Media Reporting on International Organizations: Using Machine Learning to Identify Worldwide Patterns

Michal Parizek, Charles University[1]

GLOWIN: Global Flows of Political Information (https://glowin.cuni.cz)

Paper prepared for the 15th *Political Economy of International Organizations Conference*, San Diego, United States, May 4–6, 2023

Abstract

Coverage of international organizations (IOs) in news media is essential for the long-term build-up of their legitimacy. This article argues that news reporting on IOs systematically differs between powerful and rich countries, which dominantly control global IOs, and weaker and low-income countries, which are the primary recipients of global IOs' tangible deliverables. Reporting on IOs in the former is likely to be heavily oriented to central IOs politics and processes; reporting in the latter to local implementation work of the IOs. At the core of the empirical analysis is a dataset mapping news reporting on IOs in more than 4.6 million articles across 166 states and 64 languages in the years 2018-2021 and a supervised machine learning model that classifies a sample of those articles that refer to IOs on whether they relate 1) to central IO politics, 2) to 'boots-on-the-ground' local policy implementation by IOs, or 3) to IOs' expertise and provision of information. It is based on the large pre-trained semantic model ERNIE fine-tuned with 1654 human-labeled news articles. On the core classification task, the model achieves a high accuracy of 0.89. The analysis covers 45 IOs and IO bodies jointly forming the entire United Nations System. Empirical evidence strongly supports the theoretical argument: news reporting on local IO implementation activity is more than twice as frequent in weaker and low-income states than in powerful and high-income states, *ceteris paribus.* On the other hand, reporting on central IO politics is 30% more prominent in rich and powerful states than in low-income weaker states.

Keywords: International Organizations, News Media, Machine Learning, Text Analysis

Length: approx. 8500 words

1. Introduction

International Organizations are increasingly contested by large segments of the public and elites in some of the most powerful Western states (Hooghe et al. 2019a; Lake et al. 2021; Vries et al. 2021). This is a significant phenomenon for the prospects of institutionalized cooperation, calling for a better understanding of how public attitudes towards IOs are formed in the West and beyond (Dellmuth et al. 2022; Walter 2021). This paper sheds new light on this issue by connecting it with the coverage of IOs in news media, presenting the first map of the content of reporting on 45 global IOs in media worldwide. I aim to empirically explore and theorize the most prominent patterns of everyday coverage of IOs in media. Based on this everyday media coverage and information on IOs, loyalty, trust, and 'reservoirs of confidence' in IOs are built over time (Tallberg and Zürn 2019, p. 587), and positive or negative attitudes towards IOs are formed (Brutger and Strezhnev 2022; Ecker-Ehrhardt 2018).

The contribution of this study is both theoretical and empirical. *Theoretically*, I argue that news reporting on IOs is likely to systematically differ between powerful and rich countries, which dominantly control global IOs, and weaker and low-income countries, which are the primary recipients of global IOs' tangible material deliverables. While richer and more powerful states typically provide IOs with key political backing and material support, poorer and weaker states receive more immediately tangible, intuitively understandable benefits, especially in material deliveries. I argue that the type of reporting on IOs that prevails in different countries will likely follow this principal difference in states' positions towards IOs.

I propose a conceptual scheme that classifies news articles that refer to global IOs based on whether they focus 1) on central IO politics, 2) local IO policy implementation work, or 3) information and expertise provision by IOs. This classification helps distinguish IOs' *political* and *epistemic* authority (Barnett and Finnemore 2004; Zürn et al. 2012) but also brings into the picture explicitly the operational activity of IOs, typically in the form of local 'boots-on-the-ground' implementation of IO mandates (Eckhard 2021; Honig 2019). It thus enables us to understand better *why* IOs are visible in media and what factors make them relevant in the public sphere (de Wilde 2019). I am most directly interested in variation in the prominence of the two first categories of 1) central IO politics and 2) local IO policy implementation. I hypothesize that reporting on 1) central IO politics is likely to be more prevalent in powerful and rich states, which primarily maintain and control the IOs as compared to weaker and poorer nations. At the same time, I hypothesize that reporting on 2) local IO policy implementation, or

their tangible 'boots-on-the-ground' work, is likely to be more prevalent in weaker and poorer nations and only sparse in the powerful and rich ones.

Empirically, I evaluate the theoretical argument by presenting the first dataset that systematically maps the content of news reporting on IOs globally, in 166 states accounting for 99% of the world population. It thus helps us move beyond analyses centered at a limited number of (typically Western) states. At the core of the dataset is a map of news media content worldwide introduced in a separate study that focuses on the quantitative question of how much IOs are reported on worldwide (Parizek 2022). In this article, I move beyond the *how much* question by classifying relevant reporting segments on IOs in a representative sample of news articles using a supervised machine learning model. The model is based on the large pre-trained semantic model ERNIE, fine-tuned with a sample of 1654 human-coded news articles referring to IOs. On its core classification task, my fine-tuned model achieves a very high accuracy of 89% and similar precision and recall scores, consistently high across the three classification categories. I deploy this model to classify a representative sample of more than 90 000 references to IOs identified in a sample from a large corpus of more than 4.6 million news articles across 64 languages (Parizek 2023; Parizek and Stauber 2022). Specifically, I map references to 45 IOs and IO bodies jointly forming the UN system broadly defined, consisting of UN central bodies and UN funds and programs (e.g., UNDP, UNICEF, UNHCR), all UN specialized agencies (e.g., the WHO, ILO, IMF, World Bank) and the so-called UN-related organizations (e.g., the WTO and the IAEA).[2]

Empirical evidence strongly supports my theorizing. News reporting focused on central IO politics is approximately 30% more prominent in powerful high-income countries than in weaker low-income countries. More profoundly, reporting on local IO implementation work is sparse in high-income powerful states. In low-income and weak states, this reporting is more than twice as prominent as in high-income powerful states (up to 250% of the prominence in the former compared to the latter). In rich and powerful states that control IOs, reporting is primarily focused on the central IO politics, but reporting on the tangible work of IOs, the actual implementation of their mandates on the ground, and the delivery of public services to their recipients, is rare. This type of reporting may strengthen the negative perceptions among the public of IOs as 'out-of-touch' bureaucracies and platforms for politicized debate and conflict. In contrast, the public in the primary target (recipient) countries of IOs' work does receive more

---

[2] https://www.un.org/en/about-us/un-system

information about the on-the-ground delivery work of IOs, but it only receives less information about the actual political processes at the center of IOs' work and decision-making. Such reporting may strengthen the intuitions of developing countries' publics that IOs are, to an extent, vehicles of Western influence and venues for great power conflicts.

2. Conceptual and theoretical framework: how IOs are reported on in news media

IOs are increasingly concerned with their public image (Ecker-Ehrhardt 2018). A voluminous literature on mass media and news in Media Studies and Political Communication identifies factors that make a phenomenon, a person, or an institution *newsworthy* and hence visible in media (Galtung and Ruge 1965; Hester 1973; Koopmans and Vliegenthart 2011; O'Neill and Harcup 2009; Harcup and O'Neill 2001). Building on O'Neill and Harcup (2009, p. 165), de Wilde highlights four factors that are specifically relevant for the study of IOs' media coverage: elite *status (1)* of the source body, e.g. of the international institution; *valence (2)* of the topic to be reported on; *relevance (3)* of the reported subject to the audience; and audience *identification (4)* with the reported subject (de Wilde 2019, p. 1196). In a separate study, I explore quantitatively how much IOs are reported on globally, using the key parts of this framework as the guide (Parizek 2022); in yet another study, I use the framework to motivate a systematic exploration of the intensity of media reporting on the EU, NATO and the UN in connection to the Russia-Ukraine war (Parizek 2023).

