



Coding Voting Alignment in Multilateral Assemblies With Open Text-as-Data Tools

Valentina Prescod^{1*}, Zanele Suzuki², Adaeze Grigoryan³

¹Department of Business, Congo Research Institute of Business, Bangladesh

²Department of Business, Zambezi Graduate Institute of Business, Tanzania

³Department of Business, Hejaz Graduate Institute of Business, Philippines

Article Info

Volume 1, Issue 1, 2026

[doi: 10.62715/ijir.2026.v1.i1.pp25-30](https://doi.org/10.62715/ijir.2026.v1.i1.pp25-30)

Received: 10 February 2026

Accepted: 09 April 2026

Published: 18 June 2026

Abstract

Measures of how closely states vote together in the United Nations General Assembly and similar bodies are among the most widely used indicators in the quantitative study of international relations. They serve as proxies for foreign policy preferences, for the reach of great powers and for the cohesion of regional groups. Most users take them ready-made, and few are in a position to rebuild them for a different assembly, a different period or a particular subject. This article sets out a procedure for doing so with freely available software. It has seven steps: define the assembly, period and unit of comparison; collect the vote records together with the texts of the resolutions; code each vote choice under explicit rules for abstention, absence and decisions taken without a vote; classify resolutions by topic from their text; compute one or more alignment measures; validate the result against established benchmarks and against speeches; and document and release the data and code. For each step the article explains the choice to be made, the published methods that inform it and the record that should be kept. The procedure is a proposal drawn from the literature. It is not applied to data here and has not been validated; the closing sections state what such a validation would involve.

Keywords: *Voting alignment, United Nations General Assembly, Text as data, Roll-call analysis, Foreign policy similarity, Reproducible research*

© 2026 Valentina Prescod et al. This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

1. Introduction

Scholars have counted votes in the General Assembly almost since it first met. The appeal is plain: nearly every state takes a position on the same question at the same moment, year after year, and the record is public. Lijphart (1963) criticised the early studies of bloc voting and proposed an index of agreement that is still in use. Later work asked whether the alignments of the Cold War survived its end (Kim and Russett, 1996; Voeten, 2000), and voting similarity became a routine control variable in studies of aid, trade and conflict.

The difficulty is that the measure a researcher downloads embodies decisions that someone else made:

* Corresponding author: Valentina Prescod, Department of Business, Congo Research Institute of Business, Bangladesh. E-mail: valentina.prescod@cngo.edu

© 2026 Valentina Prescod et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

which votes to include, how to treat an abstention, whether to correct for agreement that would occur by chance. Those decisions affect conclusions (Häge, 2011). They also tie the researcher to one assembly, when the Human Rights Council, the specialised agencies and regional parliaments raise the same questions (Hug and Lukács, 2014). And the standard measures say nothing about subject matter unless resolutions are first sorted by topic, a task that used to require reading each one.

Open tools for treating text as data have changed what a single researcher can do. This article proposes a stepwise procedure that joins roll-call analysis to automated text classification so that alignment can be coded transparently for any assembly that publishes its votes and documents. The procedure is described and justified; it is not applied here.

2. Background

The procedure draws on two established bodies of method. One concerns how similarity between voting records should be measured; the other, how political texts can be analysed at scale.

2.1. Measuring similarity in votes

The simplest measure is the share of votes on which two states make the same choice. Lijphart's index refines it by counting a pairing of an abstention with a yes or a no as half an agreement (Lijphart, 1963). Signorino and Ritter (1999) showed that rank correlation, then common in the field, confuses similarity with association, and proposed the S score, which is based on the distance between two states' positions scaled by the maximum possible distance. Häge (2011) pointed out that S and raw agreement are inflated when most states vote the same way on most resolutions, as they do in the General Assembly, and recommended chance-corrected coefficients of the kind used to assess agreement between coders.

A different family of methods estimates each state's position on an underlying dimension. Spatial models of roll-call voting were developed for legislatures (Poole and Rosenthal, 1985). Bailey et al. (2017) adapted the approach to the General Assembly with a dynamic model that uses resolutions repeated in identical form across sessions to make positions comparable over time, which dyadic measures cannot guarantee when the agenda changes. An overview of the data and of its uses is given by Voeten (2013).

