The **Human Rights, Big Data and Technology** Project

Identifying perpetrator information in individual-level data on killings in Syria

Anita R. Gohdes

30th October 2017

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Human Rights Centre | University of Essex

Identifying perpetrator information in individual-level data on killings in Syria

Anita R. Gohdes*

30th October 2017

1 Summary

This report presents an analysis of records of deaths that have been documented in the Syrian Arab Republic (Syria) between March 2011 and December 2015. Data for this analysis are from a total of four sources. For confidentiality reasons the sources will be referred to as Source 1, Source 2, Source 3, and Source 4.1 The report discusses processing, including cleaning, translating, canonicalising and matching of the data, as well as the procedure that was implemented to identify perpetrator information in the incident details recorded by each of the four groups.

We find that a small number of words are used to describe both circumstances of deaths, as well as the perpetrators involved in them. An analysis of the frequencies of terms reveals that the majority of perpetrator terms reference the government, or government-sponsored groups as perpetrators of violence. However, it should be noted that the majority of records analysed here do not include any reference of perpetrator at all. Further analyses, possibly making use of supervised machine-learning, will be necessary to obtain a clearer picture of perpetrator distributions in the data available on killings in the Syrian conflict.

2 A closer look at the data

2.1 Cleaning, translating, canonicalising, and matching the data

In a first step, the data is cleaned. Invalid data values are filtered from the data. For example, in many datasets the "age" variable includes a combination of ages in years as well as specific birth years; for example, ages recorded as "1970" are clearly a birth year rather than an age in years. These values are subtracted from the year of death, and the difference in years is recorded as the approximate age of the victim. Another data cleaning task is simply removing obvious typos from data values. For example, strings of unstructured text in otherwise numeric or categorical variables (such as age or sex) can usually be trimmed from those variable values.

In a second step, key analysis variables, such as sex and governorate, are translated from Arabic to English. HRDAG's Syrian expert (one of the native Arabic speakers who review records) confirms the translation of these values.[2] Other Arabic content, such as names and locations (a finer geographic description than governorate) are reviewed in their original form by the native Arabic speaking reviewers.

For other reviewers, we use Google's translation application programming interface to translate names and locations, which they then review in English. Close comparison of these decisions to those made by the native Arabic speakers confirm a high level of consistency, regardless of whether review is conducted in English or Arabic.[3]

---

* Consultant, Human Rights Data Analysis Group, and Assistant Professor, University of Zurich. I thank members of the HRBDT team for helpful comments and suggestions. This work was supported by the Economic and Social Research Council [grant number ES/M010236/1].

1 Information on the sources can be found in the appendix.

2 We translate this content to facilitate the handling of the data through English speaking staff, and to avoid encoding issues across multiple coding platforms and computers. We consult with native Arabic speakers, and specifically Syrian Arabic speakers to ensure consistency and accuracy across translations.

3 For full details, see the section on inter-rater reliability in Price et al. (2016), Appendix B.

In a third step, analysis variables are transformed to have a common structure across all of the data sources. For example, the different datasets collect a variety of infor- mation about the location of death. These locations may be recorded across numerous variables and in varying levels of precision (e.g., neighborhood, area, governorate). We match records based on governorate and compare results for different governorates, so the location variable must be standardized across data sources. In some cases, this is straightforward, in some cases we use other location information (such as city) to map to governorate. Standardization of variables is mostly necessary to ensure consistency both across datasets, but also across different versions of data we received from the individual groups. For example, location values are sometimes written in slightly different ways, or date variables come in different formats. These standardizations maintain the core information of each source but allow us to link records across databases.

This report considers documented killings that are fully identified by the name of the victim, as well as the date and location of death.[4] Each dataset considered in this study included a number of records which lacked this information and were excluded from analyses.

Lastly, as described in detail in Price et al. (2016), a computer science method called record linkage is used to remove both internal duplicates (within each individual source) and duplicates across the four sources.

After completing the aforementioned steps, a number of different versions of data are available:

Original source data This denotes the data as we received it by the individual sources, prior to any processing.

Cleaned source data This denotes the data of each individual source, after cleaning, translating, and canonicalising the records.

Deduplicated source data This denotes the data of each individual source, after removing *internal duplicates*. For each source this gives us a separate list of records, with duplicates removed.

Uniquely identified records This denotes a single, integrated dataset where each row consists of a uniquely identified victim, all available information about that victim, and which source or sources provided that information.

In the following sections, we explore how these different 'versions' of the data compare to each other.

4 Ideally, records included an unambiguous governorate of death. In some cases location was inferred from other information included in the record.

