

METHODOLOGY

Codebook: Threats and Harassment Dataset (THD)

DECEMBER 2024



**Bridging
Divides
Initiative**

BDI is a non-partisan research initiative at Princeton University that tracks and mitigates political violence in America.

bridgingdivides.princeton.edu

Threats and Harassment Dataset (THD)

Table of Contents

Introduction	2
THD Variables at a Glance	4
Methodology	5
Nature of Event	6
<i>Threat</i>	7
Threat of Death	7
Threat of Physical Harm or Violence	8
<i>Harassment</i>	8
Invasion of Privacy	9
Intimidation	10
Method of Event	10
<i>Physically Removed or Virtual</i>	11
<i>Physically Present</i>	11
Actors	12
<i>Target of Event</i>	12
Target Size and Fuzziness	12
Target Affiliation	13
Target Gender	14
<i>Perpetrator of Event</i>	15
Perpetrator Affiliation	15
Event Description	15
<i>Type of Targeting</i>	15
<i>Issue at Stake</i>	16
<i>Summary</i>	17
Temporal	17
<i>Date Fuzziness</i>	17
Geography	18
<i>Location Fuzziness</i>	18
Sources of Information	19
<i>Traditional Media</i>	20
<i>BDI OSINT</i>	20
<i>Consortium</i>	20
Dataset Management	21
<i>Deleted Events</i>	22
<i>Tracking Updates</i>	22
THD Variables and Categories/Structure	23

Introduction

In order to better make sense of the number, severity, purpose, and geographic distribution of threats targeting local, elected, appointed, or municipal officials,¹ the Bridging Divides Initiative (BDI) collects, codes, and organizes event data on threats and harassment of local officials from across the United States.

The Threats and Harassment Dataset (THD) is an event-based dataset that collects and codes reported information on targets, perpetrators, location, date, method, type and other characteristics of threats and harassment incidents. (The table presented at the end of this section offers a quick look at the variables included in the dataset and what each helps to measure; a table at the end of the Codebook offers more details around all variables and categories therein.)

Such data and research can help to inform interventions, both at the community-level, as well as for policymakers.² For example, data collected helps to shed light on a variety of research questions, including, but not limited to:

- What types of local, elected, appointed, or municipal officials face greatest risk, and where?
- Do different types of local, elected, appointed, or municipal officials face different risks? How does a local, elected, appointed, or municipal official's gender or other (perceived) demographic information affect their likelihood of facing threats/harassment?
- What types of threats or harassment are most common, and how are they executed?
- What groups are threatening/harassing local, elected, appointed, or municipal officials, and how?
- What issues/grievances seem to be motivating threats/harassment of local, elected, appointed, or municipal officials?

THD centers the experience of local, elected, appointed, or municipal officials, meaning that inclusion criteria consider, first and foremost, what activity would be perceived as threatening or harassing by such individuals. This ensures that the resulting data collection effort is more 'valid' — i.e. that the data offers a more accurate measure that truly measures what it purports to measure. In-depth, quarterly interviews with and representative surveys of such officials, along with primary data collection regarding such threats and harassments via self-reporting tools, help to ensure this lens.

¹ Election workers and/or volunteers *working in an official capacity* are also included amongst officials here, given their role in ensuring democratic processes — which aligns in spirit with the role of local, elected, appointed, or municipal officials. Partisan poll watchers are *not* considered to be working in an *official* capacity, and hence (perceived) threats or harassment against them does *not* meet the criteria for inclusion within the dataset.

² While such data and research can shed light on the outcomes of such threats and harassment (i.e. do threats and harassment of officials contribute to the resignations of such officials), it is important to consider what additional factors may also contribute to such outcomes. As such, outcomes of threats and harassment are beyond the scope of inclusion within THD; only the activity itself (i.e. the threat, or the harassment) is captured within THD. Similarly, the threats of harm that are executed (i.e. violence itself) are not included in the dataset (more on that below).

THD data at present are based on information from secondary information stemming from open-source investigation of traditional and social media; a variety of data from BDI partner organizations; and may integrate primary information gathered via interviews and surveys in the future.³ Information is coded by a team of research assistants, based both at Princeton University and across the United States, and go through multiple rounds of review prior to publication.

Following the launch of the publicly-available dataset (covering 1 January 2022 to present), new THD data are published monthly, covering reported events from the months prior. Data are freely available for download with registration on the BDI website. More detailed data, including more granular information, is available in select cases; users interested in this level of data should reach out to BDI to learn more.

³ These various sources are explored in further detail in the **Sources of Information** section below.

THD Variables at a Glance

Type of Event Information	Variable Name	Brief Description
Descriptive	Data ID	Unique numerical identified
	Event Size	How many incidents (roughly) is the event representing?
	Incident Fuzziness	How sure are we about how many incidents the event is representing?
Temporal	Date	Date that the event occurred
	Date Fuzziness	How sure are we about the coded event date?
Nature	Nature	Was this a threat or harassment?
	Nature Type	Type of threat/harassment
	Nature Detail	Specific type of threat/harassment
Geography	County	County where event occurred
	State	State where event occurred
	Location Fuzziness	How sure are we about the coded event location?
Actors	Target Size	How many targets (roughly) is the event representing?
	Target Size Fuzziness	How sure are we about how many targets the event is representing?
	Target Affiliation	Type of official targeted
	Target Gender	Gender of targeted official
	Perpetrator Affiliation	Group that perpetrator was affiliated with (if any)
Method	Method Type	How/Through what means was the threat/harassment executed?
	Method Detail	Specific method of how threat/harassment was executed
Substantive	Type of Targeting	Type of hateful language/symbols used by perpetrator (if any)
	Issues at Stake	Grievances that motivated the event (if any)
	Event Summary	Brief summary of the event
Sources	Source Type	Type of source(s) used to code event
	Source Detail	Name of source(s) used to code event
Dataset Management	Last Updated	Date (month and year) that the event was last updated and/or added to the dataset

Methodology

BDI collects and codes reported information on threats and harassment targeting locally elected, appointed, or municipal officials across the United States. It captures the geographic distribution, methods, types and purpose of such incidents, as well as information related to the target(s) and perpetrator(s).

The unit of analysis in THD is an event. Events involve an incident of a specific nature and method, involving a specific perpetrator (i.e. an individual) and a specific target (i.e. an individual), at a specific time (i.e. date) and place (i.e. geographic location).⁴ Each event is uniquely coded within THD, and is assigned a unique numerical identifier (**variable name: Data ID**).

