Comparing GDELT and ICEWS event data

Article · January 2013

CITATIONS
16

READS
1,457

6 authors, including:

Andreas Beger
Duke University
13 PUBLICATIONS **35** CITATIONS
SEE PROFILE

Cassy Dorff
University of New Mexico
7 PUBLICATIONS **81** CITATIONS
SEE PROFILE

Benjamin J Radford
6 PUBLICATIONS **20** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Proceedings of the American Statistical Association 2018 Symposium on Data Science and Statistics View project

Comparing GDELT and ICEWS Event Data

Michael D. Ward & Andreas Beger & Josh Cutler & Matthew Dickenson & Cassy Dorff & Ben Radford

October 15, 2013

In August, there were 150,000 views of a map of protest activity around the world, based on the GDELT database. These are event data, a type of data invented in the mid-1960s by Charles McClelland, who aimed at creating a way to study diplomatic history in a systematic way.[1] From WEIS, through COPDAB, CREON, and many others, event data collections have long served the policy and academic community as a working sensor, revealing details about political interactions among and within countries.[2]

In addition to global coverage, some data sets used random samples. Others focused on specific domains of behavior, such as resource nationalism. What is different now is that rather than having armies of students collect these data, it can be done automatically, using simple, but powerful, computer programs that scan text and determine the action and actors involved. Prior efforts had relied on human coding of compiled chronologies. Philip A. Schrodt was responsible for the first program (called KEDS) that automated content analysis of textual information to create event data.[3]

CAMEO–a coding scheme descendant of KEDS–serves as the coding basis for ICEWS and more recently for GDELT, a "global database of events, language, and tone." GDELT has been introduced in the past year and has generated a large amount of excitement in the policy and academic community. GDELT is well described elsewhere, and has the great benefit of being both open source and continuously updated, permitting its widespread use in academic as well as policy studies. The repository site (`http://gdelt.utdallas.edu/`) contains links to many articles covering GDELT, the complete GDELT documentation, computer programs that have been used to analyze the data, as well as the actual data. According to recent reports, GDELT now includes approximately 250 million events, dated from 1979 to the present.[4]

ICEWS is an early warning system designed to help US policy analysts predict a variety of international crises to which the US might have to respond. These include international and domestic crises, ethnic and religious violence, as well as rebellion and insurgency. This project was created at the Defense Advanced Research Projects Agency, but has since been funded (through 2013) by the Office of Naval Research.[5] ICEWS began as a 4-year DARPA program in 2007 to demonstrate the potential of using social science models and the-

This research was undertaken at `mdwardlab.com` at Duke University. We thank ICEWS colleagues Liz Boschee and Mark Hoffman who gave helpful feedback and guidance. This was partially supported by ONR contract N000014-12-C-0066 to Lockheed Martin's Advanced Technology Laboratories and by NSF Grants SES-1259190 and SES-1259266.

[1] See "The Acute International Crisis," *World Politics*, Volume 14, Special Issue 01, October 1961, pages 182-204. See also: Justin Grimmer and Brandon M. Stewart. "Text as data: The promise and pitfalls of automatic content analysis methods for political texts." *Political Analysis* 21.3 (2013):267-297; Robert C. North, Ole R. Holsti, George Zaninovich, and Dina A. Zinnes. *Content analysis: A handbook with applications for the study of international crisis*. Vol. 184. Evanston, IL: Northwestern University Press, 1963; and, Deborah J. Gerner, et al. "The analysis of political events using machine coded data." *International Studies Quarterly* 38.1 (1994): 91-119.

[2] COPDAB is introduced in Edward E. Azar, "The conflict and peace data bank (COPDAB) project." Journal of Conflict Resolution 24.1 (1980): 143-152. For CREON see Margaret G. Hermann, Barbara G. Salmore, and Stephen A. Salmore. CREON, a foreign events data set. Beverly Hills: Sage Publications, 1973

[3] Philip A. Schrodt, Shannon G. Davis, and Judith L. Weddle. "Political science: KEDS–a program for the machine coding of event data." Social Science Computer Review 12.4 (1994): 561-587. A more complete, and updated summary is available at `http://eventdata.psu.edu/utilities.dir/KEDS.History.0611.pdf`.

