PERSPECTIVES ON TERRORISM

RESEARCH ARTICLE

A Metadata Analysis of 55 Databases and Datasets on Terrorism

Marina Alonso-Villota,* and Rubén Arcos

Volume XX, Issue 3
September 2026

ISSN: 2334-3745
DOI: 10.19165/WRWT1014

Abstract: Over the last decades, the number of publicly available databases and datasets on terrorism (DDTs) has increased exponentially, facilitated by access to the internet, advancements in data collection capabilities, and the variety of supporting software. Currently, hundreds of sources provide data and information about terrorism-related issues, such as incidents, perpetrators, victims, and legislation. However, despite their growing prominence, limited attention has been paid to how these DDTs are constructed and how their underlying design choices affect research outcomes. This article presents a metadata analysis of 55 DDTs, focusing on how they document and make publicly available their definitional choices, raw-data inputs, and methodological procedures, as well as the implications for research and analysis. Results show substantial variation across DDTs. Key issues identified in the analysis include the lack of retrievable definitions of terrorism, which requires users to infer inclusion criteria; the absence of clear identification of original data sources, which limits traceability to the underlying information; and the lack of publicly available coding and methodological documentation, which constrains transparency, interpretability and reproducibility of DDTs. Overall, while some DDTs demonstrate promising practices, these aspects remain a work in progress in the field.

Keywords: terrorism, database, dataset, metadata, definition

*Corresponding author: Marina Alonso-Villota, National University of Distance Education (UNED), email: malonso1139@alumno.uned.es

Introduction

Since the last century, databases and datasets on terrorism (hereafter DDTs[1]) have been developed to provide researchers and practitioners with a wide range of variables and information to support the identification of trends, the analysis of case studies, and hypothesis testing. The past few decades have witnessed a substantial increase in research funding and the availability of quantitative data on terrorism.[2] This has been facilitated by the expansion of open access practices in digital information environments and information collection, as well as being driven by high-profile terrorist attacks such as the Oklahoma City bombing, the 9/11 attacks, the Madrid bombings of 11 March 2004, and the London bombings of 7 July 2005.[3] In 2021, at least 160 DDTs were identified,[4] capturing diverse aspects of terrorism, such as violent tactics and their characteristics, individual and group perpetrator profiles, victim-related information, and counter-terrorism-related aspects, such as operations and legislation.[5]

For the period 2007-2016, databases accounted for 8 percent of the data collection methods reported across 3,442 pieces published in nine terrorism studies journals,[6] only marginally higher than the 7 percent reported by Silke in 2001.[7] Although this represents a relatively small share of the methods employed, databases ranked as the third most frequently used method of data collection, indicating their importance within terrorism scholarship relative to other methodological approaches. This figure does not capture all research employing databases, but it nevertheless provides a useful indication of the extent to which databases are used in terrorism research. While updated empirical evidence on the prevalence of database-based research is not currently available, scholars have argued that the willingness to engage with quantitative approaches within critical terrorism studies has increased in recent years.[8]

DDTs are employed across academic and policy-oriented research. However, they should be carefully assessed before use, as seemingly similar DDTs may measure the same phenomenon in slightly different ways and therefore produce different results, even when addressing the same research question. For instance, in their attempt to provide a "comprehensive account of the activities of the terrorist organisation Euskadi Ta Askatasuna (ETA) across the whole of its 1959-2018 lifespan",[9] Loosemore and Johnson examined and combined nine different databases. They found that, while all nine recorded ETA's activities, the estimates they produced varied considerably due to a range of factors, including differences in units of analysis, sources used by the databases, temporal coverage, inclusion criteria, geographical scope, and other design choices. This highlights an important challenge for researchers: assessments of the scale and impact of terrorist activity may vary depending on the characteristics of the DDT selected, thereby affecting the comparability and interpretation of empirical findings.

These findings are not new. Prior research has shown that DDTs ostensibly addressing the same phenomenon can yield markedly divergent findings even when queried using the same parameters. Substantial discrepancies were already evident in earlier decades: for the period 1968-1977, the total number of terrorist incidents varies significantly depending on the DDT consulted, ranging from 936 incidents in RAND data, to 2,413 in CIA records, 3,027 in ITERATE and 4,091 in the U.S. State Department data, all of them monitoring the same data with almost identical definitions.[10] Similarly, a comparison between ITERATE and the U.S. State Department data for the period 1968-1998 showed differences in the number of recorded terrorist incidents,[11] and a comparison between ITERATE and RAND–MIPT revealed substantial discrepancies in quarterly terrorism event counts for the period 1993–2004, particularly in 1993 and again between 2001 and 2002, and in 2004.[12] Likewise, notable differences in the number of suicide attacks were recorded when comparing the Global Terrorism Database (GTD), RAND, and WITS for the period 2004–2008, with RAND and WITS showing much higher frequencies at almost every time point.[13] Such differences persisted in later periods, as revealed by a comparison

of terrorism incidents recorded in the United Kingdom between 2007 and 2010 across GTD, RAND, WITS, and Europol.[14]

Differences across DDTs cannot be explained by the data outputs alone. Without detailed information on how each DDT defines terrorism, selects sources, and applies coding rules, among others, it is not possible to determine whether one DDT captures a greater number of accurate events or whether the variation reflects methodological differences or misclassification. Consequently, rather than only indicating substantive variation in terrorist activity, these results may evidence differences in the DDTs' design.[15] Nevertheless, despite the analytical relevance of these design decisions and choices, comparatively little attention has been paid to examining how DDTs are constructed.

One of the most significant efforts to address this gap is Sheehan's criterial framework,[16] a seven-criterion framework that could be used to evaluate and compare a large number of terrorism datasets. The criteria were *Conceptual Clarity, Context and Immediacy of Observation, Citation Transparency, Coding Consistency, Certainty, Conflict of Interest, and Convenience and Functionality*. These criteria have proven valuable for researchers seeking to select DDTs by prioritising those dimensions most relevant to their specific studies.[17] Moreover, his framework has served as a foundation for subsequent work. For example, Johnson and Ackerman expanded upon Sheehan's approach by retaining several of his categories while integrating additional ones, consolidating a five-element framework which includes *Scope, Transparency, Error Handling, Consistency*, and *Availability*.[18]

These studies provide an important foundation for further inquiry into the metadata of DDTs. Nevertheless, each leaves relevant aspects insufficiently explored. Sheehan's criterial framework,[19] while foundational, overlooks variation in units of analysis and the availability of methodological documentation. Moreover, each criterion encompasses multiple analytical questions, rendering the criteria overly broad and limiting their analytical precision. Johnson and Ackerman,[20] by contrast, provide only a limited description of the elements they assess and do not explicitly examine the availability of information concerning source selection or the nature of the underlying data sources. Consequently, although their framework is informative, the conceptual interpretation of several criteria remains ambiguous.

As Johnson and Ackerman argue, "[n]ot all databases are of equal quality and uninformed or careless usage of these databases can lead to the drawing of incorrect inferences and thereby to poor policy guidance and specious contributions to our understanding of terrorism".[21] Accordingly, understanding the assumptions that underlie DDTs, the gaps they entail, and the extent to which they align with the research questions they are used to address is crucial for advancing knowledge in the field.

In light of this, the present study analyses the metadata of 55 DDTs and makes available the results in an open-access spreadsheet.[22] Rather than examining the results of these DDTs, the analysis focuses on their structural characteristics, including the level of accessibility and transparency of definitional elements, the raw data used, and the methods followed. By treating DDTs as analytical objects rather than merely as tools, this study seeks to clarify their design features and assumptions that guide their use. This approach makes two main contributions. First, it offers the most comprehensive metadata analysis of DDTs to date, systematically comparing 55 DDTs across multiple dimensions. Second, the study makes available the results in an open-access spreadsheet to assist researchers in aligning data selection with their theoretical frameworks and empirical questions, thereby facilitating improvements in conceptual validity, transparency, and the reproducibility of quantitative terrorism research.

This article proceeds as follows. It first sets out why the metadata of DDTs warrant systematic analysis, focusing on three recurrent challenges in the literature: definitional clarity, the management of raw data, and methodological transparency and reproducibility. It then describes the study's methodology. The subsequent section presents and discusses the results, organised around DDT's unit of analysis and the three challenges identified. The article concludes by outlining the limitations of the study and summarising the main findings and implications.

Why metadata? A rationale for the analytical approach

Metadata is often described as data about data and provides the information that enables users to understand, interpret, and appropriately use data. While data captures the content itself, metadata provides the conceptual and theoretical framework that underlies data production, including how data were collected, which definitions were applied, what sources were used, and which procedures guided the selection, coding, and organisation of variables. As Pomerantz notes, "[m]etadata is all around us all the time. In the modern era of ubiquitous electronics, nearly every device you use relies on metadata or generates it or both. But when metadata is doing its job well, it just fades into the background, unnoticed and nearly invisible."[23] It functions as "a potentially informative object that describes another potentially informative object",[24] highlighting its role as an interpretative interface rather than a mere technical supplement. In this sense, metadata allows users to understand the structure, scope, and purpose of DDTs without engaging directly with every individual data point. It makes explicit what DDTs are designed to capture, the conditions and assumptions under which they operate, and, equally important, what they systematically omit. Metadata, therefore, constitutes the conceptual layer where foundational design decisions, both implicit and explicit, become analytically accessible.