2.2. Text as data

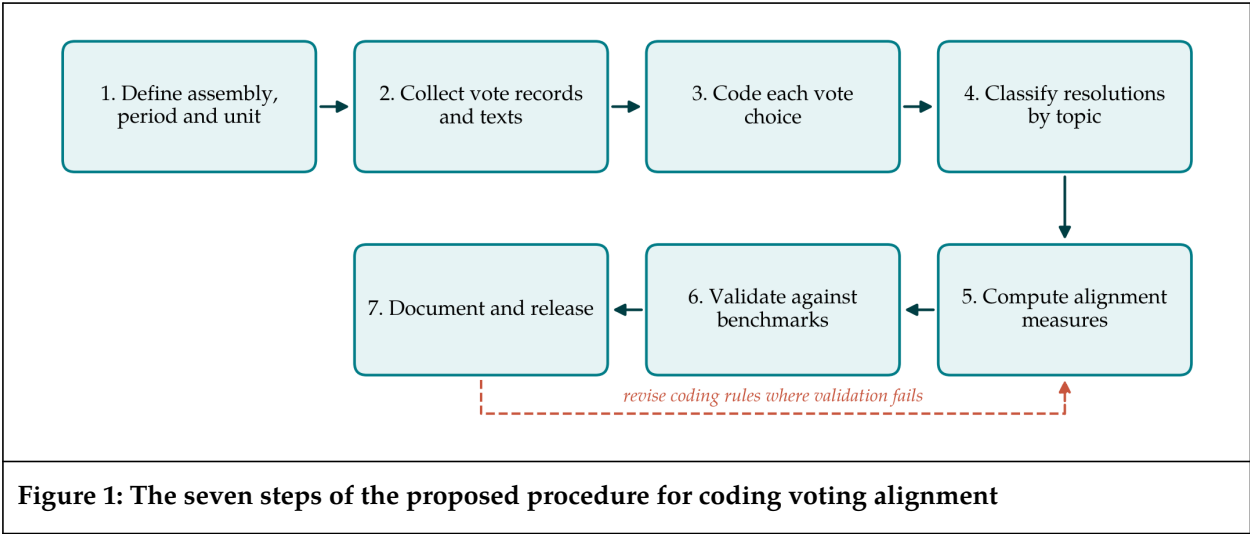
Grimmer and Stewart (2013) set out the principles that govern automated content analysis in political science: all quantitative models of language are wrong, some are useful, and none can be trusted without validation. Methods fall into those that assign documents to known categories, with dictionaries or supervised learning, and those that discover categories, such as topic models. Scaling methods place texts on a dimension using reference documents (Laver et al., 2003) or word frequencies alone (Slapin and Proksch, 2008). The corpus of General Debate statements assembled by Baturo et al. (2017) made such methods usable for the United Nations and offers a source of state positions independent of votes. Open software has matured alongside: the quanteda package handles corpus management and document features (Benoit et al., 2018), and the stm package fits structural topic models (Roberts et al., 2019).

3. The Procedure

Figure 1 shows the seven steps and the point at which the analyst may need to return to earlier choices. Each step is described below with the decision it requires.

3.1. Step 1: Define the assembly, period and unit

The analyst states which body is studied, over which sessions, and what the unit of comparison is: a pair of states in a session, a state relative to a reference state, or a state relative to a group. The rules of the body should be recorded at this point because they determine what a vote means. In the General Assembly each member has one vote, decisions on important questions need a two-thirds majority of members present and voting, and a member in arrears may lose its vote under Article 19 of the Charter.



3.2. Step 2: Collect vote records and texts

Two things are gathered for each recorded vote: the choice of every member, and the full text of the resolution with its symbol, title, date and agenda item. For the General Assembly both are published in the United Nations Digital Library, and the roll-call dataset maintained by Voeten and colleagues can serve as a check on the vote records (Bailey et al., 2017). The symbol is the key that joins votes to text. Votes on separate paragraphs or amendments should be kept and flagged, since they are often more divisive than the vote on the whole.

3.3. Step 3: Code each vote choice

Four outcomes are possible for a member on a recorded vote: yes, no, abstain and absent. The coding rule must say whether abstention is a middle category, as in Lijphart's index and the S score, or a soft no; and whether absence is treated as missing, which is the usual choice, or as informative where a state is known to stay away deliberately. A further rule is needed for the many resolutions adopted without a vote. Leaving them out overstates division, because recorded votes are requested on the contentious items; the general problem that roll calls are a selected sample of decisions was set out for legislatures by Carrubba et al. (2006). Including them as unanimous agreement pushes every similarity score towards its ceiling. The procedure asks the analyst to compute measures both ways and report the difference.

