Introducing ICBe: Very High Recall and Precision Event Extraction from Narratives about International Crises

A Preprint

Rex W. Douglass
University of California, San Diego

Thomas Leo Scherer
University of California, San Diego

J. Andrés Gannon
Vanderbilt University

Erik Gartzke
University of California, San Diego

Jon Lindsay
Georgia Institute of Technology

Shannon Carcelli
University of Maryland

Jonathan Wilkenfeld
University of Maryland

David M. Quinn
University of Maryland

Catherine Aiken
Georgetown University

Jose Miguel Cabezas Navarro
Universidad Mayor

Neil Lund
University of Maryland

Egle Murauskaite
University of Maryland

Diana Partridge
University of Maryland

July 26, 2022

Abstract

How do international crises unfold? We conceptualize of international relations as a strategic chess game between adversaries and develop a systematic way to measure pieces, moves, and gambits accurately and consistently over a hundred years of history. We introduce a new ontology and dataset of international events called ICBe based on a very high-quality corpus of narratives from the International Crisis Behavior (ICB) Project. We demonstrate that ICBe has higher coverage, recall, and precision than existing state of the art datasets and conduct two detailed case studies of the Cuban Missile Crisis (1962) and Crimea-Donbas Crisis (2014). We further introduce two new event visualizations (event icongraphy and crisis maps), an automated benchmark for measuring event recall using natural language processing (sythnetic narratives), and an ontology reconstruction task for objectively measuring event precision. We make the data, online appendix, replication material, and visualizations of every historical episode available at a companion website www.crisisevents.org and the github repository.

***K*eywords** Diplomacy · War · Crises · International Affairs · Computational Social Science

If we could record every important interaction between countries in all of diplomacy, military conflict, and international political economy, how much unique information would this chronicle amount to, and how surprised would we be to see something new? In other words, what is the entropy of international relations? This record could in principle be unbounded, but the central conceit of social science is that there are structural regularities that limit what actors can do, their best options, and even which actors are likely to survive (1, 2). If so, then these events can be systematically measured, and accordingly, massive effort is expended in social science attempting to record these regularities.[1] Thanks to improvements in natural language processing, more open-ended efforts have begun to capture entire unstructured streams of events from international news reports.[2] How close these efforts are to accurately measuring all or even most of what is essential in international relations is an open empirical question, one for which we provide new evidence here.

Our contribution is a high coverage ontology and event dataset for key historical episodes in 20th and 21st-century international relations. We develop a large, flexible ontology of international events with the help of both human coders and natural language processing. We apply it sentence-by-sentence to an unusually high-quality corpus of historical narratives of international crises (1, 29–32). The result is a new lower bound estimate of how much actually happens between states during pivotal historical episodes. We then develop several methods for objectively gauging how well these event codings reconstruct the information contained in the original narrative. We conclude by benchmarking our event codings against several current state-of-the-art event data collection efforts. We find that existing systems produce sequences of events that do not contain enough information to reconstruct the underlying historical episode. The underlying fine-grained variation in international affairs is unrecognizable through the lens of current quantification efforts.

This is a measurement paper that makes the following argument — there is a real-world unobserved latent concept known as international relations, we propose a method for systematically measuring it, we successfully apply this method producing a new large scale set of measurements, those measurements exhibit several desirable kinds of internal and external validity, and those measurements out-perform other existing approaches. The article organizes that argument into eight sections: task definition; corpus; priors/existing state of the art; ICBe coding process; internal consistency; case study selection; recall; and precision. A final section concludes.

[1] See work on crises (3, 4), militarized disputes (5–7), wars (8, 9), organized violence (10, 11), political violence (12), sanctions (13), trade (14), and international agreements (15–17), dispute resolution (17, 18), and diplomacy (19, 20).

[2] See (21); (22); (23); (24); (25); (26). On event-extraction from images and social-media see (27) and (28).

Task Definition

We consider the measurement task of abstracting discrete events about a historical episode in international relations. The easiest way to convey the task is with an example. Figure 1 shows a narrative account of the Cuban Missile Crisis (1962) alongside a mapping from each natural language sentence to discrete machine readable abstractive events. Formally, a historical episode, H, is demarcated by a period of time $[T_{start}, T_{end}] \in T$, a set of Players $p \in P$, and a set of behaviors they undertook during that time $b \in B$. International Relations, IR, is the system of regularities that govern the strategic interactions that world actors make during a historical episode, given their available options, preferences, beliefs, and expectations of choices made by others. We observe neither H nor IR directly. Rather the Historical Record, HR, produces documents $d \in D$ containing some relevant and true (as well as irrelevant and untrue) information about behaviors that were undertaken recorded in the form of unstructured natural language text. The task is to combine informative priors about IR with an unstructured corpus D to produce a series of structured discrete events, $e \in E$, that have high coverage, precision, and recall over what actually took place in history, H.

