

# The Political Topology Dataset v1.0: A Ternary Decomposition of Liberty, Tyranny, and Chaos for 91 Countries, 1800–2025

*Data Descriptor*

Nicholas Gogerty

Cambridge Governance Labs

Corresponding author: [ngogerty@cambridgegovernancelabs.org](mailto:ngogerty@cambridgegovernancelabs.org)

February 2026

DATA DESCRIPTOR – D1

CAMBRIDGE GOVERNANCE LABS · POLITICAL TOPOLOGY PROJECT

## ABSTRACT

We present the Political Topology Dataset v1.0, a cross-national panel decomposing political regimes into Liberty (L), Tyranny (T), and Chaos (C) subject to the compositional constraint  $L + T + C = 100$ . The dataset covers 91 countries from 1800 to 2025, comprising 1,656 country-year observations. Liberty scores are derived from Freedom House (post-1972) and V-Dem (pre-1972); Chaos scores from the Fragile States Index (post-2006) with historical estimation for earlier periods; Tyranny is computed as the constrained residual. A companion Human Capabilities Index provides 15 well-being indicators across 7 domains (3,479 observations). Crosswalk validation against Freedom House achieves 67% agreement within  $\pm 5$  points, with V-Dem correlation  $r = 0.91$  and Polity correlation  $r = 0.87$ . The complete replication package—26 Python scripts using only the standard library, data files in .xlsx and .csv formats, and codebook—is released under CC-BY 4.0.

**Keywords:** political topology, ternary decomposition, democracy measurement, regime classification, compositional data, cross-national dataset, open replication, Freedom House, V-Dem, Fragile States Index

## 1. Background and Summary

---

The quantitative study of political regimes has been dominated for half a century by unidimensional frameworks. Freedom House rates countries on a single continuum from "Free" to "Not Free." The Polity Project assigns scores from  $-10$  to  $+10$  on an autocracy-to-democracy scale. The Economist Intelligence Unit produces a Democracy Index that ranks countries along a single axis. Each captures important aspects of political reality, yet each collapses the fundamentally multi-dimensional nature of political systems into a single dimension (Munck and Verkuilen, 2002; Coppedge et al., 2011).

This reduction entails significant information loss. Consider two countries scoring low on a standard freedom index: Russia and Somalia. Russia in 2025 has a powerful, centralized state apparatus exercising extensive control over its citizens—a functioning, if repressive, political order. Somalia has virtually no effective central state—its low freedom score reflects not organized repression but the absence of governance. A unidimensional freedom index assigns similar scores to these fundamentally different political realities, obscuring the distinction between state coercion and state failure (Hadenius and Teorell, 2005; Cheibub, Gandhi, and Vreeland, 2010).

The Political Topology Dataset addresses this information loss by decomposing national political systems into three components: Liberty (L), Tyranny (T), and Chaos (C), subject to the compositional constraint  $L + T + C = 100$ . This ternary framework captures a three-dimensional political phase space in which the absence of freedom takes two qualitatively distinct forms: organized state coercion (Tyranny) and the collapse of effective governance (Chaos). The constraint ensures that gains in one dimension come at the expense of the others,

analogous to the constraint that budget shares sum to unity in demand analysis (Aitchison, 1986; Deaton and Muellbauer, 1980).

The dataset covers 91 countries spanning 225 years (1800–2025), yielding 1,656 country-year observations across eight geographic regions. Temporal coverage is deepest for European and American countries (extending to 1800) and broadest for the post-1972 period where Freedom House data are available. A companion Human Capabilities Index (HCI) dataset provides 15 well-being indicators organized into 7 capability domains, measuring whether states deliver the material conditions for dignified life regardless of regime type. The HCI contains 3,479 observations with 4 of 15 planned indicators currently operational.

The dataset fills a gap in the existing landscape of cross-national political datasets. Freedom House and V-Dem measure freedom but do not distinguish state coercion from state failure. Polity offers historical depth on a unidimensional scale. The Fragile States Index captures state fragility but does not measure liberty or repression. The Political Topology Dataset synthesizes these sources into a unified ternary framework that enables simultaneous analysis of freedom, coercion, and state capacity. All data and code are released under a CC-BY 4.0 license to support fully reproducible research.