3.4. Step 4: Classify resolutions by topic

Topic classification is where the text tools enter. The resolution texts are loaded into a corpus, tokenised and reduced to a document-feature matrix. Preprocessing choices such as stemming and the removal of rare terms should be made deliberately and reported, since they can change the results of unsupervised models (Denny and Spirling, 2018). Two routes are then open. Where the categories are known in advance, for example the Middle East, disarmament, human rights and development, a dictionary or a supervised classifier trained on a hand-coded sample assigns each resolution. Where they are not, a topic model proposes groupings that the analyst labels after reading the most representative documents (Roberts et al., 2019).

Either way, a random sample is coded by hand by at least two readers and compared with the machine output, with agreement reported using a chance-corrected coefficient (Krippendorff, 2018). Resolutions repeated from year to year with little change should be identified by text similarity at this stage, as they are needed for any dynamic model.

3.5. Step 5: Compute alignment measures

No single measure is correct for every purpose, and the procedure asks for at least two. Table 1 compares the main options. A reasonable default is a chance-corrected dyadic coefficient for questions about pairs

of states and an ideal point estimate for questions about change over time, each computed for all votes and separately within the topics from Step 4.

Table 1: Alignment measures available at Step 5, what each captures and the main caution in using it

Measure	What it captures	Main caution	Source
Agreement index	Share of identical votes, abstention as half agreement	Inflated when most votes are lopsided	Lijphart, 1963
S score	Distance between vote profiles, rescaled	Sensitive to the set of votes included	Signorino and Ritter, 1999
Chance-corrected agreement	Agreement beyond what marginal distributions imply	Result depends on the chance model chosen	Häge, 2011
Dynamic ideal points	Position on one latent dimension, comparable over time	Assumes a single dimension; needs repeated resolutions	Bailey et al., 2017
Text-based position	Position inferred from speeches, not votes	Speeches address different audiences than votes	Baturu et al., 2017

3.6. Step 6: Validate

Validation follows the advice of [Grimmer and Stewart \(2013\)](#) that every automated measure be checked against something outside itself. Three checks are proposed. Where a published benchmark exists, as it does for the General Assembly, the new measure is compared with it over the overlapping period and large discrepancies are traced to their source. The vote-based measure is compared with positions scaled from speeches in the same sessions; the two should be related without being identical, and persistent gaps between what a state says and how it votes are worth examining in their own right. Finally, the highest and lowest scoring pairs in each topic are inspected by a reader who knows the cases. If any check fails, the analyst returns to the coding rules.

3.7. Step 7: Document and release

The final step is to publish the vote file, the topic codes, the scripts and a codebook stating every rule adopted in Steps 1 to 5 and the software versions used. Since the inputs are public documents, there is rarely a legal obstacle to release. A measure that others can rebuild and alter is worth more to the field than a marginally better one that cannot be inspected.

4. Adapting the Procedure to Other Assemblies

The General Assembly is the easy case, because its records are complete and a benchmark exists. Elsewhere the analyst must adapt. In bodies with elected and rotating membership, such as the Human Rights Council, whose voting was studied by [Hug and Lukács \(2014\)](#), the set of voters changes each year, so dyadic measures are available only for the sessions in which both states sit, and ideal point models need observations that bridge one membership to the next. Where most business is settled by consensus, recorded votes may be too few to support any measure, and text-based positions from statements become the main source instead of a check. Documents in several languages should be analysed in one official language throughout to avoid treating translation differences as differences of substance.

5. Limitations and What Remains to Be Shown

This article proposes a procedure and does not demonstrate it. Before it could be recommended without reservation, three things would have to be shown on real data. The first is that the topic classification in Step 4 reaches acceptable agreement with human coders across sessions, since the vocabulary of resolutions drifts over decades. The second is that conclusions are stable under the alternative coding rules of Step 3; if a finding holds only when consensus decisions are excluded, it is a finding about

contested votes and should be described as one. The third is that the topic-specific measures add information beyond the overall measure, which is the main reason for the extra work.