What about the content of IO reporting, however? Several studies attempt to track key features of news reporting on IOs, specifically the associations of IOs with terms connected with protests, contestation, and politicization (Rauh and Zürn 2020; Schmidtke 2019; Sommerer et al. 2022). But what about reporting on IOs in general, not specifically focused on their contestation? In developing my scheme to map reporting on IOs, I build on the classical accounts of the diverse types of (source of) IOs' authority, highlighting why IOs at all matter in world politics. Drawing on Barnett and Finnemore's account (2004, pp. 20–29), one can distinguish three types of authority that underpin IOs' work and fill their general rational-legal authority with content: a) delegated authority, b) moral authority, and c) expert authority (2004, p. 20). I take this classification as a helpful starting point. However, I complement it with the established distinction between IOs' political and epistemic authority (Zürn 2018; Zürn et al. 2012). Furthermore, I am specifically interested in the difference between local operational

policy implementation work of IOs and central IO politics of high-level decision-making, conflicts, and centralized IO functions (Rittberger et al. 2019; see also Abbott and Snidal 1998, p. 5).

As a result, my scheme prioritizes two key dimensions of IOs' work. One is the distinction between 1) central IO politics, which will be typically concerned with how delegated and pooled authority is exercised by the central IO bodies, often in headquarters, and 2) local IO politics, concerned with the implementation work of IOs and their activity on the ground, typically outside of headquarters in the member states. The second is the distinction between the political authority of IOs, most directly associated with the two categories of central (1) and local (2) level of IO work and their epistemic authority, or recognition of their expertise and right to make authoritative statements about facts. A short description of each of the three categories follows. It closely matches the instructions used by coders (see below) in classifying news on IOs in the process of human coding.

(1) The central (global) level refers to those aspects of IO's work that pertain to highest-level IO politics and processes, the central decision-making/rule-making bodies, the central IO administration, adjudication (courts), etc. It also includes IOs' connection to international norms and law.

(2) The local, national, or regional level refers to primarily de-centralized policy implementation by IOs, mainly to work and actions of IOs in the field, the projects they run and implement in specific countries, and events relevant to specific regional or country-level IO operations. It also includes, for example, loans to governments, peace operations in countries, missions to countries, etc.

(3) Expertise, information, and policy advice refer to all the aspects of IOs' work where the primary role of the IO is to provide states with information, data, statistics, reports, rankings, expert-based evaluations, policy-relevant expertise, and advice. This category is also coded when the IO is referred to in the context of its failure to provide unique information, valid expertise, helpful advice, etc.

The specific reference to an IO in an article may correspond to any number of these categories or none. So, for example, an article about the violations of conditions of UN Security Council resolution in a specific country will be coded with (1), for UNSC as a central decision-making body, and (2), for the implementation/ field operation element. If it also refers to a report on the number of casualties by the UN, also (3) is coded. At the same time, an article may reference an IO that will not match any of the categories.

Most importantly, I expect that the type of references to IOs in media – along this three-category scheme – will differ across states. I expect it to depend on the state's position in the international system and towards the IOs. In that, I broadly distinguish between powerful high-income states that primarily staff, finance, and by and large control IOs, and weaker low-income states more on the receiving end of IOs' activities (Stone 2011; Dijkstra 2015). For the former, IOs may be seen as either irrelevant or as instruments of control and exercise of (partly legitimated) power. For the latter, IOs are relatively more important as political venues embodying at least some rules and, in practice, as often critical providers of tangible deliveries (Parizek 2022).

Based on this principal difference of the states' position toward the IOs, I hypothesize the following:

H1: News reporting on IOs is systematically more *focused on central IO politics in high-income powerful states than in low-income weaker states.*

H2: News reporting on IOs is systematically less *focused on local IO implementation work in high-income powerful states than in low-income weaker states.*

In sum, the structural position of states and the type of their dominant engagement with IOs' work is likely to drive the patterns of news reporting on IOs in them.

3. Data and coding: references to 45 IOs in news media across the world

I now turn to the presentation of my empirical data. The analysis consists of three empirical datasets. The first is the large dataset of news worldwide, presented in this section's first subsection. This is the large corpus where I identify references to IOs to be further analyzed. Second, I draw on a sizeable human coding effort in which a group of RAs and I constructed a labeled dataset from a sample of news. This is used to train my machine learning algorithm. It is presented in the second subsection of this section. Finally, third, the model developed using this labelled data is applied back to a sample of those articles from the large news corpus in which a reference to at least one of the studied IOs is made. This is the dataset used in sections 5 and 6 of this article.

The analysis maps media visibility of the entire United Nations System, with its 1) main bodies, 2) its programs and funds (e.g., UNDP, UNICEF), 3) the 15 UN specialized agencies, including

the World Bank and the IMF, but also e.g. the WHO, IMO, WIPO, WMO, as well as 4) eight IOs that are officially 'UN-related' in the classification of the UN System (e.g. the WTO, IAEA, IOM). This is not a complete list of globally relevant IOs, but it covers the core of global governance. In addition, five global IOs unrelated to the UN but covered by earlier studies by Hooghe et al. (2019b) are also included (e.g., Global Environmental Facility and the Permanent Court for Arbitration). In the coding, separate codes were devised for the central bodies of the UN: the Security Council, the General Assembly, the Secretariat, ECOSOC, and the ICJ, but in the analysis, these are all subsumed under the UN label. The same applies to World Bank's subunits. The measurement period covers the years 2018-2021.

3.1 Empirical data sources, sampling, and pre-processing

The data I use in this article come from the sizeable new dataset developed within the project Global Flows of Political Information (GLOWIN) (for a description of the data, see Parizek 2023; Parizek and Stauber 2022).[3] In the description of the data collection scheme, I turn now shortly, for several paragraphs, to the plural 'we,' as more team members were involved in the setup of the data collection infrastructure. I return to the singular 'I' afterward. In developing the dataset underpinning GLOWIN, our key media content source is GDELT, or the Global Database of Events, Language, and Tone (GDELT 2019; Leetaru and Schrodt 2013). GDELT is unique in that it covers, at least to some extent, media in virtually all countries. We collect a representative sample of 10% of GDELT's content for basic manageability of the data collection procedure. We (attempt to) access the specific URL of these news articles obtained from GDELT and download their content. Typically, this results in around 30–40 000 news articles per day of data. In the next step, we extract its full text from the downloaded HTML file of the article. To maintain full control over the data generation process, we connect raw GDELT output with systematic data on the audience geography of each website (media outlet) as estimated by Amazon Alexa Webs Services (Alexa Web Information Service 2021). We use this extensive data to drastically filter the sources covered by GDELT and only keep in the analysis media outlets that rank 500 or higher in at least one country. As a result, the majority of the around 60 000 websites tracked by GDELT are discarded, and our analysis is based on the content of articles from 3525 domains that show significant relevance in the audience countries. A random sample of articles is selected from these domains for download and

---

[3] https://glowin.cuni.cz/

processing. Applying this filter reduces the volume of data retained for analysis to approximately 30-35% of downloaded articles. After all these steps, we are typically left with around 10–15 000 downloaded and technically pre-processed (cleaned) articles per day of data.

A significant challenge for any project seeking to map news content across many countries comes with the multiplicity of languages spoken worldwide. Our data source GDELT tracks content in 64 languages. To process the downloaded data consistently, we translate 20% of the downloaded non-English content using a Google Translate API. 22% of analyzed articles were originally in English, while the remaining 78% were translated from one of 63 other languages of the media outlets. This leads to more than 4.6 million articles across the years 2018-2021.

