Constructing Theory-Informed Relational and Verifiable Protest Event Data¹

Pamela Oliver, University of Wisconsin – Madison

Pamela.Oliver@wisc.edu orcid.org/0000-0001-7643-1008

Alex Hanna, Google Research

alex.hanna@gmail.com orcid.org/0000-0002-8957-0813

Chaeyoon Lim, University of Wisconsin - Madison

chaeyoon.lim@wisc.edu orcid.org/0000-0003-4527-4390

Abstract

Protest event data are both undertheorized and riddled with inconsistencies and errors.

We call for modifying standard practices in the collection and construction of protest event data from electronic archives of news sources to both improve verifiability of data and integrate theory with data construction. The key change is a relational data structure that maintains strong links between events identified by coders and the underlying textual sources. This allows study of the relation between events and news coverage of those events. An additional change is to record data on the relations among events, including both recognition of complex events involving multiple actors, dates, action forms, or locations and explicit acknowledgement of the connection of events to specific protest campaigns or episodes of contention. We present preliminary illustrative data about Black protests from these new procedures to demonstrate the value of this approach. NOTE: This is a working paper. Comments and feedback are appreciated.

Constructing Theory-Informed Relational and Verifiable Protest Event Data

Introduction

Protest event analysis is one important method for assessing variations in movement activity and dynamics across time, place, and issue. But there are two serious problems with protest event data. One is that they are riddled with inconsistencies and errors that often can be neither detected nor corrected. The other is that they are undertheorized. As Tilly (2002) reminds us, all research involves two theories, one about the phenomenon being studied and the other about how evidence gets generated about that phenomenon. The ways we collect and store data shape what can be theorized from the data, and theory should guide decisions about how to collect data. Theory tells us protests are relational: they are structured as campaigns and episodes. Further, the meaning and impact of protests depends on how they are talked about. But most protest event datasets treat events as independent both from each other and from the texts describing them.

At the same time, protest event data is subject to the replicability crisis facing much of science. There is a need for improved verifiability of protest event datasets. Relational approaches to data construction can simultaneously address the verifiability problem and better link protest data to theory. Relational approaches provide tools for capturing the relational structure of protests into complex events, campaigns, and episodes. They also provide tools for examining the mutual relations between protests as events involving people in time and physical space and as news article mentions in discursive space.

Scholars of social movements have long recognized the mutual relationship between social movements and the news coverage they receive (Gitlin 1980, Vliegthart and Walgrave

2012), but this mutuality has rarely been a focus of quantitative protest event research. Past projects have generally focused on either constructing event catalogs from news sources, or on studying news coverage of selected protests. Relational data structures (and the technology that enables them) permit both approaches to be used in the same project and brought into dialog with each other. Relational data structures can also capture the organization of events into campaigns and episodes and permit the aggregation and disaggregation of complex events. Finally, relational data structures provide better tools for quality control and verifiability of coding for events that are inherently ambiguous, complex, and relational.

This project builds on prior work and uses automated approaches to sift through the haystack of news articles to find the 1-5% relevant to protest. Like other research teams, we have concluded that human judgment is needed for accurate coding of events within articles. Our key contribution is to pay close attention to the relation between texts and events and to create data structures that permit easier error-checking. Most projects are built around the incorrect assumption of a one-to-one correspondence between articles and events. They treat “duplication”—where the same event is described in multiple articles—as an afterthought and nuisance. However, in our data, only 7% of the instances of an article mentioning a protest involved only one event per article and only one article per event: 59% of the articles mentioned multiple events and 75% of the articles mentioned at least one “duplicate” event that was also mentioned in other articles. Methodologically, this means that constructing protest event records from news articles inherently involves disentangling reports of multiple events in one article and recognizing multiple reports of the same event in different articles. Theoretically, embracing this reality opens the door to important new research questions about how news media talk about events in relation to each other.

We begin by sketching the current state of protest event studies, focusing on the key methodological challenges and issues in collecting reliable and valid protest event data. We then explain data structures and workflow principles focused on verifiability, relationality, and linking texts and events. We present illustrative preliminary results about Black protests to demonstrate the power of our relational approach. We show the theoretical importance of the huge skew in the amount of news coverage events receive. We show the value of recognizing the clustering of events into campaigns and episodes and of examining the relation between events on the ground and events in a discursive news space. We conclude with a brief discussion on the implications for future protest event studies.

Protest Event Studies

Many have recounted the history of protest event studies (e.g., Earl et al. 2004, Hanna 2016, Hutter 2014, Olzak 1989b) as it developed from compilation of event catalogs of labor strikes and political violence (e.g., Snyder and Tilly 1972, Tilly, Tilly and Tilly 1975) to include protests and demonstrations and attention to rigorously defining events (e.g. Olzak 1989a, Olzak 1989b, Tarrow 1988). Protest event analysis remains an active method in social movement studies¹ and protocols for collecting protest event data have become somewhat standardized. Many smaller studies report following practices summarized in a widely-cited prescriptive source (Hutter 2014), itself an update of two older sources (Koopmans and Rucht 2002, Rucht and Neidhardt 1998). Researchers based in the United States (e.g., Martin, Rafail and McCarthy 2017, Rafail 2018, Ratliff 2011, Ratliff 2013) report seeking to replicate the codebook and procedures of the Dynamics of Collective Action (DCA) project (Soule et al. n.d.), which was designed in the late 1990s.

---

¹ A bibliography of recent examples is included in the on-line appendix.

Coding protest events in news articles is laborious. The Dynamics of Collective Action project was a major undertaking by four principal investigators funded by a series of NSF grants at three institutions to code protests mentioned in the *New York Times* from 1960 to 1995. It employed many graduate students working as human coders for over a decade. Patrick Rafail's ambitious project coding local newspapers from 20 cities has required years of work. More recently, Lorenzini et al. (2021) report having 35 graduate students work 642 person-days (5136 hours assuming 8-hour days) for six months reading and coding 45,680 news stories after they were selected from the 5,251,894 that had previously been selected via multiple layers of machine pre-processing. "Easy" articles that describe only one event in a straightforward way can be coded in 5-10 minutes each and screening out articles as irrelevant can usually be done in 2-3 minutes even with a careful read to be sure there are no events lurking in the text. However, complex articles that mention multiple intertwined events can require 30 or more minutes to parse. The time to code each article increases with the amount of information to be extracted about each identified event.

There have been many efforts to automate the production of event data. There are several fully-automated data collection efforts in political science for studying international and intranational conflict events, beginning with the Kansas Event Data System (KEDS) project (Bond et al. 1997, Schrodtt and Gerner 1994)². Funding for these efforts to both academic centers and defense contractors ramped up with concerns about terrorism after 2001. The well-funded Cline Center at the University of Illinois³ has spent several decades automatically compiling a huge archive of electronic news sources from around the world and developing software to classify articles as potentially relevant to conflict events. The Integrated Crisis Early Warning

---

² KEDS is now CEDS <https://eventdata.parusanalytics.com/>

³ <https://clinecenter.illinois.edu/what-we-do/resources-and-assets>

System (ICEWS) compiled by Lockheed Martin⁴ is widely used in the defense industry as well as some academic research. There are also many computational approaches to identifying recent protest events in social media. For example, Pan and Zhang (2019) use a deep learning and image recognition framework to identify protests from the text and images posted to Weibo (the Chinese equivalent of Twitter).

With the rapid development in machine learning and natural language processing in the past two decades, automated approaches have become quite good at the “haystack” problem of identifying news articles that contain some mention of protests and at the specific problems of recognizing locations and named entities (persons, organizations) and can do these jobs at least as well as humans (Hanna 2017). However, fully automated approaches typically can only capture very sparse information about events, usually date, location, named entities such as individuals or organizations, and relevant action verbs. Machines are much less accurate than humans in recognizing whether an article is describing a current or historical event, in disentangling descriptions of multiple events in the same article, in recognizing that events may have taken place in multiple locations or in different locations from the publication location, and in identifying the same event in different articles (Althaus et al. 2017, Boschee, Natarajan and Weischedel 2013, Schrodtt 2012)..

Many research teams interested in accurately capturing the characteristics of protests have concluded that the best approaches for now are semi-automated or hybrid workflows that automate the pre-processing of articles to filter out irrelevant articles but rely on human coders at the final steps of identifying and coding events within articles. Major examples include the Mass Mobilization in Autocracies project (Croicu and Weidmann 2015, Hellmeier, Rød and

---

⁴ <https://www.lockheedmartin.com/en-us/capabilities/research-labs/advanced-technology-labs/icews.html>

Weidmann 2019, Weidmann and Rød 2018), the Cline Center’s SPEED project (Nardulli, Althaus and Hayes 2015) and the Zurich-based team interested in political contestation in Europe (Lorenzini et al. 2021, Makarov, Lorenzini and Kriesi 2016). Our own project has a similar orientation to these hybrid projects and was developed concurrently with them. Our major contribution is to explicitly focus on relational data structures and the technology enabling them, which allows us to link events to multiple articles, and to capture complex and multidimensional relations between events.

The Problem of Errors and Verifiability

Our project addresses the yoked problems of giving protest data better theoretical grounding and addressing issues of verifiability. There is increasing attention across all of science to the problem of being able to check, verify, and replicate research findings. Checks of machine-coded event data generally find that they have high rates of both false positives and false negatives. For example, Wang et al. (2016) demonstrated extreme problems of duplication and mis-identification of conflict events in two widely-used automated systems such that only a small minority (about 20%) of “events” identified by the systems actually pointed to real-world events. Schrodtt (2012) cites rates of coding agreement that range from 25% to 70%. The well-funded and highly-regarded SPEED project—which trains coders for 70 hours and tests them for accuracy before putting them into production work—reports that tested coders reliably identified 72-85% of all relevant events and accurately coded the information about these events 75-89% of the time (Hayes and Nardulli 2011). The Zurich-based team (Lorenzini et al. 2021) reports that their well-trained coders agreed on event identification 60% of the time and recorded Cohen’s Kappa scores for coding attributes in the range of .45 to .57 depending on attribute, a range they say is “fair to good.” Whether this is good enough partly depends on the research goals, but we

can safely conclude that even well-funded projects that invest heavily in coder training have non-trivial rates of inconsistencies in coding, if not outright errors.

Extracting consistent information about events from texts is difficult and error-prone for either humans or machines. Theoretically, the concept of a protest event is inherently relational and contextual involving the three dimensions of actor, action, and claim or issue. Protests often use the same action forms as non-protests and are distinguished by actors and purposes: holiday parades and protest marches involve very similar behaviors, and protest rallies are physically the same as other large gatherings with speakers. Protests are generally understood as involving intention to make a claim that promotes or resists social change and as involving people who do not already have institutional power. Is an event calling for Black men to repent and care for their families and communities a protest? Many would say not, but that was the stated purpose of the 1995 Million Man March, one of the most important Black movement events of the 1990s. Actions by elected politicians are generally not considered protests, but what about a sit-in conducted in Congress by members of the Black Congressional Caucus? Editorials may advocate for social change but are not typically considered protests.

Many teams try to obtain greater agreement about whether an event qualifies for coding with extensive definitions and rules, but there are always ambiguous cases. Lorenzini et al. (2021) report dealing with the problem of ambiguity by asking coders to look for a prescribed list of actions (rather than interpreting purposes) but still report only modest inter-coder reliability. Many teams additionally impose minimum event size constraints, even though many news articles provide no explicit size estimates. Despite these challenges, most teams ask front-line coders to make these judgments in one pass through a news article. If they do not find a relevant

event or believe all the events in the article are duplicates of events they have seen before, they are typically instructed to discard the article, which then disappears from the research stream.

Difficulties in parsing texts to retrieve events by either humans or machines have often been reported in the methodological literature (e.g., Althaus et al. 2017, Boschee, Natarajan and Weischedel 2013, Lorenzini et al. 2021, Nardulli and Hayes 2011). As Nardulli and Hayes (2011:1) say, news articles about events “can be convoluted.” It is good journalism give background and context for events, but this relational content is often intertwined with event descriptions. It is often unclear from the text exactly what happened. When multiple events are described, it can be difficult to parse which sentences refer to which event. Articles that are largely about major historical events or widely covered events like the Million Man March may also mention smaller contemporary events. Specific actions and issues often do not fit unambiguously into pre-planned coding categories. Event descriptions may be vague and provide incomplete information. Articles often contain statements that summarize multiple events, what the SPEED team calls “recapitulations” (Nardulli and Hayes 2011) and we call aggregate events, discussed below. A task this difficult cannot be done consistently and with perfect accuracy. Whether machine or human generated, event data simply cannot be assumed to be accurate or consistent if there is no way to verify it.

There are different responses to these issues. Advocates of the fully-automated systems argue that imperfect automated data are better than no data and that humans simply cannot produce real-time protest event data to serve the needs of monitoring the world for conflicts that may raise security issues (e.g., Schrodtt and Van Brackle 2013). In warning against “data fundamentalism,” Taylor (2013:23-24) makes the important point that data need to be suitable for the question being asked.

However, even high-level approximations can be distorted by a failure to recognize duplicate reports of the same news story. This became a big story in the automated event-collection world in 2014 when Mona Chalabi (2014) had to retract an analysis on the high-profile FiveThirtyEight blog about the rise in kidnapping in Nigeria that relied on the automated GDELT system. Despite its attempts to automate de-duplication, the system was counting the same Boko Haram kidnapping hundreds of times a day over multiple days (Caren 2014).

When the focus is on creating event records only and verification is about documenting that the event happened, multiple citations of the same event (“duplicates”) are an annoyance to be dealt with through de-duplication. The DCA provides one citation per event; Beissinger (2002) retains up to three citations per event. Many teams report telling front-line coders simply not to code at all events they recognize as duplicates from prior coding, including not only the older DCA project but the more recent Zurich project (Lorenzini et al. 2021), which reports discarding 27.3% of their articles because the coder recognized a duplicate event and another 5.3% of articles because duplicate events were found in articles coded by different coders.

Matching up events between different texts (“de-duplication”), is even more error prone than parsing texts for events. It is impossible to do accurately without reference to the original texts. The DCA worked post hoc to address the problem of duplicate events,⁵ but the DCA is still riddled with apparent duplicate events and errors in de-duplication. A glaring example is the DCA’s coding of the Million Man March, called by Louis Farrakhan of the Nation of Islam,

---

⁵ The data cleaning and de-duplication process for the Dynamics of Collective Action dataset is described in a document available on line:

<http://www.stanford.edu/group/collectiveaction/CLEANING%20DOCUMENT%20Nov.%209,%202009.docx> .

Personal communications with PIs Susan Olzak, Sarah Soule, and John McCarthy, and then-graduate student Patrick Rafail, have established that dealing with duplicate events in DCA was a difficult problem. Initial coders were instructed to ignore duplicate events, but many were recorded, probably because there were many coders on the project who would not know what others had coded. De-duplication involved trying to match events up by date, location, and categorized issue. PDF copies of cited news sources were consulted.

which had at least 400,000 participants on October 16, 1995, and multiple smaller pre- and post-march events on October 15 and 17. The records for events 9510037 (the 16th), 9510033 (the 15th), and 9510036 (the 17th) all have codes that seem to describe the big rally. The October 16 and 17 size estimates are both 400,000; the October 15 entry lacks a size estimate but has “a million black men” in the “who” field. The multiple distinct smaller events that occurred on October 15 and 17 have been collapsed into one entry for each date and muddled with the big march. Neither Louis Farrakhan nor the Nation of Islam appears in any of these records, although the Congressional Black Caucus and the SCLC are listed for October 15. The DCA also lists a march of size 100 (9512044) on October 16 in New York that never happened. The cited article written in December describes men getting on buses in New York in October to attend the DC march and follow-up volunteer activities by groups including the NAACP, the Urban League, and the group named One Hundred Black Men. There are many other examples, including event 9401023 whose cited article unambiguously describes two different marches to New York’s City Hall by two different groups on the same day that have been jumbled into one record.

This is not an attack on the DCA, which was fielded in the late 1990s and did not anticipate the problem of duplicate events. We found these errors and many others in the DCA because we were seeking to calibrate our methods by matching our retrieved events with those in the DCA using an electronic copy of the New York Times corpus. We have found matching events in complex episodes to require constructing a timeline and comparing articles carefully to determine which events are in which articles. The assumption that a coder can accurately recognize and discard a duplicate event report in one pass of coding seems unreasonable.