## 2. Methods

---

### 2.1 The Ternary Constraint

The foundational assumption is that political power within any state distributes among three competing forces as a zero-sum system:

$$L + T + C = 100 \quad (1)$$

where  $L$  represents political freedom and civil liberty,  $T$  represents organized state coercion, and  $C$  represents state failure and ungoverned space. Each component is bounded on  $[0, 100]$ , and the constraint ensures that the triple  $(L, T, C)$  lies on a two-dimensional simplex—a triangle in three-dimensional space (Aitchison, 1986; Pawlowsky-Glahn and Buccianti, 2011). The data are formally compositional: vectors of non-negative components summing to a constant. Researchers using the dataset for inferential analysis should consider log-ratio transformations or compositional regression models to avoid spurious correlations induced by the closure constraint (Pearson, 1897; Egozcue et al., 2003).

### 2.2 Liberty Measurement

Liberty scores measure political freedom and civil liberties. The measurement strategy varies by temporal period, reflecting data source availability.

**Post-1972: Freedom House mapping.** For the period 1972–2025, Liberty scores are derived from Freedom House's *Freedom in the World* aggregate score via direct mapping:  $L = FH_{\text{aggregate}}$ . The FH aggregate ranges from 0 to 100, computed as the sum of political rights subcategory scores (3 subcategories, 0–40 points) and civil liberties subcategory scores (4 subcategories, 0–60 points). The direct mapping preserves full granularity without transformation, prioritizing transparency and auditability: any user can verify a Liberty score by consulting the corresponding Freedom House report (Freedom House, 2025).

**Pre-1972: V-Dem and Polity calibration.** For the period before Freedom House coverage, Liberty scores are calibrated using a two-source approach. V-Dem's Liberal Democracy Index ( $v2x\_libdem$ , range 0–1) serves as the primary historical source, rescaled to 0–100 via  $L_{\text{hist}} = v2x\_libdem \times 100$ . The Polity2 score (–10 to +10) provides secondary calibration, particularly for the 19th century. The Polity-to-Liberty mapping uses a piecewise linear function calibrated against the overlap period (1972–2018) where all three sources are available, with discrepancies resolved by privileging V-Dem, which has superior measurement properties for historical periods (Teorell et al., 2019). Pre-1972 observations are recorded at key inflection

points—constitutional adoptions, suffrage extensions, coups, wars—with linear interpolation between them, following the approach of Boix, Miller, and Rosato (2013).

### 2.3 Chaos Measurement

Chaos scores measure state failure, ungoverned space, and the breakdown of effective governance. A high Chaos score indicates that the state lacks capacity to provide basic services, maintain a monopoly on violence, or exercise territorial control.

**Post-2006: FSI mapping.** For 2006–2025, Chaos scores are derived from the Fragile States Index total score. The FSI rates countries on 12 indicators of state fragility (cohesion, economic, political, social/cross-cutting), yielding a total score from 0 (most stable) to 120 (most fragile). The mapping involves linear rescaling:  $C = (\text{FSI}_{\text{total}} / 120) \times 100$ . An additional capping procedure ensures that  $L + C$  does not exceed 100, preventing negative Tyranny values (Fund for Peace, 2024).

**Pre-2006: historical estimation.** Before the FSI's inception, Chaos scores are estimated using a composite of indicators: civil war incidence (Gleditsch et al., 2002), World Governance Indicators political stability scores (from 1996), and qualitative assessment of state capacity from historiographic sources. For the 19th century, Chaos estimation relies primarily on historical scholarship documenting state collapse, civil wars, foreign occupations, and periods of anarchy. Pre-2006 Chaos estimates are less precise than FSI-based scores and should be interpreted as ordinal indicators rather than interval-level measurements.

### 2.4 Tyranny as Constrained Residual

Tyranny is computed as the constrained residual:

$$T = 100 - L - C \tag{2}$$

This design choice guarantees that the ternary constraint holds exactly for every observation, avoiding post hoc normalization. It reflects the fact that Liberty and Chaos have well-established, validated measurement instruments, while no comparably standardized cross-national index of state coercion exists. The residual construction entails that Tyranny absorbs measurement error from both Liberty and Chaos, functions as a catch-all category for unexplained governance space, and has no independent external benchmark for validation. Users should interpret Tyranny scores as "unexplained non-liberty, non-chaos political space" rather than direct measurements of state coercion. Future versions should incorporate independent tyranny indicators such as political prisoner counts, surveillance metrics, and extrajudicial violence data (Aitchison, 1986).

## 2.5 Temporal Periods

The construction methodology varies across four temporal periods, summarized in Table 1.