[4] See Phil Schrodt, "GDELT: Global Data on Events, Location, and Tone," a presentation for the Conflict Research Society, Essex University, 17 September 2013 for current details and planned enhancements.

[5] Sean P. O'Brien, "Crisis early warning and decision support: Contemporary approaches and thoughts on future research." *International Studies Review* 12.1 (2010): 87-104. But especially see: Sean P. O'Brien, "A multi-method approach for near real time conflict and crisis early warning." *Handbook of Computational Approaches to Counterterrorism*. Springer New York, 2013. 401-418.

ory to forecast and understand nation-state instability across a range of countries. The program proved successful and spawned 3 component tools: iTRACE (news analytics), iCAST (instability forecasting), and iSENT (sentiment analysis and opinion propagation in social media). While it started with a test bed of twenty-five countries in the US Pacific Command, currently ICEWS gathers data on about 250 countries and territories, excluding the US. However, the forecasting effort only concerns 167 countries. ICEWS researchers decided early on not to model instability in smaller countries and territories, such as the Vatican and Pitcairn Island, even though events may be collected for them.

Four aspects of the ICEWS project are noteworthy: (1) it produces and consumes a very rich corpus of text which is analyzed with powerful techniques of automated event-data production.[6] Indeed, Schrodt was involved in the first phases of the project where the extraction techniques for ICEWS event data were developed; (2) it uses a variety of systematic (mostly statistical) models to generate predictions for five dependent variables that are created outside of the event data process: international and domestic crises, insurgency, rebellion, and ethnic and religious violence. Models, largely based on event data, make predictions for these five variables for each of the 167 countries for six months in advance. These predictions are evaluated for accuracy[7]; (3) the various predictions are averaged using ensemble methods to create an average prediction that is more accurate, with fewer false positives and false negatives, than any of the individual models[8]; and importantly (4) a version of this decision aid has been in use for several years, and has a large number of government users. The Duke team has been a participant in this research and has several recent papers related to our efforts at the models and the statistics behind them.[9]

GDELT and ICEWS are arguably the largest event data collections in social science at the moment. During their brief existence they have also been among the most influential data sets in terms of their impact on academic research and policy advice. Yet, we know little to date about how these two repositories of event data compare to each other. Given the nascent existence of both GDELT and ICEWS event data, it is interesting to compare these two repositories of event data.

[6] Boschee, Elizabeth, Premkumar Natarajan, and Ralph Weischedel. "Automatic extraction of events from open source text for predictive forecasting." *Handbook of Computational Approaches to Counterterrorism*. Springer New York, 2013. 51-67

[7] Michael D. Ward, Nils W. Metternich, Christopher Carrington, Cassy Dorff, Max Gallop, Florian M. Hollenbach, Anna Schultz, & Simon Weschle. "Geographical Models of Crises: Evidence from ICEWS," *Advances in Design for Cross-Cultural Activities, Part I*, CRC Press, edited by Dylan D. Schmorrow and Denise M. Nicholson, 2012, pp. 429-438

[8] Jacob M. Montgomery, Florian Hollenbach, Michael D. Ward. "Improving Predictions Using Ensemble Bayesian Model Averaging," *Political Analysis* 20.3 (2012): 271-291

[9] See Michael D. Ward, Nils W. Metternich, Cassy Dorff, Max Gallop, Florian M. Hollenbach, and Simon Weschle. "Learning from the past and stepping into the future: toward a new generation of conflict prediction." *International Studies Review* 15.4 (2013): in press.

A focused comparison of GDELT and ICEWS data

How to compare different databases? An important dimension when comparing databases is availability. GDELT has since the summer of 2013 been open and freely available. That is a big win for