Relational Structures: Campaigns, Episodes, and Complex Events

Scholars of social movements have long recognized the importance of relations among events. Protests tend to occur in cycles or waves that result as protests diffuse through setting examples and change the social and political conditions to encourage more protests (Andrews and Biggs 2006, Tarrow 1989, Wada 2004). Protests are often organized into sustained campaigns, such as the iconic Albany and Birmingham campaigns in the Civil Rights Movement (Morris 1984) and protest waves are often driven by campaigns (della Porta and Andretta 2002:59). McAdam, Tarrow and Tilly (2001) and Franzosi (1999) called for attention to actions by both or all “sides,” and an emphasis on episodes of contention between movements and their opponents, a call taken up in a recent paper by Kriesi, Hutter and Bojar (2019). The claims-making tradition championed by Koopmans and Statham is a variant of protest event analysis that explicitly examines the interactions between the claims by different actors as they play out in news sources (Giugni et al. 2005, Koopmans and Statham 1999a, Koopmans and Statham 1999b, Koopmans and Olzak 2004, Statham and Geddes 2006). Despite the longstanding recognition of the structuring of protest events into campaigns or episodes, most quantitative approaches to constructing protest event data fail to make this structure explicit.

The events themselves may be complex and made up of subevents. Protesters may be met by counter-protesters. A small group may split off from a large peaceful rally and break windows. A nationwide day of protest is made up of simultaneous protests in many cities. Some events last multiple days. Most projects have no way to recognizing these relations, but relational technologies can.

Beyond Selection Bias and De-Duplication: News Coverage as Data

A large “selection bias” research addresses the question of whether events get into the news at all. No news source covers everything or even a random sample of everything, and different news sources cover different things. The differences between sources are tied to their locations, their language, their political and editorial proclivities. The likelihood that any given source reports an event also depends on time-varying factors including what else is happening at the same time and its sensitivity to different kinds of news. (For comprehensive reviews, see Earl et al. 2004., Jenkins and Maher 2016, Ortiz et al. 2005)

The naïve response to the selection issue has been to construct a protest event dataset from an available source or use an existing dataset like the DCA, give some reason why it is a good source, then ignore the selection issue entirely, treating events in the data as equivalent to events on the ground. Articles in this vein are still published regularly. The claims-making tradition takes the opposite approach and asserts that what “really” happened on the ground is irrelevant, that the only meaningful events are those that appear in the news (Giugni et al. 2005, Koopmans and Statham 1999a, Koopmans and Statham 1999b, Koopmans and Olzak 2004, Statham and Geddes 2006). Amenta and collaborators (Amenta et al. 2009, Amenta, Elliott and Tierney 2016) also stay entirely within the media framework and study instances of the mention of specific social movement organizations in the news. These media-only traditions also tend to downplay the specificity of their news sources.

Another response has been to seek to expand the pool of sources. Rafail (2019), for example, codes local newspapers from 20 cities. In the extreme, huge projects like those operated by Cline Center or Rand Corporation scrape the web for millions of news stories. But even these have been criticized for using only English-language sources with their attendant

biases (Herkenrath and Knoll 2011) and for adding even more uncertainty about what is or is not being covered in the news sources (Jenkins and Maher 2016). Many who have addressed the problem have concluded that there are no perfect or comprehensive sources and argue that all media and official sources should all be treated relativistically and compared with each other. (e.g., Jenkins and Maher 2016, Maney and Oliver 2001, Strawn 2008).

Both the events-only tradition and the media-only traditions miss the opportunity to examine the interplay between events and their media coverage. As Vliegenthart and Walgrave (2012) point out, it is odd that so much social movement scholarship has treated the selectivity of media coverage as a methodological problem rather than a theoretically interesting topic in its own right. Social movements need and want media coverage and there is a how-to literature for movement activists about ways of increasing it. Cancian and Ross (1981) showed decades ago that media attention to the women's movement peaked when it was strong and under-reported earlier events. At least since Gitlin's work (1980), researchers have recognized that media coverage shapes movements. Studies of the diffusion of protest have identified media coverage as important for expanding mobilization (Andrews and Biggs 2006, Myers 2000). Vliegenthart and Walgrave (2012) review the many ways news coverage impacts movement mobilization and the policing of protests. Hellmeier, Weidmann and Rød (2018) examine the dynamic endogenous relation between movements and news coverage of them and demonstrate that widely covered protests reduce the subsequent threshold for media coverage of events. Rafail, McCarthy and Sullivan (2019) found that local news outlets printed more stories about a day of coordinated multi-city vigils when there were more vigils nearby. Seguin (2016) argues that positive feedback of mutual causation between the events and news coverage explains the phenomenon of media cascades.

Recording the relations between news articles and the events they report opens the door to new studies that move beyond the “selection bias” paradigm and examine the mutual effects of media coverage and events. There is every reason to believe that events themselves and their representations in news sources affect each other. News coverage of people, organizations, or protest events typically follows a power law: most things are mentioned only once, while a small number of things are mentioned repeatedly (Andrews and Caren 2010, Seguin 2016, van de Rijt et al. 2013, Walgrave and Vliegenthart 2010). To study the mutual impact of events and news coverage, we need to track the volume of news coverage of an events, not just its presence or absence.

The content of the texts themselves is also of interest. There is a great deal of research on how news articles describe protests (e.g., Campbell et al. 2004, Evans 2011, Martin, Rafail and McCarthy 2017, Umamaheswar 2020, Weiner 2011). Text authors do not just list random bits of information about events but rather construct narratives. News narratives are constructed across articles, not just within them. Christian Davenport (2010) found that different newspapers constructed different narratives across articles about the conflicts between the Oakland Black Panthers and the police.

In short, acknowledging directly that protest event data is simultaneously data about the interests and actions of news media turns a methodological conundrum into an arena for research. Which events are brought into relationship with each other within articles? How do different media sources vary in this? Which events are described repeatedly, and which are barely mentioned? How is the news coverage of protest events structured? Does that structure affect how protests have impact?

With these theoretical issues in mind, we now turn to addressing the second part of Tilly's provocation, that is, addressing and developing a relational methodology on how evidence ought to be generated about the complicated phenomenon of protest events and movements.

Methods

Data Source

We base this paper on data collected as part of a larger project on the dynamics of Black protests in the US. We selected newswire articles from 1994-2010, partially out of convenience, as they are archived in the Annotated English Gigaword (AGW) database available from the Linguistic Data Consortium,⁶ which we had previously acquired for other research purposes. We used the US newswire sources in this archive: the New York Times, Associated Press, and Washington Post/Los Angeles Times. Newswire stories are a good source of data about protests because they serve as the basis for many local newspapers stories and are sometimes published verbatim in multiple outlets. The data make a novel contribution because there are no systematic data on Black protests between 1994 and 2010.

We used a protest-relevant keyword search string⁷ and a Black/African American search string⁸ to retrieve a large pool of articles from the AGW database. We used MPEDS, an open-source automated system developed by Alex Hanna (2017) to select the subset of articles that were likely to have information about protests. Hanna (2017) reports that classification errors by MPEDS mimic those of human coders in this inherently difficult task. An automated system for

---

⁶ <https://catalog.ldc.upenn.edu/LDC2012T21>

⁷ boycott* OR press conference OR news conference OR protest* OR strik* OR rally OR ralli* OR riot* OR sit-in OR occupation OR mobiliz* OR blockage OR demonstrat* OR marchi* OR marche* NOT protestant*

⁸ (Black AND NOT Blacks) OR African* OR Afro*

identifying locations in news articles from place names was used to further restrict protests to those in the US.⁹

To assess the coverage of our newswire sources, we matched our events for 1994 and 1995 to events in the DCA with Black issues or actors using the DCA citation and an electronic copy of the New York Times. There were some overlaps, which each source covering events not in the other. Events in the New York area (including Connecticut and New Jersey) were slightly more likely to be covered in the DCA, while our newswire data included many more events from the rest of the country than the DCA.

This report is based on 1,333 events identified in 1,183 articles about events occurring 1994-2010. Most of these events are canonical protests such as rallies or marches, but we also include press conferences, riots, boycotts, online and petition campaigns, as well as miscellaneous actions intended to convey a protest message. Although MPEDS does not search for lawsuits, we include them if they are described in the retrieved articles. We classify 1,101 of them as Black events, 154 as anti-Black or pro-White, and 78 as non-Black. For the purposes of this methodological demonstration, we use all the events.

Relational Data Structures

Constructing a relational database is central to our method. Many protest researchers have used relational databases software such as Microsoft Access for storing and coding data,¹⁰ but few have constructed relational databases. All modern relational database management systems (e.g. MySQL, SQLite3) possess the ability to maintain relational tables. The core of the database is a table of articles, a table of events, and a table of article-event pairs linking each

---

⁹ <https://cliff.mediacloud.org/>

¹⁰ In fact, one of us was part of a project two decades ago that coded news reports of protests using forms built in Microsoft Access, but we did not construct relational tables.

event with the articles that cite it. Figure 1 is a relational database diagram of our core data structure. PK refers to primary key and FK to foreign key. There is one row in the ArticleEvent table for each instance of an article describing an event with each unique pair appearing only once. These core tables provide controlled and verifiable links between events and the articles describing them.

Figure 1 database entity relations

PK=primary key; FK=foreign key

The tables of events and article-event links are constructed in the second step of the coding workflow as sketched in Figure 2. In what we call Pass 1, human coders identify events within articles. In a subsequent step, what we call Pass 2, human coders match up events between articles and in the process generate both the event table and the article-event links table. Articles and events have a many-to-many relationship. Every event is rigorously linked to the articles that mention it and to the text marked by the Pass 1 coders. Explicit generation and rigorous maintenance of the article-event table is markedly different from projects which seek only to construct the event table and discard “duplicate” articles about the same event.

Figure 2. Coding workflow to match events between articles and creates article-event links.

Figure 2 oversimplifies the full coding process, as there may be more than one coder of an article at Pass 1 and the coder at Pass 2 may either separate or merge events identified by Pass 1 coders. In production, the input to Pass 2 is the table of marked text for each event identified by each Pass 1 coder of an article. Every article either has at least one entry in the article-event table or is tagged as irrelevant. Similarly, every coder's annotation of an article is accounted for. As discussed above, matching events between articles is one of the most difficult and error-prone steps in constructing protest event data. De-duplication errors events can count the same event multiple times, fail to count or garble the information about events, and even create non-existent events. Yet most published research protocols either provide no discussion of this step or simply assert that coders discarded duplicate events. This rigorous structure permits the Pass 2 coding also to be reviewed and corrected.

Once the database is built, queries can be used to output files for analysis by modern statistical packages. One advantage of using a relational database is that it also supports tables to capture the relations among events, as we explain below.

Coding Procedures

The appendix provides more detail about our pass 1 and pass 2 coding interfaces. Pass 1 coding uses Alex Hanna's MPEDS Annotation Interface (MAI),¹¹ an open-source tool that is specifically designed for protest event research that we modified to meet our project's needs. Coders identify events in articles, answer questions about the events, and highlight text about them. Output from Pass 1 is a table of coder annotations that is fed to Pass 2. Pass 2 is the task of matching up events between articles. We have done Pass 2 twice for these newswire stories: once by the first author using spreadsheets that included articles' full texts and again with a specialized interface built in Microsoft Access by an experienced database programmer who worked in close and ongoing consultation with the coders. The better interface allowed us to correct mistakes made the first time. We are using it in ongoing data collection with Black newspapers as sources. In addition to the rigor of the database structure, coding is facilitated with a variety of specialized forms to give the coders the information and tools they need. Liberal use of free-text description fields and human-readable event names also improve accuracy. It adds substantial work to treat "de-duplication" as a serious part of the research, not an annoyance to be dealt with by telling front line coders to ignore things, but it is the only way to have reliable and verifiable de-duplicated event data.

Complex Events and Event Counting

Buried in protest event analysis is the deep theoretical and methodological question of just what is being counted. Most protest event datasets are constructed on the implicit but false assumption that events are independent of and comparable to each other. Complex connected events like protests with counter-protests, multi-day sit-ins, or coordinated protests in multiple cities on the same day are simultaneously one event in one sense, and multiple events in a

---

¹¹ Available at <https://github.com/mpeds/mpeds-coder>.

different sense. Deciding whether and how to parse these complex events into distinct events is one of the difficult parts of coding protest events, and there are no universally agreed-upon rules. Some projects declare by fiat that everything that happens in the same city on the same day should be collapsed into one event, but there are many research purposes for which this is unreasonable. A relational database allows us to record these connections. Operationally, we give the complex event a name and create a table of event-complex event links as well as a table listing all the complex events and their descriptions and types.¹² This allows us to record the complexity so that appropriate decisions about counting events can be made at the point of analysis. Complex events can be tagged according to whether they are in one place or multiple places and one day or multiple days. For one day one place events, we instruct coders to create distinct events when identifiably different groups of people do different things. So, a hundred people first marching and then sitting-in would be one event, but a large rally with a thousand people and ten people sitting in would be two events, plus the complex event linking them.

Aggregate Event References

One kind of complex event is an aggregate reference. News stories often refer to protest events in the aggregate with phrases such as “protests last month” or “protests in 25 cities.” We record such references because we are tracking news coverage of protests, but we flag them as aggregates using the complex event table. Often the aggregate refers to events specifically reported in other articles in our data while other times the aggregate is the only reference in our data. Aggregates can be multi-place, multi-date, or both. A phrase like “protests in 25 cities including Los Angeles, Detroit, New York, and Washington” would generate five events: one for each of the four named cities and a fifth for the aggregate reference to all 25 cities with location

---

¹² Internal to our project, we call the complex events umbrellas and link events to umbrellas because it gets confusing to speak of linking events to complex events.

“multiple.” Aggregate references to past events are often in follow-up news stories and refer to protests previously reported in other newswire, but sometimes the aggregate is the only reference in our data. We consider even vague aggregate references to protests as clues that something happened as well as parts of the discursive space. When we check, we often find that the vague reference in our newswire archive points to events described more fully in a local or activist news source. In all cases, an aggregate reference is flagged as such so it can be included or excluded from data analysis as appropriate.

Event types

When we count events or participants in events, we assume they are comparable entities. Table 1 inventories all the events we identified with attention to whether it is meaningful (or possible) to count participants per specific event and whether the event is complex or simple. In this table, “physical” events are those for which countable numbers of people did something in a specified place. “Aggregate references to physical events” are claims that such events occurred on unspecified dates or in unspecified places. “Verbal-only or diffuse non-action events” involve people making claims without assembling in a place.

As table 1 indicates, 51% of the events fit the usual assumption of a stand-alone physical one-day one-place event; another 22% are one-day one-place parts of a larger complex event and 4% are multi-place or multi-day events. Another 12% of the events we record are aggregate references to physical events, and nine percent are classified as verbal or non-action events, in which we can distinguish those that can be tied to a specific date from those that occurred across multiple dates. We offer no definitive resolution to the best way to count events but believe it is appropriate to record the event type information so decisions can be made at the point of analysis.

Table 1: Event Types²

	Number	Percent	Subtotal Percent
PHYSICAL EVENTS WITH PARTICIPANTS			78
Physical One Day One Place, Stand-alone event	686	51	
Physical One Day One Place, Part of Complex Event	299	22	
Physical Multiple Specific Days One Place	52	4	
Physical Multiple Specific Days Multiple Specific Places	6	0	
AGGREGATE REFERENCES TO PHYSICAL EVENTS			12
Physical, One Day, One Place, Aggregate reference	25	1	
Physical, Multi Day, One Place, Aggregate reference	68	5	
Physical, One Day, Multi places, Aggregate reference	14	1	
Physical, Multi Day, Multi Places, Aggregate reference	20	2	
Multi-day diffuse actions (strikes, organizing efforts, volunteering)	29	2	
VERBAL AND NON-ACTION EVENTS			9
Verbal or diffuse non-actions 1 day (statements, lawsuits, agreements, boycotts)	35	3	
Verbal or diffuse non-actions multi day (complaints, petitions, multi-day boycotts, online mobilizations)	75	6	
Verbal or diffuse non-actions aggregated	7	1	
OTHER			1
Organizations formed	9	1	
Events may not have happened	8	1	
TOTAL	1,333	100	100

Issues and issue clusters

We record the relationship between events that concern the same specific issue by linking events to clusters. The term “cluster” is operational, not theoretical, and is meant to capture linkages that can include campaigns and episodes. Each event has a free-text specific issue field. Examples include discrimination by Denny’s restaurant, whether to remove the Confederate flag from the South Carolina capitol, the killing of Amadou Diallo by New York police officers, a proposed anti-begging ordinance in the historic Civil Rights area of Atlanta, celebrate the Martin Luther King holiday, protest the inauguration of George W. Bush. We group events about the same specific issue into *Issue clusters* by creating records in an event-cluster table in the relational database. Pass 2 coders create entries in the issue cluster table, describing the cluster’s

specific issue and its location or time boundaries as appropriate, and link events to issue clusters. Because articles typically describe events in relation to other events about the same issue, it is usually straightforward to make these connections while identifying events in articles. There are less straightforward cases that require researcher decisions that are documented in the descriptions in the cluster table and revisited at the point of analysis. These including the handling of events organized around abstract issues, coalitional events with long lists of issues, events centered on annual or quadrennial events like King Day celebrations or inaugurations, and nested or overlapping issues. We give examples in the online methodological appendix of how we handled such cases. The general principle is to cluster events that are discussed in together in news articles as having a common theme. The same event can logically involve multiple issue clusters. For this report, we used decision rules to eliminate overlaps due to nesting or due to issue-specific protests at King holiday events but leave 12 instances of events linked to multiple clusters.