For example, an individual who made a threatening phone call to an official and later the same day leaves several threatening voicemails for them would be coded as a single event (i.e. same day, same location, same perpetrator, same target, same method). If the individual makes *another* threatening phone call the next day, this phone call would be coded as a new, separate event (i.e. *different* day, same location, same perpetrator, same target, same nature, same method). Or, if the individual sent threatening messages to *several* officials on that same initial day, each would be coded as a separate event, assuming reporting notes who the different officials were (i.e. same day, same location, same perpetrator, *different* target, same nature, same method). Research assistants strive to find the most granular and detailed information possible so as to disaggregate reporting accordingly.

In some cases, however, this might not be possible to do. For example, there may be cases where several events *should* be coded (e.g. *several* officials with the same affiliation have been targeted,⁵ or threats were made by the same perpetrator against the same target on *multiple* occasions) yet reporting is not detailed enough to allow for disaggregated coding (e.g. no sufficient reporting on who all of the officials are, or what days/how often the threats occurred). In such cases, a single event is coded, though the (estimated) size of the event (i.e. roughly how many incidents the single event is representing) is documented (**variable name: Event Size**):

1. **'Single'**: Event represents a single incident
2. **'Small'**: Event represents a small number of incidents (i.e. ~ less than 10 incidents)
3. **'Large'**: Event represents a large number of incidents (i.e. ~ more than 10 incidents)
4. **'Unknown'**: The number of incidents that the event represents cannot be approximated from reporting

Further, an additional variable (**variable name: Incident Fuzziness**) helps to capture how accurate this estimated size (i.e. the coding of *Event Size*) may be:

⁴ Ensuring that an event has each of these components helps to ensure that the unit of analysis is substantively meaningful, rather than functionally a measure of reporting. The same incident, reported on in two different ways, must be coded in the same way to ensure a measure of substance rather than reporting.

⁵ In cases where several officials with *different* affiliations are targeted, separate events are coded, respective of *Target Affiliation* (more on that below).

1. **'Precise'**: When the specific size of an event (i.e. how many incidents the event represents) is known (e.g., reporting notes that “threats were sent to the local official five times last week”)
2. **'Approximate'**: When the approximate size of an event (i.e. how many incidents the event represents) can be deduced from reporting (e.g., reporting notes that “the commissioner received a harassing email nearly every day last month,” so the number of incidents can be deduced as nearly 30)
3. **'Unknown'**: When the approximate size of an event (i.e., how many incidents the event represents) cannot be accurately estimated and is known to be imprecise (e.g., reporting simply notes that “some officials in the state were threatened,” without any further details regarding what the number of officials could be)

Nature of Event

The nature of an event (**variable name: Nature**) captures whether a target was threatened or harassed. THD then disaggregates threats and harassments into 8 types (**variable name: Nature Type**) and 16 sub-types (**variable name: Nature Detail**). In cases where an event involves multiple types of *Nature* (e.g. 'Threat' and 'Intimidation'), multiple types of *Nature Type* (e.g. 'Threat of Death' and 'Invasion of Privacy'), or multiple types of *Nature Detail* (e.g. 'Gun Violence' and 'Home Demonstration'), the multiple relevant categories may be coded under the relevant variable(s), separated by semicolons. The table below denotes these categories and how they are nested:

Nature	Nature Type	Nature Detail
Threat	Threat of Death	Gun Violence
		Other Death Threat
		Unknown Death Threat
	Threat of Physical Harm or Violence	Sexual Assault
		Other or Unknown Threat of Physical Harm or Violence
	Other Threat	Other Threat
	Unknown Threat	Unknown Threat
Harassment	Invasion of Privacy	Home Demonstration
		Targeting of Others
		Doxing
		Stalking or Following
		Other or Unknown Invasion of Privacy

	Intimidation	Presence of Weapons
		Other or Unknown Intimidation
	Other Harassment	Other Harassment
	Unknown Harassment	Unknown Harassment

Threat

A **‘Threat’** refers to when a perpetrator communicates to a target their intention to harm them, either through physical means (e.g. threat of inflicting pain, injury, or death) or via damage or other hostile action (e.g. threat of legal action), at least in part due to that person’s role as a local, elected, appointed, or municipal official. A ‘Threat’ can be anything designed to be threatening, and that might reasonably cause the person receiving the threat to fear for their own safety or the safety of their loved ones. Only threats that are directed at a specific local, elected, appointed, or municipal official are included here, even if there is no intent of actually carrying it out.⁶ A threat need not be considered legally criminal for inclusion within THD.⁷

Threats can take on 2 primary forms (**variable name: *Nature Type***):⁸

1. **‘Threat of Death’**
2. **‘Threat of Physical Harm or Violence’**

In cases where another type of threat is issued, beyond these 2 primary forms, the *Nature Type* is coded as **‘Other Threat’** (e.g. threat of property damage, arrest, etc.).

In cases where reporting suggests that a target was threatened, yet the nature of the threat is not clear, the *Nature Type* is coded as **‘Unknown Threat.’**

Threat of Death

⁶ While Government Official John could consider violence against School Official Jane to be threatening to *all* officials and hence threatening to him, violence against School Official Jane is not included as a threat here given it is an *indirect* threat against Government Official John. Further, given the *occurrence* of violence moves beyond the *threat* of violence, the violence against School Official Jane would also not be included within THD. Other data projects tracking *incidents* of political violence — rather than *threats* of political violence — would be more appropriate for users interested in tracking such violence.

⁷ THD includes incidents of threats (and harassment) beyond those that are considered legally criminal. Legally defined threats and harassment often are judged on the extent to which an individual *knowingly* and *intentionally* engaged in such a manner (for example, see 18 U.S. Code § 876). While legally criminal threats and harassment are important, they are narrowly defined, with THD taking a broader approach.

⁸ Results from piloting and earlier iterations of data collection helped to shed light on the most appropriate sub-categories to be included here — as well as for other variables. Ensuring that sub-categories that are not only substantively salient, yet also for which reporting allows the level of disaggregation, helps to ensure analytical utility while also minimizing the risk of few data points (for under-represented sub-categories) fueling weak extrapolations.