In a random sample of these articles, I performed named entity recognition (NER) using the flairNLP tagger.[4] From the recognized named entities, I selected those identified as organizations (as opposed to persons or locations). I performed a regular expressions-enhanced dictionary search for names, in different variants, of the 45 IOs and IO bodies forming the UN system. This procedure led to identifying 94 081 references to the 45 IOs and IO bodies across 37 280 articles, forming a representative sample from the entire underlying dataset.

3.2 Human coding of IO references

A labeled dataset is needed to train a supervised machine learning algorithm to classify the news articles and references to IOs. For that, I employed a group of eleven Master's students in International Relations or Security Studies at Charles University as research assistants (RAs). These RAs labelled a random sample of 3946 references to IOs drawn from the underlying dataset from 2018-2021. An ordinal scale was offered, with the options to indicate that the specific dimension (e.g., local IO politics) is "not present" (0), "partly present" (0.5), or "present" (1) (or "impossible to judge," later discarded in the analysis).

Each article was coded by three coders independently. Hence the empirical basis consists of 11 838 IO reference codings. In all cases, the final coding result is the arithmetic mean of the three coding scores. A note on coder training is due now. In the training process, I faced a dilemma where on the one hand, extensive training is supposed to ensure that the coders share an understanding of what they are coding. For that purpose, I developed a written codebook, held

---

[4] https://github.com/flairNLP/flair/blob/master/resources/docs/TUTORIAL_2_TAGGING.md

regular sessions, and recorded commonly agreed solutions to repeatedly occurring coding issues. On the other hand, I did not want to develop a microcosmos of understanding of news on IOs specific to our group. After all, the purpose of coding with more independent coders is to ensure that we obtain a measure of replicability of the results, a degree of confidence that another competent coder from outside the group would code the references to IOs similarly (Krippendorff 2004, pp. 217–219). In this vein, though I invested serious time in developing the coding scheme and rules, I intentionally did not strive for the highest possible reliability or develop extensive definitions and highly precise coding instructions. Doing that would most likely come at the expense of validity (Krippendorff 2004, p. 213).

In *Table 1*, I report the Krippendorf alpha scores for each coded category. I report two variants. One is the strictest, where only an exact match in the codes assigned is counted as an agreement. The second is an adjusted scheme where a partial agreement counts as a match as well, so a situation where the codings differ by half a point is still considered acceptable. The strictest measure values range between 0.58 and 0.68; when a partial disagreement between coders is allowed, the scores range between 0.74 and 0.8. Taking into account that the scheme tracks references to 45 different IOs and IO bodies across close to 200 states, and hence a range of highly diverse news reporting contexts, these are very strong reliability figures. Of course, analysis with a narrower focus and, thus, a more detailed codebook mapping onto a homogeneous corpus of news should be able to achieve higher reliability scores. But to the extent to which breadth, wide geographical coverage, and diversity of news contexts are key to my analysis, I intentionally prioritized a looser coding scheme and more space for individual interpretation over the strictness of instructions and procedures.

Table 1: Inter-coder reliability

	Krippendorf alpha (three coders, ordinal scale)	Krippendorf alpha with partial disagreement allowed (three coders, ordinal scale)	Frequency (mean score)
Central IO politics	0.59	0.76	52%
Local IO politics	0.58	0.74	27%
Information provision by IOs	0.68	0.8	37%

This is all the more the case because the primary purpose of this coding exercise is to provide the material to be used to train the machine learning algorithm to label the more than 90 000 references to IOs sampled from the more than 4.6 million news articles. In other words, the ultimate criterion for evaluating the use of the scheme is whether a classification model can be built to replicate the human coding on a large scale with sufficient precision and recall.

4.   The machine learning model

I can now turn to a discussion of the construction and the key parameters of the supervised machine learning model. The model is based on fine-tuning a pre-trained semantic model ERNIE, using the human-labeled texts described above. Different pre-trained language models based on artificial neural networks have been tried, including the vastly popular BERT and its variants (cased/uncased, large/base, robust ROBERTa). By a relatively small margin, the ERNIE model used here showed the highest accuracy. The model construction and algorithm closely follow the model used by Parizek and Stauber (2022) to classify all the more than 4.6 million news worldwide according to their general topic (economy, politics, sports,…).

The typical use of artificial neural networks consists of the use of a general-purpose language model, such as FastText (Mikolov et al. 2017), BERT (Devlin et al. 2019) or ERNIE (Zhang et al. 2019), and its fine-tuning with the use of (human-) labeled data. All these models are based on semantic embeddings, with individual words (and larger text segments, such as sentences or paragraphs) embedded in a high-dimensional dense semantic space. The basic idea of embedding words in such high-dimensional space, where each word is represented by a numeric vector that captures its meaning (and words with similar meanings are located close to each other in the space) became highly popular following the innovative work by Mikolov and co-authors (Mikolov et al. 2013). Since then, progressively more complex embeddings have been developed, taking into account more contextual information and trained on much larger corpora. BERT (Devlin et al. 2019), developed by Google and used in its various applications, is probably the most popular in applied computational social sciences (Widmann and Wich 2022, p. 4). ERNIE (Sun et al. 2019; Zhang et al. 2019), which I use here, has been developed by Baidu and achieves somewhat better performance on standard tasks than BERT (and so it does

in my application).[5] Semantic (word, sentence, paragraph) embeddings use is now becoming common in IR and IO research (Chelotti et al. 2022; Rauh 2022).

In training my model, I took the following steps. First, I filtered from the human-labeled data those referring to IOs for which identification of a particular label was not clear enough (i.e., with more sizable disagreement between coders) and where coders identified more than one category. In this way, I obtained a set of labeled articles with a high degree of agreement among coders on that a specific label, *and only that one label*, describes the type of reference to the IO in the particular article. Second, because of the imbalance between the size of the categories, I down-sampled references to central IO politics and partly information provision to achieve a balanced dataset with approximately equal number of references in each of the three classification categories. Third, the set of labeled texts was divided into the train, validation, and test dataset, with 80% of labeled data used for training, 10% for validation, and 10% for testing. Because only those articles with unambiguous and unique labeling by human coders were used, the effective number of articles retained dropped dramatically from close to 4000 to 1654, resulting in 1323 references used for training, 165 for validation, and 166 for testing. For each IO reference, I randomly varied a window of 40, 60, and 80 words (i.e., +/- 20, 30, and 40 words) around the name of the IO in the text, adopting a KWIC (keyword-in-context) approach. The model does not show systematic differences in accuracy in training depending on the concrete choice of the window size.

The fine-tuning of this pre-trained language model (ERNIE) lies in the adjustment of the final layers of the neural network, modifying the language model for it to (better) fit the specific data and downstream task for which the model is used. In other words, specific word and language constructions are learned to be typical for texts carrying a particular label assigned by humans in the specific downstream application. The language model is updated with this information to better capture the nuances specific to the concrete application domain. Following our previous application (Parizek and Stauber 2022), I implement the model fine-tuning and text classification in Python, using the *Pytorch* and *transformers* architectures. The model used three epochs for training, with a monotonically decreasing learning rate.

The model achieves a strong overall accuracy of 0.89, meaning it correctly classified 89% of the previously unseen human-labeled IO references forming the test dataset. Similar results

---

[5] https://gluebenchmark.com/leaderboard

were achieved with different minor modifications of the training procedure, so an accuracy of close to 90% was achieved robustly in my training attempts. ***Table 2*** details the model's performance, highlighting two key features. *First*, an important strength of the model is that it performs consistently across all three coding categories. While it achieves the highest accuracy of 96% for the category information and expertise, the scores for the other two are 90% or higher. *Second*, as the table shows, the model has not only high accuracy levels but also (similarly) high values of the more specific metrics – precision ([0.87; 0.98]), recall ([0.87; 0.91]), and the F score ([0.87; 0.95]) across all the categories. All these values indicate that a good deal of confidence can be placed in the model results.

