

# Human Rights are (Increasingly) Plural: Learning the Changing Taxonomy of Human Rights from Large-scale Text Reveals Information Effects

BAEKKWAN PARK *East Carolina University*

KEVIN GREENE *University of Pittsburgh*

MICHAEL COLARESI *University of Pittsburgh*

*This manuscript helps to resolve the ongoing debate concerning the effect of information communication technology on human rights monitoring. We reconceptualize human rights as a taxonomy of nested rights that are judged in textual reports and argue that the increasing density of available information should manifest in deeper taxonomies of human rights. With a new automated system, using supervised learning algorithms, we are able to extract the implicit taxonomies of rights that were judged in texts by the US State Department, Amnesty International, and Human Rights Watch over time. Our analysis provides new, clear evidence of change in the structure of these taxonomies as well as in the attention to specific rights and the sharpness of distinctions between rights. Our findings bridge the natural language processing and human rights communities and allow a deeper understanding of how changes in technology have affected the recording of human rights over time.*

## INTRODUCTION

The ability to track the relative improvement or degradation of human rights protections around the globe and across time is not only an important input to decisions related to aid allocations, diplomatic engagement, and military intervention decisions (Finnemore and Sikkink 1998; Poe, Carey, and Vazquez 2001; Wood and Gibney 2010) but also crucial to citizens' and scholars' understanding of the progress and setbacks faced by liberal ideas and policies. Recent backsliding on human rights issues in previously liberal countries such as Hungary and Poland, as well as the United States, makes consistently tracking and comparing the human rights behaviors of governments one of the most important social challenges of this generation (Colaresi 2014; Levitsky and Ziblatt 2018).


The accelerating digital and computing revolutions continue to propagate more information, usually in the form of text and images, and useful algorithms to

analyze patterns in detailed evidence. The availability of higher-resolution descriptions of human rights behaviors is ultimately a positive development for the tracking of repression over time and across the globe. However, the deluge of denser data brings new measurement challenges that must be resolved in research designs, coding schemes, and future automated systems (Clark and Sikkink 2013; Fariss 2014).


For the last two decades, human-coded scoring of annual human rights reports have been used as the core repositories of these comparisons (Cingranelli and Richards 2010; Wood and Gibney 2010). Yet, because of the changing information available to the composers of these texts from the late 1970s until today, the reliability and validity of previous measures derived from these texts, if uncorrected, have been called into question (Clark and Sikkink 2013) and ignited a debate about trends in human rights protections (Cingranelli and Filippov 2018a; Fariss 2019b).

Although these measurement concerns are not new, the conceptual approach and empirical evidence we present here can help move research on information communication technologies (ICT) and human rights forward. The motivating question for our research is whether we can detect more direct and convincing evidence that changes in ICT over time have altered the composition of relevant human rights texts. To make progress on this research puzzle, we focus on the text composition process and particularly the underlying concepts that comprise human rights. We highlight an obvious but often ignored fact: human rights are plural.<sup>1</sup> Thus, the concept of human rights is itself a bundle of related

---

Baekkwon Park , Senior Data Analyst, East Carolina University, [baekkwon.park@gmail.com](mailto:baekkwon.park@gmail.com)

Kevin Greene, PhD Candidate, University of Pittsburgh, [ktg19@pitt.edu](mailto:ktg19@pitt.edu)

Michael Colaresi , William S. Dietrich II Professor, University of Pittsburgh, [mcolaresi@pitt.edu](mailto:mcolaresi@pitt.edu)

We are indebted to the participants at seminars where previous versions of this research were presented including at the University of Uppsala, the University of Pittsburgh, the St. Louis Area Methodology Meeting 2018, the University of Kentucky, the Peace Science Society Meeting 2017, and New York University. This research was supported in part by the University of Pittsburgh Center for Research Computing through the resources provided and NSF grant SES #1753528. Replication files are available at the American Political Science Review Dataverse: <https://doi.org/10.7910/DVN/CCGPUQ>.

Received: January 10, 2019; revised: March 24, 2020; accepted: April 9, 2020.

---

<sup>1</sup> Further, even high profile subsets of human rights, such as physical integrity rights, are bundles of related concepts measuring protective and repressive behaviors.

concepts. Because of this reconceptualization, we are able to explore a new set of observable implications of information effects and estimate, for the first time, the evolving taxonomy of human rights concepts that structure the systematic comparison of government behaviors in given years. We also provide a new measure of the available information density that tracks changes in ICT over time.

