Journal of Sustainable Built Environment

Review Article

https://doi.org/10.70731/mvdej997

Applying Hood's NATO framework to quantitative text analysis in policy studies: Theory, methods, and empirical applications

Daiki Yamamoto [a,*], **Ayaka Fujita** [b]

[a] *Department of Civil and Environmental Engineering, Graduate School of Engineering, Tohoku University, 6-6-06 Aramaki Aza Aoba, Aoba-ku, Sendai, Miyagi 980-8579, Japan*

[b] *Department of Urban and Regional Planning, Graduate School of Engineering, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan*

KEYWORDS

Policy Instruments;
NATO Framework;
Tools of Government;
Text as Data;
Automated Content Analysis;
Policy Design; Policy Mixes;
Supervised Learning;
Annotation;
Policy Measurement

ABSTRACT

Hood's "tools of government" framework treats governing as deploying four resource types: nodality, authority, treasure, and organization. Because it offers a stable, portable "instrument language," NATO has become influential in policy design and policy instrument research. In parallel, the "text-as-data" turn has enabled large-scale, replicable measurement of policy instruments and policy mixes from documents. This review synthesizes NATO's operationalization for quantitative text analysis and organizes the literature into three programs: (1) conceptual refinement of NATO and related taxonomies in policy design and mixes; (2) automated content analysis methods, spanning dictionaries, supervised learning, topic models, scaling models, and span-level annotation; and (3) empirical applications that code instruments from legal and policy texts, including emerging annotated corpora for training and benchmarking. Recurring challenges include construct validity (instrument intent versus effect), unit-of-analysis choice, multi-label coding of co-occurring instruments, calibration and intensity measurement, cross-jurisdiction comparability, and the reliability of human labels. The review concludes with a research agenda emphasizing transparent codebooks, model evaluation against human annotation with appropriate agreement metrics, and integration of instrument coding with causal and design-oriented policy evaluation.

INTRODUCTION

Policy instrument research asks a deceptively simple question: what do governments actually do, and through which means do they pursue their ends? NATO's enduring appeal is that it reframes this question as a resource problem. Governments govern by occupying and manipulating information positions (nodality), imposing rules (authority), spending or taxing (treasure), and acting through administrative or service-delivery organizations (organization). Hood's original treatment aimed to provide a compact "tool-kit" view of government action, enabling comparison across sectors and states while remaining agnostic about normative desirability. As later policy design scholarship expanded, NATO's categories became a scaffolding for analyz-

* Corresponding author. E-mail address: daiki.yamamoto.s8@dc.tohoku.ac.jp

Received 5 December 2025; Received in revised from 13 December 2025; Accepted 19 January 2026; Published online 31 January 2026.

ing instrument choice, combinations, and implementation pathways, especially when instrument mixes rather than single tools became the object of study.

At the same time, the empirical basis of instrument research has broadened. Early policy tool studies frequently relied on small sets of cases, expert interpretation, and hand-coded inventories. Today, policy analysis increasingly treats legal texts, policy strategies, plans, and regulatory documents as machine-readable data, enabling the measurement of policy outputs, design elements, and instrument portfolios across time and jurisdictions. The "text as data" approach argues that large-scale policy claims should increasingly be anchored in validated measurements extracted from text corpora, but it also cautions that automated methods are not substitutes for conceptual clarity and careful validation. This methodological shift creates a natural convergence with NATO: if instruments can be identified in texts, and if NATO provides a stable typology, then NATO-coded text measures can serve as a bridge between theory (how governments steer) and empirical policy measurement (what instruments are used, how intensively, and in what mixes).