the policy and academic community. Anyone, including researchers from Walmart, JPMorgan Chase, Goldman Sachs, Barclays, Expedia, the Central Intelligence Agency, the Human Rights Data Analysis Group, and even mdwardlab.com can freely use the data. This is a tremendous achievement and merits both acknowledgment and recognition. ICEWS data are not widely available. The full story of why the data are not publicly available can't be told here, but suffice it to say that the success of ICEWS within the operational community of the US government led to a reversal of policy and the contravention of extensive plans, operational as recently as 2010, to make all the ICEWS data freely available to all users. Thus, at present there are a limited number of researchers that have access to ICEWS event data. Currently, ICEWS event data are available only for government use. There are thousands of users with access to these data through ISPAN and/or the W-ICEWS servers. The real data limitation for research is that these data are being treated as for official use only (FOUO) data at this point and are therefore not available to everyone. While constraining in one sense, that limitation allows the W-ICEWS data to be linked back to the originating full story (English, Spanish, Portuguese, etc) so that the event can be examined within a textual context. This is less important for modeling but for the many users that use iTRACE to maintain situation awareness, having access to the full story is important. The ICEWS license from FACTIVA (and the Open Source Center) allow this for government consumption, but not for redistribution.

A second approach is to look at the goals of each database. The ICEWS event data collection has a traditional approach, but modern mechanisms. The collection tries to accurately reflect the activities among and within nations and their main, political actors. Thus, a fair amount of effort goes into filtering the raw stream of reported stories into a unique stream of events. Stories about the history of violence between, for example, Japan and Korea, during the 1930s are eliminated from the stream of events that apply to the current era, even if they appear in the contemporary press.[10] Also winnowed out are stories about the "war" being waged by the Bank of Japan on the Indian currency, as are the many business and sports stories that use the language of politics to describe contests that fall largely outside the realm of politics.[11] In addition, a large effort went into to refining the actor dictionaries, so that stories could be parsed into precise events among specific actors. While not perfect, this is an important aspect of correctly coding events.

The ICEWS data team improved the CAMEO ontology, largely by resolving overlaps and clarifying guidelines for each extant type of event.[12] In order to gauge the effectiveness of these changes, as

[10] See the spike in conflict found in GDELT between Russia and Afghanistan in 2011. The US was considering undertaking military action in Afghanistan and many stories about Russian involvement in the previous century surfaced.

[11] Stories involving Israel are often written with conflictual language that results in conflict events being created, even when the subject does not involve any conflict.

[12] Explained by Liz Boschee (personal communication): We expanded the codebook with additional guidelines and examples designed to clarify potential ambiguities and to resolve overlap between event codes and subcodes.

well as to provide an assessment of accuracy, an experiment was conducted for ICEWS by Liz Boschee, of BBN. As a comparison of the most recent ICEWS data gathered using advanced, graph-theoretic, natural language processing (NLP) techniques and the amplified ontology ("Serif") with the earlier vintage ("Jabari") coding system, events for four CAMEO codes (14, 17, 18, 19) were randomly selected for each system (3000 total events). These were randomly shuffled and then presented anonymously to trained coders who graded each coding as correct or incorrect. The results show a substantial jump in accuracy, illustrated in Table 1. The original coding algorithms were accurate in fewer than one-half of the randomly selected events, according to the trained coders. However, more than two-thirds of all events were correctly classified using the amplified and elaborated framework for coding to the CAMEO codes. The improved accuracy was accomplished without any loss in the number of correct events produced. Initially, only four CAMEO codes focused on conflictual events were studied, but currently ten codes are being used and all the codes in the entire CAMEO ontology are scheduled for October 2013 completion.[13]

At present, the ICEWS event data go back to 2001 and contain about 30 million "stories" that are parsed and coded using NLP techniques based on word graphs using a specially developed ontology based on CAMEO. These are gleaned from about 6000 sources, but many of these are aggregators of hundreds of other sources. So the number of sources is not really informative. What is useful to know is that these media span international, regional, national, and local sources. Importantly, these are all filtered and subjected to the developed ontology using the NLP techniques developed by BBN. The stability in the rate of collected stories, events, and stories with events is quite remarkable. A modest increase in events and stories is seen in the period from 2000 to about 2003, but thereafter the number of events is fairly constant, as shown in Figure 1. This stability does not characterize the GDELT data, as shown in Figure 2. ICEWS has contracted for data back to 1990, and these data are scheduled to be available and coded with the new ontologies by the end of 2013.

[13] One minor point, that is nonetheless important: the Jabari program was itself an improvement over the TABARI program, and that is currently being used (instead of TABARI) for those codes not coded with the SERIF approach.