Results³

We report illustrative results to demonstrate the power of this approach. Future empirical reports will be based on data that have been further refined to address specific research questions.

Events and event reports

Linking events to all the articles describing them not only improves data quality, but also provides new data about variations in the intensity of news coverage of events. We identified 1,333 events in 1,183 newswire articles and a total of 2,636 instances of an article mentioning an event (i.e., article-event pairs). Event mentions follows a power law distribution: although 70% of the events were mentioned only once, four events were each mentioned in more than 40 news

articles. As Figure 3 shows, the fourteen most frequently mentioned events (1% of the total) accounted for 17% of the news mentions of the events; 16% of the events accounted for half of the news mention of events.

Figure 3 articles per event⁴

Moreover, of the 2,636 article-event pairs, only 179 (7%) involved one article per event and one event per article. Of the 1,183 articles, 59% (696) mentioned more than one event and 75% mentioned at least one event that was also mentioned in other articles.⁵ Of the 487 articles that mentioned only one event, 63% described a duplicate event mentioned in other articles. Methodologically this means that coding protest event data research inherently involves disentangling multiple events in one article and recognizing duplicate events across articles. Substantively and theoretically, this means that some events had much more impact in the news space than others, and thus much more impact on public discourse, and that events were typically discussed in relation to other events.

Participants and News Coverage

Biggs (2016) argues that it is important to attend to the number of participants in protests, not just the number of protests. For event types for which this is meaningful, we code the minimum and maximum estimated number of participants from all available information in the news articles, including references such as “filled the council chambers” or “busloads” as well as other contextual cues. The range of estimates may be very wide if there is little information in the article. The discursive impact of events is not a simple function of their number of participants. The number of articles about an event is only weakly correlated with the minimum estimated number of participants ($r=.23$, $r^2=.05$). The largest event, the Million Man March, got a great deal of news coverage, but some very large events received little coverage.

As can be seen in Figure 4, the rise in news coverage by size is modest for most of the range of event sizes and is even weakly negative for events with fewer than 100 participants. Some very small events received a great deal of news coverage, especially the handful of daily pickets by Black street vendors that preceded the 1995 arson/murder in Harlem’s Freddy’s clothing store and the hanging of nooses on a tree at Jena high school that precipitated the mobilization for the large Jena 6 marches. These small events received a great deal of coverage because they were related to other events, a point we will develop below. Most of the events that received extensive coverage (more than 10 articles) had thousands of participants but not hundreds of thousands. Some very large events received few mentions in our data.¹³

---

¹³ The 1996 Stand for Children rally (which included Black organizations in the coalition) and 2000 Millennium March for gay rights (in our data because Black people protested its lack of diversity) as well as a 2000 King Day march in Miami by 100,000 people were mentioned only once each. The 1997 Promise Keepers rally (which prominently featured Black participants and had racial reconciliation as a theme despite being primarily White in attendance) and 2010 One Nation rally (part of a multi-racial mobilization) received only three mentions each. In the cases of large coalitional rallies, it is possible that there were other newswire stories we did not select because of the screening for Black-relevant keywords.

Figure 4. Articles per event by participants in event (log scales)⁶

$r=.23$, $R^2=.05$, $N=1049$ physical events.
Darker markers are overlapping data points; 70% of events had one article and 95% less than 5 articles.

Time Trends

Figure 5 plots the monthly counts of protests, articles about protests, and estimated numbers of protest participants. A vertical line marks 9/11/2001 as the terror attacks appear to have led to a reduction in both Black protests and articles about them, a reduction attested in qualitative accounts of the period (Fletcher and Rogers 2014, Taylor 2016). Article counts track events more than they track participants, with participation occurring in a few sharp spikes. Visually, the peaks and valleys of protests and articles about protest are similar but not identical. The correlation between articles per month and events per month is high but not perfect, $r=.80$, $r^2=.64$, while the correlation between the number of participants per month and the number of

articles per month is low, $r=.27$, $r^2=.07$.⁷¹⁴ In the next section, we show the crucial importance of the relations among events in predicting the volume of news coverage.

Figure 5 Monthly counts of articles, events, and participants⁸

Issue Clusters

Events have meaning in relation to other events. Table 2 lays out information about cluster. Of the 1333 events we identified, 83% were part of a cluster of two or more events about the same specific issue; most of the non-clustered 17% (229 events) were described in relation to other non-protest events that are not in our protest event dataset, such as policy proposals or official actions. The issue clusters vary markedly in size, both in terms of the number of events

¹⁴ The correlation between the log of the participants and the number of articles is somewhat higher, $r=.46$, $r^2=.22$, but still substantially lower than the correlation between events and articles.

and the number of participants. Although most clusters are small, a few clusters account for a large share of the events. The nine issue clusters with more than 20 events (2%) account for only 4% of the participants but 27% of the news coverage and 22% of the events.

Table 2. Proportions of events, participants, news articles by number of events in cluster⁹

Number events in Cluster	Number Clusters	Number Events	Total Participants	Total Participants	Number Articles
			Minimum	Maximum	
1	229	229	1,277,674	3,643,931	276
2	82	164	879,528	1,955,556	144
3-5	76	285	1,394,884	3,649,711	216
6-10	29	211	484,953	1,910,803	203
11-20	12	163	495,461	1,439,308	179
21+	9	295	177,968	570,631	368
Totals	437	1347	4,710,468	13,169,940	1386
PERCENTAGES					
Number events in Cluster	Clusters	Events	Participants	Participants	Articles
			Minimum	Maximum	
1	52	17	27	28	20
2	19	12	19	15	10
3-5	17	21	30	28	16
6-10	7	16	10	15	15
11-20	3	12	11	11	13
21+	2	22	4	4	27
Totals	100	100	100	100	100

* Ten events are linked to two issue clusters and two events are linked to three clusters.

Events, Participants, and Issue Coverage

What makes some specific issues more newsworthy than others? Both the number of events and the number of participants seem to matter. The correlation between the number of events in a cluster and the number of articles mentioning the cluster (both logged) is .76 ($r^2=.58$), while the correlation between the number of participants in a cluster and the number of articles mentioning it (both logged¹⁵) is somewhat lower ($r=.66$, $r^2=.43$). Table 3 presents the regression of the log number of articles mentioning an event on the log number of participants in that event

¹⁵ Clusters with no size or size 0 are recoded to 1 to prevent missing values in taking the logarithm.

and the log of the number of events in the event's issue cluster: both variables have significant and independent effects, although the event size effect is weaker.¹⁶ That is, each individual event receives more coverage on average if it is part of a larger cluster with more events.

Table 3. Regression of log number of articles about an event on logged event size and logged number of events in the issue cluster. (Standard errors in parentheses)¹⁰

	(1)	(2)	(3)
Number events in cluster (log)	0.133 (10.21)**		0.134 (10.40)**
Number participants in event (log base 10) ***		0.097 (6.82)**	0.098 (7.22)**
Constant	0.105 (3.72)**	0.189 (6.76)**	-0.052 (1.48)
R^2	0.07	0.03	0.11
N events	1,333	1,333	1,3

* $p < 0.05$; ** $p < 0.01$; *** Events with no size or size 0 are coded as 1 to avoid missing values.

To some extent the strong correlation between events and articles is an artifact, because the multiple articles about an issue cluster were more likely to distinguish among the events rather than referring to them vaguely in the aggregate. But it is not entirely an artifact: there were more articles when there were more events to report, i.e., when there was an episode of sustained contention. Issue clusters with hundreds of thousands of participants that got relatively little news coverage were generally centered on large peaceful demonstrations organized by broad coalitions around liberal policy issues. Issue clusters with fewer participants but many events that got extensive coverage tended to be ongoing conflict episodes, most often about an incident of police violence. Using the same data structure we advocate, Hellmeier, Weidmann and Rød (2018) find that “salient events” (defined as events that had at least five reports) increased the number of reports of subsequent events, especially those in the same city. Their exemplary

¹⁶ The events linked to multiple clusters are attributed to the cluster with the most events. A separate test not shown finds that also including the total participants in the cluster adds no explanatory power, with all the explanatory power of participants attached to the event not the larger cluster.

methodology would be improved if they could determine whether their finding is due to media attention to issue clusters, but their analysis does not consider this possibility.

Big Issue Clusters

Linking events into issue clusters historicizes protest event data and allows an analytic interplay between quantitative and qualitative approaches. Table A1 in the appendix lists the twelve issue clusters that were either covered in more than 20 articles or have more than 20 events in them. We sketch these briefly to show the way issue clusters enrich and historicize protest event data.

The criminal legal system loomed large in past decades as well as now. Four big clusters were responses to police killing unarmed Black men. The 1999-2000 New York episode around the killing of Amadou Diallo overlapped with several other clusters about New York police violence in the 1990s and resurfaced in discussions of protests about the killing of Sean Bell in 2006. The 2014 New York protests around the killing of Eric Garner seem to flow from this history. The 2001 “Cincinnati riot” after the police killing of Timothy Thomas had been preceded two weeks before by the filing of a lawsuit about the police and was followed by many protests, community forums, a boycott, and a settlement that led to reforms. The 2009 police killing of Oscar Grant in Oakland, California was one of the first recorded on cell phone cameras; “riots” first induced the prosecution of the officer and then followed his being found guilty of a lesser charge. Overlapping this there was a decade of periodic protests that ultimately reversed the death sentence for Mumia Abu-Jamal, who had been found guilty in 1982 of killing a Philadelphia police officer. Over-charging Black youth and under-charging White hate crime were the issues in the 2007 online ColorOfChange Jena 6 mobilization, which led to both a big march in the small town of Jena, Louisiana and solidarity events around the country.

Similarly resonating with current conflicts are the ongoing battle in the 1990s and 2000s over flying the Confederate flag on the grounds of the South Carolina capitol, the three years of contention in the 1990s over the University of California banning affirmative action, and the many protests about counting votes and the outcome of the 2000 presidential election.

The Million Man March was a huge event in Black history, both for its size and for its impact, but seems left out of mainstream White understandings of the Black movement, possibly because of its emphasis on building Black solidarity and community rather than challenging White people. The Million Youth rallies in New York also emphasized Black solidarity and were themselves peaceful but were newsworthy because the police attacked the 1998 rally when it did not end promptly and tried to prevent the 1999 rally. The arson and murder of eight people in a clothing store on 125th Street in Harlem turned a small ongoing protest by disgruntled Black Harlem vendors into a major news story about racial-ethnic and political conflict in New York.

Discussion and Conclusions

The good news is that the same research protocols that make event data verifiable put protest data on a sounder theoretical grounding and open the door to new research possibilities. Attending to which events have a lot of duplicates focuses attention on the texts. When we truly accept that the data are about protest events *as they appeared in these sources*, we can ask new questions and offer new evidence about old questions about the relations between protest events and news media coverage. Our preliminary results suggest that the ability to mount sustained protests over time about the same issue is even more important than the number of people who protest in affecting media coverage. But the ability to sustain protest is itself endogenous, as media coverage can aid the diffusion of protest and recruitment of protesters, in the positive feedback system described by Seguin (2016).

Linking events to texts led us to attend to the way that events are discussed in the context of other events. This discursive structuring of events into issue clusters is closely tied to a longstanding recognition that events are generally structured as campaigns and episodes and that news reporters are trained to follow ongoing “stories” in media attention cycles. Our preliminary results suggest that capturing issue clusters provides high explanatory power in understanding patterns of media attention. Furthermore, recognizing issue clusters puts event data into dialog with qualitative historical work. It has led us, for example, to look more closely at the major episodes of protests about police violence in the 1990s and 2000s and examine how descriptions of them may have been affected by the political contexts of the cities involved.

Deep theoretical issues are present in the mundane procedures of reading and coding texts. Attending seriously to the joint problems of verifiability and de-duplication has led us into recognizing the inherent ambiguity and relationality of protest as a concept, ambiguities that always involve edge cases tied to the ambiguity of texts and that can be resolved only by understanding the relation of one event to other events. What exactly was the issue? Does the word “protest” refer only to verbal complaints or was there some sort of physical action? Into how many distinct events should a coder parse a series of actions by a shifting pool of participants? Are the hundreds of reporters, police, and bystanders watching a confrontation between a handful of KKK and New Black Panthers part of the number of participants in the whole event? How do we count online mobilizations, petitions, lawsuits, concerted letter-writing, and telephone-call campaigns? In moving toward a quantitative analysis, a researcher must address these issues and decide which cases have numbers that are meaningful in a specific research analysis context. But they are unlikely to be adequately resolved by a coder staring at

confusing text and reviewing a 30-page memo of coding rules in their first and only encounter with the text.

Our (perhaps obsessive) desire to correct errors in our data has also led us to recognize the value of providing more information in the dataset, which simultaneously makes data correction easier and permits us to record moments of coder ambiguity that we may wish to revisit later, at the point of analysis. It has also made us appreciate the value of human-machine interfaces that facilitate searches and bringing information before a coder in an accessible format. As we have processed the data to produce some preliminary results, we have found a few more inconsistencies and errors. Because our protocols create specific descriptions of events and link events back to their source articles, we can backtrack and find and correct errors or decide how to treat ambiguities.

Methodologically, we believe it is essential to plan for the joint realities of coder error and coding ambiguities. Even the most conscientious human coders make mistakes and machine coders make even more. Without specific detail about the events and rigorous links back to the source materials, there is no way of recognizing or correcting mistakes. While there is unavoidable error in all measurement, it is still desirable to minimize error and essential to be able to check data accuracy.

Training coders and checking inter-coder reliability help improve consistency in defining events and their core characteristics within texts, but these do nothing to address the problem of matching events up between news sources. Further, coder training itself does not provide for subsequent verification of coding. Verifiability is possible only if there are links back to the underlying textual sources. In the era of electronic texts and powerful computers, it is possible to

structure datasets to maintain rigorous links between events and the full texts that describe them. This is the bare minimum for data verifiability.

These data structures must be accompanied by research protocols and coding interface tools that permit and encourage the systematic review and correction of prior coding. Failure to check and verify the data does not mean they are error-free. We thought our data were in good shape after the first author had personally reviewed the coding of every event and matched events between articles using spreadsheets. When we reviewed this work with a better interface, we found many errors. We could correct them because we had maintained the links to the original texts, but we found them because we reviewed the prior work. Relational databases themselves are the key to data integrity controls. The work of human coders is facilitated by user-friendly interfaces that bring information to the coder in a variety of accessible formats and offer tools for common tasks. A de-duplication tool that merges all the information about authorized events later deemed to be duplicates is one we use often. The appendix displays some of our interface tools.

Plans for protest event data collection should anticipate duplicate event reports and budget a significant amount of time for matching events between sources, as well as for post-coding review and correction. Protocols should allow Pass 1 coders to flag events they believe are duplicates in an article while also checking texts to see if the article mentions any new events. We oppose instructing coders to simply discard articles that appear to have duplicate events. Links should be recorded between events and all the articles in the corpus that describe them, even as there is no need to do a full coding of each new mention of an already-coded event. Protest event data collection should include protocols to permit review and error correction and methodological reports should explain how duplicate events were handled.