This review focuses on NATO's application to quantitative text analysis, with particular attention to policy tool measurement, policy mixes, and recent annotation-driven computational approaches. It is motivated by three practical demands. First, NATO-based coding is often used to summarize policy portfolios, but coding rules vary widely across studies, weakening comparability. Second, many text-based instrument measures struggle with "design versus rhetoric": texts may announce instruments without implementation, and may implement instruments without explicit, easily detectable language. Third, recent advances in supervised learning and span-level annotation make it increasingly feasible to build replicable, transportable instrument classifiers, but doing so requires clearer methodological standards in instrument conceptualization, unit selection, and evaluation.

REVIEW APPROACH AND SCOPE

Consistent with contemporary review practice, this article adopts a structured synthesis logic: define the conceptual domain (policy tools, NATO, and related instrument taxonomies), define the methodological domain (quantitative text analysis methods used to measure policy content), and then review integration studies that operationalize instrument concepts in text.

The scope includes (a) core contributions to policy instrument theory and policy design that either build on NATO directly or provide adjacent instrument frameworks; (b) foundational and policy-relevant "text as data" methods that are routinely used for policy documents; and (c) empirical studies and datasets that explicitly code instruments, instrument types, or policy design elements from text, including annotation resources designed to support supervised learning. The emphasis is on verifiable academic sources with DOIs and on research that can plausibly support SCI-indexed review standards (transparent method, cumulative argument, and a forward research agenda).

NATO AND THE EVOLUTION OF POLICY TOOL THEORY

NATO's Conceptual Core and its Comparative Logic

Hood's account of governing as a "tool-kit" remains one of the most parsimonious instrument typologies in public policy. The analytic move is to treat governance capacity as a set of deployable resources rather than as a single state attribute. Nodality is analytically distinctive because it captures governing through information position, communication, and surveillance rather than through direct coercion or spending. Authority corresponds to legal or regulatory power, including standards, prohibitions, mandates, and permissions. Treasure represents fiscal resources, including subsidies, taxes, grants, procurement, and financial incentives. Organization captures direct public provision and administrative action, including staffing, agencies, service delivery, and operational deployment. Hood's typology is now widely treated as a canonical classification scheme in policy tool research, precisely because it can accommodate both "old" instruments (regulation, spending, public provision) and "new" instruments (information campaigns, digital platforms, behavioral interventions) by treating novelty as recombination or technological transformation of underlying resource types. Hood's book provides the original tool logic and has a stable DOI-edition record.

NATO in the Broader "Policy Instrumentation" Tradition

NATO is best understood as part of a broader policy instrument tradition that emphasizes that instruments are not neutral technical devices but embody theories of social control, assumptions about target behavior, and distinctive governing relationships. Schneider and Ingram's influential account of the "behavioral assumptions" of policy tools formalized how instruments embed theories about compliance, motivation, and capacity, providing a conceptual basis for why instrument choice is not merely technical. Linder and Peters highlighted that instrument choice is mediated by decision-maker

perceptions and policy styles, not only by objective problem features, helping explain recurrent instrument patterns across systems. These perspectives complement NATO by explaining why a state might prefer, for example, nodality tools (persuasion, information) versus authority tools (mandates), even when the underlying policy goal appears similar.

Lascoumes and Le Galès' "instrumentation" approach further pushed this insight by arguing that instruments structure governing relationships and produce effects independent of declared objectives. In NATO terms, the choice among nodality, authority, treasure, and organization is also a choice among different modes of social coordination and accountability relationships. Together, these traditions underpin a key implication for text analysis: instrument coding should not reduce instruments to keywords. It must reflect the governing logic embedded in language, legal form, and administrative arrangements.

From Single Tools to Policy Mixes and Design Logics

Modern instrument research increasingly focuses on policy mixes: portfolios of instruments that jointly target complex problems. Policy design scholarship provides vocabulary for coherence, consistency, and congruence in mixes, emphasizing that mixes are constrained by legacies and institutionalized tool repertoires. Howlett and Rayner's work on policy mix design established widely used concepts for evaluating instrument portfolios, including the need to inventory instruments systematically before assessing design quality. Later work formalized design criteria and distinguished "design" from "non-design" formulation modes, clarifying that instrument portfolios often emerge from bargaining and opportunism rather than from purposive optimization. This matters for text-based measurement because instrument language may reflect compromise rather than clean theoretical categories, increasing ambiguity and multi-label overlaps.