In the case of DDTs, metadata is particularly significant because the study of terrorism often involves contested concepts, political and cultural sensitivities, and substantial variation in how theory is operationalised across data sources.[25] Moreover, discontinuities in time series are common in DDTs, due to factors such as limited or unstable funding, reliance on a limited number of researchers to maintain a database, or frequent staff turnover.[26] These conditions may introduce shifts in coding practices, norms, or analytical assumptions, increasing the likelihood of deviation from the DDT's original conceptual and methodological design.

As a result, DDTs do not merely record instances of terrorism but actively construct specific representations of the phenomenon. During their development, each DDT necessarily delineates elements like what qualifies as a terrorist act, who is recognised as a perpetrator, which targets are considered relevant, and their geographical and temporal scope. However, these design choices are rarely visible in the data outputs alone, as they are embedded in the metadata rather than in the observations themselves.[27]

Over the past decades, three key challenges have been recurrently highlighted in the literature on DDTs: definitional clarity, the management of raw data, and methodological transparency and reproducibility.

The challenge of definitional clarity

Definitional issues have posed a challenge for DDTs since the early development of the field.[28] They include the absence of publicly available definitions for key database variables, such as terrorism-related terms, as well as inconsistencies in the application of stated definitions, such as the inclusion of data entries that fall outside the DDTs' own definitional boundaries.

Given the large number of existing definitions of terrorism,[29] even seemingly minor definitional differences (e.g., defining terrorism as an act of violence for political purposes, or as an act of violence *against civilians* for political purposes) can substantially affect DDTs results.[30]

Processing terrorism data without the guidance of a specified definition of terrorism leaves inclusion and exclusion criteria unclear, increasing the likelihood of noise, unstated assumptions, and analytical error. Similarly, DDTs that do not provide a specific definition leave users unable to determine whether data were processed using a definition at all and, if so, whether that definition aligns with their own conceptual understanding. As Silke et al. have argued, a core problem in terrorism studies is not the existence of multiple definitions, but the lack of transparency regarding the definitions employed and their operationalisation.[31]

A metadata analysis of DDTs addresses this gap by making definitional choices explicit, thereby enabling researchers to assess theory-data alignment, strengthen conceptual validity and methodological rigour, and identify DDTs in which definitional, source-wise, and methodological ambiguity persists.

The challenge of managing raw data

DDTs face enduring challenges related to the availability and quality of raw data, which have direct implications for the validity and interpretation of research findings. First, data availability varies significantly across several dimensions. It differs by unit of analysis, as event-based data is generally easier to document than data on armed organisations and individuals, due to the clandestine nature of the latter.[32] Availability also varies geographically, shaped by factors such as media coverage, levels of press freedom, and broader political and democratisation contexts.[33] In addition, reporting is strongly skewed toward English-language sources, potentially creating blind spots in regions where events are covered only in local languages or not reported internationally, although efforts to monitor a wider range of languages have increased in recent years.[34]

Second, even when data are observable, reporting practices are rarely neutral. Data and reporting are influenced by persistent biases, media incentives, political environments, and levels of press freedom and censorship. As a result, events may be reported, attributed, or omitted differently across contexts, affecting the reliability of sources and the information they provide.[35]

In this context, the analytical value of metadata analysis lies in examining whether and how DDTs document their reliance on open-source material, as this enables traceability, facilitates independent assessment, and promotes more cautious and transparent interpretations of terrorism data. In particular, the presence of information on raw data sources facilitates independent verification by researchers and ensures what Johnson and Ackerman defined as "good practices" in utilising terrorism data.[36] In the absence of source documentation or direct links to original materials, such verification is not possible, limiting both transparency and analytical reliability.

The challenge of methodological transparency and reproducibility

Methodological opacity persists across many DDTs, reflected in insufficient documentation of methods, coding procedures, verification processes, and error handling, among others. Many DDTs fail to provide publicly accessible codebooks or detailed methodological descriptions, making it difficult for users to understand how data are collected, filtered, and processed.

Without transparency regarding data construction, users cannot assess the validity or reliability of the information on which their analyses are based.[37] Moreover, insufficient methodological transparency increases the risk of incorrect inferences and undermines both scholarly and policy-oriented research, particularly when combined with discontinuities in DDTs monitoring and maintenance.[38]

In this context, metadata analysis plays a crucial role by systematically documenting the availability or absence of methodological information, thereby enabling researchers to evaluate the rigour, replicability, and suitability of a DDT for specific analytical purposes before drawing substantive conclusions. Similarly, following good practices in the use of terrorism data, such as scrutinising the codebook, and detecting errors and uncertainties, can facilitate and strengthen the use of DDTs.[39]

Conducting a metadata analysis

These three challenges, i.e. the definitional clarity, the selection of raw data, and the methodological practices, converge at the level of DDT construction and are encoded in metadata. Thus, a metadata-centred approach offers a systematic way to address these challenges by shifting attention from data outputs to data construction. It enables scholars to assess theory-data alignment, select sources appropriate to specific analytical purposes, and identify potential knowledge gaps.

Our analytical approach does not seek standardisation across DDTs but instead promotes transparency, reflexivity, and methodological rigor. By systematically assessing metadata dimensions across DDTs, our study contributes to understanding the current state of the art of DDTs, identifying current weaknesses and strengths, and supporting more informed research practices.

Methodology

This study analysed the metadata of 55 databases and datasets on terrorism (DDTs). The study aimed to understand the current state of the art of DDTs and to clarify their design features and assumptions that guide their use. In addition, by making available the results in an open-access spreadsheet, it aimed to support researchers in selecting the most suitable source for conducting quantitative research. The guiding research question was as follows:

> *How do terrorism databases and datasets (DDTs) document and make publicly available their definitional choices, raw-data inputs, and methodological procedures, and how does this transparency (or its absence) shape users' ability to interpret findings and choose fit-for-purpose DDTs?*

To answer it, the research objectives focused on examining three key dimensions of DDTs: definitional clarity, transparency regarding raw data, and the availability of methodological procedures. The assessed items for each of these elements are outlined in Table 2.

The study was conducted in several stages. First, an initial data sample of DDTs was compiled. This sample consisted of 169 DDTs: the 160 DDTs listed on the inventories of Terrorism Databases and Data sets published by Bowie in four Research Notes[40] along with 9 additional DDTs known to the authors. Bowie's inventories were selected because they represent the most comprehensive compilation of DDTs known to the authors. However, some characteristics

of Bowie's inventories should be noted. The sources gathered are of three types: academic, think tank, and independent databases; commercial databases; and governmental databases. Restricted and classified DDTs were explicitly excluded from Bowie's inventories. Although the primary focus of these DDTs is terrorism-related issues, Bowie acknowledged that some of them adopt a broader scope, including political violence and armed conflicts. The 9 additional DDTs included in this study were developed by international institutions, think tanks, and victims' associations, and focus on terrorism-related issues.

Second, a preparatory analysis of the 169 DDTs was conducted. This analysis was used to identify the final selection of DDTs suitable for the analysis by assessing their accessibility and unit of analysis. Sources meeting the following criteria were included in the final analysis:

- Availability: DDTs had to be accessible through a valid link. Some links from Bowie's inventories were no longer functional. In such cases, alternative links to the same DDT were searched using the source name. If no valid link could be found, the DDT was excluded.

- Free access: DDTs had to be open access and fully accessible without requiring payment. Additionally, DDTs for which only a description was freely available were also excluded.

- Unit variable: DDTs had to use terrorism-related variables, such as terrorism events, perpetrators, victims, symbols, or counter-terrorism data, such as legislation, court cases, or terrorism designations. Some sources listed in Bowie's inventories focused on general conflict or political violence, aggregating multiple forms of violence without explicitly distinguishing terrorism. Others provided only aggregated or processed results without making the underlaying data available, such sources were excluded.

Out of the 169 initially listed sources, only 55 met the filtering criteria and were selected for the analysis. The reasons for exclusion are shown in Table 1.

Table 1. Reasons for the exclusion of DDTs from the initial list of 169 DDTs

Reasons for exclusion	Number of DDTs
NOA (No open access)	40
NVL (No valid link)	24
Unit of analysis other than terrorism (or derivate terms)	23
Displays aggregated data	18
No longer active / dismantled	8
DDT Forthcoming	1
Total of excluded sources	**114**

Most excluded sources were omitted due to their lack of open-access availability. This is not surprising, as Bowie included commercial DDTs in his inventories, which are often accessible only through paid subscriptions. The second most common reason for exclusion was the absence of a valid link: 24 DDTs could not be accessed, and no alternative links were found. This is a concerning finding, as the earliest of Bowie's inventories dates back to 2018. In less than eight years, 24 DDTs can no longer be located online. However, it is not possible to determine how many of these 24 DDTs are no longer active,[41] in contrast to the 8 DDTs that explicitly indicate they are no longer operational. This observation, identified during the preparatory phase of the analysis, highlights one of Schmid's[42] concerns: the continuity and sustainability of DDTs.

Finally, the metadata analysis was conducted on 55 DDTs. The analysis was done manually, examining each source for information relevant to each category of interest. The use of data mining and text mining tools was deliberately excluded, as they would have required individual adaptation to each DDT due to their differing formats, which could have increased the risk of error. No translation was needed for the analysis: English clearly predominates over all other languages among the DDTs examined. A total of 44 DDTs is available exclusively in English, and two additional DDTs are presented in English but include raw data in various languages. Only a small number provide multilingual access: five DDTs can be displayed in all six official United Nations languages (English, Spanish, French, Russian, Arabic, and Chinese), and one DDT provides content in Spanish, English, French, and Basque. In addition, one DDT allows access in English and French, one DDT in English and Hebrew, and one DDT is available only in Spanish.