Our research design uniquely allows us to analyze not only the changing explicit taxonomy of concepts that human rights organizations (HROs) label as being judged over time but also the implicit taxonomy of human rights that might be judged in past texts but remain unlabeled in document metadata. Adapting tools from machine learning and natural language processing to work with the concept of taxonomies, inspired by work in biology, we identify a dramatically evolving taxonomy of human rights that are being judged not only in one of the foundational human rights corpora, *Country Reports on Human Rights Practices* from 1977 to 2016 by the US State Department<sup>2</sup>, but also texts from Amnesty International from 1977 to 2016 and Human Rights Watch press releases from 1997 to 2016.<sup>3</sup> Our analysis of changes in the taxonomic structure, attention, and sharpness of human rights concepts across organizations is consistent with the explanation that the thickening scale of human-rights-relevant information is being incorporated into more specific concepts that define fine-grained distinctions between rights-violating and rights-protecting behaviors. We find that a model exploiting our available information density predicts future taxonomic change more accurately than other potential explanations do.

Further, we offer new methods and an interactive application that allows researchers to detect the specific or general human-rights concepts of interest from a given taxonomy across country reports so that comparisons could be made consistently by human coders or automated systems. We identify human rights concepts that have been consistently scored across time, providing the necessary leverage to create a valid text-based measure of human rights in the future. While both human (manual) and automated coding procedures have previously used explicit labels to guide reading and input, our analysis suggests that specific rights, such as physical integrity rights, are being judged outside of their explicit sections (locations), for example in the political-rights-labeled discussion, in previous years. Our tools can thus be used by future researchers to prioritize specific locations in reports based on the content of those sections, not potentially less accurate explicit section labels. We also show how our results and research design can reveal the different rights that are covered in reports across organizations for a given country in a given year. Finally, our evidence suggests that other concepts that are difficult to measure could be influenced by information effects.

<sup>2</sup> Available from <https://www.state.gov/j/drl/rls/hrrpt/>.

<sup>3</sup> For Amnesty International (<https://www.amnesty.org/en/search/?q=&documentType=Annual+Report>); for Human Rights Watch (<https://www.hrw.org/news>). More discussion of these corpora is provided in the Online Appendix.

## INFORMATION EFFECTS AND THE DIFFICULTY OF MEASURING HUMAN RIGHTS

A growing body of work has begun to question the naive comparison of existing measures of human rights behaviors across periods. One group of scholars suggests that the improving quantity and quality of evidence on human rights events around the world, from satellite imagery, camera phones, and internet and social media access<sup>4</sup>, has led observers to mistakenly infer trends in actual violations from changes in the evidence that is available about abuses.

In an influential article, Clark and Sikkink (2013) suggest that the information-collection tools that are available to human rights observers in government agencies and (HROs) have improved drastically over the last several decades. Using an analogy of disease and medical screening, Clark and Sikkink (2013) point out that “increased awareness” and “better information” (542) have led both to “more” and “better” data on violations over time.

Building on this idea, Fariss (2014, 2019b) suggests that new information has led to “changing standards of accountability.” Human rights organizations are now better able to detect violations, classify more acts as violations, and press governments harder for human rights reform. Thus, he argues that “[t]he standards-based data are potentially biased not because the coding procedure is biased but because the reports themselves are produced by monitoring agencies that are changing the standards that they use in the process of documenting human rights abuse” (Fariss 2019a, 19).

The core idea shared in the influential works by Clark and Sikkink (2013) and Fariss (2014) is that the information contained in reports has second-order effects on the valid interpretation of human rights scores derived from those texts.<sup>5</sup>

### The Evidence For and Debate over Information Effects: Word Counts and Harsher Scores

Yet there are a few reasons to further explore the role that evolving information technology could play in the production of human rights reports. To date, the plausible idea that there has been an increase in the

<sup>4</sup> As compared with previous decades where the technologies for collection and transmission relied on binoculars, analog newsletters, land-line phones, and analog mail, the evolution of ICT has aided in the development of human rights reporting agencies’ capabilities and spread, as we discuss below.