QUANTITATIVE TEXT ANALYSIS FOUNDATIONS RELEVANT TO NATO CODING

The "Text as Data" Paradigm and Validation Imperatives

The contemporary baseline for automated text analysis in political and policy research is the principle that text models must be validated for the task and context. Grimmer and Stewart's canonical synthesis emphasizes that automated methods reduce costs but require careful, problem-specific validation and conceptual clarity. For NATO coding, this implies that model performance must be evaluated against human judgment aligned with a clear codebook, and that researchers must distinguish measurement of "mentions" (policy talk) from measurement of enforceable commitments or implemented instruments.

Method Families Commonly Used for Policy Texts

Several families of methods are particularly relevant for NATO-based measurement.

First, dictionary methods treat instrument categories as sets of terms or phrases and measure frequency or presence. These are attractive for transparency and portability but can be brittle under domain shift and can misclassify context-dependent language (e.g., "support" may indicate treasure via subsidies, or nodality via guidance). Dictionary methods are often used as baselines or for interpretability, but they require iterative refinement and careful auditing.

Second, supervised classification methods learn mappings from text to labels using annotated examples. With enough labeled data, supervised models can capture contextual usage and reduce false positives from naive keyword matching. Recent practice increasingly uses transformer-based encoders, but the central methodological requirement is not the architecture; it is the availability of reliable labels and robust evaluation.

Third, topic models and structural topic models support discovery and measurement of thematic structure. They are useful for exploring policy agendas and framing, but they do not directly yield instrument categories unless combined with labeling strategies. Topic models are often used to discover policy domains and then connect discovered topics to instrument categories.

Fourth, scaling models such as Wordscores, Wordfish, and related approaches estimate latent positions or dimensions from word frequencies. These are useful for ideological positions or issue emphasis and can complement NATO coding by quantifying the "orientation" of policy discourse, but they are not, by themselves, instrument detectors.

Reliability and Systematic Review Standards for Coding

Instrument coding—manual or automated—depends on the quality of human labels. Contemporary practice increasingly relies on agreement metrics and transparent annotation protocols. PRISMA 2020 provides the updated standard for reporting systematic reviews, and in content analysis, reliability measures such as Krippendorff's alpha have become common. These standards are directly relevant to NATO-based text studies because they structure how corpora are assembled,

how coding rules are documented, and how results can be reproduced and audited.

OPERATIONALIZING NATO FOR QUANTITATIVE TEXT ANALYSIS

What Exactly Is an "Instrument" in Text? Construct Definition and the Rhetoric–Design Gap

A central challenge is construct validity: in NATO-based text analysis, is the target construct "instrument mention," "instrument commitment," "instrument legal form," or "instrument implementation"? A policy strategy may state an intention to subsidize, regulate, or build capacity; a law may impose enforceable mandates; an administrative circular may operationalize organizational deployment; a budget may implement treasure commitments. NATO coding from text is therefore inherently sensitive to document type. A robust operationalization typically requires (a) explicit definition of the document universe (laws, plans, strategies, regulations, budget documents), (b) hierarchical coding rules that map textual signals to instrument categories with attention to legal force, and (c) ideally, triangulation across document types (e.g., strategies plus budgets) when the research question concerns implementation rather than rhetoric.

Unit of Analysis: Document, Section, Sentence, Clause, or Span

NATO instruments often co-occur and are embedded in complex legal sentences. Document-level coding is often too coarse, because a single law can contain authority provisions, treasure allocations, and organizational mandates. Sentence-level coding improves precision but still fails when a sentence contains multiple instruments. Span-level annotation, where coders highlight the exact text that justifies a label, is increasingly considered best practice for supervised learning and interpretability. It also aligns with policy design logic, because design features (target group, conditionality, enforcement, financing mechanism) are often localized in specific clauses.