Table 1: *Coding Accuracy in Random Sample of 3000 Events, coded differently*

Category	CAMEO Code	Jabari	Serif
Protest	14	42%	86%
Coerce	17	43%	83%
Violence	18& 19	45%	74%
Mean		45%	81%

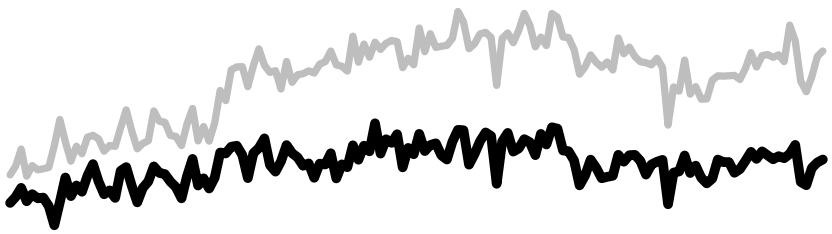


Figure 1: *Stories (in grey) in ICEWS corpus, 1 January 2001 until 30 April 2013. Events harvested from these stories are shown in black. Stories increase a bit over the period, but for the most part, the number of events is relatively stable. About 26 million stores comprise the current ICEWS corpus; there are approximately 16 million events. This averages to about 700 events per country per month.*

Unfortunately, there is no ground truth to use to gauge the accu-

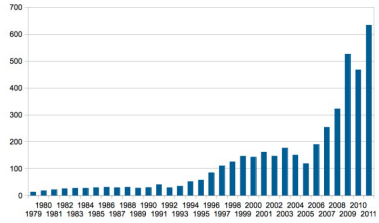
racy of these data. Each data point needs to be assessed by drilling down to the story, reading it, and figuring out if the coding is correct. To do so obviates the goal of automated event coding, but can be useful in identifying errors in that coding. While individual mileage may vary, our experience has been reasonably reassuring to us that generally ICEWS is getting at something real. Users of GDELT doubtless are also convinced that it is getting at something real. Of course, it is impossible to know what stories were not written or even suppressed, and like the well known bias in SIGACTS, events only happen when they get reported.

The GDELT data collection starts from an entirely different philosophy. Rather than trying to get to the "truth" it tries to capture an extensive picture of what is reported, both in its details (who, what, where, when) and its extensiveness (how many reports are there). Therefore GDELT has many more events per country per unit time, since it does not winnow stories extensively. GDELT has about 68,000 country-months (34 years by 167 countries) compared to about 24,000 in ICEWS. Yet, GDELT has an order of magnitude more events. Importantly, the volume of data being harvested by GDELT is growing exponentially, as are the base level of events therein–the density of data is about 100 Giga bytes in 1997 and has grown to over 600 Gb in 2011. GDELT has–at present–by design a collection mechanism that tries to actually maximize reports, but no extensive mechanism for pruning those events to eliminate the false positive reports. It does have a reduced version that we did not use, that limits to one record of each event type between actors per day. ICEWS data, on the other hand, are extensively winnowed and exhibit no corresponding exponential increase, though there is a much smaller time frame involved at present. Indeed, the number of events is relatively stable since 2001 to the present as shown in Figure 1.

We also could, for example, compare the overall correlations for all countries in all time points. If these correlation were really high, it would give to some confidence that both components were measuring the same thing. But, since the two technologies have different goals, this kind of comparison is uninformative. Scholars at Penn State have shown that in total, and for most countries in the Pacific Rim, there are more GDELT events than ICEWS events. These comparisons use an early version of the ICEWS data that is not representative of the techniques currently employed in the generation of event data by the ICEWS team.

While we have no desire to redo the massive comparisons undertaken by the PSU scholars, we found it insightful to perform a more modest comparison of results based upon GDELT and current ICEWS data for an analysis of three countries that have been the sites

Figure 2: *GDELT data density over time in Gigabytes per year. Taken from Phil Schrodt's slide presentation to the Workshop at the Conflict Research Society, Essex University, 17 September 2013 (slide 18).*



See Bryan Arva, John Beieler, Benjamin Fisher, Gustavo Lara, Philip A. Schrodt, Wonjun Song, Marsha Sowell, and Sam Stehle. "Improving Forecasts of International Events of Interest." In EPSA 2013 Annual General Conference Paper, vol. 78. 2013

of contemporary crises.