2.2 Comparing cleaned and deduplicated monthly source counts

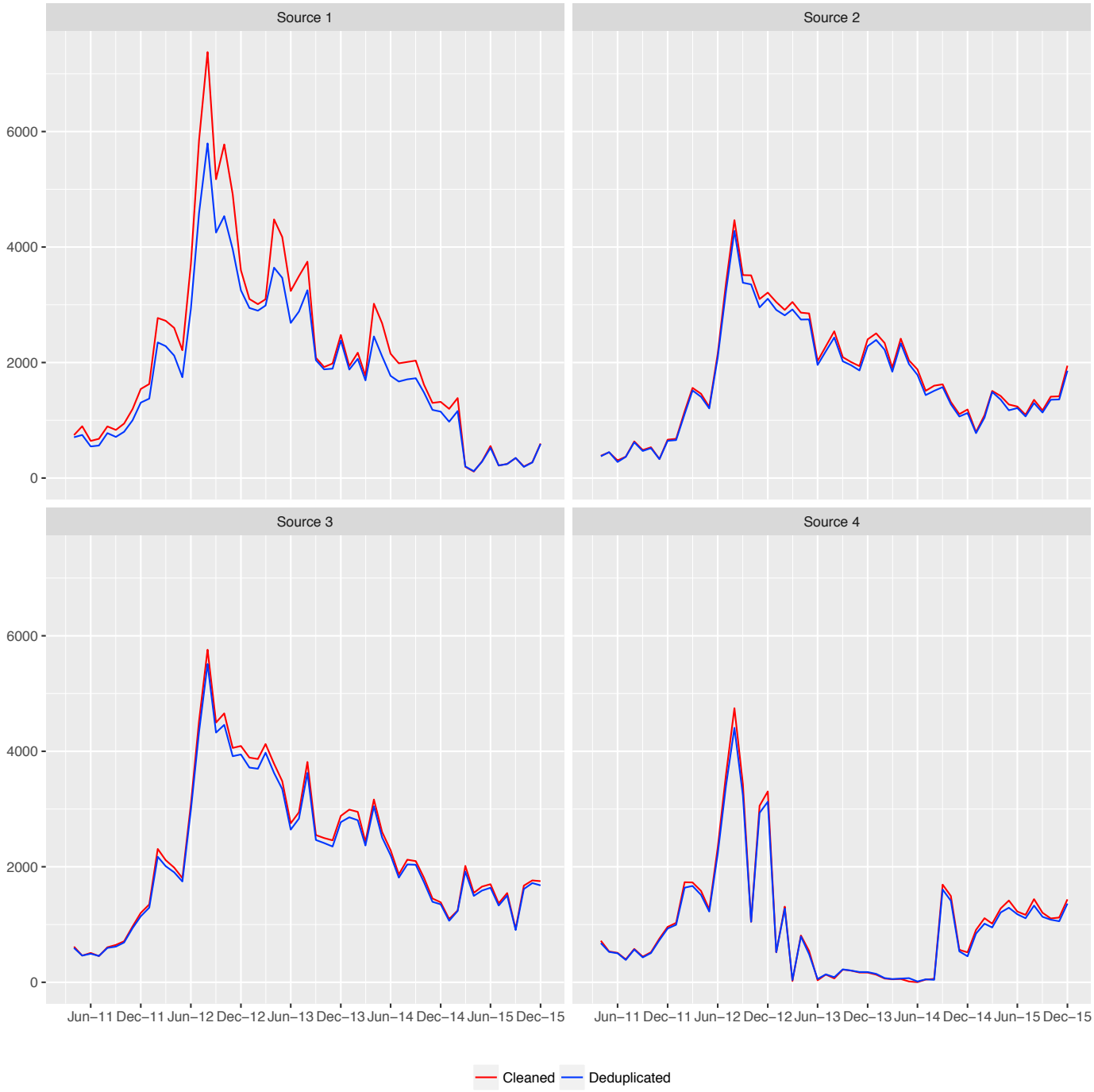


Figure 1 Cleaned and deduplicated monthly counts, by data source

Figure 1 compares the cleaned source data with the deduplicated source data. The records are aggregated to represent monthly counts to facilitate comparison. The red lines represent the monthly count of records for each data source after cleaning and further processing. The blue lines represent the monthly counts of records for each data source after internal duplicates (where the same victim is recorded more than once by the same source) are removed.

Overall, the cleaned and internally deduplicated sources follow very similar patterns. The source that provides the largest number of cleaned records (Source 1) seems to also have the largest number of internal duplicates. Furthermore, internal duplicates seem to occur at a higher frequency as the number of recorded records increases. This makes intuitive sense: where the intensity of violence increases, the probability of the same victim being documented more than once also becomes more likely.