Hierarchical Coding: NATO as a Top Layer With Subtypes

Empirical applications typically extend NATO with subtypes. For example, authority may be split into bans, mandates, standards, licensing, reporting obligations, and enforcement mechanisms; treasure may be split into subsidies, tax expenditures, grants, loans, procurement, and penalties; nodality may be split into information disclosure, guidance, consultation, monitoring, and digital communication; organization may be split into agency creation, staffing, service provision, and inter-agency coordination structures. This hierarchical practice is methodologically important for machine learning because it supports multi-task setups: models can first predict NATO category and then predict subtype, improving interpretability and allowing partial credit evaluation when subtypes are ambiguous but the NATO layer is correct.

Multi-Label Reality: Policy Designs as Composites

Most real policy provisions are multi-instrument. A regulation (authority) may also require reporting (nodality) and create an enforcement unit (organization). A subsidy (treasure) may be conditional on compliance with standards (authority). Therefore, NATO coding for text should generally be treated as a multi-label classification problem rather than a mutually exclusive labeling task. This has implications for evaluation metrics (micro/macro F1, label-based precision/recall, and calibration) and for corpus design (ensuring enough examples of co-occurrence patterns).

Measuring Intensity and Calibration, Not Only Presence

A persistent critique of text-based policy measures is that counting mentions conflates talk with strength. Recent work in policy mixes and design measurement emphasizes design features such as balance, consistency, and technology specificity, and underscores that policy intensity has temporal dynamics. A text-based NATO measure can move beyond presence by measuring calibrated features: legal stringency (e.g., "shall" with penalties), financial magnitude (when amounts are specified), target specificity, and enforcement mechanisms. However, these require either structured extraction (e.g., amounts) or enriched annotation that labels design characteristics alongside instrument type. The policy design annotations dataset (POLIANNA) exemplifies this move by providing annotated spans for multiple design elements and enabling supervised learning for policy design measurement.

EMPIRICAL APPLICATIONS LINKING NATO AND QUANTITATIVE TEXT ANALYSIS

Instrument Inventories and Portfolios in Comparative Policy Analysis

A major application is building instrument inventories that enable cross-national comparison of policy portfolios and their effects. Comparative work on policy design quality and instrument diversity demonstrates that policy portfolios can be measured and related to outcomes such as policy effectiveness and bureaucratic burden. These studies typically rely on systematic cod-

ing of policy outputs across time, often combining text with structured policy databases. NATO-based coding can serve as the classificatory backbone for such portfolios, especially when the research goal is to compare reliance on information, coercion, spending, and direct provision.

Policy Mixes in Sustainability Transitions and Climate/Energy Policy

Climate and energy policy are a particularly active domain for instrument mix analysis because policies evolve through layered mixes (targets, subsidies, standards, public investment, administrative reforms). Work on policy mix dynamics in renewable energy policy demonstrates measurement of balance and design features across countries and years. NATO provides a natural mapping for interpreting these mixes: renewable subsidies and tax credits map to treasure, renewable portfolio standards to authority, grid investments and agencies to organization, and information disclosure or labeling to nodality. Recent annotated datasets such as POLIANNA further enable scaling of such design measurement from text, offering a pathway to more standardized NATO-adjacent coding schemes.

Spatial Planning and Governance Instruments

Planning and spatial governance research has adopted NATO as a way to reconcile diverse "planning tools" under a single instrument logic. Studies conceptualizing planning tools often map consultation and information instruments to nodality, regulation and zoning to authority, infrastructure funding to treasure, and public provision or agency actions to organization. While many such studies remain qualitative or mixed-method, the conceptual mapping creates a foundation for computational coding of planning documents at scale, including cross-plan comparisons and regional governance analysis.