For each DDT, contextual information was gathered, including characteristics such as the unit of analysis, scope, or language. In addition, three major metadata categories were developed based on existing literature and informed by the key challenges previously identified: definitional clarity, raw data, and methodological procedures. Each of these three categories was operationalised through a set of specific metadata elements and corresponding guiding questions used in the analysis (see Table 2).

The categories examined are the following:

- Contextual information (C): This category is designed to provide a clear understanding of the content and scope of each DDT, enabling users to make an initial selection based on their specific needs. By examining aspects such as the unit of analysis, the geographical and temporal coverage, and the number of entries, this category allows users to efficiently filter sources of interest.

- Definitions of terrorism or derivative terms (D): This category examines the terminology and conceptual frameworks related to terrorism employed by each DDT. Having explicit and publicly accessible definitions for the variables used is crucial for all databases (see the subcategory M1: Coding transparency). However, given the multiplicity of definitions of terrorism, this category pays particular attention to whether terrorism-related definitions are explicit and publicly accessible. Some DDTs explicitly use terms such as terrorism, terrorist organisation, or terrorist attack, while others adopt alternative labels like insurgent organisation, militants, or violent non-state groups, both to describe the same groups or actions (e.g., al-Qaeda, Daesh, Hezbollah). Regardless of the label used, the availability of definitions or conceptual frameworks of reference is essential for contextualising data, understanding inclusion and exclusion criteria, and ensuring analytical consistency.

- Raw data (R): This category assesses the characteristics of the raw data included in each DDT, such as the types of data monitored, references or links to original sources, and information on data confidence and reliability. Given the persistent challenges surrounding the availability and quality of terrorism-related data, evaluating whether and how DDTs document their reliance on open-source material is essential. This supports traceability, enables independent verification, and facilitates a more transparent interpretation of the data.

- Methods transparency (M): This category evaluates whether DDTs provide information about how raw data are processed from beginning to end. Having and publicly providing a codebook is essential for knowing how variables are defined and how inclusion and exclusion decisions are made. Likewise, publicly available methodological documentation

is crucial to describing how raw data were collected, filtered, categorised, and displayed. Limited access to either the codebook or methodology descriptions creates two major problems: users cannot fully interpret the data, and neither can they assess whether the data followed systematic and replicable procedures.

Table 2. List of assessed items and their corresponding possible answers.

	Categories	Question	Response options*[43]
Contextual information			
C1	Name and website link	What is the DDT'S name?	Free text and link
C2	Source operator	Who operates the DDT?	Free text
C3	Description	What description of the DDT is provided within the source itself?	Free text
C4	Last update	When was the DDT last updated?	Free text
C5	Unit of analysis	What is the DDT's main focus (i.e., terrorism or counter-terrorism data)? What is the source's unit of analysis (e.g., events, perpetrators…)?	A. Terrorism data: • Events: Attacks; Fatalities; Bombs; Attacks and plots; Mix. • Perpetrators: Individuals; Organisations; Both. • Victims: Fatalities. • Others: Terrorist documents; Surveys and opinion data; Symbols; Chronology; Service provision; Glorification of terrorism; Propaganda; Bibliography. B. Counter-terrorism data: • Legislation. • Designation. • Convictions. • Mix.
C6	Number of entries	How many data entries does the DDT contain?	Free text
C7	Geographic scope	What locations are covered by the DDT?	Global; Regional; Various countries (VVCC), Country; Other (free text).
C8	Time scope	What is the timeframe covered by the DDT?	Free text (yyyy-yyyy)
C11	Language	In what language(s) is the DDT available?	Free text

Definitional clarity			
D1	Use of the term "terrorism" or derivates	Does the DDT use the term "terrorism" or derivative terms (e.g., terrorist attack, terrorist organisation) as a unit variable?	Options: 1; 2; and 0. 1: They use the term as a unit variable. 2: They use the term in a descriptive way but employ other term (D2) as a unit variable. 0: They do not use the term in any way and employ other term (D2) as a unit variable.
D2	Use of an alternative term	What alternative or additional terms, if any, does the DDT use instead of "terrorism" or derivative terms?	Free text
D3	Availability of definitions of terrorism	Does the DDT provide a publicly accessible definition of terrorism?	Options: 1; 2; and 0. 1: DDTs that explicitly provide a definition of terrorism within their own documentation. 2: DDTs that rely on an external definition, which they reference but do not reproduce explicitly. 0: DDTs that neither provide nor reference any definition of terrorism.
D4	Definitional approach to terrorism	What definitional approach to terrorism or related terms uses the source?	Legal; Own source's definition; Other source's definition; Multiple definitions; 0 (no approach included).
D5	Acknowledgement of definitional variety	Does the DDT acknowledge the variety of definitions of terrorism?	Binary (1= yes, 0= no)
D6	Reference to the definitions of terrorism	Does the DDT include a direct reference and/or definition of terrorism?	Free text
D7	Definitions of terrorism or derivates as included in the DDTs	What is the definition included in the DDT?	Free text
Raw data			
R1	Type of raw data's source	What type of sources does the DDT use?	Primary sources; Secondary sources; Tertiary sources; Own data; Mix; No information.
R2	Citation transparency	Does the DDT indicate/reference the reference of each data incident?	Binary (1= yes, 0= no) and N/A (when DDTs produce their own data)

R3	Confidence criteria	Does the DDT indicate the credibility/reliability of the sources used?	Binary (1= yes, 0= no) and N/A (when DDTs produce their own data)
R4	Data quality and error handling	Does the DDT indicate how they ensure data quality or proceed when handling errors in their processed data?	Binary (1= yes, 0= no) and N/A (when DDTs produce their own data)
R5	Link to data quality and error handling information	Link	Free text
Methods			
M1	Coding	Does the DDT have a publicly available codebook?	Binary (1= yes, 0= no)
M2	Methodology	Does the DDT have a publicly available methodology?	Binary (1= yes, 0= no)
M3	Link to the codebook	Link	Free text
M4	Link to the methodology	Link	Free text

Results and discussion

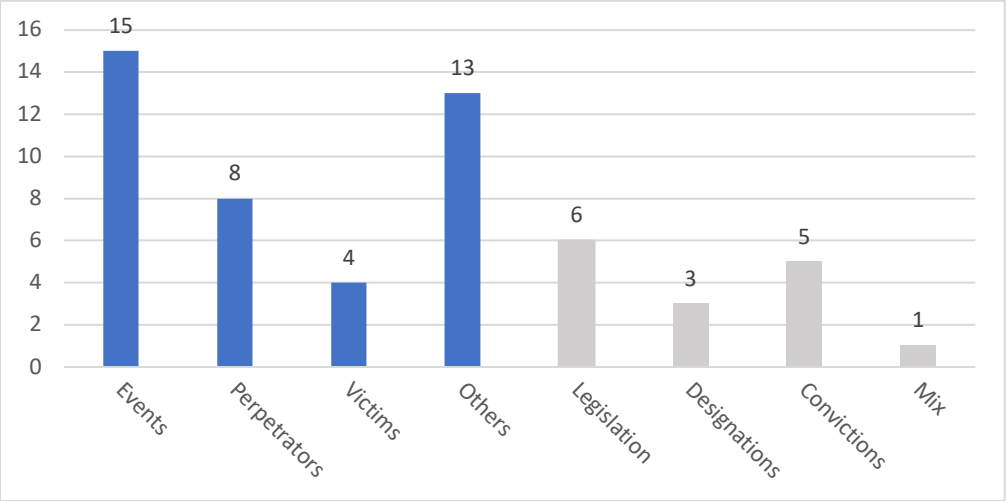
This study analysed the metadata of 55 DDTs to clarify the design features and underlying assumptions that shape how they can be used and interpreted in quantitative research. The results reflect the status of 55 DDTs as of December 2025. The analysis focused on three core dimensions: (1) how definitions and operationalisation are documented, (2) how raw data are sourced and managed, and (3) how methodological procedures are described and made accessible to users.

Contextual information: DDTs' units of analysis.

Across the 55 DDTs, the metadata reveal two analytically distinct families that produce different kinds of knowledge: terrorism-focused datasets (74.5 percent) and counter-terrorism-focused datasets (25.5 percent) (see Figure 1). This distribution matters because the unit of analysis pre-structures what becomes observable, what is aggregated, and which inferential claims can be made without slipping across levels. It is noteworthy that while LaFree[44] identifies three core units of analysis in terrorism research (i.e., events, organisations, and individuals), the results point to a more diverse analytical landscape, with DDTs capturing additional units beyond these conventional categories.

The predominance of event-based units is consistent with how, historically, major terrorism event databases have been built to support systematic descriptive reporting. As LaFree[45] notes, event-based databases "generally follow the classic journalistic format" of recording who is responsible, what happened, where and when it happened, and, where known, how it happened. In the present sample, this architecture is visible in DDTs such as the Global Terrorism Database,[46] the RAND Database of Worldwide Terrorism Incidents (RDWTI),[47] and more specific event catalogues, like the High Casualty Terrorist Bombings 1989-2017.[48] Yet the metadata also show that the label "event" masks heterogeneity: some datasets count attacks, others emphasise fatalities, and some use mixed event–actor records (e.g., plots and attacks alongside logistics or finance).