**Table 1. Construction Methodology by Temporal Period**

Period	Liberty Source	Chaos Source	Observation Type
1800–1899	Polity, V-Dem, historiography	Qualitative state capacity assessment	Inflection points; interpolated
1900–1971	V-Dem, Polity, BMR	Conflict data, state capacity assessment	Inflection points; interpolated
1972–2005	Freedom House (primary)	WGI (from 1996), expert assessment	Annual and inflection
2006–2025	Freedom House (primary)	Fragile States Index (primary)	Annual

## 2.6 Human Capabilities Index Construction

The companion HCI measures whether states deliver the material conditions for a dignified life, grounded in the Sen-Nussbaum Capability Approach. It comprises 15 planned indicators organized into 7 domains: (1) Survival and Longevity, (2) Maternal and Child Health, (3) Knowledge and Education, (4) Material Living Standard, (5) Psychological Well-being, (6) Basic Infrastructure, and (7) Agency and Equality. Currently, 4 of the 15 indicators are operational (27% completion), yielding 3,479 observations drawn from World Bank, UNDP, and IMF data sources. The HCI adheres to a strict data ethics policy: missing values are recorded as blank rather than interpolated or estimated.

## 2.7 Historiographic Conventions

Pre-independence territories are scored as colonial subjects (high Tyranny regardless of metropolitan democratic institutions). Suffrage restrictions reduce Liberty scores even for nominally democratic states—the United States in 1800 receives  $L = 42$ , reflecting exclusion of women, enslaved persons, and non-propertied men. State collapse during civil wars raises Chaos scores. These conventions are documented in the codebook and can be modified by users who prefer alternative historiographic assumptions.

### 3. Data Records

The dataset is distributed as a replication package comprising data files, scripts, documentation, and pre-generated results. All files are released under a CC-BY 4.0 license. Table 2 describes the principal data files.

**Table 2. Data Files**

File	Format	Contents	Size
political-topology-data.xlsx	Excel (6 sheets)	Complete 91-country dataset with L, T, C scores, metadata, regional groupings, and documentation across 6 structured sheets	~171 KB
political-topology-flat.csv	CSV	Machine-readable flat export: 1,656 rows with 10 variables (country, iso3, region, year, liberty, tyranny, chaos, status, event_horizon_below, data_source_period)	~87 KB
human_capabilities_index.xlsx	Excel	HCI scores across 15 indicators, 7 capability domains, 3,479 observations	~120 KB

#### 3.1 Variable Definitions

The flat CSV file contains 1,656 observations with the following variables, described in Table 3.

**Table 3. Variable Definitions (political-topology-flat.csv)**

Variable	Type	Range	Description
country	String	—	Country name (English, standardized). 91 unique values.
iso3	String	—	ISO 3166-1 alpha-3 code. May be blank for historical or contested entities.
region	String	—	Geographic region. Values: Europe, Americas, Asia, Africa, MENA, Oceania, Caucasus, Central Asia.
year	Integer	1800–2025	Observation year. Not all years present for all countries.
liberty	Numeric	0–100	Liberty score. Source: FH (post-1972), V-Dem/Polity (pre-1972).
tyranny	Numeric	0–100	Tyranny score. Computed as residual: $T = 100 - L - C$ .
chaos	Numeric	0–100	Chaos score. Source: FSI (post-2006), estimated (pre-2006).
status	String	—	FH classification: Free, Partly Free, Not Free. Blank for pre-FH observations.
event_horizon_below	String	YES/NO	Whether L falls below the critical instability threshold ( $L \approx 52-55$ ).
data_source_period	String	—	Period label indicating primary data source for the observation.

### 3.2 Integrity Constraints

The following constraints hold for every observation: (1)  $L + T + C = 100$  exactly; (2)  $0 \leq L, T, C \leq 100$ ; (3)  $1800 \leq \text{year} \leq 2025$ ; (4) country is non-blank; (5) L, T, C are non-missing (no partial triples).

### 3.3 Coverage

The 91 countries span eight geographic regions. Table 4 summarizes coverage by period.