PROTEST AND DEMONSTRATIONS in Egypt and Turkey, and fighting in Syria provide a specific, small set of interesting cases on which to compare the widely available GDELT data with the latest event data used by the ICEWS project.

We begin with an analysis of Egyptian protest in November of 2011. There were many protests in Cairo, and across the country, aimed at speeding up the reforms one the one hand, and an end to military rule on the other, ideally followed by a quick election and a new constitution. Statements by the military led to massive clashes on the 19th of November, in which many hundreds, including several deaths, were causalities of clashes with the military, especially in Tahrir Square. Clashes continued through November and into December.

Moving ahead one year to 2012, November continues to be a violent month in recent Egyptian history. Around the 18th of November secular, anti-Morsi groups abandoned the constitutional assembly in anticipation of the passage of additional anti-secular laws. Once again Tarhir Square filled with protesters on both sides. Some of these protests were to commemorate the clashes between pro and anti-Morsi forces exactly a year earlier. By the 22nd Morsi began purging judicial officials perceived to be anti-government, and by the 23rd protests and demonstrations were seen not only in Cairo, but throughout Egypt. The rest of 2012 and the first half of 2013 continued to be contentious and by June 2013 Morsi was removed from office by a military coup de état.

Looking at both event streams, GDELT and ICEWS, the signal of increasing protests is evident during the unfolding of the Egyptian Revolution and Aftermath in November of both 2011 and 2012. It is clear that GDELT has more reports of events, but this doesn't mean that there are more events–even if we know that all protests are not reported in the press. ICEWS reports also shows the evolution of protest behavior, but instead of focusing on reports, it focuses on what are purported to be events. The correlation between the two, in this case, indicates that about 2/3 of the variance in these two series is shared (actually 71%). Neither stream is perfect, nor pretends to be.

What is clear is that in 2011 both GDELT and ICEWS pick up the main protests in Egypt, with ICEWS peaking on the 21st and GDELT peaking on the 22nd, but having 200 events reported on the 20th as well. In 2012, GDELT peaks on Friday, the 23rd, and ICEWS the following week on the 27th (a Tuesday). It should be remembered that the GDELT data are growing logarithmically, yet do not appear

Table 2: *Daily Events in Egypt during November 2011 and November 2012*

	Protest Events			
	November 2011		November 2012	
Day	ICEWS	GDELT	ICEWS	GDELT
1	4	23	1	7
2	2	19	1	19
3	2	34	0	7
4	0	20	0	15
5	0	8	0	21
6	0	15	4	10
7	0	7	0	12
8	1	8	1	16
9	0	5	2	10
10	0	13	0	10
11	1	17	0	4
12	0	21	0	14
13	4	20	3	11
14	2	31	2	15
15	2	17	2	28
16	0	25	1	85
17	1	34	2	38
18	5	93	4	14
19	33	130	32	43
20	77	200	23	29
21	**104**	162	14	30
22	72	**204**	13	43
23	40	199	29	180
24	31	161	22	128
25	30	145	20	108
26	20	130	19	85
27	3	88	**40**	153
28	17	88	28	159
29	10	40	8	67
30	8	42	8	72

to be more frequent in Egypt for November 2012 than a year earlier. If we look at the series for order of magnitude changes, the picture is a little different as both GDELT and ICEWS show 2011, November 19th as a breakpoint. In 2012, ICEWS also has the 19th as a tipping point, while GDELT has double digit daily counts over much of the month, but shows a breakpoint on the 23rd.

The accompanying Web page (`http://mdwardlab.com/gdelt-and-icews`) provides a better illustration of these data. Therein you can dynamically examine protests in Egypt and Turkey over the past few years, both in terms of their timeline and geographical distribution. In addition, we have included material conflict for Syria. These displays allow one to compare the ICEWS and the GDELT data visually in these specific cases. As shown numerically in Table 3 GDELT data appear to have a wider range of geolocations than the ICEWS data. Many ICEWS events are geographical disambiguated to central locations, a characteristic that is not shared by the GDELT events. But this pattern is not uniform among all countries, nor among all categories of events. Egypt shows more geographical variance in each country, but the differences are modest, except in Turkey where GDELT shows protests happening in virtually every locale, whereas the ICEWS protest data for Turkey is more concentrated in population centers.