Figure 1. Number of DDTs per unit of analysis (N = 55).



Perpetrator-focused DDTs constitute the second largest category. Among them, five DDTs focus on organisations, three on individual perpetrators, and one captures both levels of analysis. Organisation-level DDTs often provide information on group history, ideology, leadership, tactics, capabilities, alliances, and modus operandi. Interestingly, one of them, the Mapping Militants Project,[49] also captures some forms of non-violent activities, including political engagement, governance functions, and other forms of organisational activity. Incorporating the monitoring of activities beyond the use of violence represents an important step towards a more comprehensive understanding of the *modus operandi* of terrorist organisations. Nevertheless, political activities capture only part of the broader repertoire of tactics that such organisations may employ. Including coercive, manipulative, and persuasive tactics would provide a more comprehensive picture of how terrorist organisations seek to shape behaviour, exercise influence, and pursue their objectives.[50]

Meanwhile, individual-level DDTs often provide information on the background, attributes, biographies, and legal convictions of individuals involved in terrorism offences. These DDTs, such as the BBC's 'Who are Britain's jihadist?'[51] database, may support research on radicalisation processes, behavioural trajectories, and pathways into terrorism. Among the DDTs examined, however, only the Profiles of Individual Radicalization in the United States (PIRUS)[52] explicitly records whether perpetrators are classified as lone actors. As the Global Terrorism Index notes, over the last 15 years, a growing proportion of fatal incidents have been perpetrated by individuals acting alone rather than by organised groups, with lone actors accounting for the 93 percent of fatal terrorist attacks in the West over the last five years.[53] This highlights the analytical value of distinguishing between lone actors and perpetrators operating within

broader organisational networks, as these may be associated with different pathways of radicalisation and involvement in terrorism. Furthermore, the individual-level DDTs analysed in this study are restricted to the United States and the United Kingdom, highlighting the need for further perpetrator-centred DDTs covering other geographical contexts.

The "Others" category reveals an important diversity of units of analysis of terrorism-related issues, including repositories of terrorist documents, propaganda, extremist symbols and gestures, acts of terrorism glorification, service provision, opinions related to terrorism issues, chronologies, and bibliographies. Illustrative examples include the Harmony Documents Archive,[54] which provides a repository of primary-source terrorist documents, and the Terrorist and Insurgent Organization Social Services (TIOS) Dataset (v2),[55] which captures non-violent organisational activity through data on social-service provision. This breadth reflects the multidimensional character of terrorism research and is consistent with accounts that interdisciplinary engagement in the field has been increasing, even if cross-disciplinary cooperation remains uneven.[56]

Victims constitute the least represented category among the DDTs analysed, with only 4 DDTs focusing on them. The limited availability of open-access victim-centred DDTs highlights an important gap in addressing and understanding the human impact of terrorism. As Silke et al. observe, "[a]n often ignored part of terrorism studies is the analysis of the impact of terrorism on direct and indirect victims".[57] Alarmingly, this gap remains largely overlooked. According to the survey conducted by Schmid et al., only a 4.3 percent of respondents identified the lack of research into the impact of terrorism (victim studies) as one of the greatest weaknesses in the field of terrorism studies.[58] Victim-level data may sometimes be subject to legitimate ethical, legal, privacy or data protection restrictions, particularly where sensitive personal information is involved, and therefore may not be made publicly available. At the same time, ensuring that victims are appropriately represented in terrorism research remains important to avoid rendering the human consequences of terrorism invisible. This is particularly relevant where victims themselves seek recognition, respect, and public remembrance. An example of these efforts is the Mapa del Terror,[59] developed by the Spanish victims' association COVITE to document victims and support public-facing remembrance through its victim-centred record and related memory initiatives. Addressing this gap is critical to remember and respect the victims of terrorism, and to deepen knowledge of and broaden the scope of terrorism research.

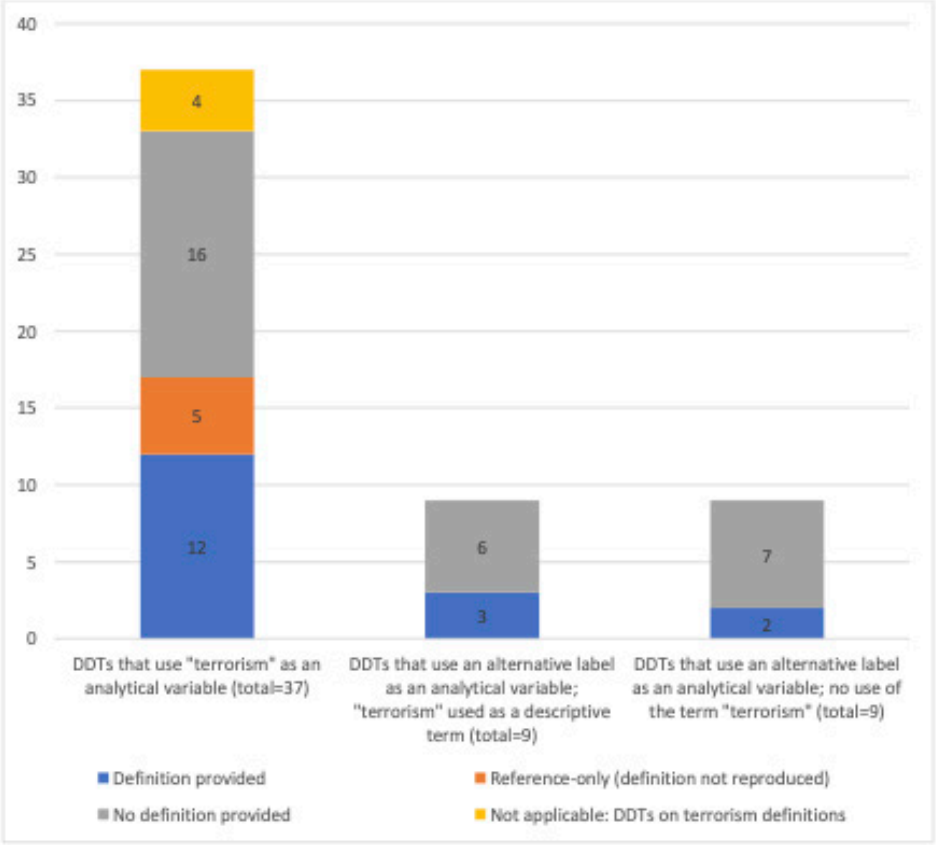
Databases and datasets solely focusing on counter-terrorism data, by contrast, typically record state actions rather than violent activity itself, and the sample illustrates how varied those state-action units can be. For example, these sources operationalise counter-terrorism through units such as terrorism-designation regimes (for individuals and organisations), legislative and policy frameworks (including definitions, treaties, and strategies), and criminal justice outputs (such as convictions and case-law). The availability of DDTs monitoring counter-terrorism data is analytically promising because it expands the empirical basis for studying governance and policy response, areas that remain comparatively under-researched, particularly with respect to evaluating counter-terrorism effectiveness.[60] At the same time, it is important to note that they capture institutional and legal activity shaped by enforcement priorities, reporting practices, and statutory thresholds, and should not be treated as direct proxies for incident prevalence.

Definitions of terrorism

Results about definitional clarity show a number of challenges, namely on the use of the term terrorism (or its derivatives) as an analytical variable and the public availability of the definition of terrorism. In the DDTs analysed, the term terrorism is not consistently specified

as an operational category. Out of 55 DDTs, 67 percent use terrorism (or derivatives) as an analytical variable; a further 16 percent employ an alternative label as the main variable (i.e., extremists, militants, hate groups, jihadists, foreign fighters, or perpetrator/motivation labels) while still using terrorism descriptively, whereas 16 percent do not use the term (see Figure 2). Interestingly, even among DDTs that do not use "terrorism" as their primary analytical variable, some still provide a definition of terrorism. This suggests that some DDT creators consider it useful to clarify their conceptual position in relation to terrorism, even when it is not the main unit of analysis.

Figure 2. Relationship between DDTs use of the term "terrorism" or derivatives as an analytical variable and the availability of the definition of "terrorism" or derivatives.



Concerns arise when examining definitional clarity. Out of the 37 DDTs that use terrorism (or derivatives) as an analytical variable, four primarily compile legal definitions of terrorism and are therefore excluded from the present assessment. Of the remaining 33 DDTs, 16 provide no definition, 12 provide a retrievable definition of terrorism on their own websites, and 5 reference an external definitional source without reproducing the definition in the dataset documentation. This means that most DDTs that use terrorism as an analytical label invite users to conduct empirical analyses of "terrorism" while leaving the construct's scope conditions implicit, pushing users to infer inclusion logic from outputs or peripheral text. Without a retrievable definition, users cannot assess construct validity, replicate selection decisions, or meaningfully compare findings across DDTs, as the conceptual boundaries governing inclusion remain implicit. This concern echoes long-standing warnings: as Sheehan argued, "in the long run, terrorism databases are only as good as the concepts they are built on".[61] Despite his advice, results show that conceptual clarity remains an unfulfilled task for most DDTs.