A ‘Threat of Death’ refers to a statement or indication of the intention to inflict harm to another person, such that the harm would lead to the target’s death (e.g. statements like “I will kill you”), or a statement or indication that harm to another person should be inflicted, such that the harm would lead to the target’s death (e.g. statements like “You deserve to die”).

One specific type of death threat — by ‘**Gun Violence**’ — is coded distinctly (**variable name: Nature Detail**). Firearms can result in a chilling effect around freedom of speech and public service through their credible signaling of violence. Further, in review of initial data collection, threats of death were most commonly issued as involving gun violence, especially given the high rates of US gun ownership, which is why they are considered specifically here. A death threat by ‘Gun Violence’ refers to a statement or indication that specifies an intent to inflict death via firearms, such as by carrying out a fatal shooting (e.g. statements like “I’ll blow your brains out.”).

All other death threats are coded as ‘**Other Death Threat**’ (e.g. public execution, hanging, etc.) or ‘**Unknown Death Threat**’ when a death threat was issued, but the details are unknown.

Threat of Physical Harm or Violence

‘Threats of Physical Harm/Violence’ refer to statements or indications of the intention to physically harm another person’s body, other than death, which is captured above (e.g. statements like “I’ll hurt you” or “I’ll tie you up.”).

One specific type of threat of physical harm — ‘**Sexual Assault**’ — is coded distinctly (**variable name: Nature Detail**). In review of initial data collection, women officials were disproportionately the targets of threats and harassment, which is why such a gendered form of threats (sexual assault) are considered specifically here. A threat of physical harm or violence by ‘Sexual Assault’ refers to a statement or indication that specifies an intent to sexually touch another person through coercion and without their consent (e.g. statements like “I’ll rape you.”).

All other threats of physical harm are coded as ‘**Other or Unknown Threat of Physical Harm or Violence**’ (e.g. hurt, kidnap, etc.).

Harassment

‘**Harassment**’ refers to when a perpetrator puts aggressive pressure on a person or intimidates another person at least in part due to that person’s role as a local, elected, appointed, or municipal official. ‘Harassment’ can be knowing and willful aggressive unwanted conduct directed at a specific person that a reasonable person would consider alarming, tormenting, or terrorizing without serving legitimate purpose. Harassment need not be considered legally criminal for inclusion within THD.

Harassment can take on 2 primary forms (**variable name: Nature Type**):

1. ‘**Invasion of Privacy**’
2. ‘**Intimidation**’

In cases where another type of harassment occurs, beyond these 2 primary forms, the *Nature Type* is coded as **‘Other Harassment’** (e.g. letter mailed to an official containing racial slurs; defacement of property or damage).

In cases where reporting suggests that a target was harassed, yet the nature of the harassment is not clear, the *Nature Type* is coded as **‘Unknown Harassment.’**

Invasion of Privacy

‘Invasion of Privacy’ refers to the unjustifiable intrusion into the personal life of another person, without their consent. Four specific types of invasion of privacy are considered (**variable name: *Nature Detail***):

1. **‘Home Demonstration’**: Demonstrations taking place at, outside of, or in front of the home of a local, elected, appointed, or municipal official
2. **‘Targeting of Others’**: Targeting of those close to a local, elected, appointed, or municipal official (e.g. their family including children, friends, neighbors, etc.).
3. **‘Doxing’**: Evidence of searching for and publishing private, sensitive, and identifiable information of another person without their consent
4. **‘Stalking or Following’**: A pattern of repeated and persistent unwanted contact or following

In regards to the first sub-category — ‘Home Demonstration’ — all demonstrations held at the homes of local, elected, appointed, or municipal officials are included here, regardless of whether such demonstrations remain peaceful or turn violent.⁹ Demonstrators that wish to congregate in order to be heard by an official directly can protest at an official’s place of work — meaning that such protesters would not be deprived of their ability to engage in democratic discourse via protests. As such, ordinances meant to keep officials free from such harassment, for example, would not result in a chilling effect over all protests.¹⁰

The one exception to this rule is during lockdown periods (e.g. COVID pandemic-related lockdowns), where demonstrators do *not* have the option of protesting at an official’s place of work, since such places of work are closed during lockdown periods. In such cases, ordinances meant to keep officials free from such harassment, for example, *would* result in a chilling effect over all protests.¹¹ During lockdown periods, home demonstrations are hence considered to be harassing (and hence *not* included in the dataset) only if:

⁹ The THD focuses on centering the experience and perceptions of officials, and feedback shared by officials during in-depth interviews has pointed to how officials view anyone, including demonstrators, showing up at their personal homes to be harassing, especially given the effect that it can have on their families, including small children.

¹⁰ Users who do not wish to include home demonstrations in their analysis (because they do not necessarily meet their threshold for harassing activity) can easily identify such events for exclusion from their analysis by filtering on the ‘Home Demonstration’ sub-category. Users who do not wish to include *peaceful* home demonstrations in their analysis (because they do not feel such events are necessarily harassing) can easily identify such events for exclusion from their analysis by searching for a “[peaceful home demonstration]” tag (stylized as such, in brackets) within the *Summary* variable (introduced later in this Codebook); these represent demonstrations without any reports of violence, vandalism, destruction of property, etc.

¹¹ For more on home demonstrations, see [this BDI Issue Brief](#) titled: *Trends in Demonstrations at Homes, May – December 2020*.

- They took place *outside of ‘work hours’* (i.e. early in the morning, late at night, on weekends) as local officials have an expectation of having a peaceful home life
- They involved *extremist groups* that are considered to be prone to violence (given their history of violence and/or intention of use of violence) and hence may seem particularly harassing
- They involved *threatening or harassing behavior* (e.g. destruction of property/vandalism, trespassing) since such behavior at a protest at an official’s place of work also would have been coded (i.e. considered threatening or harassing)

All other forms of invasion of privacy are coded as ‘**Other or Unknown Invasion of Privacy**’ (e.g. trespassing at the home of an official).

Intimidation

‘**Intimidation**’ refers to the inducement of fear to compel change (e.g. statements like “You better watch your back...”). One specific type of intimidation — by ‘**Presence of Weapons**’ — is coded distinctly (**variable name: Nature Detail**).

Intimidation via ‘Presence of Weapons’ refers to the presence of objects that are designed or used for inflicting bodily harm (e.g. bringing firearms, knives, or other objects intended to inflict harm to a demonstration).