Table 3: *Geographical Variance for ICEWS and GDELT.*

Country	Source	Lat $\hat{\sigma}$	Lon $\hat{\sigma}$
Egypt	ICEWS	0.22	0.38
Egypt	GDELT	0.74	2.25
Syria	ICEWS	0.43	1.06
Syria	ICEWS	0.82	1.21
Turkey	ICEWS	11.37	1.01
Turkey	ICEWS	22.19	1.81

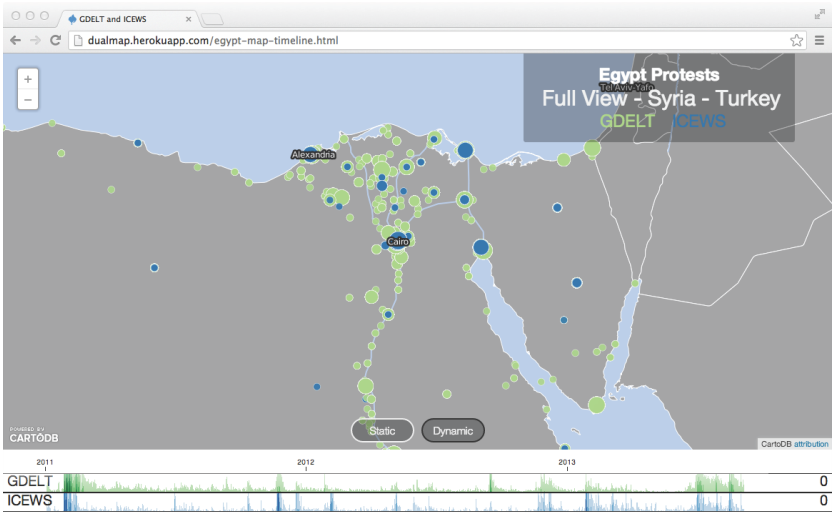


Figure 3: *Interactive comparison of ICEWS and GDELT over time and space for three countries (available at present at* `http://mdwardlab.com/gdelt-and-icews/index.html`*).*

In Turkey, the picture is similarly complicated, as shown in Figures 3&4. Recent protests were widespread, and this will have been widely reported in the Turkish press, but maybe not elsewhere. Recent government estimates have suggested that only four provinces out of 81 remained completely calm in the post-June era.

But most of these protests took place in cities or other population centers, and few events took place in smaller counties. Moreover, both GDELT and ICEWS capture the Kurdish protests (mainly in southeast Turkey), but these protests are not part of the post-June anti-government protests. For example, ICEWS shows the high level of protests in Diyarbakir. These are Kurdish protests nearly all of which took place before the post-June movement, and which voiced the demands of this ethnic minority. These protests are unrelated to the post-June movement. GDELT has few protests in Anatolia, and there were in fact some small protests there in June 2013 and afterward. It appears that ICEWS understates the geographical spread of the recent protests in Turkey, but GDELT may overstate it. Both pick up the Kurdish protests as well as the anti-government protests. The general impression provided to a small group of Turkey experts we asked to compare these two sets of data is that GDELT overstates by a lot the amount of protest, representing protests in areas that are unlikely to have been involved in the Gezi protests. That said, the ICEWS data probably understate the geographical spread of these protest. Table 4 reports these data from for four weeks around the Gezi protest. Figure 4 illustrates that both series pick up the main onset of protests in Turkey, but then ICEWS comes back to a much lower level–an order of magnitude lower–of protest counts by June 15th.

Figure 4: *ICEWS (blue) and GDELT (green) plots of protests during May and June 2013.*

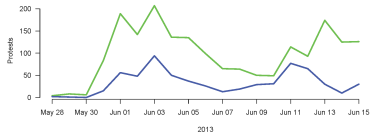


Table 4: *Pre- and Post-Gezi Protests, as reported by ICEWS and GDELT databases.*

Date	ICEWS	GDELT
May 29	1	8
May 30	0	6
May 31	15	83
June 1	56	189
June 2	48	142
June 3	94	207
June 4	50	136
June 5	37	135
June 6	26	99
June 7	13	65
June 8	19	64

What is the take-away from these comparisons?