Nevertheless, even where a definition is provided, operationalisation choices warrant scrutiny. ACLED, for example, introduced terrorism as an analytical category in 2025 as part of a broader reconfiguration of conflict tracking. As they argue, "[c]onflict is changing – and we're changing how we track it".[62] Their approach designates certain non-state actors as "terrorist" based on systematic civilian targeting under specified thresholds and then classifies all events involving those actors within a given country-year as "terrorist activity." This actor- and context-dependent rule is transparent and replicable, yet it is not intrinsically event-based: once an actor is designated, engagements with state forces or other armed groups are also captured under the terrorism label, potentially conflating heterogeneous forms of political violence. Moreover, thresholding may exclude low-frequency but high-impact actors, while cross-context comparisons may be distorted by variation in reporting and coding of civilian targeting. This illustrates why each DDT should be carefully analysed before using its data, as researchers must ensure that the definition employed by a DDT aligns with their analytical approach.

It is important to highlight that definitions of terrorism within DDTs tend to privilege non-state actors as the primary perpetrators of violence. In Bowie's inventories (n=160), only one DDT, i.e., the Political Terror Scale, systematically focuses on violent actions perpetrated by state actors. The PTS monitors *political terror*, defined as "violations of basic human rights to the physical integrity of the person by agents of the state state-based violence".[63] However, the PTS explicitly distinguishes *political terror* from *terrorism,* which they define as "the use of violence and intimidation in pursuit of political aims".[64] Consequently, as it explicitly claims not to monitor terrorism, the PTS was excluded from the present analysis.

This bias towards non-state actors reflects a broader tendency in the field: non-state actors consistently rank among the most central elements in definitions of terrorism, alongside the targeting of governments or civilian populations.[65] Meanwhile, state terrorism and state sponsorship of terrorism remain comparatively underexplored.[66] This has been argued to be the result of the field's state-centrism, in which many researchers adopt state views on what constitutes legitimate versus illegitimate uses of violence, thereby directing scholarly attention towards non-state terrorism.[67]

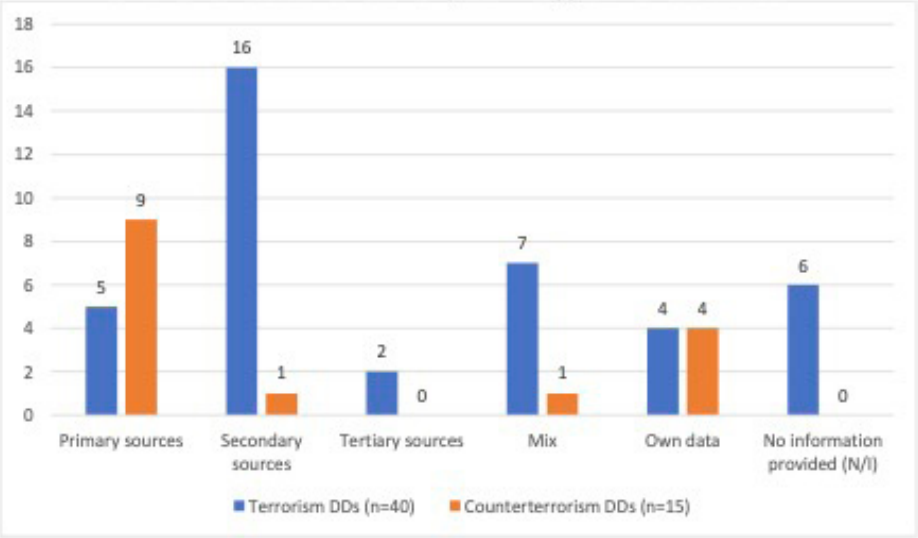
Finally, five DDTs provide only a reference to an external definitional source. It is important to distinguish them from those that do include full definitions on their websites, because reference-only definitions are inherently vulnerable to link failure, institutional closure, and changes in accessibility. The Big, Allied and Dangerous (BAAD) Database [68] case illustrates how discontinuation of public access can undermine the replicability of definitional foundations. The BAAD anchors the label *terrorist groups* to the Memorial Institute for the Prevention of Terrorism's (MIPT) reporting, yet MIPT's public access infrastructure was discontinued and is no longer available online, making the definitional basis difficult to reconstruct. In practical terms, reference-only definitions reduce reproducibility because users may be unable to retrieve the original definitional benchmark that governed inclusion decisions at the time of data construction.

Results show that most DDTs require users to infer the construct's boundaries from outputs or peripheral documentation rather than from an explicit definitional statement. It is therefore crucial that DDTs make their definitional foundations publicly available so that users can identify the definition on which a dataset rests, justify dataset selection, and assess what the resulting data can credibly support. As Silke et al. argue, the agreement on one definition may be unrealistic, but it is essential to be transparent about which definition is being used and how it is applied, so that users do not inherit definitional choices invisibly through data reuse.[69]

The management of raw data

Metadata about the raw data confirms that open-access DDTs are built on heterogeneous evidentiary pipelines, and that these pipelines vary systematically by unit of analysis (see Figure 3).

Figure 3. Relationship between DDTs' focus of information monitoring (i.e., terrorism and counter-terrorism) and the type of raw data used.



DDTs monitoring only counter-terrorism measures are primarily anchored in primary official sources and institutional "Own data", as they often record measures produced and disseminated by the implementing institutions themselves. This configuration is not surprising: when the unit of analysis is a state measure, such as a designation, a legal instrument, or a conviction, the raw data frequently consists of the formal record itself, which is designed to be public, standardised, and institutionally attributable.

DDTs monitoring only terrorism data, by contrast, draw on a wider and more heterogeneous raw-data base in which secondary sources dominate. This aligns with long-recognised constraints on terrorism data collection: unlike many forms of criminal violence, terrorism is rarely captured through the "official/victimisation/self-report" channels that structure conventional criminological data, leading researchers to rely more heavily on open-source materials.[70] This is consistent with the nature of the underlying information, particularly for event data, given the communicative dimension of terrorist attacks. Accordingly, in the sample, information on events is frequently documented through public reporting ecosystems, particularly news media, governmental sources, and communications produced by terrorist actors themselves. Similarly, information on victims and perpetrators mainly relies on secondary sources, particularly news media, governmental information, and NGOs reporting. By contrast, the "Others" category draws more heavily on primary materials, including terrorist documents, images, and video content, as well as surveys and opinion data relating to terrorism issues. This reliance on primary materials can be analytically valuable because it enables research on non-incident dimensions of terrorism.

More concerning, however, is that 6 DDTs in the sample provide no information about the origin of their raw data. Where source type is unstated, records cannot be audited, and users cannot distinguish substantive patterns from artefacts of selection and reporting. This concern

aligns with calls for stronger source-citation transparency in DDTs.[71]

In this sample, DDTs monitoring counter-terrorism (with one exception) cite their sources, yet none documents confidence criteria or data-quality/error-handling procedures. It is important to highlight that, even where inputs are official, errors and revisions remain possible. Thus, documenting uncertainty and correction practices would strengthen interpretability. As Sheehan notes, explicit acknowledgement of uncertainty and clearer documentation of coding rules are important even when inputs appear authoritative.[72]

As for DDTs monitoring terrorism events, there are concerning findings. The 4 DDTs that produced their own data are excluded from this assessment. Among the remaining 36 DDTs, 16 provide no references for the data displayed and one provides only a partial citation. Where references are absent, users cannot trace observations back to original reporting or verify contested cases. In practical terms, absent provenance prevents independent verification and makes it difficult to adjudicate disagreements across datasets that purport to measure similar phenomena.[73] Notably, a small subset of these non-referencing sources still includes confidence cues and/or data-quality handling, indicating that stronger reporting practices are feasible even under open-source constraints. Bridging this gap is crucial for auditability and cumulative research.

Among the DDTs that do cite their sources (20 plus one partial citation), only 5 also document both confidence criteria and data-quality/error-handling procedures, indicating that uncertainty reporting remains limited. Although this is a small number, it represents promising direction for improving data documentation. As Sheehan argues, confidence signalling is an important element of data documentation because sources vary in reliability. Adding structured confidence scoring could support user assessment, albeit with potential trade-offs for monitoring and processing workloads.[74]

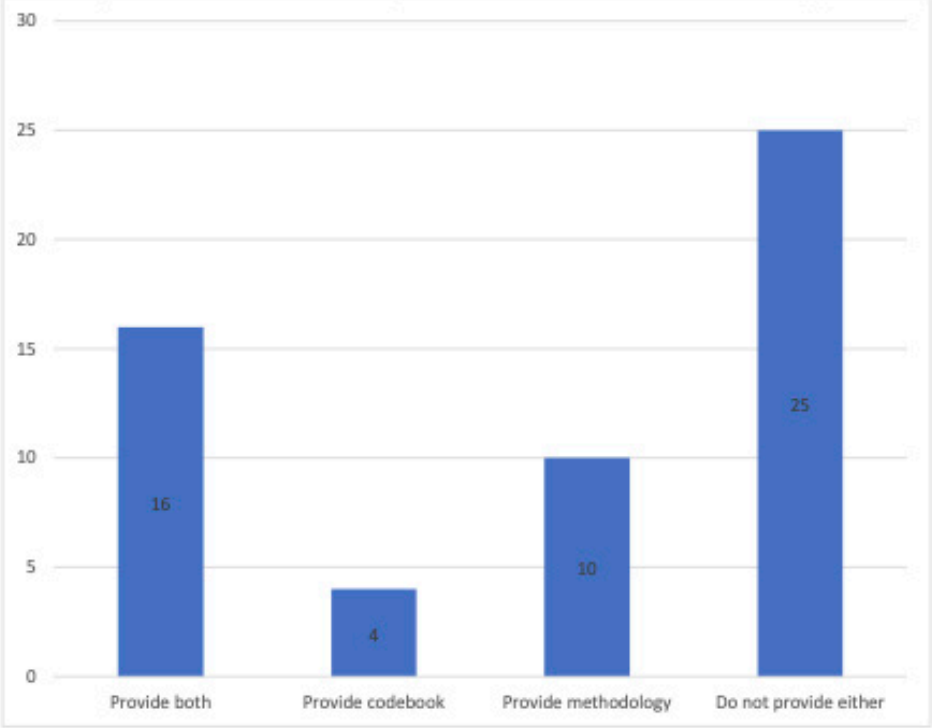
Among the DDTs analysed, the Global Terrorism Database (GTD)[75] has been described as "the most comprehensive publicly available database on acts of terrorism, covering the period 1970 to the present".[76] Notably, the GTD provides a clear definitional approach to terrorism, a transparent raw data management, and accessible documentation (including codebooks and methodological notes), supporting interpretability and reuse. Such documentation strengthens users' ability to interpret scope conditions and to conduct defensible quantitative analysis and provides a practical benchmark for improving transparency across the wider DDT ecosystem. Similarly, LaFree and Gill identify the GTD as an example of good practice in their discussion of how open-source terrorism databases can be improved.[77] In particular, they highlight the GTD's use of source-quality rankings, its transparent inclusion criteria, and its provision of filtering mechanisms that allow users to adapt the dataset to different analytical objectives. These features go beyond documentation by enabling users to assess data quality, evaluate uncertainty, and make informed decisions about case selection.