Table 2: Test results for the trained ML model

Category (label) of news:	Central IO politics (1)	Local IO politics (2)	Information and expertise (3)
Test data results (*N=166*)			
True positives	47	48	53
True negatives	112	108	112
False positives	7	6	1
False negatives	7	6	5
Precision	0.87	0.89	0.98
Recall	0.87	0.89	0.91
F-score	0.87	0.89	0.95
Accuracy	0.92	0.9	0.96

In deploying the model on the more than 90 000 items of the unlabelled dataset, I only count a particular observation as labeled with the specific category if the probability assigned to that label surpasses 67%. That means only those observations categorized by the model with a high degree of certainty are counted; the remaining approximately 25% are labeled with a residual category "other."

In addition, I estimate sentiment attached to each reference to IOs using a pre-trained semantic models based on BERT, specifically the RoBERTa-base-SST-2 model fine-tuned for sentiment analysis, using the Stanford Sentiment Treebank.[6]

---

[6] https://huggingface.co/textattack/roberta-base-SST-2

5.   Exploring variation in reporting on IOs

Once trained, the model can be deployed on the 94 081 news article IO references sampled for this analysis. In this section, I present some interesting descriptive observations from this exercise pertinent to the overall prominence of the three reporting categories and to variation across IOs and their classification. In the following section, I turn to cross-national variation in reporting on IOs and the evaluation of the theoretical propositions embodied in hypotheses H1 and H2.

First of all, it is interesting to note the relative prominence of the three types of reporting on IOs. *Table 3* summarizes the frequencies of the three categories. The most frequent type of reporting is that on central IO politics, which includes all IO activities connected to central processes and outcomes and references to international law and norms related to IOs. In total, 34% of coded references belong to this category. Information provision and expertise have been identified as the dominant label in 21%, while local IO implementation work is the label for 20% of references. The remaining 25% are unassigned, with label assignment probability lower than 67%. These 94 081 individual references appeared in 37 280 different articles. At the article level, category (1) central IO politics appeared at least once in 39% of articles, (3) information and expertise provision in 31%, and local IO implementation work in 23% of articles.

**Table 3: Frequencies of news reporting types**

Category (label) of news:	Number of coded references	% of coded references	Number of articles with the type of reference	% of articles with the type of reference[7]
Central IO politics (1)	32002	34%	14368	38.5%
Local IO politics (2)	18459	19.6%	8667	23.2%
Information and expertise provision (3)	20085	21.3%	11452	30.7%
Other (residual)	23535	25%	13875	37.2%
Total	94081	100%	48362	

[7] Numbers add up to more than 100% because a reference n article can contain more than one (type of) reference.

Second, about variation across IOs, one can expect that operational IOs, whose important task is to implement policy on the ground in member states, e.g., development agencies/programs such as the UNDP or the World Bank, will be systematically more associated with reporting focusing on local IO politics. In contrast, program IOs like the WTO should be connected with reporting on central IO politics. IOs with prominent epistemic authority, such as the WHO, should be associated with reporting on information and expertise provision.

Figure 1 shows the location of IOs in a news reporting space. In the left chart, I depict the position of the IOs based on the prominence of their reporting focused on central IO politics (horizontal axis) and local implementation work (vertical axis). The diagonal distinguishes IOs with primarily central IO politics-focused reporting from those with reporting centered on their local implementation activity. This can be seen as a useful proxy measure distinguishing program and operational IOs, as perceived by the public (Rittberger et al. 2019). I highlight in red the UN as a body primarily reported on in the context of central IO politics, and the IMF more prominently reported on in connection to its implementation work in the partner countries. In the right-hand chart, I provide an analogous depiction, highlighting the prominence of the information and expertise on the vertical axis. IOs above the diagonal can be seen as bodies reported on primarily in the context of their epistemic authority (Zürn et al. 2012). The WHO is a highly visible (and highlighted) example here.

Figure 1: Reporting on IOs: central IO politics, local implementation, and information
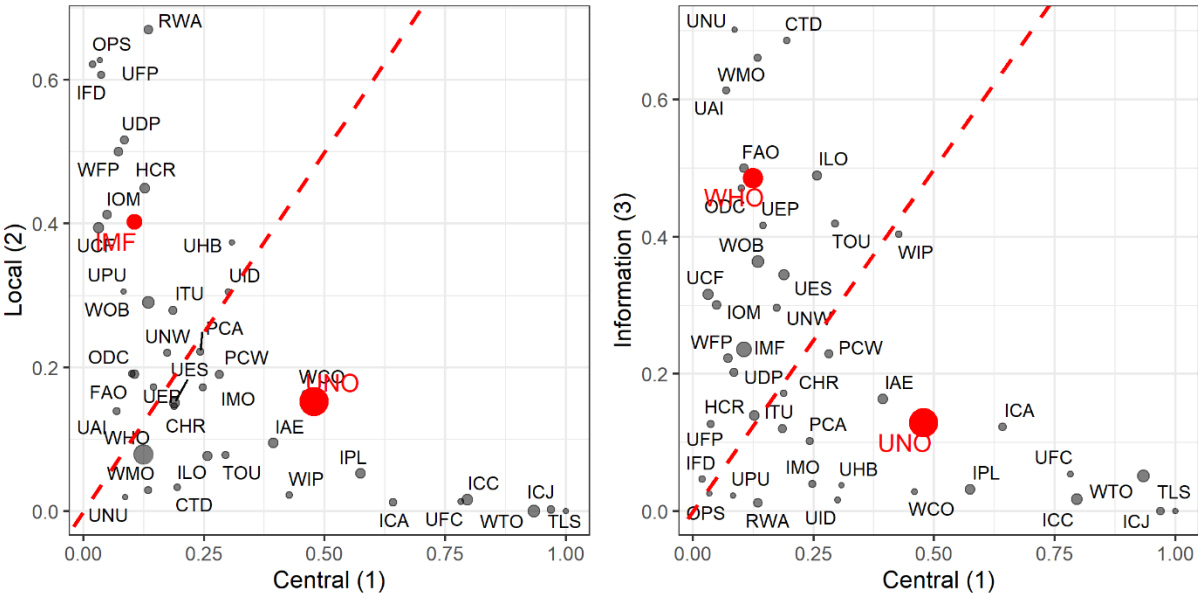
Such visualization can be extended to capture all three types of reporting on IOs and thus localize IOs in a 3D space of central IO politics, local IO implementation work, and information provision, such as in *Figure 2*. In the left-hand chart, the UN is an example of a body strongly associated with central IO politics reporting. In the central chart, IMF is most prominently associated with local policy implementation work reporting. Finally, in the right-hand chart, the WHO's profile is strongly associated with reporting on its provision of expertise and information.