First, most of the shortcomings of the GDELT data are well known and well established–even if they are ignored by many users and pushers alike. They are well known by the community that creates and uses these data, but largely overlooked by the community that uses creations based on these data.[14] The community that has created these data, and stewards their growing use is well aware of the shortcomings of these data, as well as the strengths. Many different client communities will be able to write filters–perhaps in the form of user friendly widgets–that focus only at some feature of these data. In this way, the GDELT approach of collating and encoding all the printed news, may also serve as a data source for event data encodings that have specific substantive foci, such as human rights abuses or disappearances of political actors. These filters will get good, in short order, at elimination of some of the false positives as well–the historical references that often confuse NLP text encoding. Schrodt noted these data are in BETA, but many treat them as fully finished.

However, it is one thing to have a great data set that is newly avail-

[14] As an example of the wisdom of the community, see Philip Schrodt's analysis: `http://asecondmouse.wordpress.com/2013/05/03/seven-remarks-on-gdelt/`.

able, but has a high rate of error. It is quite another to have to explain to General Dempsey why you woke him up, and find out upon further inspection that it was because of a false positive generated by the data collection algorithm. Thus, it is important to have some sense of the error bands on whatever uses the data are employed to accomplish. Our sense is that the uncertainty on events in ICEWS is less than GDELT, a judgement presaged by the goals of each collection, but validated in research as well. These data serve the modeling goals of the ICEWS research project, at present. That said, the availability of GDELT data is terrific, and we have little doubt that these data can be utilized for similar purposes.

SECOND, even automated approaches to text processing need an ontology from which to construct meaning. The CAMEO framework is a very good one, one that has been improved on considerably over time, and according to Phil Schrodt–the originator of CAMEO–will shortly be supplanted by a new one, PETRARCH. The ICEWS elaboration of the CAMEO ontology undertaken for ICEWS by Elizabeth Boschee is superb, and along with the introduction of advanced NLP techniques produced a substantial improvement in the quality of the data over the prior CAMEO framework we used. Insofar as we know, no other automated coding framework has been examined against the "ground truth" in this way. Without that improvement the accuracy of the coding system as gauged by trained human coders was less than 50% in correctly identifying the type of event. A fifty percent improvement in accuracy is substantial and affects not only false positives, but also false negatives. This evidence undergirds much of our confidence in the ICEWS data.

`https://github.com/eventdata/PETRARCH`

Figure 5: *Map of GDELT Protests in Turkey*

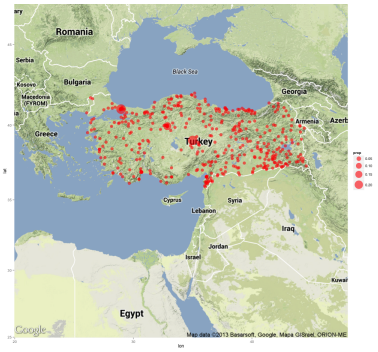


Figure 6: *Map of ICEWS Protests in Turkey*

THE WRONG QUESTION TO ASK is whether ICEWS or GDELT is superior. But more sensible is the question about which data can be usefully applied to what kinds of questions. Are the data complementary? Is one database better at addressing under-reported parts of the world, such as three of the largest countries in the world: China,

India, and Indonesia? And most importantly, what can each database be used to accomplish in an academic as well as policy setting? It is clear to us that both databases pick up major events remarkably well. The volume of GDELT data is very much larger than the corresponding ICEWS data, but they both pick up the same basic protests in Egypt and Turkey, and the same fighting in Syria. GDELT may have 553 protests in Egypt on January 27, 2011 and ICEWS reports only 95, but both give a similar message. Which is correct? Users would like to know the whether erring on the side of of false positives (GDELT) is than the ICEWS strategy of avoiding false positives. Which gets more events correct? Unfortunately, we don't know the answer to this question, but it should be possible to answer.[15] It seems clear, however, that GDELT over-states the number of events by a substantial margin, but ICEWS misses some events as well.

Characteristic of many decision-making problems, the choice is between willingness to be wrong and desire to be right.

[15] We have designed such a study, for which we hope to have results soon.