Methodological transparency and reproducibility

A major finding from this section is the methods transparency gap: almost half of the DDTs (45 percent) are not reproducible, as neither coding information nor a methodology documentation is publicly available (see Figure 4). This distribution is not confined to a single observation level: DDTs that provide both coding and methodology appear across event, victim, perpetrator, other incidents, and counter-terrorism units of analysis.
The availability of coding and methodological information is crucial. As Moravcsik argues,

transparency requires making available "the data, theory and methodology on which [the] conclusions rest".[78] Without this information, users cannot know which conceptual and procedural filters were applied at the point of classification and, therefore, confidence in the research may be compromised.[79]

Figure 4. Shares of DDTs that provide codebook and/or methodology.



Without coding documentation, definitional boundaries become implicit and users may treat an indicator as *terrorism* when it actually reflects a narrower (or broader) inclusion logic, or when key exclusions (uncertain attribution, ambiguous motive, duplicate reports) were handled inconsistently. Where coding and methodology are missing, additional problems arise with versioning and maintenance. DDTs frequently depend on staff turnover, changing sourcing environments, and periodic redesigns; without a stable methodological record, users cannot identify whether apparent trends reflect real-world change or shifting rules, coverage, or verification thresholds.[80]

Most of the analysed DDTs did not provide access to their codebook and/or methodology. In our analytical scheme, we captured non-availability rather than non-existence of such information, indicating that the relevant items could not be located online, not that the information is absent from the DDTs' internal practices. Accordingly, whether these DDTs have indeed developed such documentation but have not made it publicly available remains unknown. Nevertheless, from the perspective of users and reproducible research, what matters is public accessibility: if decision rules cannot be retrieved, they cannot be evaluated, replicated, or meaningfully compared across sources. This makes transparency itself a substantive condition of interpretability and a pre-analytic criterion for dataset selection.

A constructive implication is that several DDTs illustrate promising practices worth highlighting: transparent inclusion criteria, clear unit definitions, and publicly available codebooks or methodological notes that make selection and classification decisions reproducible. These examples demonstrate that meaningful methods transparency is feasible even under the

practical constraints of terrorism data collection. They also provide a concrete benchmark against which less documented DDTs can be evaluated and improved.

Limitations

Some limitations of the study should be considered, particularly the selection and content of the DDTs analysed, the exclusion of funding analysis, and the methodology employed in the study.

First, the DDTs analysed do not encompass all the sources available in the field of terrorism. Bowie's inventories accounted for the largest selection of sources on terrorism known to the authors, providing a practical starting point for the analysis, in addition to the other 9 sources known to the authors. However, the resulting sample is linguistically concentrated: most DDTs were available either in English or Spanish, with only 8 sources available in two or more languages (with English or Spanish among them) and 1 source only in Spanish. Future work incorporating more non-English sources could therefore yield a fuller picture of the state of the art. Likewise, regarding geographical coverage, the sample does not necessarily reflect the full diversity of DDTs available globally. Accordingly, the present findings should be interpreted as applying to this sample rather than being generalised to the field as a whole.

Second, although Bowie's inventories were used to reduce potential selection biases, the inventories themselves do not document the criteria or procedures by which sources were identified and included. As shown in the preparatory analysis, some of the sources gathered by Bowie cover broader conflict and political violence rather than focusing solely on terrorism. Bowie does not make the selection process transparent; therefore, it is not possible to understand the criteria underlaying their inclusion, whether additional relevant sources were omitted, or how comprehensive the underlying inventory is.

Third, analysing the funding of DDTs was considered at the early stages of the analysis. As Sheehan argues, it is important to understand the extent to which sponsors, funders, or other actors influence the scope of a database on terrorism, since it is a national security concern.[81] Nevertheless, this element was ultimately excluded, as analysing the funding of DDTs would require a dedicated study capable of disentangling multiple stakeholders' roles in design, maintenance, hosting, and operational control. These roles, however, are not often disclosed with enough clarity to support systematic coding. For these reasons, funding analysis falls outside the scope of the present study but remains an important line of research for future studies.

Finally, regarding the methodology, the analysis of DDTs was done manually. Automated processes were initially explored to improve the standardisation of the analysis. However, the diversity of DDT formats, terminology, and web architectures across DDTs rendered software-supported extraction impractical. The results were double checked by December 2025. However, as exemplified with ACLED,[82] DDTs also change their parameters and measures over time, meaning that some metadata information and results can become outdated. Finally, the frequent use of dispersed hyperlinks within individual websites complicated the retrieval of necessary information and increased the risk that relevant documentation might be difficult to locate even where exists. As illustrated by the BAAD database,[83] reliance on external hyperlinks can create a fragility problem, and some information may no longer be retrievable when referenced links are discontinued or become inaccessible.

Conclusions

This article analysed the metadata of 55 open-access databases and datasets on terrorism (DDTs) seeking to clarify the design features and underlying assumptions that guide their use. It shifts attention away from DDTs' outputs towards the design choices documented in metadata, particularly those relating to definitional clarity, raw data management, and methodological transparency and reproducibility. By doing this, the study shows that many divergences between DDTs' results can be traced to differences in how these DDTs are constructed, documented, and maintained, rather than to substantive disagreement about terrorism alone. To operationalise these insights for users, the full results are made available as a downloadable spreadsheet,[84] intended to facilitate DDTs selection by enabling researchers to compare DDTs systematically and to assess their alignment with specific research designs.

Several strengths emerge. First, the open-access landscape is more diverse than is often assumed: alongside event catalogues, the sample includes DDTs centred on perpetrators, victims, counter-terrorism measures, and other units (e.g., documents, propaganda, symbols, surveys and opinion data, and service provision). Second, many DDTs provide at least some traceability through citations and source descriptions, offering a basis for audit and verification. Third, the presence of counter-terrorism-focused DDTs expands the empirical infrastructure for studying legal and administrative measures, an area that previous assessments of the field continue to regard as comparatively under-developed, particularly in relation to evaluating effectiveness.

At the same time, the analysis identifies persistent gaps that limit interpretability and cumulative research. Victim-centred measurement remains under-represented, constraining systematic quantitative insight into harm and impact. Definitional transparency is frequently insufficient: a substantial share of DDTs that use "terrorism" (or derivative terms) as unit of analysis do not make their definitional scope conditions retrievable, forcing users to infer inclusion criteria from outputs or peripheral text and weakening construct validity. Raw-data provenance is uneven, and in a subset of cases not documented at all, which undermines auditability and increases the risk that selection and reporting artefacts are mistaken for substantive patterns. Finally, methods transparency remains a central bottleneck: many DDTs do not make coding rules and/or methodology publicly accessible. When available and depending on their level of detail, codebooks and methodological documentation can provide important information about how DDTs are constructed and maintained, thereby contributing to traceability, interpretability, and transparency. Nevertheless, such materials remain unavailable for a substantial proportion of DDTs, limiting reproducibility and making it difficult to determine whether observed trends reflect substantive developments or changes in update routines, verification thresholds, and maintenance practices over time.

The practical implication is that metadata transparency should function as a pre-analytic filter for DDT selection. First, users should verify whether the DDT's terminology and definitional scope align with the concepts and boundaries required by their research question. Second, they should assess raw-data traceability: whether the DDT provides sufficient provenance for users to consult original source material, so that, in cases of doubt or contestation, analysis is not confined to processed outputs alone. Third, they should confirm that coding rules and inclusion/exclusion criteria are accessible, including the methodological procedures that govern classification, revision, and (where relevant) updates, since these features shape interpretability, comparability, and reproducibility.

Without these elements, even sophisticated modelling risks becoming an artefact of opaque construction rather than a defensible statement about terrorism-related phenomena.

Improving quality does not necessarily require complex redesign but rather relatively straightforward steps. Actions such as publishing retrievable definitions, providing codebooks and methodological notes, and documenting provenance and revision practices would strengthen interpretability, comparability, and the reliability of quantitative terrorism research.