Data collection protocols should plan for the inherent ambiguity and relationality of protest events. Instead of immediately classifying events in texts into pre-planned categories, protocols should record more specific detail about events through marking text and writing open-ended descriptions that can include comments about ambiguity. Greater specificity up front greatly improves the ability to match events between sources, recognize issue clusters, and acknowledge ambiguity and edge cases. Coders should be able to flag cases that seem ambiguous and require further discussion.

Data collection protocols should plan for multi-place, multi-action, multi-actor, multi-date events and provide some way for Pass 1 coders to report the connections between the components of complex events, as well as their recognition that multiple events are about the same issue. The well-funded SPEED project reports having a coding interface that presents coders with different choices depending on their responses to initial screening questions. Our own project recognized these relations early but initially had only a primitive way of recording them, so we have had to do most of this work post hoc. Similarly, providing Pass 1 coders with a list of already-recognized events and issue clusters would speed Pass 2 coding.

These mundane recommendations about how to do the initial data collection are deeply substantive and theoretical. They flow from both a recognition of the inherent relationality of protest and the way everything we know about protest is filtered through texts that describe protests. To return to Tilly, we must integrate our theories of the phenomenon of protest with our theories of how evidence about protests is generated. Embracing and not discounting the difficulty of the data collection task will deepen theory and allow us to generate richer data that will tell us more about how protests work.

Table A1. Summary of the ten largest clusters that were covered in more than 20 articles.

Cluster Name	Description	Issue group	# events	# articles	Min. Est. People	Dates	Places
Diallo	Near daily protests in Feb & March around the killing of Amadou Diallo by New York Police. More protests in Jan & Feb 2000 around the trial of the police; memorial in 2003.	Policing & CrimJust	62	74	38,203	1999-2000, 2003	New York City & Albany
Cincy01	Protests and disruption after police killing of Timothy Thomas on 7 April 2001. Includes intertwined lawsuit, community forums, and a boycott.	Policing & CrimJust	50	73	5,343	2001-2	Cincinnati
Jena6	Online mobilization and protests about over-charging Black youth and under-charging symbolic hate crimes. Includes a big march and many sympathy rallies around the country.	Policing & CrimJust	32	61	23,982	2007	Jena LA and multi
MillionMan95	Million Man March Oct 1995 and related events	Black Solidarity	11	58	404,025	1995	Washington, DC and elsewhere
SCFlag00	Many events 1999-2003 including protests and a boycott about removing the Confederate flag from the South Carolina state capitol.	Confederate Symbols	35	46	56,819	1999-2003	South Carolina
MillionYouth	Three Million Youth Marches in New York City 1998-2000 organized by Khalid Muhammad of New Black Panthers + attempts to block the rallies & a related 1998 Atlanta rally. Police attacked the 1998 rally.	Black Solidarity	6	42	8,151	1998 - 2000	New York City

Cluster Name	Description	Issue group	# events	# articles	Min. Est. People	Dates	Places
125thStreet	In Dec 1995 in Harlem New York City a man who had been part of a small group of Black pickets outside Freddy's clothing store set the store on fire and killed 8 people. After the arson/murder, extensive coverage tied the violence to ongoing conflicts about street vendors and non-Black businesses.	NY125thSt	24	37	2,155	1994-5	New York City
SeanBell	Multiple protests in New York City about the police killing of Sean Bell in 2006 including many events in 2007 and especially after police acquittal in 2008.	Policing & CrimJust	21	24	3,280	2006-7	New York City
OGrant	Protests in 2009 and 2010 about the police killing of Oscar Grant in Oakland California and subsequent light punishment for the officer who killed him.	Policing & CrimJust	16	24	2,653	2009-2010	Oakland, Calif.
Mumia	Protests in many places and some counter-protests across multiple years about not executing Mumia Abu-Jamal who was found guilty of killing a police officer.	Policing & CrimJust	29	23	19,677	1994-2008	multiple
Vote2000	Events in many cities protesting the 2000 presidential election in November 2000 through early January 2001.	Politics	21	16	11,635	2000-1	Multi
UCAffAct	Protests about banning affirmative action in University of California	Affirmative Action	21	14	16,874	1995-7	California

Cluster Name	Description	Issue group	# events	# articles	Min. Est. People	Dates	Places
TOTAL Big Clusters			328	462	592,797		
Total All Clusters			1,333	1,386	4,710,468		
Percent of All Clusters			25%	35%	13%		

References

- Althaus, Scott L., Joseph Bajjalieh, John F. Carter, Buddy Peyton and Dan A. Shalmon. 2017. "Cline Center Historical Phoenix Event Data Variable Descriptions (V.1.0.0)." Vol. <http://www.clinecenter.illinois.edu/data/event/phoenix/>: Cline Center for Democracy.
- Amenta, Edwin, Neal Caren, Sheera Joy Olasky and James E. Stobaugh. 2009. "All the Movements Fit to Print: Who, What, When, Where, and Why Smo Families Appeared in the New York Times in the Twentieth Century." *American Sociological Review* 74(4):636-56.
- Amenta, Edwin, Thomas Alan Elliott and Amber C Tierney. 2016. "Political Reform and Newspaper Coverage of Us Movements in Depression, Recession, and Historical Perspective." *Mobilization: An International Quarterly* 21(4):393-412.
- Andrews, Kenneth T. and Michael Biggs. 2006. "The Dynamics of Protest Diffusion: Movement Organizations, Social Networks, and News Media in the 1960 Sit-Ins." *American Sociological Review* 71(5):752-77.
- Andrews, Kenneth T. and Neal Caren. 2010. "Making the News: Movement Organizations, Media Attention, and the Public Agenda." *American Sociological Review* 75(6):841-66. doi: 10.1177/0003122410386689.
- Beissinger, Mark R. 2002. *Nationalist Mobilization and the Collapse of the Soviet State*. New York: Cambridge University Press.
- Bond, Doug, J. Craig Jenkins, Charles L. Taylor and Kurt Schock. 1997. "Mapping Mass Political Conflict and Civil Society: Issues and Prospects for the Automated Development of Event Data." *Journal of Conflict Resolution* 41(4):553-79. doi: 10.1177/0022002797041004004.
- Boschee, Elizabeth, Premkumar Natarajan and Ralph Weischedel. 2013. "Automatic Extraction of Events from Open Source Text for Predictive Forecasting." Pp. 51-67 in *Handbook of Computational Approaches to Counterterrorism*, edited by V. S. Subrahmanian. New York, NY: Springer New York.
- Campbell, Shannon, Phil Chidester, Jason Royer and Jamel Bell. 2004. "Remote Control: How Mass Media Delegitimize Rioting as Social Protest." *Race, Gender & Class* 11(1):158.
- Cancian, Francesca M. and Bonnie L. Ross. 1981. "Mass Media and the Women's Movement: 1900-1977." *Journal of Applied Behavioral Science* 17(1 (March)):9-26.
- Caren, Neal. 2014. "It Is Time to Get Rid of the E in Gdelt." *BadHessian*. Retrieved 2021 (<https://badhessian.org/2014/05/it-is-time-to-get-rid-of-the-e-in-gdelt/>).
- Chalabi, Mona. 2014. "Mapping Kidnappings in Nigeria." *FiveThirtyEight*. Retrieved 2021 (<https://fivethirtyeight.com/features/mapping-kidnappings-in-nigeria/>).
- Croicu, Mihai and Nils B Weidmann. 2015. "Improving the Selection of News Reports for Event Coding Using Ensemble Classification." *Research & Politics* 2(4):2053168015615596. doi: 10.1177/2053168015615596.
- della Porta, Donatella and Massimiliano Andretta. 2002. "Changing Forms of Environmentalism in Italy: The Protest Campaign on the High Speed Railway System." *Mobilization* 7(1):59-77.
- Earl, Jennifer, Andrew Martin, John D. McCarthy and Sarah A. Soule. 2004. "The Use of Newspaper Data in the Study of Collective Action." *Annual Review of Sociology* 30(1):65-80.
- Evans, Erin. 2011. "Mass Media and "Moral Shocks": The Radical Flank Effect in Media Coverage of Protests." *Conference Papers -- American Sociological Association*:1062-62.
- Fletcher, Bill, Jr. and Jamala Rogers. 2014. "No One Said That It Would Be Easy." *Black Scholar* 44(1):86-112. doi: 10.5816/blackscholar.44.1.0086.
- Franzosi, Roberto. 1999. "The Return of the Actor. Interaction Networks among Social Actors During Periods of High Mobilization (Italy, 1919-1922)." *Mobilization* 4(2):131-49.

- Gitlin, Todd. 1980. *The Whole World Is Watching: Mass Media in the Making and Unmaking of the New Left*. Berkeley, CA: U California Press.
- Giugni, Marco, Ruud Koopmans, Florence Passy and Paul Statham. 2005. "Institutional and Discursive Opportunities for Extreme-Right Mobilization in Five Countries." *Mobilization: An International Journal* 10(1):145-62.
- Hanna, Alex. 2016. "Automated Coding of Protest Event Data: Development and Applications." Ph.D., The University of Wisconsin - Madison, Ann Arbor. Retrieved from Dissertations & Theses @ CIC Institutions; Dissertations & Theses @ University of Wisconsin at Madison; ProQuest Dissertations & Theses Global; Sociological Abstracts; Technology Collection, 10190827.
- Hanna, Alex. 2017. "Mpeds: Automating the Generation of Protest Event Data." *Deposited at SocArXiv* <https://osf.io/preprints/socarxiv/xugmy>.
- Hayes, Matthew and Peter Nardulli. 2011. "The Quality and Reliability of Data Generated by Speed's Societal Stability Protocol: Mechanisms and Tests."
- Hellmeier, Sebastian, Nils B. Weidmann and Espen Geelmuyden Rød. 2018. "In the Spotlight: Analyzing Sequential Attention Effects in Protest Reporting." *Political Communication* 35(4):587-611. doi: 10.1080/10584609.2018.1452811.
- Hellmeier, Sebastian, Espen Geelmuyden Rød and Nils B Weidmann. 2019. "Coding Instructions for the Mass Mobilization in Autocracies Database, Version 2.0." http://cybis.univ-grenoble-alpes.fr:8082/dataset/08a3a90e-0b6a-435e-b566-135c8f3b9f34/resource/c542f8d9-2499-4e5a-8bba-8f68ca53a033/download/mmad_codebook_v2-20190821.pdf.
- Herkenrath, Mark and Alex Knoll. 2011. "Protest Events in International Press Coverage: An Empirical Critique of Cross-National Conflict Databases." *International Journal of Comparative Sociology (Sage Publications, Ltd.)* 52(3):163-80. doi: 10.1177/0020715211405417.
- Hutter, Swen. 2014. "Protest Event Analysis and Its Offspring." Pp. 335-67 in *Methodological Practices in Social Movement Research*, edited by D. della Porta. Oxford: Oxford University Press.
- Jenkins, J. Craig and Thomas V. Maher. 2016. "What Should We Do About Source Selection in Event Data? Challenges, Progress, and Possible Solutions." *International Journal of Sociology* 46(1):42-57. doi: 10.1080/00207659.2016.1130419.
- Koopmans, Ruud and Paul Statham. 1999a. "Political Claims Analysis: Integrating Protest Event and Political Discourse Approaches." *Mobilization* 4(2):203-21.
- Koopmans, Ruud and Paul Statham. 1999b. "Challenging the Liberal Nation-State? Postnationalism, Multiculturalism, and the Collective Claims Making of Migrants and Ethnic Minorities in Britain and Germany." *American Journal of Sociology* 105(3):652-96. doi: 10.1086/210357.
- Koopmans, Ruud and Dieter Rucht. 2002. "Protest Event Analysis." Pp. 213-59 in *Methods of Social Movement Research*, edited by B. Klandermans and S. Staggenborg. Minneapolis: University of Minnesota Press.
- Koopmans, Ruud and Susan Olzak. 2004. "Discursive Opportunities and the Evolution of Right-Wing Violence in Germany[1]." *American Journal of Sociology* 110(1):198-230.
- Kriesi, Hanspeter, Swen Hutter and Abel Bojar. 2019. "Contentious Episode Analysis." *Mobilization* 24(3):251-73.
- Lorenzini, Jasmine, Hanspeter Kriesi, Peter Makarov and Bruno Wüest. 2021. "Protest Event Analysis: Developing a Semiautomated Nlp Approach." *American Behavioral Scientist* 0(0):00027642211021650. doi: 10.1177/00027642211021650.
- Makarov, Peter, Jasmine Lorenzini and Hanspeter Kriesi. 2016. "Constructing an Annotated Corpus for Protest Event Mining." Pp. 102-07 in *Proceedings of the First Workshop on NLP and Computational Social Science*.
- Maney, Gregory M. and Pamela E. Oliver. 2001. "Finding Collective Events: Sources, Searches, Timing." *Sociological Methods and Research* 30(2):131-69.
- Martin, Andrew W., Patrick Rafail and John D. McCarthy. 2017. "What a Story?". *Social Forces* 96(2):779-802. doi: 10.1093/sf/sox057.

- McAdam, Doug, Sidney Tarrow and Charles Tilly. 2001. *Dynamics of Contention*. New York: Cambridge University Press.
- Morris, Aldon. 1984. *The Origins of the Civil Rights Movement: Black Communities Organizing for Change*. New York: Free Press.
- Myers, Daniel J. 2000. "The Diffusion of Collective Violence: Infectiousness, Susceptibility, and Mass Media Networks." *American Journal of Sociology* 106(1, July):173-208.
- Nardulli, Peter F. and Matthew Hayes. 2011. "Transforming Textual Information on Events into Event Data within Speed." University of Illinois, Urbana, <https://uofi.app.box.com/s/y8nhzmzfim774flnvzro26rooz3l5tdx>.
- Nardulli, Peter F., Scott L. Althaus and Matthew Hayes. 2015. "A Progressive Supervised-Learning Approach to Generating Rich Civil Strife Data." *Sociological Methodology* 45(1):148-83. doi: 10.1177/0081175015581378.
- Olzak, Susan. 1989a. "Labor Unrest, Immigration, and Ethnic Conflict in Urban America, 1880-1914." *American Journal of Sociology* 94(6):1303-33.
- Olzak, Susan. 1989b. "Analysis of Events in the Study of Collective Action." *Annual Review of Sociology* 15:119-41.
- Ortiz, David G., Daniel J. Myers, N. Eugene Walls and Maria-Elena D. Diaz. 2005. "Where Do We Stand with Newspaper Data?". *Mobilization: An International Journal* 10(3):397-419.
- Pan, Jennifer and Han Zhang. 2019. "Casm: A Deep-Learning Approach for Identifying Collective Action Events with Text and Image Data from Social Media." *Sociological Methodology* 49: (forthcoming).
- Rafail, Patrick. 2018. "Protest in the City: Urban Spatial Restructuring and Dissent in New York, 1960–2006." *Urban Studies (Sage Publications, Ltd.)* 55(1):244-60.
- Rafail, Patrick. 2019. "Who Is Doing What and Where? Patterns of Contention in 20 U.S. Cities, 1996–2006." Paper presented at the Social Science History Association, November 2019, Chicago.
- Rafail, Patrick, John D. McCarthy and Samuel Sullivan. 2019. "Local Receptivity Climates and the Dynamics of Media Attention to Protest*." *Mobilization: An International Quarterly* 24(1):1-18. doi: 10.17813/1086-671x-24-1-1.
- Ratliff, Thomas N. 2011. "On the Stage of Change: A Dramaturgical Approach to Violence, Social Protests, and Policing Styles in the U.S." PhD, Sociology, Virginia Polytechnic Institute and State University.
- Ratliff, Thomas N. 2013. "A Dramaturgical Approach to Protest Policing in the United States: Actors, Enemies, the Stage, and Performance." *Conference Papers -- American Sociological Association*:1-43.
- Rucht, Dieter and Friedhelm Neidhardt. 1998. "Methodological Issues in Collecting Protest Event Data: Units of Analysis, Sources and Sampling, Coding Problems." Pp. 65-89 in *Acts of Dissent: New Developments in the Study of Protest*, edited by D. Rucht, R. Koopmans and F. Neidhardt. Berlin: Edition Sigma Rainer Bohn Verlag.
- Schrodt, Philip A. and Deborah J. Gerner. 1994. "Validity Assessment of a Machine-Coded Event Data Set for the Middle East, 1982-92." *American Journal of Political Science* 38(3):825-54. doi: 10.2307/2111609.
- Schrodt, Philip A. 2012. "Precedents, Progress, and Prospects in Political Event Data." *International Interactions* 38(4):546-69. doi: 10.1080/03050629.2012.697430.
- Schrodt, Philip A. and David Van Brackle. 2013. "Automated Coding of Political Event Data." Pp. 23-49 in *Handbook of Computational Approaches to Counterterrorism*, edited by V. S. Subrahmanian. New York, NY: Springer New York.
- Seguin, Charles. 2016. "Cascades of Coverage: Dynamics of Media Attention to Social Movement Organizations." *Social Forces* 94(3):997-1020. doi: 10.1093/sf/sov085.
- Snyder, David and Charles Tilly. 1972. "Hardship and Collective Violence in France, 1830 to 1960." *American Sociological Review* 37(5):520-32.