<sup>5</sup> Where Clark and Sikkink (2013) clearly highlight changing informational contexts, the connection between Fariss’s (2014) argument and the availability of information is more subtle but important. Unless (a) human rights accusations and judgments are arbitrary and not supported by information and facts or (b) human rights report composers were holding out on evidence of identified violations in the past, a changing standard of accountability is only possible and credible when there is the increasing availability of information to support judgments on more specific rights. We return to this idea of more specific rights below.

information available to the composers of textual human rights records over the last four decades has only been supported by indirect evidence through word counts and measuring the harshness of human-coded scores over time.<sup>6</sup> Both Clark and Sikkink (2013) and Fariss (2014) count the number of words in sets of texts over time, arguing that more information could be translated into longer reports. It could be countered, however, that conceptual redundancy or attempts to signal a commitment to human rights by writing verbose reports could also explain longer texts, without a deepening information base. Richards (2016) uses counts of subsections to make this case, stating that overall word counts are not definitive evidence of bias in the human-coded rights measures or information effects more generally.

Additional indirect evidence of information effects has been offered by linking report length and external sources of information to the harshness of the judgments found in human-coded scores. For example, Clark and Sikkink (2013, 547) suggest the possibility that “more detailed information drives harsher coding” and present an empirical result that longer reports for a country in a given year are assigned worse scores, controlling for the previous human-assigned score. Fariss (2014) uses a measurement-model approach and finds that, when a dichotomous event indicator of at least one violation in a given year is assumed as ground truth, human coders appeared to give harsher scores in later years than in earlier years. This research has generated a set of influential corrected scores as well as some controversy about whether the correction is necessary (Cingranelli and Filippov 2018b; Richards 2016).

However, it is important to note that none of the criticism of Fariss (2014) engages with his core *theoretical* argument and instead focuses on his measurement strategy and necessary identifying assumptions. In the next two sections, we take up the charge of how to move the literature on information effects and changes in human-rights reporting further forward, first by offering an empirical puzzle and then by engaging specifically with the theoretical mechanism posited by the information effects and changing standards of accountability conceptualizations.

### Missing Information about Information Effects

Three extant gaps in this literature have thus far blocked progress on resolving this debate and pointing a way forward. First, previous work has not systematically measured changes in relevant information and communication technology. Improvements in (a) open-source satellite imagery, (b) internet-connected camera phones, (c) the adoption of social media, and (d) internet access have not followed a linear trajectory. We fill this gap by estimating the latent available

<sup>6</sup> However, see Greene, Park, and Colaresi (2018) and Bagozzi and Berliner (2018) for work that provides evidence of potential instability in composition of human rights reports.

information density underlying increases in indicators of these four technologies.<sup>7</sup>

Second, the explicit harshness of the language in the texts has not been directly analyzed. To date, only the scalar human-coded scores, not the underlying texts, have been analyzed for a trend in worsening features over time and a potential correction offered by using an explicit measurement model. A skeptic could argue that the event-based measures used by Fariss (2014) to correct text-based scores may also be generated and affected by the increasing availability of information over time. If that was the case, the estimated harshness of the human rights scores would not necessarily be a reflection of the tone of the underlying texts, but simply a reflection of the relative measurement error in the event data as compared with the human rights scores themselves. On the other hand, as his measures use scalar summaries that are updated through time, it is possible he mitigates the reporting bias that Baum and Zhukov (2015) and Weidmann (2015) identify in event data.<sup>8</sup> Thus, the pattern of relative judgments over time in the reports is an open question.<sup>9</sup> Below, we map the relationship between available information density and the harshness of language in the US State Department Reports to explore this potential linkage.

Third, and even more fundamentally, previous research on information effects has largely ignored potential changes in the definition of the concept of human rights. Previous empirical work, in particular, has assumed either that human rights is a singular concept or that the relevant set of plural human rights being judged in these texts is fundamentally stable over time. However, if we conceptually disaggregate human rights, as suggested by Richards (2016), Bagozzi and Berliner (2018), and others, information effects are likely to lead to potentially dramatic shifts in types of behaviors about which human rights monitoring organizations can collect systematic evidence. Clark and Sikkink (2013) hinted at this type of potential change where they write that the recording of “on-the-ground changes in types of violations” (557) and “a wider range of rights” (560) could be evident in more recent reports than in older texts.

With a greater density of information available, more specific rights can be viewed as being violated or protected within reports. If human rights are plural, then

<sup>7</sup> The Bayesian dynamic latent variable model is described in the Online Appendix.