Digital Governance, Nodality, and the Renewed Centrality of Information Tools

Recent scholarship revisits nodality in the context of digital platforms, algorithmic governance, and data infrastructures. Margetts' work argues that the digital environment transforms nodality's practice and relevance, motivating renewed attention to how governments use information network position as a governing resource. This strand is particularly important for text analysis because many digital-era policy instruments are implemented through guidance, standards, data-sharing rules, and platform-based communication, which may be textually subtle compared to classic "regulate/spend/build" verbs. Text coding in this domain must therefore address concept drift: the language of nodality shifts over time (e.g., from "public information campaigns" to "data governance," "platform moderation," "open data," or "algorithmic transparency").

Procedural Policy Tools and Instrument Sequencing

A related development is the growth of procedural policy tool research, emphasizing tools that shape policy processes rather than directly altering substantive outcomes. Procedural tools often manifest in texts as consultation rules, committees, coordination mandates, reporting procedures, or deliberative mechanisms. NATO can be extended here by treating procedural nodality and organization as central (consultation, coordination, administrative process), while authority and treasure appear as procedural constraints (rule-making authority, funding for process). Empirical work in this area provides design concepts and case-based evidence that can be linked to text coding, especially where procedural tools are embedded in statutes and administrative orders.

Instrument Coding From Text Beyond NATO: Institutional Grammar and Design Annotation

While NATO provides a high-level typology, other frameworks offer complementary coding primitives. The Institutional Grammar approach parses institutional statements into components and has been used for automated coding of policy texts, demonstrating how structured annotation can support machine learning for policy statement extraction. Such work is relevant to NATO coding because it suggests a division of labor: institutional grammar can identify the action structure (who must do what, under what conditions), while NATO labels the governing resource type used in that action. Integrating these approaches can improve both precision (better statement segmentation) and validity (clearer mapping between linguistic form and instrument logic).

METHODOLOGICAL CHALLENGES AND BEST-PRACTICE RECOMMENDATIONS FOR NATO-BASED TEXT CODING

Transparency and Codebooks: From "Interpretive Mapping" To Replicable Annotation Rules

Because NATO categories are broad, many studies apply them interpretively. For quantitative text analysis, this is inadequate unless the interpretive logic is made explicit. A best-practice codebook should specify: (a) definitional criteria for each NATO category and subcategory; (b) positive and negative examples; (c) decision rules for multi-label cases; (d) rules for handling aspirational language versus enforceable provisions; and (e) rules for ambiguous verbs (e.g., "support," "promote,"

"ensure"). Without this, supervised models learn coder idiosyncrasies rather than instruments.

Evaluation: Beyond Accuracy to Calibration, Robustness, and Domain Transfer

NATO coding is often used for cross-sector or cross-country comparison, so evaluation must include transfer tests. Models trained on one sector (e.g., climate policy) may fail in another (e.g., health or education) because instrument language differs. Researchers should report out-of-domain performance, error typologies, and sensitivity to document type. Where NATO measures are used for longitudinal inference, researchers should test temporal robustness (lexical drift). For policy mixes, evaluation should also test whether derived portfolio measures (shares, diversity indices) are stable under alternative coding thresholds and model seeds.

Reliability: Agreement Metrics Aligned With the Unit of Analysis

Reliability must match the coding unit. If coding is span-based, agreement should be measured both on label assignment and on span overlap. If sentence-based, agreement should be measured per sentence. Multi-label settings require label-wise agreement reporting. Standard reliability measures remain important, but they should be interpreted carefully: low agreement may reflect real ambiguity in instrument language rather than coder negligence, which in turn motivates refining codebooks or adding hierarchical labels.

The "Instrument Versus Outcome" Inference Problem

A frequent analytical error is treating instrument presence as policy impact. Text coding typically measures policy outputs (what instruments are adopted or announced). Outcomes depend on implementation capacity, enforcement, political economy, and context. Therefore, NATO-coded text measures should be integrated with complementary data (budgets, administrative capacity proxies, enforcement records, or implementation indicators) when the research question concerns effectiveness. Where causal claims are desired, researchers should use NATO coding as an input to identification strategies rather than as evidence of impact by itself.