- Soule, Sarah, Doug McAdam, D. McCarthy John and Susan Olzak. n.d. "Dynamics of Collective Action." <http://web.stanford.edu/group/collectiveaction/cgi-bin/drupal/>.
- Statham, Paul and Andrew Geddes. 2006. "Elites and the 'Organised Public': Who Drives British Immigration Politics and in Which Direction?". *West European Politics* 29(2):248-69. doi: 10.1080/01402380500512601.
- Strawn, Kelley D. 2008. "Validity and Media-Derived Protest Event Data: Examining Relative Coverage Tendencies 114 Mexican News Media." *Mobilization* 13(2):147-64.
- Tarrow, Sidney. 1988. *Democracy and Disorder: Politics and Protests in Italy, 1965-1975*. New York: Oxford University Press.
- Tarrow, Sidney. 1989. *Struggle, Politics and Reform: Collective Action, Social Movements, and Cycles of Protest*. Ithaca: Center for Institutional Studies, Cornell.
- Taylor, Charles Lewis. 2013. "Data Quality for Measuring Political Protest and Government Change." *All Azimuth: A Journal of Foreign Policy and Peace* 2(2):23-29.
- Taylor, Keeanga-Yamahatta. 2016. *From #Blacklivesmatter to Black Liberation*. Chicago, Illinois: Haymarket Books.
- Tilly, Charles, Louise Tilly and Richard Tilly. 1975. *The Rebellious Century, 1830-1930*. Cambridge: Harvard University Press.
- Tilly, Charles. 2002. "Event Catalogs as Theories." *Sociological Theory* 20(2):248-54. doi: 10.1111/1467-9558.00161.
- Umamaheswar, Janani. 2020. "Policing and Racial (in)Justice in the Media: Newspaper Portrayals of the 'Black Lives Matter' Movement." *Civic Sociology* online first(March). doi: <https://doi.org/10.1525/001c.12143>.
- van de Rijt, Arnout, Eran Shor, Charles Ward and Steven Skiena. 2013. "Only 15 Minutes? The Social Stratification of Fame in Printed Media." *American Sociological Review* 78(2):266-89. doi: 10.1177/0003122413480362.
- Vliegthart, Rens and Stefaan Walgrave. 2012. "The Interdependency of Mass Media and Social Movements." in *The SAGE Handbook of Political Communication*, edited by H. A. Semetko and M. Scammell. London: SAGE Publications Ltd.
- Wada, Takeshi. 2004. "Event Analysis of Claim Making in Mexico: How Are Social Protests Transformed into Political Protests?". *Mobilization: An International Journal* 9(3):241-57.
- Walgrave, Stefaan and Rens Vliegthart. 2010. "Why Are Policy Agendas Punctuated? Friction and Cascading in Parliament and Mass Media in Belgium." *Journal of European Public Policy* 17(8):1147-70. doi: 10.1080/13501763.2010.513562.
- Wang, Wei, Ryan Kennedy, David Lazer and Naren Ramakrishnan. 2016. "Growing Pains for Global Monitoring of Societal Events." *Science* 353(6307):1502-03.
- Weidmann, Nils B and Espen Geelmuyden Rød. 2018. "Coding Instructions for the Mass Mobilization in Autocracies Database, Version 1.0." Vol.
- Weiner, Melissa F. 2011. "All the News That's Fit to Print? Silence and Voice in Mainstream and Ethnic Press Accounts of African American Protest." *Research in Social Movements, Conflicts & Change* 31:297-327. doi: 10.1108/S0163-786X(2011)0000031012.

EndNotes

NOTE: These notes are for the purpose of ensuring that we can replicate the sources of all information in this article. This is a working paper and comments are appreciated.

¹ This research was funded by National Science Foundation Grants SES1423784 and SES1918342.

² Table 1 is Excel Spreadsheet EventTypes.xlsx in the 2021newswire_pc folder, tab format 7-26, created from events_types_forms_202106.do on 7/26/21 at 10 am. The evtypeX variable after further coding and distinction of complex events

³ Documentation for all results is recorded in an anonymized public document

https://docs.google.com/document/d/1eoZJfsZ1yPEA89pD_DIHG3z4VdaDIZkgwE7fqFOky2E/edit?usp=sharing

⁴ Figure 3 derived from a table calculated in events_overview_20210528.do and copied to spreadsheet results_events_2021-5-28.xlsx. Same number of events as the 5/31 export which changed some cluster links and some event names but not the number of articles or events. Graph is P_events_articles which reads sheet nart. I made a .png version of this graph and put it in the methods paper folder.

⁵ These numbers were calculated in xwwalk_20210529.do

⁶ nartev_size_2021-6-30.png created by events_20210528_overview.do

⁷ correlations calculated in months_2021_arts_events.do

⁸ months_ev_art_partic.png generated by timeplot_art_ev_partic.do on 6/30/21

⁹ From include file table2_2021-6-8.do which was run from clusters_analysis_201206.do on 6/8/21 to create table2_clustersizes_2012-6-8.xlsx The bivariate correlations in the first paragraph in Events, Participants, and Issue Coverage before table 3 on the cluster-level correlations between number of events and number of articles comes from clusters_only_202106.do re-run on 7/22/21

¹⁰ Document regnart_2021-7-26.doc created by outreg in events_issues_201216.do on 7/26/21

On-Line Appendix to “Constructing Theory-Informed Relational and Verifiable Protest Event Data”

Although our ongoing data collection involves undergraduate coders who were trained and tested for accuracy, the original Pass 1 coding of these newswire data was conducted without external funding and made use of less well-trained undergraduates working for research experience and course credit. All Pass 1 coding was reviewed and often re-done by the first author and was reviewed and corrected again in the Pass 2 coding by two of the authors and an experienced graduate student research assistant. Although well-trained coders are desirable and reduce the need for repeated returns to the original source materials, the rigorous relational database structures provide a way to check and correct the work of all coders, even in low-budgets projects relying on student labor or the researcher’s own labor.

Screen shots of our Pass 1 and Pass 2 interfaces are at the end of this document.

Overall Workflow Plan

We code data in three passes. In Pass 1, events are identified within texts. In Pass 2, events are linked between articles and a database of confirmed events with event-article links is created. In what we call Pass 3, we return to the database and code more variables as we need them for analysis, pulling in marked text from the coder annotations and testing automated approaches to coding specific event characteristics.

Pass 1 Coding: Text Marking Within Articles

For Pass 1 coding, we used the [MPEDS Annotation Interface \(MAI\)](#) written by Alex Hanna, an open-source tool that can be modified to meet a particular project’s needs. This software permits identifying article-level characteristics including publication information, a short article description, and yes/no flags for whether there are codable events in the article. Coders must explicitly state why they are not coding events: no protest, historical events only, non-US events only. (In the newswire data, the coders also screened for whether an event was a Black event and were instructed to state why they thought an event was not a Black event. This proved to be extremely error-prone, and all these decisions were later reviewed and in many cases corrected.) Articles flagged as not containing codable events are reviewed at Pass 2 for quality control.

The MAI permits multiple events to be identified within articles. For each event, coders provide a short event description, date, and time; answer several closed-ended questions about a few other closed-ended options and comment options, and mark text for each event, marking article text relevant to event features such as actors, actions, issues, named organizations, police actions,

size-relevant text, and location-relevant text. The software can be readily customized for exactly which items should be coded in a given project.

Pass 1 coders are instructed to write informative free-text descriptions of events that include decisions they made about subdividing complex events, to break event sequences into components if different groups or numbers of people were involved in different actions, always to create a separate event for counter-protesters, to follow the rule “when in doubt, create an event,” and to mark text relevant to key concepts including issues, actors, actions, organizations, police actions, specific locations, and size. Because a few events are described in many articles and coders can become bored and fatigued marking the same information repeatedly in multiple articles, our protocol includes a “duplicate event” flag that allows coders to provide less detail for events they know they have coded before.

Output from the coding interface is a comma- or tab-delimited text file for each action by a coder that can be processed and read as an input to other software, in our case the Pass 2 interface.

As far as we know, only the MAI is able explicitly to generate two-dimensional tagging to associate marked text with different events within the same article, although it may be possible to adapt other products to this task. Other available coding interfaces include commercial products such as ATLAS.ti, NVIVO, and Dedoose or open-source projects developed by social and political scientists for human annotation of events in text sources including [Yoshikoder](#) developed by Will Lowe; [CIVET](#) developed by Phil Schrodtt, and [TextThresher](#) developed by Nick Adams. [Taguette](#) is an open-source text annotation product developed by information scientists associated with libraries.

Pass 2 Coding: Connecting Events Across Articles and Establishing Relational Tables

There are two elements to the Pass 2 system: the underlying relational tables and an interface for interacting with the tables. Both are important. The tables are constructed with controls for data integrity, including primary and foreign keys. The interface is designed to facilitate the work of coders.

The output from the MAI is a table with one record for each action a coder takes, so there are many records for each article and each event within the article. This output is pre-processed to reshape it into one record for each article read by a given coder, and one record for each event identified within an article by a given coder. These reshaped records are the inputs from the MAI into the Pass 2 database coder line tables. A separate article table with full article text and publication data is also an input. A unique article ID links the article table to the coder line table.

Our current Pass 2 process was built in Microsoft Access and Visual Basic by an experienced database programmer who was familiar with Access and worked closely with the Pass 2 coding team (two of the authors and a graduate research assistant) to design coding forms that would facilitate their work. Significant development time was also allocated to specialized forms for reviewing and correcting work previously done with inferior tools

The key task of Pass 2 is to define confirmed events and connect events across articles in a many-to-many relational system that maintains links between events and the articles describing them. The event table is built as Pass 2 coders create confirmed events and link coder lines to events and thus articles to events in the event-article link table. Pass 2 coders can also link articles directly to events, thus correcting a Pass 1 coder's failure to identify an event. Every article reviewed in Pass 1 is either flagged as not relevant if it does not contain codable events or it is associated with one or more events in the article-event table. A firm link is established at Pass 2 between text marked by a Pass 1 coder and the appropriate event confirmed by a Pass 2 coder.

Pass 2 coding is article based. The Pass 2 coder can view both Pass 1 coder annotations and the full article text. The workflow involves searching for previously created confirmed events and creating new confirmed events, then linking the coder lines to these events. Because the Pass 2 coder may parse events differently from the Pass 1 coder, the same annotations may be applied to more than one confirmed event or the same confirmed event may be associated with multiple annotations from the same Pass 1 coder. Pass 2 coders can edit previously created events in light of new information and are encouraged to write descriptions that will distinguish each event from other related events and note any ambiguities in event identification. While working with the texts, Pass 2 coders also link events to each other in issue clusters. (We are currently linking events to complex event umbrellas in a backfill process, but plan to bring this into Pass 2.) The Pass 2 coder also identifies issue clusters of related events and gives events human-readable names such as Jena_bigrally070820 or Diallo_990215_motorcade to supplement system ID numbers and to provide a human-readable indicator of relations among events.

Pass 2 coders are encouraged to write full descriptions that distinguish each event from related events and that include comments about ambiguity if needed. A 255-character limit in the initial implementation proved too confining, and descriptions can be longer if needed. Having an adequate event description is important for being able to recognize duplicate events across sources. Pass 2 coders can search for previously confirmed events (and issue clusters) by date, location, and keywords in the description. Flags can be thrown for difficult cases and the free-field description can be used to indicate doubts about whether two events are the same or different. Pass 2 coders also review Pass 1 coders' judgments that there were no codable events and can mark articles or coder annotations as not relevant. It is possible to query and review these Pass 2 decisions.

The Pass 2 coder also estimates the number of participants. Even though many news articles lack explicit estimates of event size, size-relevant information is present in most news articles, including phrases such as "filled the city council chambers," "filled ten blocks," "busloads," "a handful" and other contextual information. Even vague event mentions occur in contexts that provide some clues about orders of magnitude of event sizes. Pass 1 annotators mark text relevant to event size. Pass 2 coders both indicate the amount of information given about event size (numerical, partial or conflicting, none) and provide minimum and maximum size estimates, thus indicating the degree of uncertainty about their estimates. A news report that 300 marched is really an estimate, which we record as 280 to 320. Conflicting size reports are reflected in the

minimum and maximum. A very large range such as 50 to 5000 can express the coder's high uncertainty about event size. Future or planned events that may not have occurred have a minimum size of 0. Event types such as lawsuits or petitions are not given size estimates, nor are vague aggregates of other events that are in the data.

Pass 2 coders flag characteristics relevant to complex events: time type (one day, consecutive days, periodic days, unknown); place type (one place, multiple places, unknown place, no place e.g. Internet); whether the event is an aggregate; whether the event is part of a larger event. These flags allow us to identify the events to be associated with complex event umbrellas. These will allow users of the data to expand multi-date records into one record per day, avoid double-counting the participants in complex events, and include or exclude aggregate events that lack specific dates or locations.

Pass 2 coding can be thought of as involving two different kinds of cases. The majority of events are described only once in one article, in which case an efficient workflow involves sending Pass 1 coding directly into the database, unless there are multiple coders of the same article. However, matching events up between texts and de-duplicating can be very complicated and requires a wide variety of tools. Matching up events between articles is difficult and error prone. It requires complex search functions, the ability to access the underlying texts, and the ability to correct mistakes. When we used better tools, we were able to find and correct errors in Pass 2 coding previously done by the first author using inferior tools based on spreadsheets. The system also includes lookup tools and error-correction tools including a duplicate event tool that can merge two confirmed events that are later recognized to be the same event. The extant Visual Basic code and the database design specifications are available from the authors. Here are screen shots of the interface. One of our experienced undergraduate coders called it an "old fashioned" interface, and it may be possible to perform the same work with a more modern interface with drag and drop functions. However implemented, Pass 2 coding requires the ability to retrieve information in many different forms.

Identifying Issue clusters

Each confirmed event has a free field describing its issue. Issue clusters have the same specific issue. In most cases, this is straightforward as events are typically described together in texts that are about an issue. Some cases are less straightforward and reveal the ways in which researchers must make decisions about the inherent ambiguities in coding protest issues and the value of having data systems that permit the coding decisions to be revisited. Our general principle is to cluster events that are described together within articles as being about the same specific issue. Examples of specific issues are allegations of racial bias by the Adams Mark hotel chain, demanding punishment of the police officer who killed Timothy Thomas in Cincinnati, opposition to a proposed anti-begging ordinance in the historic Civil Rights district of Atlanta, or demands for the removal of the Confederate flag from the South Carolina capitol. This is usually unproblematic. Long-lasting episodes of contention often occur in phases, and one decision is whether to create separate clusters for each phase or one large cluster for the whole episode. Examples are the three years of contention over abolishing affirmative action in the University of California system, the decade of protests against imposing the death penalty on Mumia Abu-

Jamal, and the distinct legal phases in cases where a police officer is prosecuted for killing someone. It is possible to create nested issue clusters to capture these phases, or more specific issues within the larger issue. For example, there was focused contention in 1996 at Cal State Northridge about inviting KKK leader David Duke to debate affirmative action with Civil Rights activist Joe Hicks; College Republicans went to court to try to block Duke from speaking because they believed that Duke's opposition to affirmative action would imply that support for Proposition 209, the anti-affirmative action ballot initiative, was racist. We created an issue cluster for this very specific issue as well as associating these events with the larger California affirmative action cluster. In general, however, we created only the broader issue cluster if the core issue remained essentially the same over time.