2.3 Comparing deduplicated source data and uniquely identified records

Moving from individual sources to an integrated count of uniquely documented records of killings, **Figure 2** on the next page compares the four (internally) deduplicated sources with the list of uniquely identified records that is based on the four sources. The blue lines, like in **Figure 1** on the preceding page, show the internally deduplicated monthly source counts, while the red line here represents the uniquely identified records. Because this list in based on all four sources the number of records is substantially higher. This list represents the number of all documented records we have for this period of the Syrian conflict.It does not include an estimate of the undocumented records, those victims who were never reported to any of the documentation groups. Therefore, this list gives us a baseline or lower bound number for the number of people killed during this time period.

It is important to note that the comparison of individual sources is in no way a criticism of the tireless work and effort these documentation groups put into recording as much information as possible on atrocities committed in this ongoing conflict. On the contrary, it should be noted that the level of clarity, detail, and technical sophistication displayed by documentation groups in Syria is unparalleled. The fact that we can analyse individual- level data on killings throughout the entire country in such a short time is remarkable. As will be discussed in further detail in the next section, beyond victim characteristics and event date and location information, the groups also collect incident details as well as information on the causes of death, where possible.

3 Identifying perpetrator information

Identifying perpetrators of violence in large-scale individual-level data on violence is challenging. It is therefore remarkable that all four groups included in this analysis also collect information on the circumstances under which each individual was killed. This information is collected in an open text field in Arabic. All sources record incident details, two groups also collect specific information on the cause of death, and one groups records additional notes surrounding the event circumstances.

In a first step of the analysis, we translate all of this open text information into English. After a review of 500 randomly drawn observations, we conclude that event circumstances, as well as perpetrator information and references to the victim are represented in the incident details, cause of death, as well as notes variables. Instead of analysing them separately, all three variables are concatenated and treated as a single corpus of text per observation.[5] The three variables are concatenated because they all include the same kind of information with respect to the details and circumstances of death.[6]

5 Perpetrator information may be conflicting if a) the record linkage procedure linked two records that are not true duplicates, or b) different data sources recorded different causes of death. In the current analysis, we do not take possible conflicting information into account, but plan to do so in future iterations of the project.

6 One exception is the variable 'cause of death' as it is provided by the Violations Documentation Center.This variable has 15 unique values, the most frequent of which are *shooting*, *shelling*, and *aerial bombardment*. We include this information in the analysis but do not analyse it separately as it is only available for one data source.