Figure 2: Location of IOs in the 3D media reporting space: UN, IMF, and the WHO
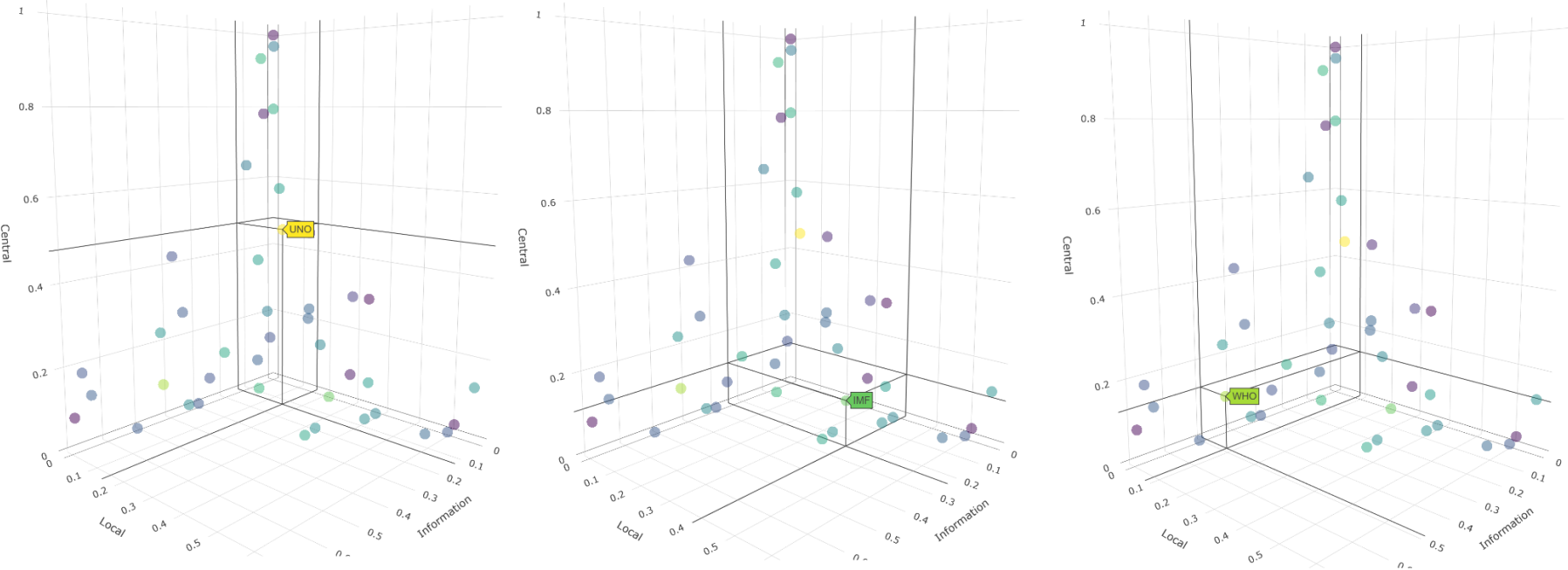
Descriptively, such mapping may provide important systematic insight into the patterns of reporting on individual IOs worldwide in a comparative cross-IO perspective. While rich quantitative datasets have become available lately, e.g., on the design and authority of IOs (Hooghe et al. 2019b; Zürn et al. 2021) and, related, their coverage in media in general (Parizek 2022; Parizek and Stephen 2021) and in connection to protests (Rauh and Zürn 2020; Sommerer et al. 2022), systematic data on public information and the content on news reporting on individual IOs of this kind is still missing. Deploying the kind of analysis presented here will enable us to analyze how different IOs are primarily presented in media and in connection to what types of activity and source of their authority.

6.   Reporting on IOs across the world

In this section, I will use the data aggregated cross-nationally to test the propositions embodied in hypotheses H1 and H2. While in a related paper (Parizek 2022), I show that IOs are generally significantly more frequently reported on in the developing world than in economically developed countries, the data presented here allows us to explore whether the reporting content differs. As I argued above in formulating hypothesis H1, we should expect central IO politics reporting to be more prominent in powerful and high-income countries than in weaker and low-income countries. In contrast (H2), local IO implementation work reporting can be expected much more prominently in weaker, low-income states.

Empirical evidence seems to support this theorizing. In *Figure 3*, I plot the share of news referring to central IO politics against countries' level of economic development, measured by their GNI per capita (left) and against their overall economic power (GNI, both variables logged). As theorized above, central IO politics are systematically more frequently reported on in richer states, at around 34%, than in low-income states, at about 27%. Similarly, central IO politics reporting is systematically more frequent in powerful states, at about 37% of news, than in weaker states, at around 28% (right). Descriptively, central IO politics reporting is thus approximately 30% more frequent in rich and powerful states likely to dominate IO decision-making and exercise control over the bodies (Dijkstra 2015; Stone 2011).

Figure 3: Country development and power and reporting on central IO politics
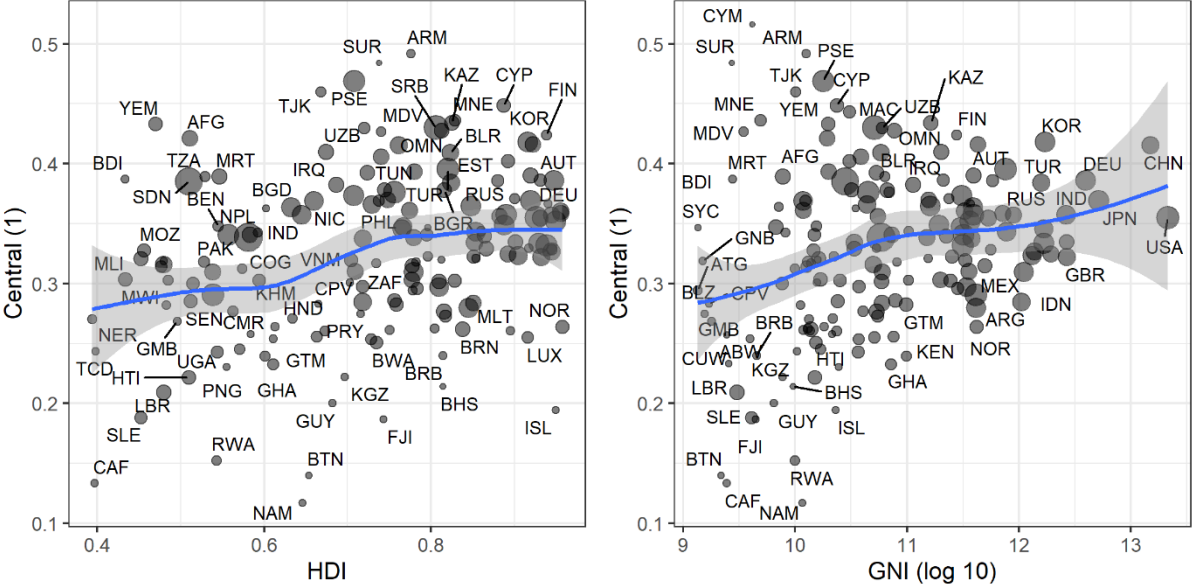
A similar but even more pronounced picture is obtained when we relate local IO implementation work reporting against countries' level of economic development, again measured by GNI per capita (left) and their overall economic power (GNI, both variables logged). ***Figure 4*** visualizes the association. I expect local IO politics news to be more prominent in weaker, low-income states. Descriptive evidence in ***Figure 4*** supports these contentions. In low human development (left) and weaker (right) states, local IO activity is reported on in approximately 30% of all IO news. In contrast, in high human development and power states, this share is only between 10 and 15%, so between one third and one half of the former values.