Episodes of contention sometimes involved specific issues that had become intertwined, in which case we defined one issue cluster to include both, but also made operational decisions about putting boundaries on cases. The police killing of Timothy Thomas in Cincinnati happened two weeks after a lawsuit had been filed about racial bias by Cincinnati police; the NAACP joined the lawsuit after the killing and the subsequent legal proceedings and settlement became intertwined with ongoing protests about the specific killing. After one of a handful of picketers outside Freddy's clothing store in Harlem in 1995 burned the store and killed eight people in the fire, extensive news coverage linked the arson-murder to ongoing conflicts about Black vs non-Black vendors in Harlem and ethnic conflict in New York more broadly. The Jena 6 case involved both the original issue of under-charging White students who hung nooses on a high school tree after Black students sat there and over-charging Black students who fought with a White student. After the large Jena 6 rally and nationwide solidarity rallies, there was a wave of protests about noose hangings and calls for greater prosecution of symbolic hate crimes; we created a separate issue cluster for noose protests happening in the months following the Jena 6 rallies. There was a great deal of mobilization around the South Carolina Confederate flag struggle from 1999 onwards which we put into one issue cluster; there were many protests about Confederate symbols elsewhere in this period, which we grouped together into one multi-place cluster because they were often discussed together in articles, except for the Georgia cases, which were frequent enough to merit their own cluster.

We also made some relatively arbitrary decisions about how to define issue clusters. (1) We do not cluster for events organized around abstract issues such as police violence or racism in general unless events are described as being connected to each other. (2) We do not itemize the list of issues in coalitional events about a broad list of issues. If there are multiple events around the same coalition, we create a cluster for the coalition. (3) We create clusters for the events centered on external events, including the annual King Day celebrations and the quadrennial presidential nomination conventions and inaugurations, again because these protests are typically described in articles about the larger event. (4) We created a cluster for relatively unconnected events that all addressed the issue of Black-Hispanic relations in 2006, the year of major immigrant protests, because they were often discussed in the same articles.

Events can logically belong to more than one issue cluster, even though cluster overlaps present difficulties for analysis. Some issues are nested within other issues. The clearest example is a

series of events at Cal State Northridge in 1996 about inviting former KKK leader David Duke to debate affirmative action (the Republicans thought the Duke invitation unfairly portrayed opposition to affirmative action as racist) that was part of a larger cluster of events opposing Proposition 209 which, in turn, were part of the ongoing struggle over affirmative action in the California public universities between 1995 and 1997.

By retaining the original information about an event's specific issue, including decision details in issue cluster definitions, and maintaining the data relationally, we acknowledge the role of the researcher in making the choices about categorization, especially in ambiguous cases, and provide information that permits this to be revisited and done differently depending on research purposes.

In the data reported in this paper, 42 events were coded as being connected to more than one issue cluster. Six cases involving 20 events involved a specific issue cluster nested within a larger issue. For this paper, we eliminated overlaps due to nesting and kept only the overarching cluster. Seven overlaps involving nine events arose because we created issues clusters for the annual events commemorating Martin Luther King, Jr.'s birthday and the anniversary of his death. Sometimes the larger commemoration would include a protest about some other issue, generally discussed as part of the commemoration events. In this paper, we attribute these events to the specific issue cluster rather than the commemoration cluster.

Another 12 events addressed multiple issues. A few events in the fall of 2007 after the big Jena 6 rally addressed both the Jena 6 events themselves and subsequent noose crimes and, in one case, also referenced the torture and rape of a Black woman in West Virginia (which was the specific issue for several West Virginia protests). One Confederate symbol protest referenced the specific Confederate flag campaigns in South Carolina and Georgia as well as other states and so was part of three clusters. In New Orleans there were protests in 2006 calling for the right of people displaced by hurricane Katrina to vote in a local election that raised issues about racism in post-hurricane development policies that became central in 2007 protests; some events addressed both issues. An anti-Obama rally held by Glenn Beck on the 47th anniversary of the March on Washington is part of two clusters, one that includes the counter-protests the day of the rally and a follow-up rally led by John Stewart and Stephen Colbert on October 30 and the other a series of mobilizations for a rally on October 10 by a multi-racial One Nation coalition.

MAI Coding Interface Screenshots

Black Protest Project

Black Protest Project: Testing VersionHome

Logged in as [user]Logout

Racial Justice In Memphis.... A Long, Hard Struggle

The Tennessee Tribune; Nashville, Tenn. | 2006-03-02 | The-Tennessee-Tribune; Nashville, Tenn. | 2006-03-02 | 16213f1ee9a5078b42858206d07930a5

The nation mourned when civil rights leader Dr. Martin Luther King was shot and killed on the balcony of the Lorraine Motel in Memphis, Tenn., in 1968.

Many viewed the heightening unrest of the city sanitation workers' strike and conflicts stemming from the Civil Rights Act of 1964 for setting the stage for this act. But, according to University of Texas at Austin Assistant Professor Laurie Green, Black Memphians had experienced severe racial strife on numerous fronts for decades before King walked with the city's sanitation workers whose "I AM A MAN" proclamation rallied males and females, poor and middle class during this historic struggle for equality.

"Blacks in Memphis had been battling for freedom from discrimination in all facets of their everyday lives for decades," said Green whose research is chronicled in an upcoming book on race, gender and freedom in the civil rights era in Memphis. "Memphis was dominated by corrupt, racist city officials since the 1930s and Blacks struggled to overcome what many referred to as a 'plantation mentality' that kept them relegated to menial, low-paying jobs."

Although issues of desegregating public services and gaining equality in the workplace were significant in Memphis, police brutality, poverty, censorship and harassment were part of Blacks' everyday lives for decades. In the fall of 1940, for instance, Memphis city boss Edward H. Crump started what came to be known as the "reign of terror." His efforts were enforced by a brutal police force and sought to repress Black Republicans, subdue labor activism and exert social control of migrants, Green said.

During World War II, when Memphis' manufacturing sector grew from just more than 27,000 to nearly 49,000 jobs in two years, many Blacks were denied jobs in the burgeoning defense industry. Thousands of Black men and women were bused each day from their homes in Memphis back to the rural farmlands surrounding the area and forced to pick cotton, considered a wartime necessity.

Black women, responding to ads calling on females to do their part for the wartime effort as "Rosie the Riveter," were turned away. In fact, most Black women who moved from rural towns to Memphis quickly found the only jobs they could get were as servants in white households.

"Black men and women were denied the opportunity to work in skilled positions," said Green. "The jobs they could obtain paid extremely low wages."

The resulting poverty that overwhelmed the Black community caused many families to live in slums where poor sanitation led to widespread health problems. White Memphians then blamed the spread of disease on Blacks.

In addition to limited economic opportunities for Blacks, Memphis' police force, led by Public Safety Commissioner Joseph Boyle, brutalized Black men and women. In the early 1940s, Boyle initiated a campaign to rid Memphis of "idle" Blacks, making it clear that migrant Blacks should take jobs as laborers or return to the fields outside Memphis, Green said.

Memphis police also regularly picked up Black women and sexually assaulted them, Green said. In 1945, in response to one particularly egregious instance where two white police officers raped two young, Black cafeteria workers, public outrage finally exploded. In a break from tradition, the local newspapers covered the case and called for an end to police brutality. The resulting publicity galvanized efforts by Blacks and moderate whites for reform during and after the war.

Beale Street was a cultural center for Black Memphians during the 1940s, 1950s and 1960s. Here, hundreds gather to enjoy a parade.

"Memphis was a city filled with contradictions during the 1940s and 1950s," said Green. "For instance, Blacks in Memphis could vote while those in neighboring Mississippi were denied this right. However, as most Black Memphians could not afford the poll tax that was in effect until 1951, their votes were controlled by the Crump machine, which paid the tax on their behalf."

The cultural landscape of Memphis also was a mass of contradictions. While access to most theaters, restaurants and businesses was segregated if not completely barred, Beale Street flourished as an urban cultural center for the Black community. Blacks would dress up and head to Beale Street, still an active center of culture in Memphis, on the weekends and enjoy cafes, movie theaters, dance halls and other entertainment.

The first all-Black-oriented radio station in the nation also was founded in Memphis during the 1940s. This radio station became an important voice in the racial struggle both during and after the war. Battling against this social achievement, however, was one of the nation's strictest media censors. Films such as "Lost Boundaries," which featured a light-skinned Black doctor who "passed" as white so he could practice medicine, were banned in the city. In addition, any movie scenes showing Blacks in roles other than servants, were edited out of the feature.

By the early 1950s, a representative from the regional office of the National Association for the Advancement of Colored People (NAACP) noted in a memorandum that the efforts of the local chapter had been largely squashed by Memphis' white establishment. The Crump machine, backed by the corrupt police department, pressured middle-class Blacks and other community leaders to maintain the status quo.

Event descriptions

Provide a short description of this article.

If no events were coded, why not?

☐ No protests

☐ No events in the US

☐ No events after 1990

Which if any of the above items needs special review? Please add and code below any events in doubt.

Add event +

Mark Done and Go to Next »

Opens New event menus

Opening screen.
Article text can always be seen while coding.
Coders give brief description of the article and either create events or mark reason for no events (no protest, only protests before 1990, only non-US protests)

MAI Interface after events have been created using “add event.” Showing the “basic info” for highlighted event: description, date, location.

Black Protest Project: Testing VersionHome

Logged in as [redacted]Logout

Protest March Planned in Chattanooga Against Racism and Police Brutality

The Tennessee Tribune; Nashville, Tenn. | 2000-03-22 | The-Tennessee-Tribune;-Nashville,-Tenn._2000-03-22_85f0fa87152524813650854c3b3e0f76

Protest March Planned in Chattanooga Against Racism and Police Brutality

To protest the burning of a cross at the home of a black family and other acts of terror against black people, a march against racism and police brutality will be held Saturday, March 18.

"Some people think that racism in the South is a thing of the past," said Lorenzo Lomboa Ervin, acting secretary of the Hoc Coalition Against Racism and Police Brutality, which is sponsoring the march. "There is no new South. The local politicians and police still allow the Ku Klux Klan and other white racists to go unpunished for sowing terror against people of color."

On February 6, a still unidentified black man allegedly killed a while man, Stephen Davis, in a fight in the racially tense East Lake area of Chattanooga, a predominantly white area to which blacks and Latinos have recently made open threats against the family of the suspected killer, and unidentified whites burned a cross in the yard of the black family, Ervin said.

Last year, 450 black students, supported by their parents, boycotted classes at Central High School after white racists sent threatening letters to several black students at the school. No one was arrested or charged for sending the letters.

In 1996, the homes of the Rev. Amos Baker and Ralph Williams, two black activists who had protested against racism in the city, were burned down in two separate incidents. Complaints were made to the Chattanooga police and the FBI, but no was arrested and no serious investigation was undertaken, Ervin said.

Chattanooga has a long - term reputation for police brutality and murder of poor and black people, according to Ervin. He cited statistics from the U. S. Department of Justice which show that since 1980, the city has ranked first in the nation for police brutality complaints in cities with less than 200,000 people.

"Not one police officers has ever been prosecuted by the local government, even when clearly guilty. The local federal courts have covered up deaths that have taken place in police custody," Ervin charged.

Ervin, Damon McGee and Mikail Musa Muhammad face six months in state prison for protesting against police brutality at a meeting of the Chattanooga City Council on May 19, 1998. The three black activists, known as the Chattanooga 3, were arrested following a demonstration at city hall where over 150 people protested the killings of Montrail Collins and Kevin McCullough. The two black men were shot to death by police in separate incidents on April 28 and may 7, 1998.

The march will assemble at noon at Miller Park on Market Street, then proceed to the Hamilton County Courthouse for a rally and speeches. For information, call (423) 622-7614.

Event descriptions

804

march against racism and police brutality will be held Saturday, March 18, response to cross burning at home of Black family

839

protest against police brutality at a meeting of the Chattanooga City Council on May 19, 1998. where over 150 people protested the killings of Montrail Collins and Kevin McCullough. The Chattanooga

840

Last year, 450 black students, supported by their parents, boycotted classes at Central High School after white racists sent threatening letters to several black students at the school.

Provide a short description of this article.

There will be a protest in response to the burning of a cross at the home of a Black family.

If no events were coded, why not?

☐ No protests
☐ No events in the US
☐ No events after 1990

Which if any of the above items needs special review? Please add and code below any events in doubt.

Add event +

Basic InfoYes/NoText Select

Provide a short description of this event.

protest against police brutality at a meeting of the Chattanooga City Council on May 19, 1998. where over 150 people protested the killings of Montrail Collins and Kevin

Start date (yyyy-mm-dd)

1998-05-19

☐ Start date estimated
☐ Article describes this as a future event

End date (yyyy-mm-dd)

1998-05-19

☐ End date estimated
☐ Article describes this as an ongoing event

State/Province/Territory

Tennessee

City/Town

Chattanooga

☐ No location specified

Duplicate event

☐ I have previously coded this exact event

Which if any of the "Basic Info" items needs special review?

Save

Mark Done and Go to Next »

Report any errors with this interface by raising an issue on [GitHub](#).

Middle tab. Yes/no questions about events. Whether an aggregate or part of a larger event, whether a counter-event, whether anti-Black event, whether coder is uncertain about classification. Article description repeats automatically

Black Protest P: XBlack Protest P: XBlack Protest P: XBlack Protest P: XMAI - Google D: XMAI Logins - G: X+ - X

←→↺↻🏠

🔍📄⋮🔒🌐🔍☰

Black Protest Project: Testing VersionHomeLogged in as █████ Logout

Protest March Planned in Chattanooga Against Racism and Police Brutality

The Tennessee Tribune; Nashville, Tenn. | 2000-03-22 | The-Tennessee-Tribune;-Nashville,-Tenn._2000-03-22_85f0fa87152524813650854c3b3e0f76

Protest March Planned in Chattanooga Against Racism and Police Brutality

To protest the burning of a cross at the home of a black family and other acts of terror against black people, a march against racism and police brutality will be held Saturday, March 18.

"Some people think that racism in the South is a thing of the past," said Lorenzo Lomboa Ervin, acting secretary of the Hoc Coalition Against Racism and Police Brutality, which is sponsoring the march. "There is no new South. The local politicians and police still allow the Ku Klux Klan and other white racists to go unpunished for sowing terror against people of color."

On February 6, a still unidentified black man allegedly killed a white man, Stephen Davis, in a fight in the racially tense East Lake area of Chattanooga, a predominantly white area to which blacks and Latinos have recently made open threats against the family of the suspected killer, and unidentified whites burned a cross in the yard of the black family, Ervin said.

Last year, 450 black students, supported by their parents, boycotted classes at Central High School after white racists sent threatening letters to several black students at the school. No one was arrested or charged for sending the letters.

In 1996, the homes of the Rev. Amos Baker and Ralph Williams, two black activists who had protested against racism in the city, were burned down in two separate incidents. Complaints were made to the Chattanooga police and the FBI, but no was arrested and no serious investigation was undertaken, Ervin said.

Chattanooga has a long - term reputation for police brutality and murder of poor and black people, according to Ervin. He cited statistics from the U. S. Department of Justice which show that since 1980, the city has ranked first in the nation for police brutality complaints in cities with less than 200,000 people.

"Not one police officers has ever been prosecuted by the local government, even when clearly guilty. The local federal courts have covered up deaths that have taken place in police custody," Ervin charged.

Ervin, Damon McGee and Mikail Musa Muhammad face six months in state prison for protesting against police brutality at a meeting of the Chattanooga City Council on May 19, 1998. The three black activists, known as the Chattanooga 3, were arrested following a demonstration at city hall where over 150 people protested the killings of Montrail Collins and Kevin McCullough. The two black men were shot to death by police in separate incidents on April 28 and may 7, 1998.

The march will assemble at noon at Miller Park on Market Street, then proceed to the Hamilton County Courthouse for a rally and speeches. For information, call (423) 622-7614.

Event descriptions

804

✎

📄

march against racism and police brutality will be held Saturday, March 18, response to cross burning at home of Black family

839

✎

📄

protest against police brutality at a meeting of the Chattanooga City Council on May 19, 1998. where over 150 people protested the killings of Montrail Collins and Kevin McCullough. The Chattanooga

840

✎

📄

Last year, 450 black students, supported by their parents, boycotted classes at Central High School after white racists sent threatening letters to several black students at the school.

Provide a short description of this article.

There will be a protest in response to the burning of a cross at the home of a Black family.

If no events were coded, why not?

☐ No protests

☐ No events in the US

☐ No events after 1990

Which if any of the above items needs special review? Please add and code below any events in doubt.

Add event +

Basic Info

Yes/No

Text Select

Characteristics of this event (check all that apply)...