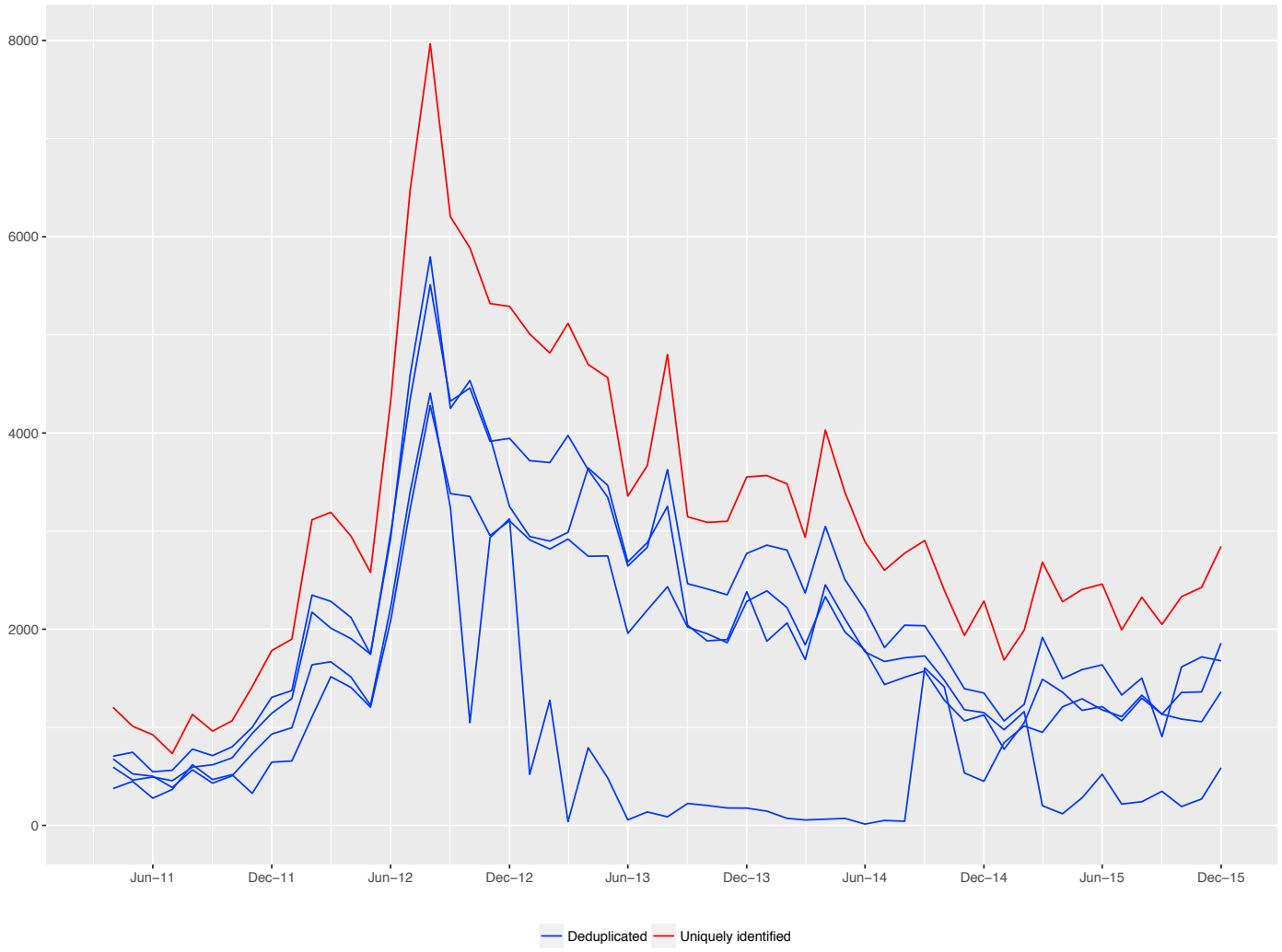


Figure 2 Deduplicated source data and uniquely identified records, monthly count

In a second step, we look at the most frequently used words, defined as words that appeared in at least 1,000 observations, after removing stop words. This list includes 219 words, of which many are names (such as Ahmed, or Mohammed), place names, and words that refer to the victim (e.g. women, brother), not the perpetrator. After removing names and other words that do not point towards incident circumstances, we arrive at a list of 100 frequent words.[7] **Figure 3 on the next page** shows a wordcloud of these words, whereby size and color of the words are dependent of the frequency of observations where the respective word was used to describe the circumstances.

The circumstantial information used here usually is written in full sentences. For example, one record includes the following information from one source:

7 See appendix.

Figure 3 Frequency of words referring to circumstances of violence

He was killed during the clashes between the regime and the forces of the PKK

Another record includes information from two sources (separated by "| "):

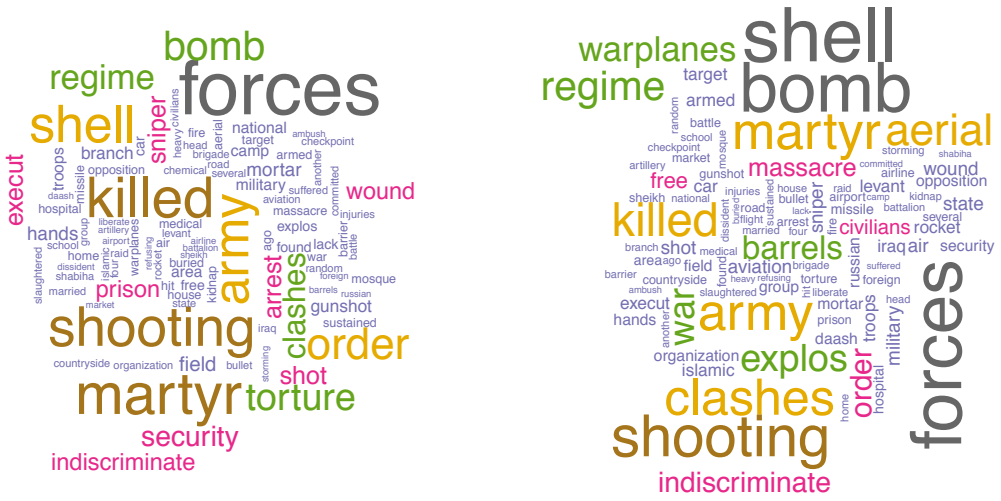
Martyred as a result of the bombing explosive barrels | died as a result of the indiscriminate shelling of the city

Another record includes information from multiple sources that describe the same event with different words (separated by "| "):

Martyred as a result of the bombing explosive barrels | By shelling explosive barrels on the neighborhood | Cited by the bombing of warplanes system of explosive barrels | Aerial bombardment

As the results below show, the majority of records refer to shelling and bombings. However, a significant proportion refers to incidences of torture. This record includes information from multiple sources:

Martyred under torture in intelligence branch | Under torture | Martyred under torture at the Air Force Intelligence Branch in Damascus | Arrest—Torture



(a) Frequencies for Damascus (b) Frequencies for Aleppo

Figure 4 Incident details referring to circumstances of violence

These examples show that the type of information provided by the different sources is very similar in nature, usually identifying the circumstances in which a victim was killed. The level of detail varies by record and can also vary by source.