All other intimidation is coded as ‘**Other or Unknown Intimidation**’ (e.g. forcefully entering an official’s work office in a restricted area in an intimidating manner).

Method of Event

The method of an event captures the means through which a threat or harassment occurs. THD disaggregates event method into 3 types (**variable name: Method Type**) and 8 sub-types (**variable name: Method Detail**). In cases where an event involves multiple types of *Method Type* (e.g. ‘Physically Removed or Virtual’ and ‘Physically Present’), or multiple types of *Method Detail* (e.g. ‘Verbal in Person’ and ‘Demonstration’), the multiple relevant categories may be coded under the relevant variable(s), separated by semicolons. The table below denotes these categories and how they are nested:

Method Type	Method Detail
Physically Removed or Virtual	Virtual Medium
	Non-Electronic Medium
	Audio Medium
	Other or Unknown Physically Removed or Virtual
Physically Present	Demonstration
	Verbal in Person

	Other or Unknown Physically Present
Unknown Method	Unknown Method

In cases where reporting suggests that an official was threatened or harassed, yet the method of how that event occurred is not clear, the *Method Type* is coded as ‘**Unknown Method.**’

Physically Removed or Virtual

Threats or harassment that are executed through a physically removed means, or through virtual means, are those in which the perpetrator(s) is not in the physical presence of the target(s) at the time of the event. Three primary types of ‘Physically Removed/Virtual’ events are considered (**variable name: Method Detail**):

1. ‘**Virtual Medium**’: A threat or form of harassment that is conveyed in an electronic format, whether written or otherwise, which involves something virtually reaching the official (e.g. email containing words and/or photos; social media post containing words and/or photos)
2. ‘**Non-Electronic Medium**’: A threat or form of harassment that is conveyed in a non-electronic format, whether written or otherwise, which involves something physically reaching the official (e.g. letter containing words and/or photos; postcard; a ‘gift’ left at an official’s home or office)
3. ‘**Audio Medium**’: A threat or form of harassment that is conveyed in a verbal fashion without the perpetrator(s) being present (e.g. phone call; voicemail)

Of course, not all forms of electronic correspondence (e.g. emails), non-electronic correspondence (e.g. letters), nor verbal engagement from afar (e.g. phone calls) are considered to be threatening or harassing. Individuals have a right to voice their concerns to local, elected, appointed, or municipal officials, as a part of a healthy democracy. In cases where this is done in a peaceful or non-threatening way, such incidents are *not* included here. Cases in which such correspondence includes a threat, or in which the manner is meant to put pressure or intimidate an official, *are* included here (e.g. an email including a threat to kill an official).

All other methods of issuing threats or harassment, done in a physically removed or virtual fashion, are coded as ‘**Other or Unknown Physically Removed or Virtual**’ (e.g. mailing an official an orange prison jumpsuit).

Physically Present

Threats or harassment that are executed when a perpetrator(s) *is* in the physical presence of the target(s) at the time of the event are categorized into two primary types of ‘Physically Present’ events (**variable name: Method Detail**):

1. ‘**Demonstration**’: A public expression of objection, disapproval, or dissent towards an idea, action, or person, regardless of whether it remains peaceful or turns violent (e.g. protest; riot)

2. **'Verbal in Person'**: A threat or form of harassment that is conveyed in a verbal fashion while the perpetrator(s) is physically present (e.g. yelling outside of an official's office; town hall)

Again, of course, not all demonstrations nor verbal engagements in person (e.g. town hall engagement) are considered to be threatening or harassing. Individuals have a right to voice their concerns to local, elected, appointed, or municipal officials, as a part of a healthy democracy. In cases where this is done in a peaceful or non-threatening way, such incidents are *not* included here. Cases in which such correspondence includes a threat, or in which the manner is meant to put aggressive pressure or intimidate an official, *are* included here (e.g. a peaceful protest including signs depicting racial slurs against an official; a violent riot outside the home of an official; a town hall at which extremist groups are present with the intent of intimidating officials).

All other methods of issuing threats or harassment, done in a physically removed or virtual fashion, are coded as **'Other or Unknown Physically Present'** (e.g. destruction of an official's property/vandalism).

Actors

Events in THD are dyadic; as such, each has an associated target and perpetrator, both of which are coded by THD. While details regarding both the targets and perpetrators of threats and harassment against officials are captured — such as what type of official was targeted and their gender; or what the affiliation of a perpetrator may be, and whether it is an extremist group — more personal details, such as the names of such individuals, are not included within the dataset. This is in line with BDI's effort to 'Do No Harm', by both minimizing a further spotlight on officials already be at heightened risk of being a continued target of a threat or harassment, as well as avoiding inadvertently elevating the actions of individual perpetrators who may enjoy further notoriety and support from other like-minded individuals.

Target of Event

The individual or group that is targeted by a threat or by harassment is coded as the target within each event. For targets, THD captures a target's affiliation, gender, and how many targets the event represents in cases where reporting is not granular enough to allow for adequate disaggregation.

Target Size and Fuzziness

As the unit of analysis within THD is an event, involving a single target, research assistants strive to find the most granular and detailed information possible so as to disaggregate reporting accordingly. In some cases, however, this might not be possible to do. For example, there may be cases where several events should be coded (e.g. *several* officials have been targeted) yet reporting is not detailed enough to allow for disaggregated coding (e.g. no sufficient reporting on who each of the many officials is, meaning individual events cannot be coded). In such cases, a single event is coded, though the (estimated) size size of the event's target(s) (i.e. roughly how many targets the single event is representing) is documented (**variable name: Target Size**):

1. **'Single'**: Event represents a single target
2. **'Small'**: Event represents a small number of targets (i.e. ~ less than 10 targets)
3. **'Large'**: Event represents a large number of targets (i.e. ~ more than 10 targets)
4. **'Unknown'**: The number of targets that the event represents cannot be approximated from reporting

Further, an additional variable (**variable name: Target Size Fuzziness**) helps to capture how accurate this estimated size (i.e. the number of targets) may be:

1. **'Precise'**: When the specific number of targets of an event is known (e.g. reporting notes that “threats were sent to 5 local officials”)
2. **'Approximate'**: When the approximate number of targets of an event can be deduced from reporting (e.g. reporting notes that “the commissioner of most counties in the state were sent a harassing email,” so the number of targets can be approximated by knowing the number of counties in the state)
3. **'Unknown'**: When the approximate number of targets of an event cannot be accurately estimated and hence is known to be imprecise (e.g. reporting simply notes that “some officials in the state were threatened,” without any further details regarding what the number of officials could be)

Target Affiliation

Target Affiliation specifies the position affiliation of the local, elected, appointed, or municipal official that was threatened or harassed.