☐ MULTIPLE EVENTS (The "event" is a reference to multiple events e.g. "protests last month" or "protests in 10 cities")

☐ NOT SURE IF A BLACK PROTEST

☐ COUNTER-PROTEST (A response to some other group's protest)

☐ CAMPAIGN (This event is part of a larger campaign or episode)

☐ OPPOSED TO BLACK INTERESTS (This event's participants oppose the interests of black protesters)

☐ NONE OF THESE

Which if any of the "Yes/No" items needs special review?

Save

Mark Done and Go to Next »

Report any errors with this interface by raising an issue on [GitHub](#).

Coding interface third tab, coders mark text in the article with a cursor then click the + by the correct item to mark it as belonging to that field. Text is associated with the highlighted event.

Black Protest P...Black Protest P...Black Protest P...Black Protest P...MAI - Google D...MAI_logins - G...+--

Black Protest Project: Testing VersionHome

Logged in a...Logout

Protest March Planned in Chattanooga Against Racism and Police Brutality

The Tennessee Tribune; Nashville, Tenn. | 2000-03-22 | The-Tennessee-Tribune;-Nashville,-Tenn._2000-03-22_85f0fa87152524813650854c3b3e0f76

Protest March Planned in Chattanooga Against Racism and Police Brutality

To protest the burning of a cross at the home of a black family and other acts of terror against black people, a march against racism and police brutality will be held Saturday, March 18.

"Some people think that racism in the South is a thing of the past," said Lorenzo Lombo Ervin, acting secretary of the Hoc Coalition Against Racism and Police Brutality, which is sponsoring the march. "There is no new South. The local politicians and police still allow the Ku Klux Klan and other white racists to go unpunished for sowing terror against people of color."

On February 6, a still unidentified black man allegedly killed a white man, Stephen Davis, in a fight in the racially tense East Lake area of Chattanooga, a predominantly white area to which blacks and Latinos have recently made open threats against the family of the suspected killer, and unidentified whites burned a cross in the yard of the black family, Ervin said.

Last year, 450 black students, supported by their parents, boycotted classes at Central High School after white racists sent threatening letters to several black students at the school. No one was arrested or charged for sending the letters.

In 1996, the homes of the Rev. Amos Baker and Ralph Williams, two black activists who had protested against racism in the city, were burned down in two separate incidents. Complaints were made to the Chattanooga police and the FBI, but no was arrested and no serious investigation was undertaken, Ervin said.

Chattanooga has a long - term reputation for police brutality and murder of poor and black people, according to Ervin. He cited statistics from the U. S. Department of Justice which show that since 1980, the city has ranked first in the nation for police brutality complaints in cities with less than 200,000 people.

"Not one police officers has ever been prosecuted by the local government, even when clearly guilty. The local federal courts have covered up deaths that have taken place in police custody," Ervin charged.

Ervin, Damon McGee and Mikail Musa Muhammad face six months in state prison for protesting against police brutality at a meeting of the Chattanooga City Council on May 19, 1998. The three black activists, known as the Chattanooga 3, were arrested following a demonstration at city hall where over 150 people protested the killings of Montrail Collins and Kevin McCullough. The two black men were shot to death by police in separate incidents on April 28 and may 7, 1998.

The march will assemble at noon at Miller Park on Market Street, then proceed to the Hamilton County Courthouse for a rally and speeches. For information, call (423) 622-7614.

Event descriptions

804

march against racism and police brutality will be held Saturday, March 18, response to cross burning at home of Black family

839

protest against police brutality at a meeting of the Chattanooga City Council on May 19, 1998. where over 150 people protested the killings of Montrail Collins and Kevin McCullough. The Chattanooga

840

Last year, 450 black students, supported by their parents, boycotted classes at Central High School after white racists sent threatening letters to several black students at the school.

Provide a short description of this article.

There will be a protest in response to the burning of a cross at the home of a Black family.

If no events were coded, why not?

☐ No protests
☐ No events in the US
☐ No events after 1990

Which if any of the above items needs special review? Please add and code below any events in doubt.

Add event +

Basic InfoYes/NoText Select

Highlight text and click + for the relevant category.

Actions by protesters +

protesting x
demonstration x

Actors +

Ervin, Damon McGee a ... Mikail Musa Muhammad x
Chattanooga 3 x

Bystanders +

Issues +

killings of Montrail ... and Kevin McCullough x
police brutality x

Location +

Chattanooga City Council x
city hall x

Named organizations +

Police actions during protest +

arrested x

Size-relevant text +

over 150 people x

Target +

City Council x

Time/duration +

Which if any of the "Text Select" items needs special review?

Save

Mark Done and Go to Next »

Report any errors with this interface by raising an issue on GitHub.

Pass 2 Coding Screenshots Black Protest Project

Form design optimized for being viewable on the
primary coder's laptop.

Task Picker

Production

Navigation Pane

Black Protest Project Tasks

Black Newspaper Pass 2 Coding

Newswire Post Pass 2 Review

Umbrella Backfill

View Links

Duplicate Event Tool

Edit this Event

Find Event

Create Event

Edit this Cluster

Find Cluster

Create Cluster

Find Article

Quit

Record: 1 of 1

No Filter

Search

Form View

Num Lock

Home screen for pass 2 coding. Coder can view this article, find other articles, examine coder annotations, create a new event or cluster, find previously-recorded events or clusters. If one or more events are associated with the article, those events can be viewed in detail and edited.

Black Protests -- Production

Task PickerPass 2 Coding Management

Pass 2 Coding: Black Newspaper Articles

Article

Students March In Support of Dr. Coleman

Show article detail

Published

4/12/1994

Dateline

SOLR_ID

Columbus-Times;-Columbus,-Ga._1994-04-12_fc3274815b67

Find other articles

Pass 1 Annotations

alana

4/12/1994

Columbus

Pending review

Students march to the Muscogee County school board meeting at the Bradley Library in support of their embattled Principal Dr. Coleman.

Show annotation detail

Check Links

alana

4/12/1994

Columbus

Pending review

A meeting to discuss strategy and to air students' reasons for supporting Dr. Coleman.

Record: 1 of 5

No Filter

Search

Link selected annotation to selected event

Selected annotation is not of interest

All annotations not of interest

Associated Events

Show event detail

Find other events

New event

Find issue clusters

New issue cluster

Record: 1 of 1926

No Filter

Search

Edit selected event

Inactivate event

Unlink from article

Prior Article

Next Article

Done

New Flag

Form View

Num LockFiltered

Full text of the article is available.

Black Protests -- Production

Task Picker Pass 2 Coding Management

Article Student

Published

Pass 1 Annotations

alana 4/12

alana 4/12

Record: 14 of 1 of

Link selected an

Associated Events

Record: 14 of 1 of

Edit select

Prior Article Next Article Done

Detail for article Columbus-Times;-Columbus;-Ga_-1994-04-12_fc3274815b674390f8bad121f87b7c22

SOLR_ID Columbus-Times;-Columbus;-Ga_ Date 4/12/1994 Dateline

Headline Students March In Support of Dr. Coleman

Students March In Support of Dr. Coleman.
By Susie Chandler
Columbus, Ga. -- From the outside looking in, you would have thought it was an event for celebration. Students filed out classrooms,gathered in groups and quietly talked among themselves. However, the event was anything, but one of celebration. The students were making last minute preparations for a march that would take them to a 1:30 p.m. Muscogee County school board meeting at the Bradley Library.
The students had meet the previous night to discuss strategy and to air their reasons for supporting Dr. Coleman. The meeting, scheduled to be held at Benning Hills Park, was changed to a new location because they felt they were being harassed by the media.
Some 45 to 50 people meet at the Modern Free Complex on 5th Avenue. Rev. Richard Jessie, the spearheader of the Save Dr. Coleman campaign, was the main speaker during the pre-planning event.
The students, surrounded by Columbus Police, were escorted to the library. The demonstration was peaceful. Thursday, March 31, was the first day students gave an inkling as to how they would demonstrate their support for their embattled principal. Students left class in twos and gathered in front of their school. The demonstration was peaceful. However when Columbus police and firemen were called in, some students were arrested and at least four firefighters claimed to have been injured from rock throwing.
One female sophomore was allegedly pushed in the side with a night stick by a policeman and other students received minor injuries. Students claim they were attempting to leave the campus, when fire trucks rolled onto campus blocking their path. When the students refused to return to class, the truck drove through the students and the firemen began unrolling the fire hose.
Mr. Leonard Leavell, president of the Columbus Chapter NAACP, warned the firemen against spraying the students with the hose. Dr. Coleman and Mr. Ronzell Buckner were responsible for getting the students back into the school and back to class.
"We had no intentions of doing anyone or anything any harm," said one student. "All we wanted to do was show our principal that we stand behind him and that were were willing to take a stand to let the community know we love Dr. Coleman and want him to remain at Carver. But when the police came in, instead of being cooperative, they were provoking students to respond to them with violence," he continued.
Mr. Buckner, part of the Carver

Close

article detail

other articles

notation detail

check Links

event detail

other events

ew event

issue clusters

issue cluster

New Flag

Record: 14 of 1 of 1926 No Filter Search

Form View Num Lock Filtered

View of basic information from pass 1 coder annotations

Black Protests -- Production

Detail for pass 1 annotation 16043

Coder: Start: Estimated: ☐ End: Estimated: ☐

☐ Duplicate Event ☐

Event:

City: State: Location: Unspecified: ☐

AntiBlack: ☐ Maybe Not Black: ☐ Counter: ☐ Multiple: ☐ Ongoing: ☐ Future: ☐ Campaign: ☒

Article:

No Protests: ☐ Not US: ☐ Pre 1990: ☐

Issue:

Size:

Actor Text:

Actions Text:

Record: 1 of 1926 No Filter Search

Form View Num Lock Filtered

Event creation form. Coders can copy/paste information from pass 1 coding or the article. Open-ended description permits coder to differentiate this event from related events or to explain ambiguities. The coder can always flag a problem that requires further review.

Black Protests -- Production

Task Picker Pass 2 Coding Management

Pass 2 Coding: Black Newspaper Articles

frmNewEvent

Create New Event

Event Name State Event Type

Date Estimated Date ☐ City Event Form

This event -- time

- ☒ All on one day
- ☐ Consecutive days
- ☐ Periodic days
- ☐ Unknown days

Note

This event -- place

- ☒ Single place
- ☐ Multiple places
- ☐ Unspecified places
- ☐ No place -- internet, etc.

Note

Relation to other events

- ☐ Aggregates other events in the data
- ☐ Aggregate, but events might not be in data
- ☐ One part of a larger event on same day in same place
- ☐ One part of larger event on same day in multiple places

Note

Issue

Description

Size Class Minimum Size Maximum Size

Record: 1 of 2312 No Filter Search

Form View Num Lock

Search screen for finding previously-identified events

Black Protests -- Production

Find other Events

Find Events

Begin Date State Reset Criteria Search Now

End Date City

DescLong contains:

Mumia_9902_townhall	2/26/1999	Protest at Town Hall last Friday evening to demand new trial for Mumia Abu-Jamal. Attendees and speakers included Monica	Philadelphia
Mumia_9902_townhall_counter	2/26/1999	Policemen's Benevolent Association held a counter rally in support of the slain officer outside of Town Hall where pro-Mumia	Philadelphia
MillionsforMumia99	4/24/1999	Millions for Mumia campaign slated for April 24 in Philadelphia (coordinated by Monica Moorehead of Workers World)	Philadelphia
Busch99_protests	8/30/1999	Hundreds of people protested after after an apparently mentally ill Hasidic man Gideon Busch was killed by police in	NewYork

Record: 1 of 2312 No Filter Search

Link Event to Selected Annotation Link Event to Selected Article Edit Selected Event Close

Show article detail
Find other articles
Show annotation detail
Check Links
Show event detail
Find other events
New event
Find issue clusters
New issue cluster
New Flag

Record: 1 of 1926 No Filter Search

Form View Num Lock

Finding events from the utility menu.

Production

Task Picker

Find other Events

Find Events

Begin Date State

End Date City

descstd contains:

DescLong contains:

Mumia_170104_hearingprotest	1/4/2017	Protesters gathered outside Philadelphia's criminal justice building after a hearing for the convicted Mumia Abu-Jamal was held	Philadelphia
Mumia000507_NY	5/7/2000	A massive protest at Madison Square Theatre Sunday afternoon for freedom for Mumia Abu-Jamal. There was also a	NewYork
Mumia_9902_townhall_counter	2/26/1999	Policemen's Benevolent Association held a counter rally in support of the slain officer outside of Town Hall where pro-Mumia	Philadelphia
MillionsforMumia99	4/24/1999	Millions for Mumia campaign slated for April 24 in Philadelphia (coordinated by Monica Moorehead of Workers World).	Philadelphia

Record: 1 of 114 Filtered Search

Record: 1 of 1 No Filter Search

Form View

Num Lock Filtered

Articles linked to an event can be viewed from the edit event form. “Show selected article” button permits immediate access to full text. Additional articles can be searched for. “Disassociate selected article” is used to correct erroneous links in review or error-correction.

frmEditEvent

Edit Event

Event

Mumia_950703_Philadelphia

State

Pennsylvania

Event Type

Black Event

Date

7/3/1995

Estimated Date

☐

City

Philadelphia

Event Form

protest

Enter New City Record

Enter New Event Form

This event -- time

☒ All on one day

☐ Consecutive days

☐ Periodic days

☐ Unknown days

Note

This event -- place

☒ Single place

☐ Multiple places

☐ Unspecified places

☐ No place -- internet, etc.

Note

Relation to other events

☐ Aggregates other events in the data

☐ Aggregate, but events might not be in data

☐ One part of a larger event on same day in same place

☐ One part of larger event on same day in multiple places

Note

Issue

stay of execution for Abu-Jamal

Description

Protest in Philadelphia, PA against the execution of Mumia Abu-Jamal. Protesters also demanded that Judge Albert Sabo be removed from the case.

Size Class

4

Minimum Size

50

Maximum Size

500

Articles

Sun-Reporter;-San-Francisco,-Calif._1995-07-13_c99bbfffd46c82e3d14c2f6969b4

7/13/1995

Drums In The Global Village: Mumia Abu-Jamal Update

Drums In The Global Village: Mumia Abu-Jamal Update.
developments in

*

Record: 14

1 of 1

No Filter

Search

Show selected article

Search for other articles

Disassociate selected article

Save and Close

Cancel

Find article form linked from edit event form permits linking article to event. This option is generally used for correcting Pass 2 coding when errors are found in a later state of review, processing, or analysis.

Find other articles

Find Articles

Begin Date

End Date

Headline contains

Full text contains:

Reset Criteria

Search Now

AGW_APW_ENG_19941112.0003	11/12/1994	Protesters Urge Clinton To Order Cleanup at Former U.S. Bases		About 100 Fill U.S. Embassy President Bill
AGW_APW_ENG_19941112.0114	11/12/1994	Anti-U.S. Protesters Stopped By Police in Manila		Shouting "Clinton Home!" hundreds toward the pr
AGW_APW_ENG_19941216.0472	12/16/1994	Panel Recommends Intervention To Boost Currency; Unions Set Strike	PARAMARIBO, Suriname (AP) (BC-Suriname-Economy)	A panel recor government i weakening lo
AGW_APW_ENG_19941227.0091	12/27/1994	U.S. Soldiers Occupy Main Plaza Following Deadly Firefight		U.S. soldiers a occupying a n

Show Full Text

Link article to selected event

Close

Record: 1 of 1

No Filter

Search

Clusters are defined and the coder can search for events to add to the cluster. The remove event button allows the work to be corrected.

The screenshot shows a software interface with a window titled "Black Protests -- Production". Inside, a modal form titled "Edit Issue Cluster" is open. The form contains the following fields and sections:

- Name:** DetroitPublicSchools99
- Place:** Detroit Michigan
- Time:** 1999
- Description:** Series of protests and public statements against Gov Engler's proposed takeover of the Detroit Public Schools which are failing.
- Specific Issue:** Gov Engler's proposed takeover of the Detroit Public Schools which are failing.
- Contained Events:** A list of events with columns for event name, location, state, and date.