Figure 4: Country development and power and reporting on local IO work
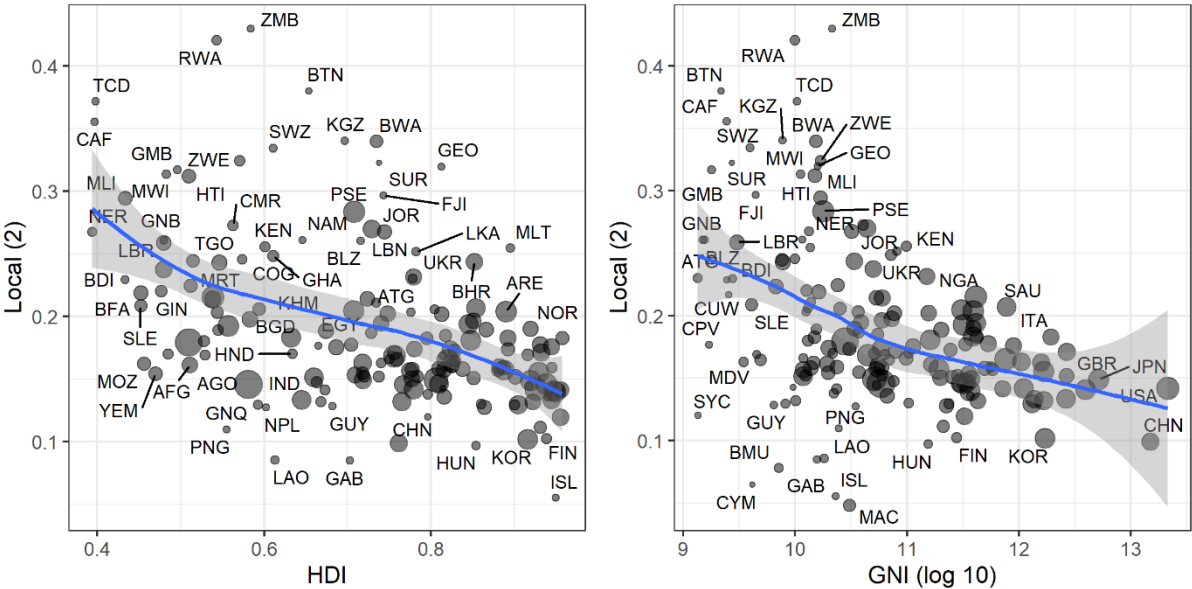
A series of OLS models in **Table *4*** demonstrates this in a multivariate setting, controlling for further factors. These controls include political regime, measured with the V-DEM Liberal democracy score of each country (Coppedge et al. 2019), the overall exposure of states to foreign, as opposed to domestic news (Parizek and Stauber 2022), average visibility of the mapped media in each country, estimated using Amazon Alexa Webs Services (Alexa Web Information Service 2021), and the share of the mapped articles in English. In all the models, standardized coefficients are reported.

The results strongly and systematically support hypotheses H1 and H2. In ***Model 1***, a one-SD rise in country HDI is associated with a 0.3-SD rise in the share of central IO articles, and a one-SD rise in GNI (log 10) with a 0.23-SD rise, other things equal. Substantively, the predicted share of central IO politics-focused articles rises from around 28% to about 35% across the empirical range of the two predictors, other things equal. ***Model 2*** shows even stronger effects for the share of local IO implementation work articles as the outcome variable, with standardized coefficients of -0.3 and -0.34, respectively. Substantively, as we move from low-HDI states (HDI ≈ 0.4) to high-HDI states (HDI>0.9), the predicted share of local IO work-focused articles halves from 25% to 13%. It decreases even more strongly, from 25% to 10%, as we move from the weakest countries (GNI ≈ 1 bill. USD) to the strongest (GNI > 10 trills. USD). These are very sizable effects.

In **Model 3**, I combine information on the prominence of central and local reporting by formulating the dependent variable as the dominance of local over central reporting, using a simple formula for the construction of the dependent variable: 'Local dominance over central' = (local-central)/(local+central). The variable ranges from 1 (all reporting focuses on local IO work) to -1 (all reporting focuses on central IO politics). The predicted scores decrease from around 0 (parity) to around -0.4 as we move from lowest to highest predictor values. That means, substantively, that for low-HDI weak countries, local and central IO reporting is equally frequent, while for high-HDI and powerful states, the ratio is approximately 2:1 in favor of central IO reporting. Finally, in **Model 4**, I explore the share of information-oriented IO articles as the dependent variable. I identify a statistically significant negative effect for HDI and a positive effect for GNI (log 10). This lies outside of my theoretical framework, though, and the overall predictive power of the model is below 10%.

Table 4: OLS regression results

	Central (%)	Local (%)	Local dominance over central	Information (%)
	(1)	(2)	(3)	(4)
HDI	**0.332*** (0.097)**	**-0.292** (0.091)**	**-0.363*** (0.090)**	-0.274** (0.094)
GNI (log 10)	**0.227* (0.095)**	**-0.337*** (0.092)**	**-0.346*** (0.083)**	0.286** (0.109)
Liberal democracy (V-DEM)	-0.299*** (0.066)	-0.021 (0.060)	0.123* (0.059)	0.118 (0.075)
Foreign news coverage	0.047 (0.111)	-0.049 (0.111)	-0.039 (0.098)	0.009 (0.104)
Avg. article Alexa rank	-0.009 (0.081)	0.113 (0.079)	0.114' (0.067)	-0.037 (0.104)
Share of art. in English	-0.352*** (0.068)	0.216** (0.079)	0.322*** (0.072)	-0.040 (0.084)
Constant	-0.014 (0.070)	0.048 (0.069)	0.038 (0.062)	-0.096 (0.080)
Observations	158	158	158	158
R^2	0.316	0.348	0.440	0.071
Adjusted R^2	0.288	0.322	0.418	0.034

Note: 'p<0.1; *p<0.05; **p<0.01; ***p<0.001

A natural next step for the analysis is to explore the stability of these findings across IOs. At this stage, though, exploring differences between IOs is hampered by insufficient volume of

data processed (number of detected IOs references) for the majority of IOs in the dataset –a sufficiently sizable number of references is only available for a handful of IOs, in particular the UN (approx. 190 000 observations at IO-state level), the WHO (70 000), the IMF (32 000), the World Bank and the WTO (both around 15 000 IO-state level observations).[8] When each of the IOs is analyzed individually, the preliminary evidence suggests that while the hypothesized patterns of systematic differences across countries persist, for different IOs, they are primarily associated with different country features. Variation in news media reporting on the UN is systematically associated with the country's HDI but not its power. The same applies to reporting on the World Bank. In contrast, a systematic association of news media reporting and country power (GNI) is visible for the WHO, the IMF, and the WTO. No association is visible, for these IOs, with country HDI. In other words, the aggregate patterns reported in **Table *4*** are driven by different IOs for each of the two variables (and partly for H1 and H2). This area calls for further exploration once more sizable data is processed and available for analysis.

These results are essential not only because they highlight the significant differences in reporting focus across countries. They are important mainly because of the implications of local, as opposed to central, IO reporting for debates about the public image of IOs. Reporting on local IO activity is connected with IO deliverables, actual implementation work, and tangible work of IOs. Central IO politics reporting centers on abstract public goods provided by IOs, general norms, and the political conflicts surrounding IO decision-making. This may also translate into differences in the overall tone or sentiment of reporting on IOs. In *Figure 5*, I show how the sentiment of reporting on IOs varies across countries. The left chart shows no systematic relationship between reporting sentiment to IO countries' HDI. However, the right chart shows that the sentiment of reporting on IOs is systematically less positive in more powerful countries. Interestingly, China is a clear outlier from the overall pattern here. But in the overall pattern, approximately 54% of references to IOs have positive sentiment in low-GNI countries; for high-GNI countries, this share of positive sentiment references declines to around 42%. To be sure, this is not a measure of expressed (critical) attitude towards the IOs; the measure merely reflects the overall positive or negative semantic associations with the text of the window around the IO reference. And in many areas, implementation of 'on-the-ground' work by IOs may be as contested as their central decision-making processes and political

---

[8] Since each IO observation can be associated with several audience states, depending on the news outlet audience geography data, these figures add up to more than the overall number of unique references to IOs mentioned earlier.

authority. Once again, these matters warrant further exploration. Yet the evidence further highlights the differences in reporting patterns across the world.