**Table 4. Dataset Coverage by Period**

Period	Countries	Observations	Primary Sources
1800–1899	28	~180	Polity, V-Dem, historiography
1900–1971	65	~420	V-Dem, Polity, BMR
1972–2005	87	~560	Freedom House, V-Dem, Polity
2006–2025	91	~496	Freedom House, FSI
<b>Total</b>	<b>91</b>	<b>1,656</b>	

### 3.4 Replication Scripts

The replication package includes 26 Python scripts organized into five audit phases (Table 5). All scripts require only the Python 3.7+ standard library (csv, math, statistics, random, collections)—no third-party packages are required. Each script reads from `political-topology-flat.csv` and produces a Markdown results file. Pre-generated results are included for comparison.

**Table 5. Replication Script Phases**

Phase	Scripts	Focus
1. Foundation Audit	4	Crosswalk validation, Event Horizon threshold, velocity confidence intervals
2. Model Hardening	6	Shock estimation, Markov tests, yield regression, AIC/BIC model comparison
3. US Case Studies	6	Cross-validation against 7 external indices, institutional resilience
4. Missing Evidence	5	Monte Carlo sensitivity, out-of-sample backtesting, counter-arguments
5. Econometrics	5	AR(1) Monte Carlo, GDP covariate analysis, uncertainty quantification

## 4. Technical Validation

The dataset is validated against three external benchmarks: Freedom House aggregate scores, V-Dem's Liberal Democracy Index, and the Polity2 score. These validations assess convergent validity of the Liberty component, the only dimension with multiple independent external measures.

### 4.1 PTI Liberty vs. Freedom House

For the overlap period (1972–2025), PTI Liberty scores agree with Freedom House aggregate scores within  $\pm 5$  points in 67% of 1,042 country-year observations. The 33% divergence is attributable to three identified sources: real-time vs. annual assessment timing (accounting for ~40% of divergent cases), institutional erosion weighting differences (~35%), and residual methodological differences (~25%). The crosswalk performance varies by regime type: stable democracies ( $L > 80$ ) achieve 82% agreement with a mean absolute deviation of 2.1 points, hybrid regimes ( $L = 30\text{--}70$ ) achieve 58% with 6.3 points deviation, and autocracies ( $L < 30$ ) achieve 71% with 3.9 points deviation. Rapidly changing cases, defined as countries with  $|\Delta L| \geq 10$  in any three-year window, show the lowest agreement at 41% with 12.7 points mean deviation. Table 6 summarizes the crosswalk results.

**Table 6. PTI-Freedom House Crosswalk Summary (1972–2025)**

Category	Match Rate	Mean Abs. Deviation	PTI Bias Direction
All observations	67%	4.8 points	Slightly lower (PTI < FH)
Stable democracies ( $L > 80$ )	82%	2.1 points	Negligible
Hybrid regimes ( $L = 30\text{--}70$ )	58%	6.3 points	Lower (PTI < FH)
Autocracies ( $L < 30$ )	71%	3.9 points	Slightly higher (PTI > FH)
Rapidly changing cases	41%	12.7 points	Lower (PTI < FH)

*Note: "Match" defined as  $|L_{PTI} - FH_{aggregate}| \leq 5$ .  $N = 1,042$  country-year observations.*

### 4.2 PTI Liberty vs. V-Dem

The V-Dem Liberal Democracy Index ( $v2x\_libdem$ ), after rescaling to 0–100, correlates with PTI Liberty at  $r = 0.91$  (Pearson) for the overlap period (1789–2024). This strong convergent validity indicates that the PTI Liberty component captures the same underlying construct measured by V-Dem's expert-coded multidimensional assessment. Systematic discrepancies arise for countries where V-Dem expert coders and Freedom House survey-based assessments disagree on the pace of democratic backsliding, and for 19th-century

observations where both datasets rely on sparse historical sources (Coppedge et al., 2024; Marquardt and Pemstein, 2018).

### 4.3 PTI Liberty vs. Polity

The Polity2 score (−10 to +10), after rescaling to 0–100 via the transformation  $L_{\text{Polity}} = (\text{Polity2} + 10) \times 5$ , correlates with PTI Liberty at  $r = 0.87$ . The weaker correlation relative to V-Dem reflects Polity's coarser measurement scale and its documented limitations in coding transition and interregnum periods (Gleditsch and Ward, 1997; Vreeland, 2008; Marshall and Gurr, 2020).

### 4.4 Illustrative Face Validity

Table 7 presents ternary profiles for five countries spanning the full range of political configurations, demonstrating face validity of the decomposition.