RESEARCH AGENDA: TOWARD A MATURE NATO-BASED MEASUREMENT PROGRAM

First, the field would benefit from shared, open NATO codebooks that specify hierarchical subtypes and multi-label rules, enabling comparability. Second, more span-annotated corpora are needed across sectors and languages, because supervised learning performance depends on the breadth and quality of labels. Third, instrument coding should expand from "type detection" to "design characterization," including conditionality, enforcement, magnitude, and target specificity. Fourth, NATO-based coding should be linked to policy design evaluation frameworks, allowing researchers to test whether certain NATO portfolios predict outcomes under specified conditions, rather than assuming universal effects. Fifth, digital-era nodality demands conceptual updating: instrument language increasingly references data infrastructures and platform governance, and NATO-based dictionaries and models must be periodically recalibrated.

CONCLUSION

NATO remains a uniquely durable framework for instrument analysis because it provides a stable, resource-based language that travels across sectors and governance systems. The rise of quantitative text analysis creates an opportunity to operationalize NATO at scale, enabling systematic measurement of instrument portfolios, design features, and policy mixes across time and place. Realizing this opportunity, however, requires methodological discipline: clear construct definitions, carefully chosen units of analysis, explicit codebooks, multi-label modeling, and robust validation including transfer and temporal tests. Emerging annotated datasets and machine coding approaches show that instrument coding can become more standardized, but they also highlight that NATO-based measurement is as much a conceptual task as a computational one. The next stage of research should therefore treat NATO not merely as a convenient classification, but as a measurement theory that links governing resources to textual signals through transparent, testable rules.

References

1. Bali, A. S., Capano, G., & Ramesh, M. (2019). Anticipating and designing for policy effectiveness. Policy and Society, 38(1), 1–13. https://doi.org/10.1080/14494035.2019.1579502
2. Bali, A. S., Howlett, M., Lewis, J. M., & Ramesh, M. (2021). Procedural policy tools in theory and practice. Policy and Society, 40(3), 295–311. https://doi.org/10.1080/14494035.2021.1965379
3. Benoit, K., & Laver, M. (2007). Estimating party policy positions: Comparing expert surveys and hand-coded content analysis. Electoral Studies, 26(1), 90–107. https://doi.org/10.1016/j.electstud.2006.04.008
4. Bruinsma, B., & Gemenis, K. (2019). Validating Wordscores: The promises and pitfalls of computational text scaling. Communication Methods and Measures. Advance online publication. https://doi.org/10.1080/19312458.2019.1594741
5. Capano, G. (2020). The knowns and unknowns of policy instrument analysis. SAGE Open. https://doi.org/10.1177/2158244019900568
6. Capano, G., Pritoni, A., & Vicentini, G. (2020). Do policy instruments matter? Governments' choice of policy mix and higher

education performance in Western Europe. Journal of Public Policy, 40(3), 375–401. https://doi.org/10.1017/S0143814X19000047

7. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding (arXiv:1810.04805). arXiv. https://doi.org/10.48550/arXiv.1810.04805
8. Fernández-i-Marín, X., Knill, C., & Steinebach, Y. (2021). Studying policy design quality in comparative perspective. American Political Science Review, 115(3), 931–947. https://doi.org/10.1017/S0003055421000186
9. Grimmer, J., & 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
10. Hannah, A. (2021). Procedural tools and pension reform in the long run: The case of Sweden. Policy and Society, 40(3), 362–378. https://doi.org/10.1080/14494035.2021.1955487
11. Hase, V. (2022). Automated content analysis. In The international encyclopedia of communication research methods. https://doi.org/10.1007/978-3-658-36179-2_3
12. Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. Communication Methods and Measures, 1(1), 77–89. https://doi.org/10.1080/19312450709336664
13. Hjorth, F. (2015). Comparing automated methods for estimating party positions. Research & Politics. https://doi.org/10.1177/2053168015580476
14. Hood, C. C. (1983). Tools of government. Palgrave Macmillan. https://doi.org/10.1007/978-1-349-17169-9
15. Howlett, M. (2018). The criteria for effective policy design: Character and context in policy instrument choice. Policy and Politics. https://doi.org/10.1080/17516234.2017.1412284
16. Howlett, M. (2019). Procedural policy tools and the temporal dimensions of policy design: Resilience, robustness and the sequencing of policy mixes. International Review of Public Policy. https://doi.org/10.4000/irpp.310
17. Howlett, M., & Mukherjee, I. (2014). Policy design and non-design: Towards a spectrum of policy formulation types. Politics and Governance, 2(2), 57–71. https://doi.org/10.17645/pag.v2i2.149
18. Howlett, M., & Rayner, J. (2007). Design principles for policy mixes: Cohesion and coherence in "new governance arrangements." Policy and Society, 26(4), 1–18. https://doi.org/10.1016/S1449-4035(07)70118-2
19. Howlett, M., Giest, S., Mukherjee, I., & Taeihagh, A. (2025). New policy tools and traditional policy models: Better understanding behavioural, digital and collaborative instruments. Policy Design and Practice. https://doi.org/10.1080/25741292.2025.2495373
20. Isoaho, K., Gritsenko, D., & colleagues. (2021). Topic modeling and text analysis for qualitative policy research. Policy Studies Journal. https://doi.org/10.1111/psj.12343
21. Jin, Z., & Mihalcea, R. (2023). Natural language processing for policymaking. In Handbook of computational social science for policy (pp. 141–162). Springer. https://doi.org/10.1007/978-3-031-16624-2_7
22. John, P. (2013). All tools are informational now: How information and persuasion define the tools of government. Policy & Politics, 41(4), 605–620. https://doi.org/10.1332/030557312X655729
23. Klemmensen, R., Binzer Hobolt, S., & Hansen, M. E. (2007). Estimating policy positions using political texts: An evaluation of the Wordscores approach. Electoral Studies, 26(4), 746–755. https://doi.org/10.1016/j.electstud.2007.07.006
24. Lascoumes, P., & Le Galès, P. (2007). Understanding public policy through its instruments: From the nature of instruments to the sociology of public policy instrumentation. Governance. https://doi.org/10.1111/j.1468-0491.2007.00342.x
25. Laver, M., Benoit, K., & Garry, J. (2003). Extracting policy positions from political texts using words as data. American Political Science Review, 97(2), 311–331. https://doi.org/10.1017/S0003055403000698
26. Linder, S. H., & Peters, B. G. (1989). Instruments of government: Perceptions and contexts. Journal of Public Policy, 9(1), 35–58. https://doi.org/10.1017/S0143814X00007960
27. Lowe, W. (2008). Understanding Wordscores. Political Analysis, 16(4), 356–371. https://doi.org/10.1093/pan/mpn004
28. Margetts, H. (2024). How rediscovering nodality can improve democratic governance in a digital world. Public Administration. https://doi.org/10.1111/padm.12960
29. Mukherjee, I., Çoban, M. K., & Bali, A. S. (2021). Policy capacities and effective policy design: A review. Policy Sciences, 54(2), 243–268. https://doi.org/10.1007/s11077-021-09420-8