Marina Alonso-Villota is a Ph.D. candidate in International Security at University Institute General Gutiérrez Mellado, National University of Distance Education (IUGM-UNED). She was a researcher at University of the Bundeswehr Munich while conducting this study. Her main research interests include terrorism and hybrid threats.

Rubén Arcos is an Associate Professor of Communication at University Rey Juan Carlos. He is a Visiting Professor of Intelligence Analysis at the College of Europe in Bruges, and faculty of the PhD program in International Security at IUGM-UNED. Arcos is also a visiting professor at the Preparedness and Resilience Research Platform of University of Vassa. He serves as Chair of the Intelligence Studies Section at the International Studies Association (ISA).

Declaration of generative AI and AI-assisted technologies: During the preparation of this work, the authors used ChatGPT solely for grammar correction and for improving the language and readability of certain sentences. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Annex I. List of the 55 DDTs analysed

	Name and link
1	Armed Conflict Location and Event Data Project (ACLED) Explorer
2	Big, Allied and Dangerous (BAAD) Database 1 - Lethality Data, 1998-2005.
3	BBC News Database of Jihadists
4	Chronological List of Armenian Terrorist Activities from 1918 to 1989
5	Combating Terrorism Center (CTC) Harmony Database
6	Domestic Terrorist Victims Dataset (Version 1)
7	Southeast Asia Militant Atlas (SEA Militant Atlas)
8	Global Terrorism Database (GTD)
9	Hate Symbols Database
10-	Iraq Body Count
11	Replication Data: Revolutionary Terrorist Victims Dataset
12	South Asia Terrorism Portal
13	Victims of ETA dataset (2007)
14	Counter-Terrorism Statistics (Operation of Police Powers under the Terrorism Act 2000) United Kingdom
15	Listed Terrorist Entities
16	The RAND Database of Worldwide Terrorism Incidents (RDWTI)
17	High Casualty Terrorist Bombings, 1989-2017
18	Muslim Public Opinion on U.S. Policy, Attacks on Civilians, and al Qaeda
19	The Nigeria Security Tracker (NST)
20	Nuclear Facilities Attack Database (NuFAD)
21	Pew Research Center - Data Surveys on Terrorism
22	Profiles of Individual Radicalization in the United States (PIRUS)
23	Profiles of Perpetrators of Terrorism in the United States (PPT-US)
24	Extremists, groups & ideologies
25	Terrorism Against Israel (1920-Present)
26	Gallup Polls on Terrorism-related Topics
27	Terrorism Most Wanted Lists
28	Sanctions List Search Tool
29	International Crimes Database (ICD) > Foreign Fighters
30	RTV Dataset
31	Sources of Blame Attribution: Citizen Attitudes Towards Public Officials after 9/11 [Data Set]
32	The Global CVE Data set
33	The International Crimes Database (ICD) > Terrorism
34	SHERLOC > Case Law Database
35	H.E.A.T. Map (Hate, Extremism, Antisemitism, Terrorism)
36	Database of Legislation on the Definition of Terrorism
37	Islamist Terror Attacks in the World 1979-2019

38	Lebanese Hezbollah Select Worldwide Activities Interactive Map and Timeline
39	The Mapping Militants Project (MMP)
40	9/11 Timeline
41	Suicide Attacks Database (2020)
42	Terrorists and Extremists Database
43	The Terrorism and Foreign Fighters Database
44	Terrorist and Insurgent Organization Social Services (TIOS) Dataset (v2)
45	Typology of Terror
46	Foreign Terrorist Organisations (FTO's) List
47	SHERLOC > Database of Legislation
48	SHERLOC > Bibliography Database
49	SHERLOC > Treaties Database
50	SHERLOC > Strategies Database
51	The Islamic State, Select Worldwide Activity
52	Mapa del terror
53	Observatorio de la radicalización
54	Militant Imagery Project
55	Terrorist and Extremist Groups

Endnotes

1 DDTs, unless otherwise specified, are used throughout this article as a broad term referring to databases and datasets on terrorism without presupposing whether they monitor terrorist activity or counter-terrorism measures. Where explicitly indicated, this distinction is applied.

2 Gary LaFree, 'The Evolution of Terrorism Event Databases', in *The Oxford Handbook of Terrorism*, ed. Erica Chenoweth et al. (Oxford University Press, 2019), 414; Gary LaFree and Laura Dugan, 'Research on Terrorism and Countering Terrorism', Crime and Justice 38, no. 1 (2009): 413–77, https://doi.org/10.1086/599201.

3 LaFree, 'The Evolution of Terrorism Event Databases', 414; LaFree and Dugan, 'Research on Terrorism and Countering Terrorism'

4 Neil G. Bowie, 'Terrorism Events Data: An Inventory of Databases and Data Sets, 1968-2017', *Perspectives on Terrorism* 11, no. 4 (2017): 50–72; Neil G. Bowie, '30 Terrorism Databases and Data Sets: A New Inventory', *Perspectives on Terrorism* 12, no. 5 (2018): 51–61; Neil G. Bowie, 'A New Inventory of 30 Terrorism Databases and Data Sets', *Perspectives on Terrorism* 14, no. 1 (2020): 54–66; Neil G. Bowie, '40 Terrorism Databases and Data Sets: A New Inventory', *Perspectives on Terrorism* 15, no. 2 (2021): 147–61.

5 Bowie's inventories (2017, 2018, 2020, and 2021) compile a wide range of databases and datasets on terrorism, although he acknowledges that some of these sources have a broader scope, encompassing political violence or general armed conflict.

6 Bart Schuurman, 'Research on Terrorism, 2007–2016: A Review of Data, Methods, and Authorship', *Terrorism and Political Violence* 32, no. 5 (2020): 1011–1026.

7 Andrew Silke, "The Devil You Know: Continuing Problems with Research on Terrorism," Terrorism and Political Violence 13, no. 4 (2001): 5, quoted in Ibid., 1020..

8 Lee Jarvis, 'Critical Terrorism Studies and Numbers: Engagements, Openings, and Future Research', *Critical Studies on Terrorism* 16, no. 4 (October 2023): 720–43, https://doi.org/10.1080/17539153.2023.2267280.

9 Kathryn Loosemore and Matthew Johnson, 'The Challenge of Establishing the Impact of Terrorist Kathryn Loosemore and Matthew Johnson, 'The Challenge of Establishing the Impact of Terrorist Organisations: Development of a Database on ETA's Activities', *Critical Studies on Terrorism* 16, no. 2 (April 2023): 402, https://doi.org/10.1080/17539153.2023.2170514. It is important to note that Losemoore and Johnson refer here only to violent activities perpetrated by the terrorist organisation ETA, such as terrorist attacks, kidnappings, assassinations, and related forms of violence. A comprehensive analysis of the activities of the terrorist organisation, however, would also need to consider non-violent and legal activities, including those of political, social, and economic nature. For a more detailed discussion of this distinction and its implications for the study of terrorist organisations, see Alonso-Villota and Arcos (2024) and Arcos, Alonso-Villota, and Milani (2026).

10 Alex Schmid, 'Statistics on Terrorism: The Challenge of Measuring Trends, Global Terrorism', in *Forum on Crime and Society* (2004), 4:54–55, https://doi.org/10.18356/9789211554199. Schmid explains that discrepancies between terrorism databases stem not only from who collects the data, but also from how incidents are defined and counted. For example, a wave of 40 bombings in one night was recorded as 40 separate incidents by the CIA but as a single event by RAND, illustrating how coding choices can significantly affect totals even when definitions are similar

11 Ibid., 55–56.

12 Ivan Sascha Sheehan, 'Assessing and Comparing Data Sources for Terrorism Research', in *Evidence-Based Counterterrorism Policy*, ed. Cynthia Lum and Leslie W. Kennedy (New York, NY: Springer New York, 2012), 15–16, https://doi.org/10.1007/978-1-4614-0953-3_2.

13 Ibid., 16–17.

14 Sheehan, 'Assessing and Comparing Data Sources for Terrorism Research'.

15 Ibid

16 André Python et al., 'A Bayesian Approach to Modelling Subnational Spatial Dynamics of Worldwide Non-State Terrorism, 2010–2016', *Journal of the Royal Statistical Society Series A: Statistics in Society* 182, no. 1 (January 2019): 325, https://doi.org/10.1111/rssa.12384.

17 Johnson and Ackerman, 'Terrorism Databases: Problems and Solutions'.

18 Python et al., 'A Bayesian Approach to Modelling Subnational Spatial Dynamics of Worldwide Non-State Terrorism, 2010–2016'.

19 Johnson and Ackerman, 'Terrorism Databases: Problems and Solutions'.

20 Ibid., 160.

21 Marina Alonso-Villota and Rubén Arcos, 'Results: A Metadata Analysis of 55 Databases and Datasets on Terrorism.', version V1, Mendeley Data, 2026, https://doi.org/10.17632/2p659d7tk3.1.

22 Jeffrey Pomerantz, *Metadata* (Cambridge, Massachusetts; London, England: The MIT Press, 2015), 1.

23 Ibid., 16..

24 Schmid, 'Statistics on Terrorism: The Challenge of Measuring Trends, Global Terrorism'; LaFree and Dugan, 'Research on Terrorism and Countering Terrorism'; LaFree, 'The Evolution of Terrorism Event Databases'.

