la combinación de variables de países que pueden catalizar el problema. Esta es una cuestión importante para las naciones que desean desarrollar estrategias sociopolíticas eficaces para reducir la composición de combatientes extranjeros a grupos extremistas o, al menos, para reducir la penetración de sus mensajes radicales. Por lo tanto, para estudiar este fenómeno, realizamos un estudio a través de la práctica de la minería de datos para descubrir estándares nacionales que podrían influir en la inclusión de estos jóvenes en grupos terroristas, utilizando como ejemplo el caso del llamado Estado Islámico. Después de un procedimiento de preselección, las 950 variables seleccionadas inicialmente se redujeron a 22, y posteriormente se utilizaron en árboles de decisión. Los resultados revelan la existencia de seis grupos específicos de países, caracterizados por variables espaciales, estructurales (económicas y políticas) y sociales que crean un entorno favorable para la aparición de este fenómeno. Más tarde, discutimos recomendaciones académicas y prácticas.

Palabras-clave: teoría de la complejidad; minería de datos; combatientes extranjeros; ISIS; descubrimiento de conocimientos en bases de datos.