30. Pacheco-Vega, R. (2020). Environmental regulation, governance, and policy instruments: Where are we now, 20 years after the stick, carrot and sermon typology? Journal of Environmental Policy & Planning. https://doi.org/10.1080/1523908X.2020.1792862
31. Page, M. J., McKenzie, J. E., Bossuyt, P. M., & colleagues. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, Article n71. https://doi.org/10.1136/bmj.n71
32. Restemeyer, B., van den Brink, M., & Arts, J. (2024). A policy instruments palette for spatial quality: Lessons from Dutch flood risk management. Journal of Environmental Policy & Planning, 26(3), 249–263. https://doi.org/10.1080/1523908X.2024.2328072
33. Rice, D., Siddiki, S., Frey, S., Kwon, J. H., & Sawyer, A. (2021). Machine coding of policy texts with the Institutional Grammar. Public Administration, 99(2), 248–262. https://doi.org/10.1111/padm.12711
34. Roberts, M. E., Stewart, B. M., & Airoldi, E. M. (2016). A model of text for experimentation in the social sciences. Journal of the American Statistical Association, 111(515), 988–1003. https://doi.org/10.1080/01621459.2016.1141684
35. Roberts, M. E., Stewart, B. M., & 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
36. Roberts, M. E., Stewart, B. M., Tingley, D., & colleagues. (2014). Structural topic models for open-ended survey responses. American Journal of Political Science, 58(4), 1064–1082. https://doi.org/10.1111/ajps.12103
37. Rogge, K. S., & Reichardt, K. (2016). Policy mixes for sustainability transitions: An extended concept and framework for analysis. Research Policy, 45(8), 1620–1635. https://doi.org/10.1016/j.respol.2016.04.004
38. Ruedin, D. (2013). The role of language in the automatic coding of political texts. Swiss Political Science Review. https://doi.org/10.1111/spsr.12050
39. Schneider, A. L., & Ingram, H. (1990). The behavioral assumptions of policy tools. The Journal of Politics. https://doi.org/10.2307/2131904
40. Schoenefeld, J. J., Schulze, K., Hildén, M., & Jordan, A. J. (2021). The challenging paths to net-zero emissions: Insights from the monitoring of national policy mixes. The International Spectator, 56(3), 24–40. https://doi.org/10.1080/03932729.2021.1956827
41. Slapin, J. B., & Proksch, S.-O. (2008). A scaling model for estimating time-series party positions from texts. American Journal of Political Science, 52(3), 705–722. https://doi.org/10.1111/j.1540-5907.2008.00338.x
42. Stead, D. (2021). Conceptualizing the policy tools of spatial planning. Journal of Planning Literature, 36(3), 297–311. https://doi.org/10.1177/0885412221992283
43. Steinebach, Y., Hinterleitner, M., Knill, C., & Fernández-i-Marín, X. (2024). A review of national climate policies via existing databases. npj Climate Action, 3, Article 80. https://doi.org/10.1038/s44168-024-00160-y
44. Uyarra, E., Flanagan, K., & Laranja, M. (2011). Reconceptualising the "policy mix" for innovation. Research Policy, 40(5), 702–713. https://doi.org/10.1016/j.respol.2011.02.005
45. Vabo, S. I., & Røiseland, A. (2012). Conceptualizing the tools of government in urban network governance. International Journal of Public Administration, 35(14), 934–946. https://doi.org/10.1080/01900692.2012.691243
46. Viehmann, C., Beck, T., Maurer, M., Quiring, O., & Gurevych, I. (2023). Investigating opinions on public policies in digital media: Setting up a supervised machine learning tool for stance classification. Communication Methods and Measures, 17(2), 150–184.

https://doi.org/10.1080/19312458.2022.2151579

47. Xie, M., Liao, X., & Yaguchi, T. (2025). The policy spatial footprint: Causal identification of land value capitalization using network-time exposure. Land, 14(11), 2240. https://doi.org/10.3390/land14112240
48. Young, L., & Soroka, S. (2012). Affective news: The automated coding of sentiment in political texts. Political Communication, 29(2), 205–231. https://doi.org/10.1080/10584609.2012.671234
49. Zhao, J., & Li, C. (2022). Research on the classification of policy instruments based on BERT model. Discrete Dynamics in Nature and Society, 2022, Article 6123348. https://doi.org/10.1155/2022/6123348