<sup>8</sup> We see the theory for a changing standard of accountability as complementary to our focus on a changing taxonomy of rights. Where Fariss (2014) is attempting, in part, to measure the threshold for a human coder’s judgment, for example, to score a country as a 3 instead of 4 on the Political Terror Scale, we are attempting to measure the underlying rights on which judgments are being expressed in the underlying texts that those human coders were using. We return to these synergies in the next section and the conclusion.

<sup>9</sup> Recent work analyzing the text of these reports, while finding evidence consistent with information effects, does not find that human rights reports are scored systematically harsher in later years (Greene, Park, and Colaresi 2018), but it does suggest the potential for a “topical shift” (Bagozzi and Berliner 2018).

this opens the possibility that the set of rights being judged in relevant texts could shift over time in observable ways. To be clear, the argument here is not that standards for encoding a violation of a fixed right have changed but that the very set of rights being judged across countries can evolve and grow.

Yet, as we detail further below, simply counting the number of rights being judged is insufficient to test information effects. We believe information effects are likely to be visible not only in the number of the rights being judged but also, more importantly, in the organization of those rights. Here we investigate this potentially complementary information effect, as it has gone unexplored empirically.<sup>10</sup> We are the first to ask whether the underlying taxonomy of human rights protections and violations being judged in reports has changed as available information density has increased. Understanding the proposed mechanism of the changing standards of accountability could be limited without a reference to the aspects of human rights that are being judged across countries because the changing expectations might consist of not only more stringent judgments on a fixed set of rights but also the addition of judgments on new forms of human rights protections and violations. Further, even if we were to find constant judgments over time, this could simply be a reflection of the different rights being judged in later years, as compared with earlier reports. For example, it may be the case that newly added rights are more difficult or costly to protect than older rights are. Thus, the comparison of countries, regions, or global behavior over time might be biased towards more negative language overall, but it may be more positive over time on specific rights that have been judged consistently over a longer period.

### Only Harsher Judgments?

One way to probe whether more theorizing on the nature of information effects in human rights reports is needed is to return to the underlying texts and measure the underlying sentiments/judgments in the reports as information density has increased over time. This will help us move beyond the counting of words or indirect measurement of tone from aggregate manual scores back into the content of the reports.

If expectations and the ability to collect evidence of violations and protections have grown at a faster rate than states' practices have improved, we would expect

<sup>10</sup> Our focus on the changes in the taxonomy of human rights does not imply that there are not potential changes in standard for whether a violation on a fixed right is triggered. We only point out that it is incoherent to discuss the relative judgments on a given right if you do not first identify what that right is. Similarly, it is invalid to compare an aggregate judgment over a set of rights, again, if those rights are not comparable. In sum, our argument is that the rights being judged are logically prior to the expressed judgments on those rights. Thus, first, by exploring the set of rights that define systematic textual comparisons of countries' human rights behaviors, we are aiding research that attempts to compare judgments on those rights in the future.

to see overall negativity from reports increase as information density accelerated. Although preliminary, this would potentially supply more direct evidence consistent with information effects and the changing standard of accountability.