The most frequently used words are forces and shooting, both of which are referred to in over 60,000 records. Note that we deduplicate the terms before including them in the wordcloud, so as to ensure that each word is only counted once per record. Highly used words also include shell (short for shelling), bomb (short for bombing, bombs, and bombardment), army, and martyr. Among the most used references are also killed, clashes, and regime. While some of these words clearly refer to specific perpetrators, such as regime or military, others only refer to the causes of death.[8]

To see if variations in the frequency of words used show us meaningful differences depending on context we compare the frequency of words used in two different locations, Damascus and Aleppo, as presented in **Figure 4**. **Figure 4a** shows us that the most frequently referenced word in Damascus is forces, followed by killed, martyr and shooting, and then followed by shell, army, and then order, bomb, torture, and regime. The pattern looks somewhat different in Aleppo, presented in **Figure 4b**: shell is used much for frequently when referring to victims killed in Aleppo than in Damascus. We also see a much more frequent mention of clashes, warplanes, explos (short for explosions or explosives), and barrels. Based on the auxiliary information provided on each incident, we find significant differences in the documented circumstances of violence in Damascus and Aleppo.

Not all words included here refer to specific perpetrators. For the next step of the analysis we therefore specifically only select words that refer to perpetrators. The following words are included that refer to the government or government-sponsored actors as perpetrator: assad, regime, shabiha, military, army troops, security. The following words are included that refer to the Islamic State as perpetrator: islamic, levant/sham, daash. Lastly the word russian was also among the most frequently mentioned words, which is why it is also included in the analysis. Three further terms are included that likely refer to perpetrators, but are not quite as straightforward. These include refusing to fire/refused to fire, sniper, and martyr. Groups refer to regime defectors as victims who were killed because they refused to fire. The word martyr is generally used to refer to victims killed by government forces.[9]

8 Other approaches might make sense here. Variation in words could addressed by using similarity measures or fuzzy matching. Alternatively, a dictionary could be used to extract relevant information.

9 For example, VDC's classification on martyrs is: 'The Revolution's Martyrs: who were killed by regime forces and thugs [. . .]'