Affiliations are divided into 4 primary categories:

1. **'Elected or Appointed Government Official'**: Officials who have been elected by the public; appointed by an executive or legislative body to an office, agency, department or other institution; or anyone authorized to act on behalf of a municipality (e.g. mayors; council members; board members; supervisors; city, town, or county commissioners; city or county administrators or managers; other local government workers)¹²
2. **'School Official'**: Educational authorities at the local-level (e.g. members of school boards, boards of education, boards of directors, or school committees; school district board presidents; school principals or superintendents¹³)
3. **'Election Official or Poll Worker'**:¹⁴ Election authorities and workers, including those who have been elected, appointed, or hired (whether paid or unpaid) (e.g. clerks or assistant clerks; poll managers or assistant managers; elections administrators or administrative assistants; equipment operators or technicians; election inspectors, deputies, commissioners, judges, directors or supervisors, poll workers, or election volunteers; elections chairs)

¹² While most judicial officials are elected or appointed, they are grouped separately in the data under 'Judicial Official' rather than here under 'Elected or Appointed Government Official.'

¹³ School principals or superintendents are appointed by school boards or other school district entities and are hence included here.

¹⁴ As noted above, election workers and/or volunteers *working in an official capacity* are included here, given that their role in ensuring democratic processes aligns in spirit with the role of local, elected, appointed, or municipal officials. Partisan poll watchers are *not* considered to be working in an *official* capacity, and are *not* included here.

4. **'Judicial Official':** Officials at the municipal, city, county, or lowest level of court structures (i.e. county judges, magistrates, county prosecutors, district attorneys)

In cases where multiple officials may have been targeted in a single event (i.e. cases where *Target Size* is *not* 'Single'), the event is disaggregated by target affiliation, if relevant (e.g. if both school officials *and* judicial officials are targeted, 2 events are coded: one in which the target affiliation is 'School Official,' and the other where the target affiliation is 'Judicial Official'). In unique cases where a single targeted official has *multiple* affiliations (e.g. a school official who is *also* a judicial official), then all relevant categories are coded under *Target Affiliation*, separated by semicolons.

In cases where a target has a different affiliation, beyond these 4 primary forms, the *Target Affiliation* is coded as **'Other Target Affiliation'** (e.g. city, town, or county health officials; law enforcement, etc.). Motivating issues may change over time, with new issues often emerging, in turn motivating threats and harassment of various officials. As such, the THD team regularly monitors the 'Other Affiliation' category here, so as to identify when new types of officials are systematically targeted; if/when such trends emerge, such new categories are added to the data, with events back-coded and updated to reflect such changes.

In cases where reporting suggests that a local, elected, appointed, or municipal official target was threatened or harassed, yet the affiliation of the target is not clear, the *Target Affiliation* is coded as **'Unknown Target Affiliation.'**

Target Gender

Target Gender captures the gender of the official at whom the threat or harassment is targeted (based on the self-reported gender of the individual¹⁵). This variable is divided into 3 primary categories:

1. **'Male'**
2. **'Female'**
3. **'Other Target Gender'**

'Other Target Gender' includes cases where a target may identify as non-binary or some other categorization other than male or female. Targets described as 'transgender' without any indication of their gender identity are also included here; in cases where a target is described as 'transgender male' or 'transgender female,' the target is categorized per their gender identity (i.e. male or female, respectively). Details, such as them being transgender, would be captured within the event *Summary*.

In cases where reporting does not offer any distinction regarding the gender of a target, *Target Gender* is coded as **'Unknown Target Gender.'**

In cases where an event represents more than one individual (i.e. when *Target Size* is *not* coded as 'Single'), if the gender make-up of the group is known to be all male or all female,

¹⁵ In cases where a target may have been misgendered in reporting by journalists or in police reports, such inaccurate gender categorizations are not used in coding.

Target Gender is coded as ‘Male’ or ‘Female,’ respectively. In cases where the gender make-up of the group is known to be mixed, *Target Gender* is coded as ‘Other Gender.’ In cases where the gender make-up of the group is unknown, *Target Gender* is coded as ‘Unknown Target Gender.’

Perpetrator of Event

The individual or group that threatens or harasses a target is coded as the perpetrator within each event. For perpetrators, THD captures the perpetrator’s affiliation (**variable name: *Perpetrator Affiliation***), and whether or not the perpetrator is part of or affiliated with an extremist group (**variable name: *Extremist Group***).

Perpetrator Affiliation

Perpetrator Affiliation specifies the affiliation of the individual or group that threatened or harassed a local, elected, appointed, or municipal official, in cases where they have such an affiliation. The affiliations that THD captures are groups that are named and organized (i.e. a group that is assembled for more than a single incident) towards a political purpose. Such groups include, but are not limited to, paramilitary groups, armed community defense groups, hate or extremist groups, as well as entities like social movements, work-related groups, or unions, to name a few. Once a group appears within THD, its name appears consistently within the dataset, allowing for monitoring of patterns of behavior over time.

In cases where multiple groups may have been present or involved in a single event, all relevant groups are coded under *Perpetrator Affiliation*, separated by semicolons.

In cases where the perpetrator(s) are not part of a named and organized group (e.g. parents, Democrats, Trump supporters),¹⁶ or where no information is known about a perpetrator (e.g., “A letter was sent to an official...”), *Perpetrator Affiliation* is coded as ‘**Unknown or No Perpetrator Affiliation.**’ Descriptive information (e.g., whether they are parents, Democrats, or Trump supporters) is noted in the *Summary* (see below).

Event Description

For each event, 3 substantive components are captured: the type of targeting that a threat or harassment involves (**variable name: *Type of Targeting***), the specific issue(s) at stake behind a threat or harassment (**variable name: *Issue at Stake***), and a summary of the event itself (**variable name: *Summary***).

Type of Targeting

Type of Targeting captures whether specific types of language or symbols (e.g. hateful language or symbols) were used as part of a threat or harassment.