Figure 5: IO reporting sentiment and country development level and power
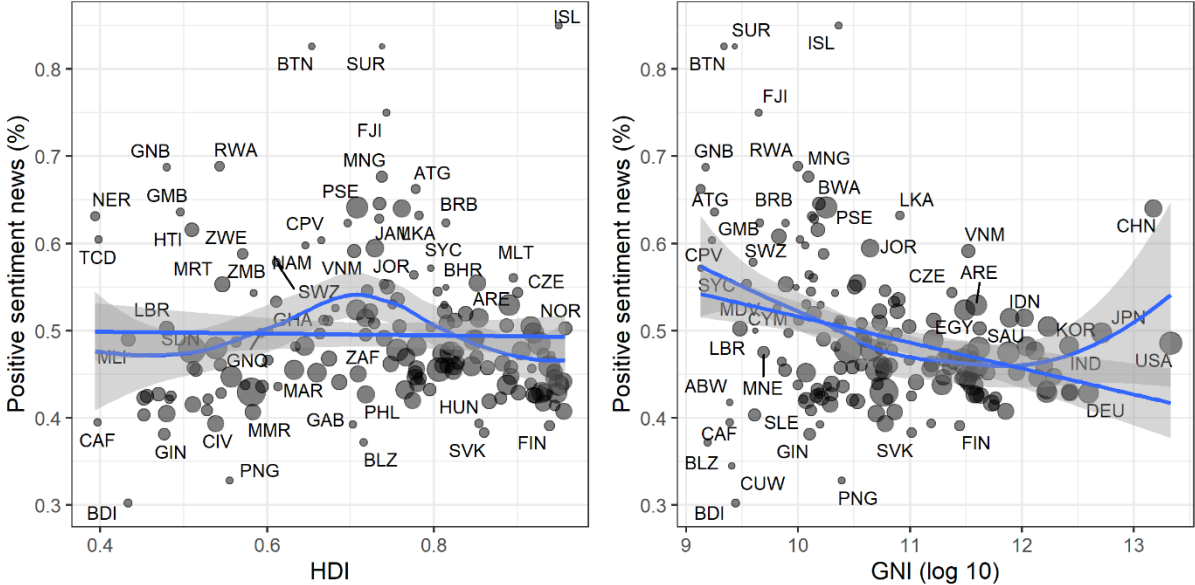
7.   Conclusions

This article aims to contribute to the literature on IOs and their legitimacy in several ways. *First*, it advances the literature on how IOs are portrayed in media and what position they enjoy in the public sphere of their member states (Parizek 2022, 2023; Rauh and Zürn 2020; Schmidtke 2019; Sommerer et al. 2022). It goes beyond the existing studies in geographical and cross-lingual media coverage and the simple volume of material analyzed. *Second*, methodologically, I show that a supervised machine learning model can be trained to classify news articles that refer to IOs along a theoretically derived classification scheme that captures key dimensions of IOs' authority. *Third*, I provide descriptive results on the prominence of different types of reporting on IOs across IOs. *Fourth*, these results are highly interesting when explored from a cross-national perspective. Reporting on IOs appears to differ systematically in content between developed and developing countries and between powerful and weaker countries. Reporting in rich and powerful states seems systematically less focused on local IO implementation work and more focused on central IO politics.

The empirical results presented here point to significant avenues for further research. They suggest possible challenges to the legitimacy of IOs, especially in the eyes of the public in the rich and powerful Western states. The results show that reporting on IOs is systematically less connected to people's everyday experiences and tangible deliverables in richer and more powerful states. Tentative evidence also suggests that in these states reporting on IOs is less positive in sentiment. Irrespective of the ultimate cause, if reporting is systematically less frequent (Parizek 2022) and less tangible (more abstract) in Western democracies than in the developing world, a serious legitimacy problem may emerge among the Western states. It is entirely plausible to hypothesize that IO media reporting more focused on central IO politics and less on the more tangible local IO implementation work, and possibly more negative in sentiment, translates into more negative attitudes to IOs and their lower legitimacy over time (Dellmuth and Schlipphak 2020; Dellmuth and Tallberg 2015). Opportunities for the politicization of IOs are larger if political information the public receives about IOs is systematically associated with more negative issues (see Brutger and Strezhnev 2022).

Historically, the largely asymmetrical positions of Western 'developed' states and the 'developing' world of the Global South in IOs may have been unproblematic for much of the public in the West. The massive differences in per capita income levels between the developed and the developing world, combined with a strong focus on the bipolar conflict with the Soviet Union, may have shielded key global IOs with robust implementation (and spending) mandates from critique. Yet in the world where 'the Rest' is increasingly recognized as a direct competitor of the West, growing constituencies in the West may question the very purpose of maintaining and disproportionately financing the key IOs over which they are losing control. If news reporting on the bodies focuses more prominently on the central political level of IOs and the conflicts associated with it and less on the tangible deliverables of their work, especially the budget-heavy IOs may find it increasingly hard to justify their mission and relevance in the eyes of large segments of the Western public.

8. References

Abbott, K. W., & Snidal, D. (1998). Why States Act through Formal International Organizations. *Journal of Conflict Resolution*, *42*(1), 3–32. https://doi.org/10.1177/0022002798042001001

Alexa Web Information Service. (2021). Alexa Web Information Service. https://aws.amazon.com/marketplace/pp/B07Q71HJ3H

Barnett, M., & Finnemore, M. (2004). *Rules for the World: International Organizations in Global Politics*. Ithaca: Cornell University Press.

Brutger, R., & Strezhnev, A. (2022). International Investment Disputes, Media Coverage, and Backlash Against International Law. *Journal of Conflict Resolution*, *66*(6), 983–1009. https://doi.org/10.1177/00220027221081925

Chelotti, N., Dasandi, N., & Jankin Mikhaylov, S. (2022). Do Intergovernmental Organizations Have a Socialization Effect on Member State Preferences? Evidence from the UN General Debate. *International Studies Quarterly*, *66*(1), sqab069. https://doi.org/10.1093/isq/sqab069

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., & others. (2019). *V-Dem [Country-Year/Country-Date] Dataset v9*. Varieties of Democracy (V-Dem) Project. https://doi.org/10.23696/vdemcy19

de Wilde, P. (2019). Media logic and grand theories of European integration. *Journal of European Public Policy*, *26*(8), 1193–1212. https://doi.org/10.1080/13501763.2019.1622590

Dellmuth, L. M., & Schlipphak, B. (2020). Legitimacy beliefs towards global governance institutions: a research agenda. *Journal of European Public Policy*, *27*(6), 931–943. https://doi.org/10.1080/13501763.2019.1604788

Dellmuth, L. M., Scholte, J. A., Tallberg, J., & Verhaegen, S. (2022). The Elite–Citizen Gap in International Organization Legitimacy. *American Political Science Review*, *116*(1), 283–300. https://doi.org/10.1017/S0003055421000824

Dellmuth, L. M., & Tallberg, J. (2015). The social legitimacy of international organisations: Interest representation, institutional performance, and confidence extrapolation in the

United Nations. *Review of International Studies*, *41*(3), 451–475. https://doi.org/10.1017/S0260210514000230

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv:1810.04805 [cs]*. http://arxiv.org/abs/1810.04805. Accessed 24 April 2022

Dijkstra, H. (2015). Shadow bureaucracies and the unilateral control of international secretariats: Insights from UN peacekeeping. *The Review of International Organizations*, *10*(1), 23–41. https://doi.org/10.1007/s11558-014-9203-7