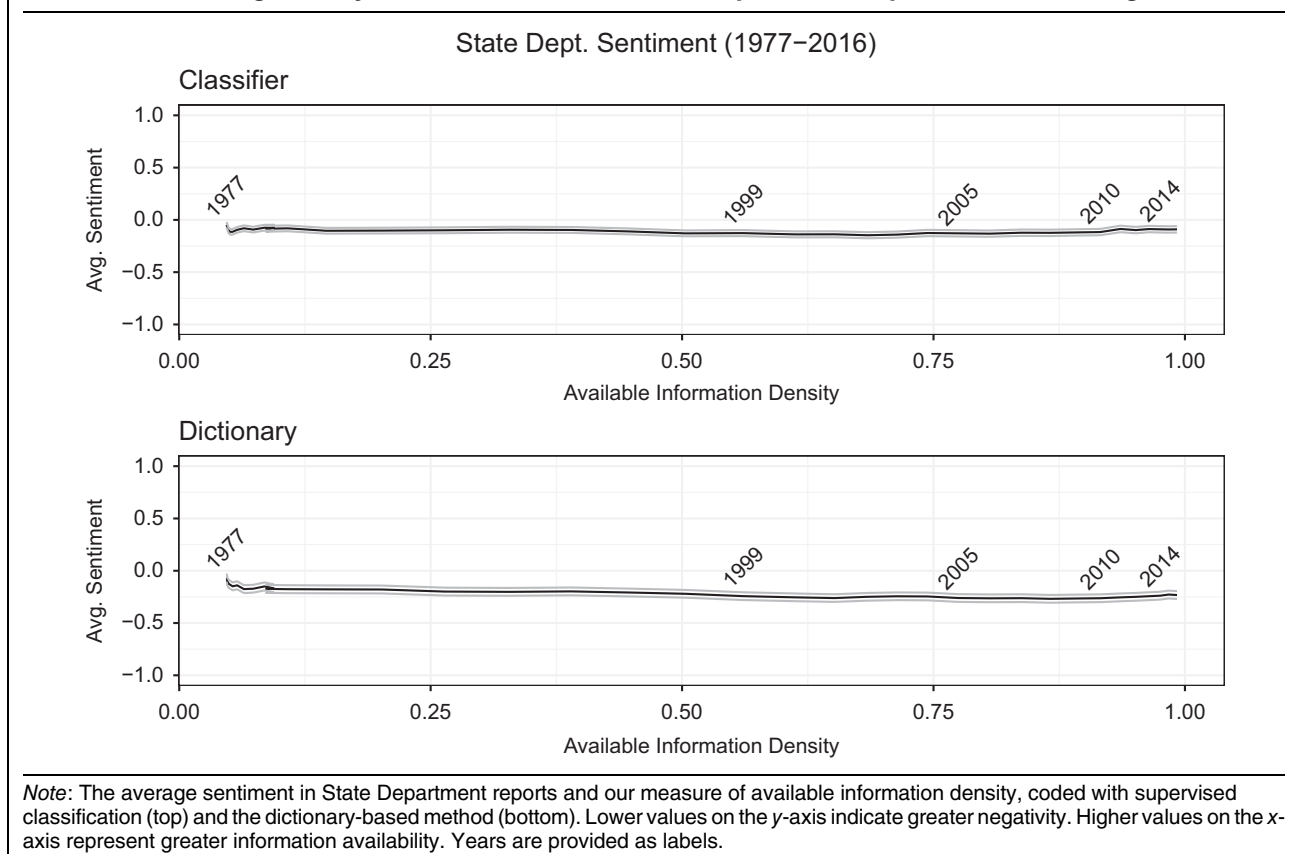
Figure 1 illustrates the average sentiment coded through a sentiment dictionary (bottom panel) and a supervised learning approach (top panel) in the reports from 1977 to 2016 (See Online Appendix for details). Interestingly, we do not see the expected trend where average judgment/sentiment becomes more negative when available information density increases in more recent years. While the dictionary approach shows a very slight decline, most of the change occurs by 1985, which is before the World Wide Web was invented (Gillies and Cailliau 2000). Moreover, the same pattern is not apparent in the more accurate classifier approach.

If the country reports do not contain more negative judgments/sentiment on states' human rights practices, on average, as available information increased, how do information effects present themselves in the reports? Our answer to this question lies in the previously unexplored hint in Clark and Sikkink (2013) that there have been "on-the-ground changes" that are conceptually independent from judgments but involve the definition of human rights themselves. This focus on the rights being judged as compared with the judgments themselves also builds on Fariss's (2019a, 19) as yet unexplored suggestion that we need more research on "the process of documenting human rights abuse." We accelerate and deepen this line of research below by exploring whether it is the taxonomy of human rights being judged over time that have changed (and how) rather than the positive or negative judgments themselves.

In fact, if we can identify changes in the underlying rights being judged over time, the relatively flat pattern in judgments in Figure 1 could be explained as Fariss's (2014, 2019a) theory of changing standards suggests. As one set of human rights was criticized and then potentially improved, another set of rights, where improvement was needed, would receive more attention. Thus, this would lead to more rights being judged, but with the overall harshness remaining relatively consistent.

### A GROWING TAXONOMY OF HUMAN RIGHTS ACROSS REPORTS?

It is important to justify our reconceptualization of human rights beyond an aggregate fixed set of concepts and then argue how increasing information is likely to be incorporated and observable in the text of human rights reports. The distinctions and connections between individual rights are central to our research design. Foremost, the concept of human rights is itself plural. Landman (2006, 8) explicitly defines human rights as "a set of individual and collective rights that have been formally promoted and protected through international and domestic law . . ." This is mirrored in other influential work (Finnemore and Sikkink 1998; Sikkink 2011). Thus, by definition, the concept of human rights nests a taxonomy of related but distinct

**FIGURE 1. Average Yearly Sentiment in the US State Department Reports on Human Rights**

protections such as political, worker, civil, and physical integrity rights (Finnemore and Sikkink 1998; Landman 2006). Core legal texts are often dedicated to defining specific rights, such as the International Covenant on Civil and Political Rights and the International Covenant on Economic, Social, and Cultural Rights (Forsythe 2017, 54).