Four primary types of hateful language or symbols are captured:

¹⁶ In this example, parents are not a cohesive group (i.e. they are not a monolith); Democrats, including those who happen to vote Democrat, are not an organized group as anyone can vote Democrat; Trump supporters might share a similar goal (supporting Trump), but they are not an organized entity as anyone can support Trump.

1. **'Racist'**: When the race or ethnicity of the target is evoked in targeting (e.g. use of racial slurs; symbolism associated with hate towards specific races, such as lynching imagery; ethnic slurs)
2. **'Sexist'**: When the gender of the target is evoked in targeting (e.g. use of misogynistic language)
3. **'Anti-LGBTQ+'**: When the (perceived) sexual orientation and/or gender presentation of the target is evoked in targeting (e.g. use of homophobic slurs or transphobic language)
4. **'Anti-Religious'**: When the religion of a target is evoked in targeting (e.g. use of anti-Semitic or Islamophobic language)

In cases where multiple forms of hateful language or symbols are used in a single event, all relevant forms of hateful language or symbols are coded under *Type of Targeting*, separated by semicolons.

In cases where other forms of hateful language or symbols are used, outside of the categories noted above, *Type of Targeting* is coded as **'Other Targeting.'**

In cases where multiple types of hateful language/symbols may have been used in a single event, all relevant categories are coded under *Type of Targeting*, separated by semicolons.

In cases where the specific language/symbols used by a perpetrator are not known, or if no specific hateful language is used (e.g., reporting only says that a threat was issued, without any context on the substance of that threat), *Type of Targeting* is coded as **'Unknown Targeting or No Targeting Language.'**

Issue at Stake

What type of issue (i.e. grievance) may be motivating a threat or harassment is captured within *Issue at Stake*. To allow for meaningful analysis for users, 4 primary categories of issues at stake are captured:

1. **'Elections'**: Grievances associated with elections specifically (e.g. fraud concerns)
2. **'Education'**: Grievances associated with school administration, curriculum, or policies (e.g. teaching of Critical Race Theory; school mask policies)
3. **'Hyperlocal'**: Grievances associated with decisions made by local government bodies or associated with community specific issues (e.g. local jobs, resolutions passed by local officials, reactions to local criminal activity)
4. **'Personal Legal'**: Grievances associated with legal decisions, arrests, or other elements of the criminal justice system experienced on a personal level (e.g. family court decisions, trials)

In cases where multiple grievances may be motivating a threat or harassment, all relevant issues are coded under *Issue at Stake*, separated by semicolons.

All other issues at stake are coded as **'Other Issue at Stake.'** Specific motivating issues come and go, changing over time; new issues often emerge, motivating threats and harassment of local, elected, appointed, or municipal officials, or are branded as such. To ensure that THD captures these ebbs and flows, the THD team regularly monitors the 'Other' category; as new salient issues emerge (i.e. issues at stake that appear to regularly

motivate threats or harassment against local, elected, appointed, or municipal officials), they are added as a new category, with events back-coded and updated to reflect such changes.

In cases where multiple types of issues may have been at stake in a single event, all relevant categories are coded under *Issue at Stake*, separated by semicolons.

In cases where the types of issues that may have been at stake are not known (e.g. reporting only says that a threat was issued, without any context on the substance of that threat), *Issue at Stake* is coded as **'Unknown Issue at Stake.'**

Summary

The event *Summary* offers a short summary of the main features of an event, specific details about an event based on variables noted above (e.g. what specific grievance motivated an event coded with 'Education' as one of its *Issue at Stake*), as well as any additional important details not captured by other variables (e.g. demographic information about the perpetrator or target, etc.). More granular details about the target of the event (e.g. role/title, name, etc.) are not included here, in order to ensure enough anonymity to 'Do No Harm.'

Temporal

For each event in the dataset, 2 temporal components are coded:

1. The date that an event took place (**variable name: *Date***)
2. The level of confidence in the accuracy of the date coded (**variable name: *Date Fuzziness***)

The *Date* of an event refers to the day, month, and year on which it took place (*not* to be confused with the date that an event was *reported*). It appears in the data structured as YYYY-MM-DD.

In cases where *Event Size* is *not* coded as 'Single' because a single coded event represents numerous incidents (as a function of reporting not allowing for further event disaggregation), the *Date* refers to the date on which the first incident within the string of events occurred. For example, if threats were made by the same perpetrator against the same target on *multiple* occasions, yet reporting is not detailed enough to allow for disaggregated coding (e.g. no sufficient reporting on what days/how often the threats occurred), the date of the first incident is recorded — i.e. when the targeting first began.

Date Fuzziness

While research assistants strive to identify the precise date on which an event occurred, in some cases this is not possible, as reporting associated with an event may be too vague. In such cases, the event is still coded — allowing THD to make use of vague reporting — though the fuzziness is documented by the *Date Fuzziness* variable via 3 categories:

1. **'Precise'**: When the exact date on which an event occurred is known
2. **'Approximate'**: When the approximate date that the event occurred can be deduced from reporting with relative accuracy (e.g. when reporting notes that an event occurred "last weekend" or "in mid-August")

3. **'Imputed'**: When the date that an event occurred can be estimated or imputed based on clues in reporting, though is known to be imprecise (e.g. when reporting notes that an event occurred “last month” yet it is not known whether it was the beginning, middle or end of the month; or if reporting a few months after an election notes that an event occurred “after the election,” meaning that there exists a few month window during which the event could have occurred)

N.B. Event dates with more fuzziness than that which is noted here — i.e. cases where an event date cannot be approximated or imputed within a few-month window — are not included within THD.

In cases where *Date Fuzziness* is coded as ‘Approximate’ or ‘Imputed,’ the middle date within the approximated or imputed date window is coded as the event’s *Date* (e.g. 15 August 2022 is coded if reporting suggests an event occurred in “mid-August 2022”).

In cases where a single coded event represents numerous incidents (as a function of reporting not allowing for further event disaggregation), and where the *Date* refers to the date on which the first incident within the string of events occurred (as noted above), *Date Fuzziness* is coded as ‘Approximate’ to denote that while the date of the first incident is known, the dates of the other incidents are approximated.