25 Schmid, 'Statistics on Terrorism: The Challenge of Measuring Trends, Global Terrorism', 51.

26 Sheehan, 'Assessing and Comparing Data Sources for Terrorism Research', 17.

27 Schmid, 'Statistics on Terrorism: The Challenge of Measuring Trends, Global Terrorism', 52–54.

28 Alex P. Schmid and A. J. Jongman, Political Terrorism: *A New Guide to Actors, Authors, Concepts, Data Bases, Theories, and Literature,* Rev., expanded and updated ed (New Brunswick (USA) and London (UK): Transaction Publishers, 1988); Alex P. Schmid, *Defining Terrorism* (International Centre for Counter-Terrorism, 2023), https://doi.org/10.19165/2023.3.01.

29 LaFree, 'The Evolution of Terrorism Event Databases', 57.

30 Andrew Silke, Lara A. Frumkin, and John F. Morrison, 'Chapter 1. Contemplating a Research Agenda for Terrorism Studies', in *A Research Agenda for Terrorism Studies*, ed. Lara Frumkin, John Morrison, and Andrew Silke (Cheltenham, UK: Edward Elgar Publishing, 2023), 1–20, https://doi.org/10.4337/9781789909104.00005.

31 Gary LaFree, 'Chapter 7. Terrorism Open Source Databases', in *Contemporary Terrorism Studies (1st Edn)*, ed. Diego Muro and Tim Wilson (Oxford: Oxford University Press, 2022), 114, https://doi.org/10.1093/hepl/9780198829560.003.0007.

32 Schmid, 'Statistics on Terrorism: The Challenge of Measuring Trends, Global Terrorism', 56–57; LaFree, 'The Evolution of Terrorism Event Databases', 57–58.

33 LaFree, 'The Evolution of Terrorism Event Databases', 58.

34 Laura Dugan and Michael Distler, 'Measuring Terrorism', in *The Handbook of the Criminology of Terrorism*, ed. Gary LaFree and Joshua D. Freilich (2016), 187–205, https://doi.org/10.1002/9781118923986.

35 Johnson and Ackerman, 'Terrorism Databases: Problems and Solutions', 168-171.

36 Sheehan, 'Assessing and Comparing Data Sources for Terrorism Research'.

37 Schmid, 'Statistics on Terrorism: The Challenge of Measuring Trends, Global Terrorism', 51.

38 Johnson and Ackerman, 'Terrorism Databases: Problems and Solutions'.

39 Bowie, 'Terrorism Events Data'; Bowie, '30 Terrorism Databases and Data Sets: A New Inventory'; Bowie, 'A New Inventory of 30 Terrorism Databases and Data Sets'; Bowie, '40 Terrorism Databases and Data Sets: A New Inventory'.

40 For instance, RAND is no longer active, but its data remains accessible online. Thus, RAND was included for the analysis.

41 Schmid, 'Statistics on Terrorism: The Challenge of Measuring Trends, Global Terrorism', 51.

42 For all the criteria, the following responses are an option: N/A (Not Applicable), said when the nature of the source does not align with the information required by the criterion and therefore excludes it; and N/I (No Information), used when the source is expected to provide such information, but it is not publicly available.

43 Alonso-Villota and Arcos, 'Results: A Metadata Analysis of 55 Databases and Datasets on Terrorism.'

44 Gary LaFree, 'Chapter 7. Terrorism Open Source Databases', in *Contemporary Terrorism Studies*, ed. Diego Muro and Tim Wilson (Oxford University Press, 2022), 114–15.

45 LaFree, 'The Evolution of Terrorism Event Databases', 50.

46 National Consortium for the Study of Terrorism and Responses to Terrorism (START), 'Global Terrorism Database', https://www.start.umd.edu/gtd/.

47 RAND Corporation, 'RAND Database Scope', 2009, https://www.rand.org/nsrd/projects/terrorism-incidents/about/definitions.html.

48 Center for Systemic Peace, 'High Casualty Terrorist Bombings, 1989-2021', 2022, https://www.systemicpeace.org/inscrdata.html.

49 Martha Crenshaw and Kaitlyn Robinson, Mapping Militants Project (Rice University, 2025), https://doi.org/https://doi.org/10.25613/G0K4-WF70.

50 Marina Alonso-Villota and Rubén Arcos, 'The Coercion-Manipulation-Persuasion Framework: Analyzing the Modus Operandi of Systems of Non-State Actors', *Terrorism and Political Violence* 37, no. 5 (2025): 632–50, https://doi.org/10.1080/09546553.2024.2357082.

51 'Who Are Britain's Jihadists?', UK, *BBC News*, 25 March 2015, https://www.bbc.com/news/uk-32026985.

52 National Consortium for the Study of Terrorism and Responses to Terrorism (START), 'Profiles of Individual Radicalization in the United States (PIRUS)', https://www.start.umd.edu/data-tools/profiles-individual-radicalization-united-states-pirus.

53 Institute for Economics & Peace, *Global Terrorism Index 2026: Measuring the Impact of Terrorism* (Sydney, 2026), 66, https://www.visionofhumanity.org/wp-content/uploads/2026/03/Global-Terrorism-Index-2026-Report.pdf.

54 Combating Terrorism Center, 'Harmony Documents Archive', 2026, https://ctc.westpoint.edu/harmony-document/.

55 One Earth Future, 'Terrorist and Insurgent Organization Social Services (TIOS) Dataset', n.d., https://oefresearch.org/datasets/tios.

56 Alex P. Schmid, James J. F. Forest, and Timothy Lowe, 'Terrorism Studies: A Glimpse at the Current State of Research', *Perspectives on Terrorism* 15, no. 3 (2021): 146–47.

57 Silke, A. Frumkin, and F. Morrison, 'Contemplating a Research Agenda', 11.

58 Schmid, Forest, and Lowe, 'Terrorism Studies: A Glimpse at the Current State of Research', 147.

59 COVITE, 'Mapa del terror', Mapa del terror, 2024, https://mapadelterror.com.

60 Schmid, Forest, and Lowe, 'Terrorism Studies: A Glimpse at the Current State of Research', 147–50.

61 Sheehan, 'Assessing and Comparing Data Sources for Terrorism Research', 3.

62 'Conflict Categories', ACLED, 30 July 2025, https://acleddata.com/conflict-data/conflict-categories.

63 Mark Gibney et al., 'The Political Terror Scale 1976-2021 Codebook', Political Terror Scale Website, 2022, 1, http://www.politicalterrorscale.org/.

64 Ibid.

65 Schmid, Forest, and Lowe, 'Terrorism Studies: A Glimpse at the Current State of Research', 143–44.

66 Ibid., 148–49.

67 Bart Schuurman, 'Topics in Terrorism Research: Reviewing Trends and Gaps, 2007-2016', *Critical Studies on Terrorism* 12, no. 3 (July 2019): 466, https://doi.org/10.1080/17539153.2019.1579777.

68 Victor Asal, R. Karl Rethemeyer, and Ian Anderson, 'Big Allied and Dangerous (BAAD) Database 1 - Lethality Data, 1998-2005', version 3.0, Harvard Dataverse, 6 June 2011, https://doi.org/10.7910/DVN/GPEUFH.

69 Silke, A. Frumkin, and F. Morrison, 'Contemplating a Research Agenda'.

70 LaFree and Dugan, 'Research on Terrorism and Countering Terrorism', 416.

71 Sheehan, 'Assessing and Comparing Data Sources for Terrorism Research', 23.

72 Ibid.

73 Ibid.

74 Ibid.

75 National Consortium for the Study of Terrorism and Responses to Terrorism (START), 'Global Terrorism Database'..

76 Schmid, Forest, and Lowe, 'Terrorism Studies: A Glimpse at the Current State of Research', 144.

77 Gary LaFree and Paul Gill, 'Strengths and Weaknesses of Open Source Data for Studying Terrorism and Political Radicalization', *Studies in Conflict & Terrorism*, 5 June 2024, 12-13, https://doi.org/10.1080/1057610X.2024.2361948.

78 Andrew Moravcsik, 'Transparency: The Revolution in Qualitative Research', *PS: Political Science & Politics* 47, no. 01 (January 2014): 48, https://doi.org/10.1017/S1049096513001789.

79 Ibid.

80 Schmid, 'Statistics on Terrorism: The Challenge of Measuring Trends, Global Terrorism'.

81 Sheehan, 'Assessing and Comparing Data Sources for Terrorism Research', 24.

82 ACLED, 'Conflict Categories'.

83 Asal, Rethemeyer, and Anderson, 'Big Allied and Dangerous (BAAD) Database 1 - Lethality Data, 1998-2005'.

84 Alonso-Villota and Arcos, 'Results: A Metadata Analysis of 55 Databases and Datasets on Terrorism.'

About

Perspectives on Terrorism

Established in 2007, *Perspectives on Terrorism* (PT) is a quarterly, peer-reviewed, and open-access academic journal. PT is a publication of the International Centre for Counter-Terrorism (ICCT), in partnership with the Institute of Security and Global Affairs (ISGA) at Leiden University, and the Handa Centre for the Study of Terrorism and Political Violence (CSTPV) at the University of St Andrews.

Copyright and Licensing

Perspectives on Terrorism publications are published in open access format and distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License, which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, the source referenced, and is not altered, transformed, or built upon in any way. Alteration or commercial use requires explict prior authorisation from the International Centre for Counter-Terrorism and all author(s).

© 2023 ICCT

Contact

E: pt.editor@icct.nl
W: pt.icct.nl

Universiteit Leiden

HANDA CSTPV