There are limits no validation will remove. A vote records a public position taken under instructions, group pressure and occasionally inducement, and it is evidence of alignment in the assembly and not of a state's private preferences (Voeten, 2013). Abstention and absence are strategic and their meaning varies by state. Automated classification of formulaic diplomatic prose can be confidently wrong. These are reasons to report choices openly and to read the documents, and the procedure is built to make both easier.

6. Conclusion

Voting alignment can be coded from public records with open software by any researcher prepared to make a small number of decisions explicit. The seven steps set out above pair the established measures of vote similarity with text-based classification of what is being voted on, and they place validation and release inside the workflow instead of after it. The procedure awaits application. Its promise is a family of alignment measures that are specific to topic, portable across assemblies and open to correction by those who use them.

Conflict of interest

To be declared by the author(s) before publication.

Funding

To be declared by the author(s) before publication.

References

- Bailey, M.A., Strezhnev, A. and Voeten, E. (2017). [Estimating dynamic state preferences from United Nations voting data](#). *Journal of Conflict Resolution*, 61(2): 430-456.
<https://doi.org/10.1177/0022002715595700>
- Baturo, A., Dasandi, N. and Mikhaylov, S.J. (2017). [Understanding state preferences with text as data: Introducing the UN General Debate corpus](#). *Research & Politics*, 4(2).
<https://doi.org/10.1177/2053168017712821>
- Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S. and Matsuo, A. (2018). [quanteda: An R package for the quantitative analysis of textual data](#). *Journal of Open Source Software*, 3(30): 774.
<https://doi.org/10.21105/joss.00774>
- Carrubba, C.J., Gabel, M., Murrah, L., Clough, R., Montgomery, E. and Schambach, R. (2006). [Off the record: Unrecorded legislative votes, selection bias and roll-call vote analysis](#). *British Journal of Political Science*, 36(4): 691-704.
- Denny, M.J. and Spirling, A. (2018). [Text preprocessing for unsupervised learning: Why it matters, when it misleads, and what to do about it](#). *Political Analysis*, 26(2): 168-189.
- Grimmer, J. and Stewart, B.M. (2013). [Text as data: The promise and pitfalls of automatic content analysis methods for political texts](#). *Political Analysis*, 21(3): 267-297. <https://doi.org/10.1093/pan/mps028>
- Hug, S. and Lukács, R. (2014). [Preferences or blocs? Voting in the United Nations Human Rights Council](#). *The Review of International Organizations*, 9(1): 83-106.
- Häge, F.M. (2011). [Choice or circumstance? Adjusting measures of foreign policy similarity for chance agreement](#). *Political Analysis*, 19(3): 287-305.
- Kim, S.Y. and Russett, B. (1996). [The new politics of voting alignments in the United Nations General Assembly](#). *International Organization*, 50(4): 629-652.
- Krippendorff, K. (2018). [Content Analysis: An Introduction to Its Methodology](#), 4th edn. Sage, Thousand Oaks, CA.

- Laver, M., Benoit, K. and Garry, J. (2003). [Extracting policy positions from political texts using words as data](#). *American Political Science Review*, 97(2): 311-331.
- Lijphart, A. (1963). [The analysis of bloc voting in the General Assembly: A critique and a proposal](#). *American Political Science Review*, 57(4): 902-917.
- Poole, K.T. and Rosenthal, H. (1985). [A spatial model for legislative roll call analysis](#). *American Journal of Political Science*, 29(2): 357-384.
- Roberts, M.E., Stewart, B.M. and Tingley, D. (2019). [stm: An R package for structural topic models](#). *Journal of Statistical Software*, 91(2): 1-40. <https://doi.org/10.18637/jss.v091.i02>
- Signorino, C.S. and Ritter, J.M. (1999). [Tau-b or not tau-b: Measuring the similarity of foreign policy positions](#). *International Studies Quarterly*, 43(1): 115-144.
- Slapin, J.B. and Proksch, S.-O. (2008). [A scaling model for estimating time-series party positions from texts](#). *American Journal of Political Science*, 52(3): 705-722.
- Voeten, E. (2000). [Clashes in the Assembly](#). *International Organization*, 54(2): 185-215.
- Voeten, E. (2013). [Data and analyses of voting in the United Nations General Assembly](#). In: Reinalda, B. (ed.), *Routledge Handbook of International Organization*, Routledge, London.