Geography

For each event in the dataset, 3 geographic components are coded:

1. The county where an event occurred (**variable name: County**)
2. The state where an event occurred (**variable name: State**)
3. The level of confidence in the accuracy of the location coded (**variable name: Location Fuzziness**)

Location information is coded down to the county level, even in cases where more granular information may be available. This is to ensure enough anonymity to ‘Do No Harm’ (e.g. even without target names noted, in small enough cities with few local officials, details around an event may render attempts to anonymize targets useless.¹⁷

In cases where *Method Type* is coded as ‘Physically Present,’ the target(s) and perpetrator(s) of the event are in the same location, which is the geographic information that is coded. However, in cases where *Method Type* is coded as ‘Physically Removed/Virtual,’ the target(s) and perpetrator(s) of the event are *not* in the same location; in such cases, the geographic information coded refers to the location of where the event occurred *in relation to the target*, given the target-centered lens that THD employs. For example, if an individual in Philadelphia, PA, threatens a school board official in Trenton, NJ, the geographic information coded would be in reference to Trenton, NJ, where the school board official is located, *not* Philadelphia, PA, where the perpetrator is located.

Location Fuzziness

¹⁷ Users interested in more granular information beyond the county-level — i.e. down to the city-level — can reach out to BDI to discuss the possibility of such data sharing.

While research assistants strive to identify the specific location in which an event occurred, in some cases this is not possible, as reporting associated with an event may be too vague. In such cases, the event is still coded — allowing THD to make use of vague reporting — though the fuzziness is documented by the *Location Fuzziness* variable via 2 categories:

1. **‘Precise’**: When the specific city location of an event is known
2. **‘Approximate’**: When the specific city location of an event is not known, though approximated

For events coded with *Location Fuzziness* as ‘Approximate’: if the specific city in which an event occurred is not known yet the county is known, the county seat of the county (i.e. where county-level government buildings and county government leaders tend to work) is coded as the location; if the county seat of the county is also unknown yet the state is known, the capital of the state (i.e. where state-level government buildings and state government leaders tend to work) is coded as the location. *N.B. Event locations with more fuzziness than that which is noted here — i.e. cases where an event location cannot be approximated beyond the state-level — are not included within THD.*

Sources of Information

THD collects information from 3 types of sources (**variable name: Source Type**), explored in further detail in the following subsections. THD also offers more specific details about each source (**variable name: Source Detail**). The table below denotes these 3 categories, along with a brief description and an indication of the type of more detailed information that is offered regarding each.

Source Type	Brief Definition	Source Detail Information
Traditional Media	<i>Secondary information via traditional media</i>	Name of traditional media outlet
BDI OSINT	<i>Information from trusted social media accounts, monitoring conducted by BDI</i>	Name of social media platform; name of partner organization
Consortium	<i>Data from BDI partner organizations</i>	Name of partner organization

In future versions of this data collection, self-reported incidents may be considered for inclusion.

Every event is based on at least one source; in cases where more than one source is used in coding, all sources (both *Source Type* and *Source Detail*) are listed, separated by a semicolon. *N.B. The number of sources listed is not equivalent to the number of sources in existence reporting on a specific event; the latter is outside of the scope of what can be*

*captured by THD. Also important to note is that more sources should not be equated to more reliability.*¹⁸

Traditional Media

The core of THD information collection comes from traditional media. [Nexis Uni](#) — an academic search engine, bringing together a variety of local, regional, national and international newspapers, magazines, online journals, television and radio broadcasts, newswires and blogs, transcripts, and legal research, as well as federal and state cases and statutes — is a helpful tool, used by research assistants in their information collection, offering an initial basis of information.

To ensure that local news (e.g. newspapers, TV, radio) is well-integrated — especially given the increased likelihood of such local outlets to report on incidents of threats and harassment against officials — unique, state-specific sourcing profiles have also been developed. For each state (and the District of Columbia), 8 additional state-specific news sources are identified and regularly reviewed by research assistants. Going forward, BDI plans to code additional state-specific news sources per state, to be added after events stemming from the source have been back-coded (to ensure artificial spikes are not introduced into the dataset).

Beyond its work on THD, BDI engages in regular open-source intelligence (OSINT) monitoring, which helps to inform its work around action-oriented, responsive research; enabling cross-sector collaboration; and driving policy and community response. As a part of this work, research assistants monitor a variety of publicly available sources, including traditional and social media. Information *stemming from traditional media* that captures incidents of threats against or harassment of officials is also integrated here.

BDI OSINT

In addition to traditional media (and monitoring work which yields traditional media stories), information uncovered via OSINT monitoring by BDI is also integrated into THD. Such OSINT monitoring also draws on information coming from social media platforms — which is subject to vetting prior to integration. A diverse sample of social media accounts are followed, which include, but are not limited to, local elected officials and reporters, government agencies, law enforcement, and other relevant individuals — which together help to capture entities that are ‘in the know’ within specific contexts.

Consortium

Collaboration is core to BDI's approach to data collection and research. As discussions with the broader US research sector continue, BDI may add new or additional consortium partners; references to partners will be added to this codebook to indicate new sources of collaboration on data sharing or self-reported incident collection.

¹⁸ For more on this, and what factors contribute to more reliability in sourcing, see Miller, Kishi, Raleigh, and Down (2022).

The [Center on Extremism at Anti-Defamation League \(ADL\)](#) tracks extremist trends in the United States, including information on incidents of threats and harassment targeting officials across the country. Information from ADL is especially helpful in identifying instances of anti-religious types of targeting.

The [Atlantic Council's Digital Forensic Research Lab \(DFRLab\)](#) tracks disinformation, fake news stories, and other incidents that involve attempts against democracy. Information from DFRLab is helpful in shedding light on such activities.

The [Brennan Center for Justice](#) provides research and policy recommendations on election-related issues, among other issue areas. As part of their monitoring and research efforts, the Brennan Center provides information on threats, harassment, and resignations of election officials across the United States, which helps to complement information gathered by other THD sources.

The [Armed Conflict Location and Event Dataset \(ACLED\)](#) tracks data on political violence, demonstrations, and important strategic developments around the world, including in the United States. Information collected by ACLED comes from traditional media, reports, their local partners, and targeted and verified new media. THD uses data from ACLED to identify incidents of demonstrations (e.g. demonstrations at the homes of officials; demonstrations involving threats against officials), as well as specific strategic developments (e.g. extremist groups flyering the house of an official). Such incidents, especially the latter, are often under-reported by traditional media.