Event Name	City	State	Date
DetroitSchoolsBoardProtest990414	Detroit	Michigan	4/14/1999
Thirty or more Detroit students marched in front of Murray Wright High School Wednesday, April 14 before the first session of Mayor Dennis Archer's appointed School Board scheduled to meet. As they marched, they chanted, "This corporate board of the Detroit schools"			
DetroitPublicSchool9902_MACstatement	Detroit	Michigan	2/1/1999

At the bottom of the form, there are buttons for "Find Events", "Add Event from Parent Form", "Remove Event from List Above", "Save and Close", "Save and Stay", and "Cancel". The background shows a sidebar with "Navigation Pane" and a main area with "Article" and "Associated Events" sections.

Searching for events to add to the cluster.

Task Picker

Find other Events

Article

Published

Pass 1 Anno

alan

alan

Record:

Link sel

Associated Ev

Record:

Ed

Prior A

Find Events

Begin Date State

End Date City

DesLong contains:

Mumia_9902_townhall

2/26/1999

Protest at Town Hall last Friday evening to demand new trial for Mumua Abu-Jamal. Attendees and speakers included Monica

Philadelphia

Mumia_9902_townhall_counter

2/26/1999

Policemen's Benevolent Association held a counter rally in support of the slain officer outside of Town Hall where pro-Mumia

Philadelphia

MillionsforMumia99

4/24/1999

Millions for Mumia campaign slated for April 24 in Philadelphia (coordinated by Monica Moorehead of Workers World)

Philadelphia

Busch99_protests

8/30/1999

Hundreds of people protested after after an apparently mentally ill Hasidic man Gideon Busch was killed by police in

NewYork

Record: 1 of 2312

No Filter

Search

Add Event to Cluster

Edit Selected Event

Close

Record: 1 of 1

No Filter

Search

Save and Close

Save and Stay

Cancel

ew Elag

Form View

Num Lock

Searching for previously-defined clusters and viewing cluster information. Clusters can be edited or inactivated.

Black Protests -- Production

Task Picker Pass 2 Coding Manager Search for Issue Clusters

Find Issue Clusters

Any field contains:

Search Now

Search Results

DetroitPublicSchools99 Detroit Michigan 1999	Series of protests and public statements against Gov Engler's proposed takeover of the Detroit Public Schools which are failing. Gov Engler's proposed takeover of the Detroit Public Schools which are failing. DetroitSchoolsBoardProtest990414 DetroitPublicSchool9902_MACstatement DetroitPublicSchool
NationalGunViolence96 United States; Washington, D 1996	National protests and pledges against gun violence in 1996 gun violence in the U.S., especially in schools and among youth NationalGunViolence960930_DCShoesMarch NationalGunViolence961010_pledge

Record: 1 of 212 No Filter Search

Edit Selected Cluster Inactivate cluster Close

Record: 1 of 1 No Filter Search

Prior Article Next Article Done New Elag

Form View Num Lock

Navigation Pane

Article: Students March In
Published: 4/12/1994
Pass 1 Annotations
alana Column
4/12/1994 Pendi
alana Column
4/12/1994 Pendi
Record: 1 of 5
Link selected annotation to
Associated Events
Record:
Edit selected event

Coders can edit the issue cluster definition and add or remove events from an issue cluster

frmEditCluster

Edit Issue Cluster

Name	NYSchoolsCellPhones		
Place	New York, NY	Time	April and May 2006
Description	Several protests and vague references to future protests in response to the New York public school system banning cell phones in all of their schools and allowing NYPD to step in for some discipline.		
Specific Issue	New York public school system's "scared straight" policies		

Contained Events

NYSchoolsCellPhoneBan060502	NewYork	New York	5/2/2006
Protest by high school students in Brooklyn, including members of the Trust Us Coalition and with support of members of 100 Blacks in Law Enforcement, against the unilateral cell phone ban in NY public schools.		a citywide rule banning cell phones in public schools	
NYSchoolsCellPhoneBan060412	NewYork	New York	4/12/2006

Record: 1 of 3 No Filter Search

Find Events Add Event from Parent Form Remove Event from List Above

Save and Close Save and Stay Cancel

Umbrella tool is used for linking events that are parts of complex events. (Using the term umbrella for complex events makes communication within the project clearer.) Workflow begins with an event, which can be linked to an existing umbrella or be the basis for creating a new one. In this screenshot, the event is not yet linked to an umbrella, but was tagged in prior coding as being part of a larger event so enters this workflow.

Task PickerEvent and Cluster CurationUmbrella Workflow

Production

Create and Edit Event Umbrellas

Event

Louima020303

City

NewYork

State

New York

Show more event detail

Description

About 150 marchers gathered Sunday at a federal courthouse to protest an appeals court's decision to overturn the convictions of three white

Edit this event

Date

3/3/2002

Multi date

Estimated Date

Event Type

Black Event

Event Form

march rally

Create umbrella from this event

This event -- time

☒ All on one day

☐ Consecutive days

☐ Periodic days

☐ Unknown days

Note

This event -- place

☒ Single place

☐ Multiple places

☐ Unspecified places

☐ No place -- internet, etc.

Note

Relation to other events

☐ Aggregates other events in the data

☐ Aggregate, but events might not be in data

☒ One part of a larger event on same day in same place

☐ One part of larger event on same day in multiple places

Note

Articles

AGW_APW_ENG_20020303.1088	3/3/2002	Marchers protest appeals court's action in Louima torture case AP Photos	
Philadelphia-Tribune;-Philadelphia,-Pa.,_2002-03-	3/5/2002	Leaders react to Louima judgment	

Show selected article

Events linked to this article

Search for other articles

Umbrellas containing this event

Search for other umbrellas

Create new umbrella

Records: 1 of 343

No Filter

Search

Edit this umbrella

Remove event from this umbrella

Inactivate this umbrella

Next Event

Prior Event

Singlet

Done

New Flag

Form View

Num Lock

Define a New Umbrella

Name

Type

Corresponding Event

Inactive

☐

Place info

Time info

Description

Coding Notes

Save Umbrella and Close

Save Umbrella and Stay

Cancel

Edit Umbrella

Name

Louima020303

Type

Same day, same place

Corresponding Event

Inactive

☐

Place info

New York City

Time info

3/3/2002 March 3, 2002

Description

Multiple protests on March 3, 2002, about appeal court overturning convictions of police who stood by while police officer tortured Abner Louima

Coding Notes

Contained Events

Record: << < > >> >>> No Filter Search <

Find Events

Add Event from Parent Form

Remove Event from List Above

Done

Cancel

Umbrellas are tagged by time type (one day, consecutive days, periodic days, unknown) with a note field, a place type (single, multiple, unspecified, no place), and may correspond to aggregate events in the data.

Edit Umbrella

Edit Umbrella

Name

125thStVendors_94

Type

Multiple days, same place

Corresponding Event

125thStVendors_94

Inactive

☐

Place info

New York Harlem

Time info

October 1994 Tuesday 10/18/94 thru Friday 10/21/94

Description

Picketing Tuesday 10/18 thru Friday 10/21 by street vendors displaced from 125th Street in Harlem against stores on the street.

Coding Notes

Contained Events

	10/18/1994	NewYork	New York	Inactive link	☐	ID	1
for boycott by street vendors on 125th Street blocking access to stores in NY in October of 1994 after 1000+ vendors were							
	10/21/1994	NewYork	New York	Inactive link	☐	ID	1268
cking access to stores and call for boycott on Friday, by street vendors on 125th Street in NY in October of 1994 4th day							

Record: 1 of 2No FilterSearch

Find Events

Add Event from Parent Form

Remove Event from List Above

Done

Cancel

Umbrella Search Form

Search Umbrellas

Find Umbrellas

Name or description contains:

Search Now

Same day, same place ☒

Multiple days, same place ☒

No physical place ☒

Same day, multiple places ☒

Multiple days, multiple places ☒

Place info contains:

New York

Time info contains:

125thStVendors_94

Multiple days, same place

Umbrella Event:

1

Inactive? ☐

Coder

peo

Picketing Tuesday 10/18 thru Friday 10/21 by street vendors displaced from 125th Street in Harlem against stores on the stre

New York Harlem

October 1994 Tuesday 10/18/94 thru Friday 10/21/94

BETMTVSatProtsets07

Multiple days, multiple pla

Umbrella Event:

Inactive? ☐

Coder

peo

Weekly Saturday protests outside the homes of Viacom executives in Washington DC and New Rochelle New York

Washington DC and New Rochelle New York

every Saturday between September and November 2007

Binghamton961015

Same day, same place

Umbrella Event:

Inactive? ☐

Coder

nen

Record: 1 of 32

Filtered

Search

Include Selected Event in this Umbrella

Edit Selected Umbrella

Inactivate Umbrella

Close

Record: 1 of 1

No Filter

Search

Specialized tool merges the records of two events determined to be duplicates, associating all article links and pass 1 coding from the inactivated event with the surviving event.

Task Picker

Production

frmDuplicateEventTool

Duplicate Event Tool

First Event

ID

descstd

Search for Event

Inactivate this event

Edit this event

Description

Linked articles

Record: 1 of 1

No Filter

Search

Second Event

ID

descstd

Search for Event

Inactivate this event

Edit this event

Description

Linked articles

Record: 1 of 1

No Filter

Search

Done

Navigation Pane

Form View

Num Lock

Utility for obtaining linked lists. Used for sorting out complicated episodes.

Production

Task PickerViewer

Navigation Pane

Linked Record Viewer

Cluster -> Events -> Articles

Cluster IDCluster NameShow Report

Events -> Articles

Event IDdescstdShow Report

Events -> Clusters

Event IDdescstdShow Report

Article -> Events -> Clusters

SOLR_IDShow TextShow Report

Quit

Record: 1 of 1No FilterSearch

Form ViewNum Lock

The database can be queried, and files exported for analysis.

Post Pass 2 Review

Search...

- tblCoderLines
- tblCrosswalk
- tblDispositionCodes
- tblEventToUmbrella
- tblFormCodes
- tblIssueCodes
- tblPass1Coders
- tblPass1ToEvent
- tblSizeAnalysis
- tblSizeCategories
- tblSnapshot
- tblStates
- tblTimeStampActions
- tblTypeCodes
- tblUmbrellas
- tblUmbrellaTypes
- Copy Of qryCClustersFromArticles
- q_crosswalk_descstd
- q_events_newswire
- q_newswire_clusters**
- q_newswire_xwalk_events

Task Picker

q_newswire_clusters

q_events_newswire

- EventID
- descstd
- locationstd
- EventDate
- masterev

tblEventToCluster

- fkEvent
- fkCluster
- Inactive

tblClusters

- ID
- ClusterName
- ClusterType
- ClusterDescription
- SpecificIssue

Field: descstd

Table: q_events_newswire

Sort:

Show:

Criteria:

or:

Field:	Table:	Sort:	Show:	Criteria:	or:
descstd	q_events_newswire		☒		
	tblClusters.*		☒		
	tblClusters		☒		
	ClusterType		☒		
	tblClusters		☒		
	EventDate		☒		
	q_events_newswire		☒		
	fkType		☒		
	q_events_newswire		☒		
	DescLong		☒		
	q_events_newswire		☒		
	MinSizeEst		☒		
	q_events_newswire		☒		
	MaxSizeEst		☒		
	q_events_newswire		☒		
	IssueSpec		☒		
	q_events_newswire		☒		

Form View

Num Lock

SQL

Post Pass 2 Review

Search...

- tblSizeCategories
- tblSnapshot
- tblStates
- tblTimeStampActions
- tblTypeCodes
- tblUmbrellas
- tblUmbrellaTypes
- Copy Of qryCClustersFromArticles
- q_crosswalk_descstd
- q_events_newswire
- q_newswire_clusters**
- q_newswire_xwalk_events
- q_xwalk_descstd_newswire_active
- qrrUmbrellaTest
- qryActiveClusters
- qryActiveEvents
- qryActiveIssueClusters
- qryActivelyFlaggedBNPCoderLines
- qryAllEventsCluster
- qryArticleEventCluster
- qryArticlesFromUnresolvedFlags
- qryArticlesWithBNPFlags

Task Picker

q_xwalk_descstd_newswire_active

q_newswire_xwalk_events

q_newswire_clusters

descstd	ID	ClusterName	tblClusters.	ClusterDesc	SpecificIssue	ClusterTime	ClusterPlace	tblClusters.I
125thStVendors_94	274	125thStreet	1	In Dec 1995 Har	Conflicts in Har	1994-5	New York	☐
125thStVendorsLawsuit	274	125thStreet	1	In Dec 1995 Har	Conflicts in Har	1994-5	New York	☐
AdamsMark_boyco	275	AdamsMark	1	2000-2001 multi	Adams Mark hc	2000-2001	Multi	☐
AdamsMark_Memphis	275	AdamsMark	1	2000-2001 multi	Adams Mark hc	2000-2001	Multi	☐
AdamsMarkDaytona01	275	AdamsMark	1	2000-2001 multi	Adams Mark hc	2000-2001	Multi	☐
AffActCA96	281	UCAffAct	1	campaign about	Whether to ab			☐
AffActCA96_picketgov	281	UCAffAct	1	campaign about	Whether to ab			☐
AffActCA96_RNC	281	UCAffAct	1	campaign about	Whether to ab			☐
AnnArbor0106_conf	1228	AnnArbor_Con	1	A student activ	Student activis	2001	Ann Arbor Mich	☐
AntiCuban00_bananas	330	ElianGonzales	1	Two protests b	Opposition to c	2000	Miami	☐
AntiCuban000506	330	ElianGonzales	1	Two protests b	Opposition to c	2000	Miami	☐
AntiGay_Atlanta04	276	AntiGaySatlan	1	Atlanta Georgi	Pro- & anti-gay	2004	Atlanta	☐
AtlantaAntiGay04	276	AntiGaySatlan	1	Atlanta Georgi	Pro- & anti-gay	2004	Atlanta	☐
AtlantaAntiGay04_counter	276	AntiGaySatlan	1	Atlanta Georgi	Pro- & anti-gay	2004	Atlanta	☐
AntiGiuliani000420	328	Dorismond	1	Series of prote	Police killing o	2000	New York City	☐
antiimmig_SantaClarita100116	384	SantaClarita10	1	City official cal	City official in	2010	Santa Clarita Cr	☐
antiimmig_SantaClarita100126	384	SantaClarita10	1	City official cal	City official in	2010	Santa Clarita Cr	☐
MLKday_CT_anti	314	CT_MLK00	1	Rally by White	White suprem	2000	Wallingford CT	☐
AntiMosque0607	721	FL_AntiMosque	1	A couple of prc	against a mosq	2006	Pompano Beac	☐
antiObama_racist100320	624	ACA10_racist	1	This is racist im	Racist language	2010	multiple	☐
LosAngeles05_commission2	603	LosAngeles05_	1	Protests about	police killing o	2005	Los Angeles CA	☐
Ashe_memorial	285	Ashe	1	Ceremony hon	Confederate m	1995	Richmond VA	☐
Ashe_memorial_counter	285	Ashe	1	Ceremony hon	Confederate m	1995	Richmond VA	☐
Atlanta_beg05	278	Atlanta_beg05	1	Series of prote	Proposal to bai	2005	Atlanta	☐
Atlanta_beg050814_camp	278	Atlanta_beg05	1	Series of prote	Proposal to bai	2005	Atlanta	☐
Atlanta_beg05_council	278	Atlanta_beg05	1	Series of prote	Proposal to bai	2005	Atlanta	☐
Atlanta_beg050815	278	Atlanta_beg05	1	Series of prote	Proposal to bai	2005	Atlanta	☐
AtlanticCity000308	362	NJPProfiling	1	Several protes	Racial profiling	1999-2000	New Jersey	☐

Record: 1 of 1242

No Filter

Search

Num Lock

SQL

Exported spreadsheets can be read by statistical packages. Analysis files can have different units of measurement: events, articles, clusters, article-event records.

Export - Excel Spreadsheet

?

×

Select the destination for the data you want to export

Specify the destination file name and format.

File name:

\\Client\CS\q_newswire_clusters.xlsx

Browse...

File format:

Excel Workbook (*.xlsx)

▼

Specify export options.

We will not import table relationships, calculated columns, validation rules, default values, and columns of certain legacy data types such as OLE Object.

Search for "Import" in Access 2016 Help for more information.

☐ Export data with formatting and layout.

Select this option to preserve most formatting and layout information when exporting a table, query, form, or report.

☐ Open the destination file after the export operation is complete.

Select this option to view the results of the export operation. This option is available only when you export formatted data.

☐ Export only the selected records.

Select this option to export only the selected records. This option is only available when you export formatted data and have records selected.

OK

Cancel