**Table 7. Ternary Profiles for Selected Countries (2025)**

Country	L	T	C	FH Score	V-Dem LDI	Interpretation
Finland	96	2	2	100	0.93	Near-maximum liberty; stable, effective state
Russia	10	80	10	13	0.07	Minimal liberty; dominant state coercion
Somalia	8	15	77	7	0.03	Near-zero liberty; dominant state failure
Hungary	63	23	14	69	0.54	Declining liberty; rising institutional capture
India	62	18	20	66	0.39	Moderate liberty; mixed coercion and fragility

*Note: FH Score = most recently published Freedom House aggregate (2024 report year). V-Dem LDI = Liberal Democracy Index (v14).*

### 4.5 HCI Validation

The Human Capabilities Index composite correlates with the UNDP Human Development Index at  $r = 0.92$  for overlapping country-year observations, confirming convergent validity with the established measure of human well-being. The HCI differs from the HDI in its domain structure (7 domains vs. 3), its grounding in capability theory, and its strict no-interpolation data ethics policy.

### 4.6 Replication Audit Findings

The five-phase replication audit produced several findings relevant to data quality. A simple AR(1) model (next year's Liberty score equals this year's score plus noise) outperforms theoretically motivated stage-based models with  $\Delta\text{AIC} > 300$ , indicating high persistence in Liberty scores. Data-driven shock probability estimates ( $\sigma = 0.45\text{--}4.45$ ) are 2–7x lower than initial stipulated values. The mean reversion parameter  $k$  is approximately 0, indicating

weak attractor dynamics at annual frequency. These findings are fully documented in the pre-generated results files included with the replication package.

## 5. Usage Notes

---

### 5.1 Known Limitations

Users should be aware of the following limitations when working with this dataset.

**Tyranny as residual.** Tyranny absorbs measurement error from both Liberty and Chaos and functions as a catch-all category. If Freedom House systematically overrates liberty for certain country categories, the dataset will systematically underrate their tyranny. Validation of Tyranny is limited to face validity rather than external benchmark comparison.

**Real-time vs. annual divergence.** The PTI's emphasis on leading institutional indicators can produce scores that diverge from published annual indices during rapid change. Users should evaluate rapidly changing cases under both PTI scores and established indices.

**Historical data sparsity.** Pre-1972 data rely on inflection-point coding with linear interpolation. The 19th-century data (1800–1899) have some countries with only 4–6 observations across a century. Users should apply appropriate caution and treat historical observations as approximate regime characterizations.

**Standard library statistical methods.** All replication scripts use Python's standard library without third-party statistical packages. Implementations of bootstrap confidence intervals, GMM estimation, and panel regression are written from scratch, which limits sophistication compared to validated library implementations.

**Country coverage.** The current 91 countries represent approximately 47% of the world's sovereign states. Selection may introduce bias in cross-national studies.

### 5.2 Recommended Usage

The compositional nature of the data ( $L + T + C = 100$ ) requires attention to statistical method. Standard OLS regression, correlation analysis, and PCA can produce spurious results due to the induced negative correlations among components. Researchers should consider isometric log-ratio (ILR) transformations (Egozcue et al., 2003) or compositional regression models (van den Boogaart and Tolosana-Delgado, 2013). The dataset is suitable for regime transition modeling, comparative political economy, institutional erosion detection, and cross-national panel analysis, with appropriate compositional methods.

### 5.3 Citation

Users of this dataset should cite this data descriptor. Users of the Human Capabilities Index should additionally cite the HCI companion paper. The underlying source data (Freedom House, V-Dem, FSI, Polity) should be cited according to each project's citation requirements.

## 6. Code Availability

---

The complete replication package is released under a CC-BY 4.0 license and comprises 26 Python scripts (Python 3.7+, standard library only), 3 data files (.xlsx and .csv), pre-generated results for all audit phases, and full documentation including a replication README with variable definitions, mapping decisions, and a reproducibility checklist. The package is designed for zero-dependency replication: any system with a Python 3.7+ interpreter can execute all scripts without installing additional packages. Monte Carlo scripts use `random.seed(42)` for reproducibility. Each script runs in under 30 seconds. Hardcoded paths in Phase 1–4 scripts must be updated for the user's filesystem; Phase 5 scripts use relative paths. Pre-generated result files are included for comparison; stochastic variation between runs is expected for bootstrap and Monte Carlo procedures, but qualitative conclusions should be identical.