The [Crowd Counting Consortium \(CCC\)](#) records data on demonstrations across the United States. Information collected by CCC comes from publicly available sources. THD uses data from CCC to identify incidents of demonstrations (e.g. demonstrations at the homes of officials; demonstrations involving threats against officials) that may not have been tracked elsewhere.

[The Prosecution Project \(tPP\)](#) collects information on felony criminal cases involving political violence in the United States. Information collected by tPP comes from primary sources such as court documents, government press releases, reports, as well as secondary sources, such as news or academic articles, for triangulation. THD uses data from tPP to identify cases of threats or harassment that resulted in prosecution. Such information from tPP helps to offer more detail around the (arguably) most contentious cases.

Dataset Management

The THD is publicly updated on a monthly basis to include: (1) events that occurred in the previous calendar month; (2) events that occurred at any point within the inclusion dates for which information is newly available (January 2022 – present day, month prior); or (3) to make updates to or to delete existing events (based on newly released details, revisions to variables, etc.). The THD notes when events are no longer included in the data collection (included in additional *Deleted Data ID* sheet within the data file) and when new events are added and/or changes are made to existing events (variable name: **Last Updated**).

Deleted Events

Events may be *removed* from the THD if new information becomes available that indicates that the event does *not* meet THD’s inclusion criteria, or if the THD research team finds additional reason(s) to remove the event from the collection (e.g. duplicate coding, aggregating events). Deleted IDs should be removed from local files to ensure up-to-date data.

Deleted events are coded separately from the main collection *on the* ‘Deleted Data ID’ sheet of the full data file, noting both the respective **Data ID** and the **Month Deleted**, following the format of “Year - Month” using numerics (e.g. 2024-05 to note removal during the May 2024 data release). The **Data ID** of deleted events will not be reassigned to new events in the dataset.

Tracking Updates

The **Last Updated** variable is coded for each event with the month and year that the information was last updated, following the format of “Year - Month” using numerics (e.g. 2024-05). The month coded coincides with the monthly update in which the changes were made public (e.g. 2024-05 means that the event was last updated during the May 2024 release).

This variable coding is intended to help users track updates and additions to the dataset across data releases so that important changes to data points are not missed. Previous data downloads will become out of date and historical data may change as the THD incorporates new information. In other words, merely appending events from May 2024 to a local data file following that month’s release, for example, may mean that a user’s local data file is out of date, as *newly coded* events occurring between January 2022 and April 2024 will be missed, as will any *updates* to events already recorded between January 2022 and April 2024.

To ensure an up-to-date data file, users can either:

- Re-download the full THD data file (covering January 2022 to the latest available month); or
- Update their previously downloaded local datafile (which may be of interest to users who have added any additional coding for personal analysis). To do this, users should:
 - [Download a full THD data file](#) from the [BDI website](#)
 - Filter by the **‘Last Updated’** variable to identify new or updated events
 - Append any new events to their local data file (i.e. those with a new **Data ID**)
 - Replace any updated events in their local data file (i.e. replacing those without a new **Data ID**, per this identifying variable)
 - Delete any events that should no longer be in the dataset, by identifying the relevant **Data ID** on the **‘Deleted Data ID’** sheet, and removing relevant events

THD Variables and Categories/Structure

Type of Event Information	Variable Name	Categories (bulleted) or Structure (<i>italics</i>)
Descriptive	Data ID	<i>Unique numerical value</i>
	Event Size	<ul style="list-style-type: none"> · Single · Small · Large · Unknown
	Incident Fuzziness	<ul style="list-style-type: none"> · Precise · Approximate · Unknown
Temporal	Date	<i>Year, month, date (YYYY-MM-DD)</i>
	Date Fuzziness	<ul style="list-style-type: none"> · Precise · Approximate · Imputed
Nature	Nature	<ul style="list-style-type: none"> · Threat · Harassment
	Nature Type	<ul style="list-style-type: none"> · Threat of Death · Threat of Physical Harm or Violence · Other Threat · Unknown Threat · Invasion of Privacy · Intimidation · Other Harassment · Unknown Harassment
	Nature Detail	<ul style="list-style-type: none"> · Gun Violence · Other Death Threat · Unknown Death Threat · Sexual Assault · Other or Unknown Threat of Physical Harm or Violence · Other Threat · Unknown Threat
Geography	County	<i>Name of county (standardized)</i>
	State	<i>Name of state (standardized)</i>
	Location Fuzziness	<ul style="list-style-type: none"> · Precise · Approximate
Actors	Target Size	<ul style="list-style-type: none"> · Single · Small · Large · Unknown
	Target Size Fuzziness	<ul style="list-style-type: none"> · Precise · Approximate · Unknown

	Target Affiliation	<ul style="list-style-type: none"> · Elected or Appointed Government Official · School Official · Election Official or Poll Worker · Judicial Official · Other Target Affiliation · Unknown Target Affiliation
	Target Gender	<ul style="list-style-type: none"> · Male · Female · Other Target Gender · Unknown Target Gender
	Perpetrator Affiliation	<ul style="list-style-type: none"> · <i>Name(s) of group(s) (standardized)</i>
Method	Method Type	<ul style="list-style-type: none"> · Physically Removed or Virtual · Physically Present · Unknown Method
	Method Detail	<ul style="list-style-type: none"> · Electronic Medium · Non-Electronic Medium · Audio Medium · Other or Unknown Physically Removed or Virtual · Demonstration · Verbal in Person · Other or Unknown Physically Present · Unknown Method
Substantive	Type of Targeting	<ul style="list-style-type: none"> · Racist · Sexist · Anti-LGBTQ+ · Anti-Religious · Other Targeting · Unknown or No Targeting Language
	Issue at Stake	<ul style="list-style-type: none"> · Elections · Education · Hyperlocal · Personal Legal · Other Issue at Stake · Unknown Issue at Stake
	Event Summary	<i>Brief qualitative description</i>
Sources	Source Type	<ul style="list-style-type: none"> · Traditional Media · BDI OSINT · Consortium
	Source Detail	<i>Name(s) of source or partner (standardized)</i>
Dataset Management	Last Updated	<i>Year, month (YYYY-MM)</i>



Bridging Divides Initiative

The Bridging Divides Initiative (BDI) is a non-partisan research initiative based at Princeton University that tracks and mitigates political violence in the United States.

bridgingdivides.princeton.edu

Copyright © 2024 BDI. All rights reserved.