Conclusion

We investigated event abstraction from narratives describing key historical episodes in international relations. We synthesized a prior belief about the latent unobserved phenomena that drive these events in international relations and proposed a mapping to observable concepts that enter into the observed historical record. We designed an ontology with high coverage over those concepts and developed a training procedure and technical stack for human coding of historical texts. Multiple validity checks find the resulting codings have high internal validity (e.g. intercoder agreement) and external validity (i.e. matching source material in both micro-details at the sentence level and macro-details spanning full historical episodes). Further, these codings perform much better in terms of recall, precision, coverage, and overall coherence in capturing these historical episodes than existing event systems used in international relations.

We release several open-source products along with supporting code and documentation to further advance the study of IR, event extraction, and natural language processing. The first is the International Crisis Behavior Events (ICBe) dataset, an event-level

aggregation of what took place during the crises identified by the ICB project. These data are appropriate for statistical analysis of hard questions about the sequencing of events (e.g. escalation and de-escalation of conflicts). Second, we provide a coder-level disaggregation with multiple codings of each sentence by experts and undergrads that allows for the introduction of uncertainty and human interpretation of events. Further, we release a direct mapping from the codings to the source text at the sentence level as a new resource for natural language processing. Finally, we provide a companion website that incorporates detailed visualizations of all of the data introduced here (www.crisisevents.org).

Acknowledgements

We thank the ICB Project and its directors and contributors for their foundational work and their help with this effort. We make special acknowledgment of Michael Brecher for helping found the ICB project in 1975, creating a resource that continues to spark new insights to this day. We thank the many undergraduate coders for their patience and dedication. Thanks to the Center for Peace and Security Studies and its membership for comments. Special thanks to Rebecca Cordell, Philip Schrodt, Zachary Steinert-Threlkeld, and Zhanna Terechshenko for generous feedback. Thank you to the cPASS research assistants that contributed to this project: Helen Chung, Daman Heer, Syeda ShahBano Ijaz, Anthony Limon, Erin Ling, Ari Michelson, Prithviraj Pahwa, Gianna Pedro, Tobias Stodiek, Yiyi 'Effie' Sun, Erin Werner, Lisa Yen, and Ruixuan Zhang. This project was supported by a grant from the Office of Naval Research [N00014-19-1-2491] and benefited from the Charles Koch Foundation's support for the Center for Peace and Security Studies.

Author Contributions

Conceptualization: R.W.D., E.G., J.L.; Methodology: R.W.D., T.L.S.; Software: R.W.D.; Validation: R.W.D., T.L.S.; Formal Analysis: R.W.D., T.L.S.; Investigation: S.C., R.W.D., J.A.G., C.K., N.L., E.M., J.M.C.N., D.P., D.Q., J.W.; Data Curation: R.W.D., D.Q., T.L.S., J.W.; Writing - Original Draft: R.W.D., T.L.S.; Writing - Review & Editing: R.W.D., J.A.G., E.G., T.L.S.; Visualization: R.W.D., T.L.S.; Supervision: E.G.; Project Administration: S.C., R.W.D., J.A.G., D.Q., T.L.S., J.W.; Funding Acquisition: E.G., J.L.