Ecker-Ehrhardt, M. (2018). Self-legitimation in the face of politicization: Why international organizations centralized public communication. *The Review of International Organizations*, *13*(4), 519–546. https://doi.org/10.1007/s11558-017-9287-y

Eckhard, S. (2021). Bridging the citizen gap: Bureaucratic representation and knowledge linkage in (international) public administration. *Governance*, *34*(2), 295–314. https://doi.org/10.1111/gove.12494

Galtung, J., & Ruge, M. H. (1965). The Structure of Foreign News: The Presentation of the Congo, Cuba and Cyprus Crises in Four Norwegian Newspapers. *Journal of Peace Research*, *2*(1), 64–90. https://doi.org/10.1177/002234336500200104

GDELT. (2019). The GDELT Project. https://www.gdeltproject.org/. Accessed 31 March 2019

Harcup, T., & O'Neill, D. (2001). What Is News? Galtung and Ruge revisited. *Journalism Studies*, *2*(2), 261–280. https://doi.org/10.1080/14616700118449

Hester, A. (1973). Theoretical Considerations in Predicting Volume and Direction of International Information Flow. *Gazette (Leiden, Netherlands)*, *19*(4), 239–247. https://doi.org/10.1177/001654927301900404

Honig, D. (2019). When Reporting Undermines Performance: The Costs of Politically Constrained Organizational Autonomy in Foreign Aid Implementation. *International Organization*, *73*(1), 171–201. https://doi.org/10.1017/S002081831800036X

Hooghe, L., Lenz, T., & Marks, G. (2019a). Contested world order: The delegitimation of international governance. *The Review of International Organizations*, *14*(4), 731–743. https://doi.org/10.1007/s11558-018-9334-3

Hooghe, L., Lenz, T., & Marks, G. (2019b). *A Theory of International Organization*. Oxford: Oxford University Press.

Koopmans, R., & Vliegenthart, R. (2011). Media Attention as the Outcome of a Diffusion Process—A Theoretical Framework and Cross-National Evidence on Earthquake Coverage. *European Sociological Review*, *27*(5), 636–653. https://doi.org/10.1093/esr/jcq032

Krippendorff, K. (2004). *Content Analysis: An Introduction to Its Methodology* (2nd ed.). London: SAGE Publications.

Lake, D. A., Martin, L. L., & Risse, T. (2021). Challenges to the Liberal Order: Reflections on International Organization. *International Organization*, *75*(2), 225–257. https://doi.org/10.1017/S0020818320000636

Leetaru, K., & Schrodt, P. A. (2013). GDELT: Global Data on Events, Location and Tone, 1979-2012. Presented at the International Studies Association Annual Convention, San Francisco, USA.

Mikolov, T., Grave, E., Bojanowski, P., Puhrsch, C., & Joulin, A. (2017). Advances in Pre-Training Distributed Word Representations. *arXiv:1712.09405 [cs]*. http://arxiv.org/abs/1712.09405. Accessed 24 April 2022

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed Representations of Words and Phrases and their Compositionality. *arXiv:1310.4546 [cs, stat]*. http://arxiv.org/abs/1310.4546. Accessed 1 February 2022

O'Neill, D., & Harcup, T. (2009). News Value and Selectivity. In K. Wahl-Jorgensen & T. Hanitzsch (Eds.), *The Handbook of Journalism Studies* (pp. 161–174). New York: Routledge. Accessed 15 September 2021

Parizek, M. (2022). Media Visibility of International Organizations Worldwide: Politicization of International Authority or 'Boots-on-the-Ground' Work? *Paper presented at the 14th Political Economy of International Organization Annual Conference, July 7-9, 2022, Oxford*. https://www.peio.me/wp-content/uploads/PEIO14/PEIO14_paper_122.pdf

Parizek, M. (2023). Worldwide Media Visibility of NATO, the European Union, and the United Nations in Connection to the Russia-Ukraine War. *Czech Journal of International Relations*, *58*(1), 15–44. https://doi.org/10.32422/mv-cjir.60

Parizek, M., & Stauber, J. (2022). Mapping Global Flows of Political Information: Data, Methods, and First Insights. *Paper presented at the COMPTEXT 20222 Conference, May 6–7, 2022, Dublin*.

Parizek, M., & Stephen, M. D. (2021). The Increasing Representativeness of International Organizations' Secretariats: Evidence from the United Nations System, 1997–2015. *International Studies Quarterly*, *65*(1), 197–209. https://doi.org/10.1093/isq/sqaa088

Rauh, C. (2022). Supranational emergency politics? What executives' public crisis communication may tell us. *Journal of European Public Policy*, *29*(6), 966–978. https://doi.org/10.1080/13501763.2021.1916058

Rauh, C., & Zürn, M. (2020). Authority, politicization, and alternative justifications: endogenous legitimation dynamics in global economic governance. *Review of*

International Political Economy, *27*(3), 583–611. https://doi.org/10.1080/09692290.2019.1650796

Rittberger, V., Zangl, B., Kruck, A., & Dijkstra, H. (2019). *International Organization* (3rd ed.). London: Red Globe Press.

Schmidtke, H. (2019). Elite legitimation and delegitimation of international organizations in the media: Patterns and explanations. *The Review of International Organizations*, *14*(4), 633–659. https://doi.org/10.1007/s11558-018-9320-9

Sommerer, T., Agné, H., Zelli, F., & Bes, B. (2022). *Global Legitimacy Crises: Decline and Revival in Multilateral Governance*. Oxford, New York: Oxford University Press.

Stone, R. W. (2011). *Controlling Institutions: International Organizations and the Global Economy*. Cambridge: Cambridge University Press.

Sun, Y., Wang, S., Li, Y., Feng, S., Tian, H., Wu, H., & Wang, H. (2019). ERNIE 2.0: A Continual Pre-training Framework for Language Understanding. https://doi.org/10.48550/arXiv.1907.12412

Tallberg, J., & Zürn, M. (2019). The legitimacy and legitimation of international organizations: introduction and framework. *The Review of International Organizations*, *14*(4), 581–606. https://doi.org/10.1007/s11558-018-9330-7

Vries, C. E. D., Hobolt, S. B., & Walter, S. (2021). Politicizing International Cooperation: The Mass Public, Political Entrepreneurs, and Political Opportunity Structures. *International Organization*, *75*(2), 306–332. https://doi.org/10.1017/S0020818320000491

Walter, S. (2021). The Backlash Against Globalization. *Annual Review of Political Science*, *24*(1), 421–442. https://doi.org/10.1146/annurev-polisci-041719-102405

Widmann, T., & Wich, M. (2022). Creating and Comparing Dictionary, Word Embedding, and Transformer-Based Models to Measure Discrete Emotions in German Political Text. *Political Analysis*, 1–16. https://doi.org/10.1017/pan.2022.15

Zhang, Z., Han, X., Liu, Z., Jiang, X., Sun, M., & Liu, Q. (2019). ERNIE: Enhanced Language Representation with Informative Entities. *arXiv:1905.07129 [cs]*. http://arxiv.org/abs/1905.07129. Accessed 24 April 2022

Zürn, M. (2018). *A Theory of Global Governance: Authority, Legitimacy, and Contestation*. Oxford: Oxford University Press.

Zürn, M., Binder, M., & Ecker-Ehrhardt, M. (2012). International Authority and Its Politicization. *International Theory*, *4*(1), 69–106. https://doi.org/10.1017/S1752971912000012

Zürn, M., Tokhi, A., & Binder, M. (2021). The International Authority Database. *Global Policy*, *12*(4), 430–442. https://doi.org/10.1111/1758-5899.12971