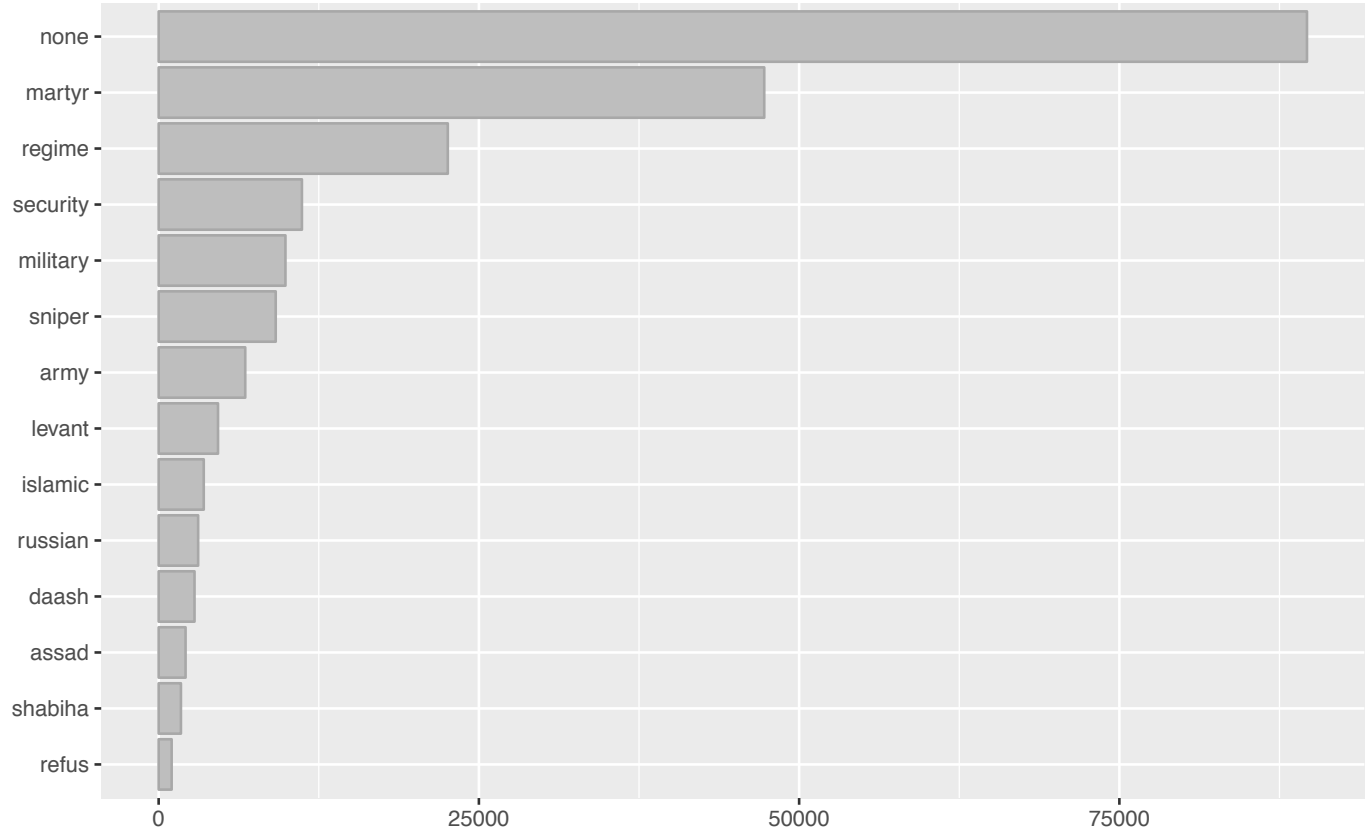


Figure 5 Frequency of words referring to perpetrator

Figure 5 shows the frequency distribution of keywords related to perpetrators for all uniquely identified records.[10] The Figure includes a bar for all records that don't include any of the just-mentioned keywords (denoted as none). The bar graph shows that the majority of records have none of the perpetrator-identifying keywords mentioned in their incident details.[11] Following this, the most frequently found word indicating the perpetrator is the word martyr — referring to the government as perpetrator. Further government-related terms also show up with high frequency; these include *regime, security, military,* and *army troops*.

The words referring to the Islamic state as perpetrator are much less frequent. This should not be taken as evidence for the frequency of killings committed by the Islamic State. All documentation groups have confirmed that documenting information on violence committed by the Islamic State is even more dangerous than in other circumstances, therefore these violent events are likely to be systematically under-reported.

Figure 6 on the following page breaks down the frequency of words referring to whereas most deaths perpetrated by the Shabiha were recorded in Homs. The highest number of records without any perpetrator reference are documented in Aleppo. And the highest number of defectors who refused to fire at civilians or opposition groups were documented in Idlib.

(see http://www.vdc-sy.info/index.php/en/about).

10 The present analysis uses a simple bag-of-words approach. Future analyses could consider implementing more complex procedures, such as capturing multi-word expressions, or using dependency parsing.

11 A close reading of these cases will be necessary to classify as many of these cases as possible. Supervised machine-learning techniques might be useful for this endeavour, as well.