## References

---

- Aitchison, J. (1986). *The Statistical Analysis of Compositional Data*. Chapman and Hall.
- Boix, C., Miller, M., and Rosato, S. (2013). A complete data set of political regimes, 1800–2007. *Comparative Political Studies*, 46(12), 1523–1554.
- Bush, S. S. (2017). The politics of rating freedom: Ideological affinity, private authority, and the Freedom in the World ratings. *Perspectives on Politics*, 15(3), 711–731.
- Cheibub, J. A., Gandhi, J., and Vreeland, J. R. (2010). Democracy and dictatorship revisited. *Public Choice*, 143(1), 67–101.
- Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., et al. (2011). Conceptualizing and measuring democracy: A new approach. *Perspectives on Politics*, 9(2), 247–267.
- Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., et al. (2024). *V-Dem Codebook v14*. Varieties of Democracy Institute, University of Gothenburg.
- Deaton, A. and Muellbauer, J. (1980). *Economics and Consumer Behavior*. Cambridge University Press.
- Diamond, L. (2015). Facing up to the democratic recession. *Journal of Democracy*, 26(1), 141–155.
- Egozcue, J. J., Pawlowsky-Glahn, V., Mateu-Figueras, G., and Barceló-Vidal, C. (2003). Isometric logratio transformations for compositional data analysis. *Mathematical Geology*, 35(3), 279–300.
- Freedom House (2025). *Freedom in the World 2025: The Rise of Election Manipulation*. Freedom House.
- Fund for Peace (2024). *Fragile States Index Annual Report 2024*. The Fund for Peace.
- Gleditsch, N. P., Wallensteen, P., Eriksson, M., Sollenberg, M., and Strand, H. (2002). Armed conflict 1946–2001: A new dataset. *Journal of Peace Research*, 39(5), 615–637.
- Gleditsch, K. S. and Ward, M. D. (1997). Double take: A reexamination of democracy and autocracy in modern polities. *Journal of Conflict Resolution*, 41(3), 361–383.
- Hadenius, A. and Teorell, J. (2005). Assessing alternative indices of democracy. *Political Concepts: Committee on Concepts and Methods Working Paper Series*, 6, 1–23.
- Haggard, S. and Kaufman, R. R. (2021). *Backsliding: Democratic Regress in the Contemporary World*. Cambridge University Press.
- Huntington, S. P. (1991). *The Third Wave: Democratization in the Late Twentieth Century*. University of Oklahoma Press.
- Marquardt, K. L. and Pemstein, D. (2018). IRT models for expert-coded panel data. *Political Analysis*, 26(4), 431–456.
- Marshall, M. G. and Gurr, T. R. (2020). *Polity5: Political Regime Characteristics and Transitions, 1800–2018*. Center for Systemic Peace.
- Munck, G. L. and Verkuilen, J. (2002). Conceptualizing and measuring democracy: Evaluating alternative indices. *Comparative Political Studies*, 35(1), 5–34.
- Pawlowsky-Glahn, V. and Buccianti, A. (2011). *Compositional Data Analysis: Theory and Applications*. John Wiley and Sons.
- Pearson, K. (1897). Mathematical contributions to the theory of evolution: On a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proceedings of the Royal Society of London*, 60, 489–498.
- Teorell, J., Coppedge, M., Lindberg, S., and Skaaning, S. E. (2019). Measuring polyarchy across the globe, 1900–2017. *Studies in Comparative International Development*, 54(1), 71–95.
- van den Boogaart, K. G. and Tolosana-Delgado, R. (2013). *Analyzing Compositional Data with R*. Springer.
- Vreeland, J. R. (2008). The effect of political regime on civil war: Unpacking anocracy. *Journal of Conflict Resolution*, 52(3), 401–425.
- Waldner, D. and Lust, E. (2018). Unwelcome change: Coming to terms with democratic backsliding. *Annual Review of Political Science*, 21, 93–113.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018.

---

**Data Availability:** All data and code are available under CC-BY 4.0 from Cambridge Governance Labs. The replication package includes 26 Python scripts, 3 data files (.xlsx, .csv), pre-generated results, codebook, and documentation.

**Competing Interests:** The author declares no competing interests.

**Correspondence:** [ngogerty@cambridgegovernancelabs.org](mailto:ngogerty@cambridgegovernancelabs.org)

**Acknowledgments:** The dataset draws on data from Freedom House, the Varieties of Democracy Institute, the Fund for Peace, the Polity Project, the World Bank, the United Nations Development Programme, and the International Monetary Fund. The author thanks these organizations for making their data publicly available.