Bibliography

1. Brecher M (1999) International studies in the twentieth century and beyond: Flawed dichotomies, synthesis, cumulation: ISA presidential address. *International Studies Quarterly* 43(2):213–264.
2. Reiter D (2015) Should We Leave Behind the Subfield of International Relations? *Annu Rev Polit Sci* 18(1):481–499.
3. Brecher M, Wilkenfeld J (1982) Crises in World Politics. *World Politics* 34(3):380–417.
4. Beardsley K, James P, Wilkenfeld J, Brecher M (2020) The International Crisis Behavior Project. doi:10.1093/acrefore/9780190228637.013.1638.
5. Palmer G, et al. (2021) The MID5 Dataset, 2011–2014: Procedures, coding rules, and description. *Conflict Management and Peace Science*:0738894221995743.
6. Gibler DM (2018) *International Conflicts, 1816-2010: Militarized Interstate Dispute Narratives* (Rowman & Littlefield) Available at: `https://books.google.com?id=_4VTDwAAQBAJ`.
7. Maoz Z, Johnson PL, Kaplan J, Ogunkoya F, Shreve AP (2019) The Dyadic Militarized Interstate Disputes (MIDs) Dataset Version 3.0: Logic, Characteristics, and Comparisons to Alternative Datasets. *Journal of Conflict Resolution* 63(3):811–835.
8. Sarkees MR, Wayman F (2010) *Resort to war: 1816-2007* (CQ Press).
9. Reiter D, Stam AC, Horowitz MC (2016) A Revised Look at Interstate Wars, 1816–2007. *Journal of Conflict Resolution* 60(5):956–976.
10. Ralph Sundberg, Mihai Croicu (2016) *UCDP GED Codebook version 5.0* (Department of Peace and Conflict Research, Uppsala University).
11. Pettersson T, Eck K (2018) Organized violence, 1989–2017. *Journal of Peace Research* 55(4):535–547.
12. Raleigh C, Linke A, Hegre H, Karlsen J (2010) Introducing ACLED: An armed conflict location and event dataset: Special data feature. *Journal of peace research* 47(5):651–660.
13. Felbermayr G, Kirilakha A, Syropoulos C, Yalcin E, Yotov YV (2020) The global sanctions data base. *European Economic Review* 129:103561.
14. Barari S, Kim IS Democracy and Trade Policy at the Product Level: Evidence from a New Tariff-line Dataset. 16.
15. Kinne BJ (2020) The Defense Cooperation Agreement Dataset (DCAD). *Journal of Conflict Resolution* 64(4):729–755.
16. Owsiak AP, Cuttner AK, Buck B (2018) The International Border Agreements Dataset. *Conflict Management and Peace Science* 35(5):559–576.

17. Vabulas F, Snidal D (2021) Cooperation under autonomy: Building and analyzing the Informal Intergovernmental Organizations 2.0 dataset. *Journal of Peace Research* 58(4):859–869.
18. Frederick BA, Hensel PR, Macaulay C (2017) The Issue Correlates of War Territorial Claims Data, 1816–20011. *Journal of Peace Research* 54(1):99–108.
19. Moyer JD, Turner SD, Meisel CJ (2020) What are the drivers of diplomacy? Introducing and testing new annual dyadic data measuring diplomatic exchange. *Journal of Peace Research*:0022343320929740.
20. Sechser TS (2011) Militarized Compellent Threats, 1918–2001. *Conflict Management and Peace Science* 28(4):377–401.
21. Li Q, et al. (2021) A Comprehensive Survey on Schema-based Event Extraction with Deep Learning. Available at: `http://arxiv.org/abs/2107.02126` [Accessed September 10, 2021].
22. Halterman A (2020) Extracting Political Events from Text Using Syntax and Semantics.
23. Brandt PT, DOrazio V, Holmes J, Khan L, Ng V (2018) Phoenix Real-Time Event Data. Available at: `http://eventdata.utdallas.edu`.
24. Boschee E, et al. (2015) ICEWS coded event data. *Harvard Dataverse* 12.
25. Hegre H, Croicu M, Eck K, Högbladh S (2020) Introducing the UCDP Candidate Events Dataset. *Research & Politics* 7(3):2053168020935257.
26. Grant C, Halterman A, Irvine J, Liang Y, Jabr K (2017) OU Event Data Project. Available at: `https://osf.io/4m2u7/` [Accessed September 1, 2021].
27. Zhang H, Pan J (2019) CASM: A Deep-Learning Approach for Identifying Collective Action Events with Text and Image Data from Social Media. *Sociological Methodology* 49(1):1–57.
28. Steinert-Threlkeld ZC (2019) The Future of Event Data Is Images. *Sociological Methodology* 49(1):68–75.
29. Brecher M, James P, Wilkenfeld J (2000) Crisis escalation to war: Findings from the International Crisis Behavior Project. *What Do We Know About War*.
30. Wilkenfeld J, Brecher M (2000) Interstate crises and violence: Twentieth-century findings. *Handbook of war studies II*:282–300.
31. James P (2019) What do we know about crisis, escalation and war? A visual assessment of the International Crisis Behavior Project. *Conflict Management and Peace Science* 36(1):3–19.
32. Iakhnis E, James P (2019) Near crises in world politics: A new dataset. *Conflict Management and Peace Science*:0738894219855610.