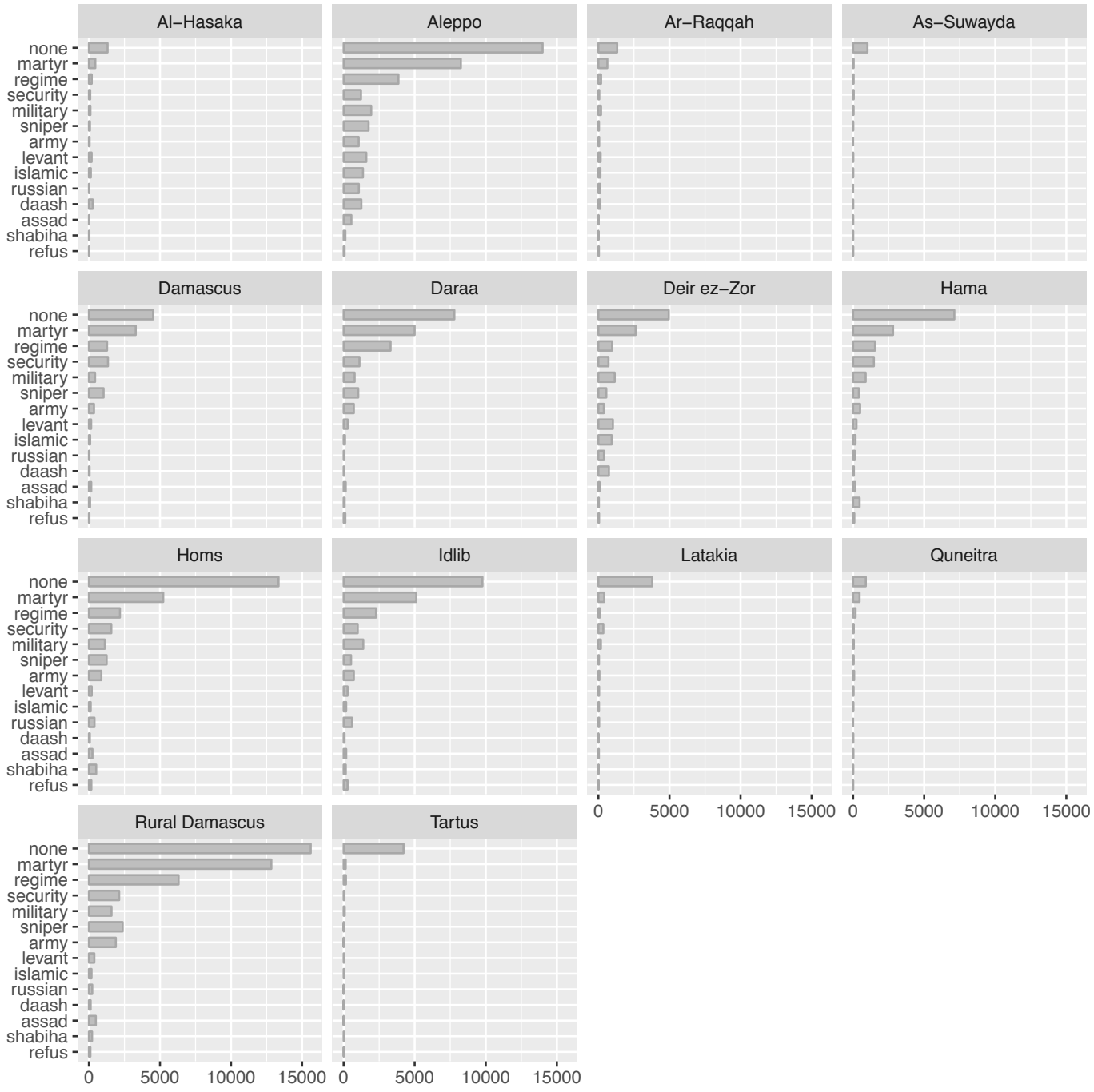


Figure 6 Frequency of words referring to perpetrator, by governorate

4 Discussion

This report discusses a first examination of perpetrator characteristics in individual- level violence data in the Syrian conflict. Documentation groups are recording detailed information on individual killings that were perpetrated throughout Syria, and this information forms the basis of this analysis.

A Sources

Note that the listing of data sources does not correspond to the numbering of sources used throughout this report.

- Syrian Center for Statistics and Research (CSR-SY).[12] This list was initially provided to HRDAG in November and December 2013. Subsequent updates to their files were shared with HRDAG in June 2014, October 2014, May 2015, and March 2016. As described on its website, "The center includes a local network of reporters and a team of researchers and academics inside and outside of Syria."

- Damascus Center for Human Rights Studies (DCHRS).[13] This list was provided to HRDAG in April 2015 and updated records were shared with HRDAG in January and February 2016. The Damascus Center for Human Rights Studies maintains several documentation projects in addition to lobbying and advocating for Syrian human rights and working to draw attention to the situation in Syria.

- Syrian Network for Human Rights (SNHR).[14] This list was initially provided to HRDAG by OHCHR in August 2012. Beginning in February 2013, HRDAG estab- lished a direct relationship with SNHR. SNHR conducts monthly reviews of their records and subsequently updates their dataset with newly discovered or verified victims. SNHR shared their list and subsequent updates with HRDAG in Febru- ary 2014, June 2014, October 2014, March 2015, and June 2016. SNHR maintains a website where they describe that they "adopt the highest approved documentation principles by the international bodies." Also available on their website is a descrip- tion of their three phase documentation process and the six categories of victims they document.

- Violations Documentation Center (VDC).[15] This list was initially provided to HRDAG by OHCHR in February 2012. Subsequently HRDAG scraped[16] the website several times between 2012 and 2016 to obtain updated data. This process captures two of the lists maintained by VDC, "Martyrs" and "Regime fatalities." The "About" page of their website describes the data classification methods and three-stage data verification process implemented by the VDC.

12 http://www.csr-sy.org/
13 http://www.dchrs.org/english/news.php?aboutus
14 http://sn4hr.org/
15 http://www.vdc-sy.info
16 Using a computer program to extract information from websites.

B Most frequent words referring to circumstances of violence

Term	Freq.	Term	Freq.	Term	Freq.	Term	Freq.
forces	69,695	field	9,151	hospital	3,372	hit	1,758
shooting	63,783	sniper	9,139	russian	3,084	shabiha	1,740
shell	55,255	hands	8,386	countryside	2,970	several	1,718
bomb	51,064	air	7,564	daash	2,800	storming	1,710
army	49,833	barrels	7,531	home	2,772	airport	1,677
martyr	47,276	car	7,263	ago	2,680	four	1,667
killed	41,078	prison	6,935	group	2,647	airline	1,651
clashes	37,001	aviation	6,392	raid	2,640	ambush	1,645
regime	35,760	target	5,891	national	2,612	lack	1,575
order	27,030	gunshot	5,821	organization	2,513	battalion	1,548
indiscriminate	22,638	battle	5,606	kidnap	2,502	foreign	1,545
aerial	19,685	civilians	5,571	road	2,475	slaughtered	1,504
war	17,490	state	4,868	head	2,460	heavy	1,490
warplanes	15,672	area	4,210	house	2,307	checkpoint	1,455
wound	13,851	fire	4,181	branch	2,266	dissident	1,452
execut	13,717	mortar	4,118	injuries	2,208	medical	1,439
shot	13,640	barrier	3,916	mosque	2,153	suffered	1,312
explos	13,371	rocket	3,874	opposition	2,092	committed	1,278
free	12,669	iraq	3,611	another	2,048	chemical	1,264
massacre	11,310	missile	3,598	bullet	2,033	sheikh	1,251
security	11,185	islamic	3,522	brigade	2,012	liberate	1,156
torture	10,765	armed	3,487	market	1,900	school	1,106
military	9,899	sustained	3,436	random	1,897	buried	1,059
troops	9,761	levant	3,418	camp	1,887	flight	1,040
arrest	9,178	found	3,387	artillery	1,869	refusing	1,033

Table 1 most frequent words referring to circumstances of violence

References

Price, M. E., Gohdes, A. R., and Ball, P. (2016). Technical Memo for Amnesty International Report on Deaths in Detention.

This paper was written by Anita Gohdes on behalf of HRDAG for The Human Rights, Big Data and Technology Project (HRBDT).

About HRBDT

The Human Rights, Big Data and Technology Project (HRBDT) began in 2015, funded by the Economic and Social Research Council and the University of Essex. One of the largest of its kind in the world, the Project is based at the Human Rights Centre at the University of Essex with over 30 researchers, and additional researchers based at Cambridge University, the Geneva Academy and Queen Mary University. The team addresses human rights and technology issues across a range of disciplines including communication studies, computer science, economics, law, philosophy, political science and sociology.

HRBDT identifies and assesses the risks and opportunities for human rights posed by big data, artificial intelligence and smart technologies and proposes solutions to ensure that new and emerging technologies are designed, deployed and regulated in a way that promotes, rather than threatens, human rights. HRBDT's research assesses the adequacy of existing ethical and regulatory approaches to big data and new and emerging technologies from a human rights perspective. The research also demonstrates how human rights standards are capable of adapting, and offering solutions to, rapidly evolving technological landscapes. HRBDT engages with responses to the risks and opportunities posed by data and technology at the multilateral and multi-stakeholder level, as well as within specific sectors, such as law enforcement, health and humanitarian responses, and at the national level. More information is available on HRBDT's website (www.hrbdt.ac.uk).

About HRDAG

The Human Rights Data Analysis Group is a non-profit, non-partisan organization[1] that applies scientific methods to the analysis of human rights violations around the world. This work began in 1991 when Patrick Ball began developing databases for human rights groups in El Salvador. HRDAG grew at the American Association for the Advancement of Science from 1994–2003, and at the Benetech Initiative from 2003–2013. In February 2013, HRDAG became an independent organization based in San Francisco, California; contact details and more information is available on HRDAG's website (https://hrdag.org) and Facebook page.[2]

HRDAG is composed of applied and mathematical statisticians, computer scientists, demographers, and social scientists. HRDAG supports the protections established in the Universal Declaration of Human Rights, the International Covenant on Civil and Political Rights, and other international human rights treaties and instruments. HRDAG scientists provide unbiased, scientific results to human rights advocates to clarify human rights violence. The human rights movement is sometimes described as "speaking truth to power:" HRDAG believes that statistics about violence need to be as true as possible, with the best possible data and science.

The materials contained herein represent the opinions of the authors and editors and should not be construed to be the view of HRDAG, any of HRDAG's constituent projects, the HRDAG Board of Advisers, the donors to HRDAG or to this project. The content of this analysis does not necessary reflect the opinion of OHCHR.

1 Formally, HRDAG is a fiscally sponsored project of Community Partners, (see http://www. communitypartners.org/).
2 https://www.facebook.com/HumanRightsDataAnalysisGroup.

The Human Rights, Big Data and Technology Project

Human Rights Centre,
University of Essex,
Colchester CO4 3SQ
+44 (0)1206 872877

@HRBDTNews
www.hrbdt.ac.uk