Even more important than the number of rights being judged in reports is the organization of those rights. Many human rights reports explicitly organize the rights being focused on into hierarchies or taxonomies, where the overall concept of human rights is the root and more specific rights are defined downwards, providing explicit distinctions to classify ever more specific subsets of rights (Forsythe 2009). Thus, one can think of the organization of human rights concepts as analogous to biological taxonomies used to classify life into related but distinct bins of increasing specificity as it descends the hierarchy. As we will describe in this section, denser information should allow finer-grained distinctions between rights.

In fact, the influential State Department reports include section and descending subsection headings that organize each country report in a given year. For example, each country report in 1977 included four explicit sections “Governmental Attitude Regarding International and Nongovernmental Investigation of Alleged Violations of Human Rights,” “Respect for Civil and Political Liberties,” “Respect for the Integrity of the Person,” and “Governmental Policies Relating to the Fulfillment of Such Vital Needs.” The “Respect for

the Integrity of the Person” section was then itself organized into four subsections “Cruel Inhuman or Degrading Treatment or Punishment,” “Arbitrary Arrest or Imprisonment,” “Denial of Fair Public Trial,” and “Invasion of the Home.” Thus, a specific country report in a given year is usefully conceived of as not one report but as a set of reports on an unknown number of nested human-rights concepts. To date, the debate over information effects and standards has not grappled with what we argue is a key indicator of these technologically influenced processes: as technology makes relevant information more available to monitoring agencies, the taxonomy of human rights that structure the judgments of human rights reports will evolve.<sup>11</sup>

### A Theory of Information-induced Taxonomic Shifts in Human Rights Reports

As noted earlier, there has been an increase in the availability of dense information about human rights over the last four decades (Clark and Sikkink 2013; Fariss 2014). It is difficult to ignore the dramatic change in

<sup>11</sup> Although see Bagozzi and Berliner (2018); Clark and Sikkink (2013); Fariss (2014); Greene, Park, and Colaresi (2018); and Poe, Carey, and Vazquez (2001) for work analyzing changes in the underlying reports more generally. Our suggestions are distinct from the topical shifts in Bagozzi and Berliner (2018) because we focus on changes in the organization of concepts that comprise human rights.

the technology of observing human rights, moving from observers using binoculars on a hilltop to satellite images and camera phones. For example, Amnesty International noted in 2013 that social media and other tools meant that “investigators now have hundreds of potential crime scenes at their fingertips.”<sup>12</sup> Similarly, satellite images from DigitalGlobe were used by Human Rights Watch and Amnesty International for evidence in the Myanmar, Burundi, and Iraq<sup>13</sup> initiatives that were run by the US State Department to call on citizens to explicitly “use your phone to stop conflict and abuse,” highlighting the documentation that can be gathered by the proliferation of internet-connected camera phones.<sup>14</sup> The United Nations Special Rapporteur on Summary Executions wrote in 2015 that “we have all seen how the actions of policy officers and others who use excessive force are captured on cell phones.”<sup>15</sup> Propelled by these technologies, the network of HROs and volunteers available to wield these tools and amplify their messages has grown in scale (Carpenter 2014; Keck and Sikkink 1998; Murdie 2014a, 2014b; Wilson, Davis, and Murdie 2016).

Because human-rights monitors have, on one hand, incentives to use as much information as possible to avoid making false accusations or missing actual violations and, on the other hand, to mitigate information overload and efficiently process the available signals on violations and protections, we posit that composers will deal with increases in information by systematically altering the way they organize the concepts that are judged in human rights reports. Specifically, denser information will allow more and sharper attention to deeper taxonomies of rights.

### *Incentives for Increasing Taxonomic Complexity*

As denser, higher-resolution information flows into human rights monitors such as the US State Department, Amnesty International, and Human Rights Watch, there are two possible responses: the new data can be categorized within an existing taxonomy of rights that is used to create a new taxonomy of rights from scratch, or it can spur more nuanced changes in the underlying taxonomies. Research on the organiza-

tion of knowledge suggests that the latter option is the most efficient for sorting and analyzing this new information.

Processing new and complex information within an unchanged, existing taxonomy that was optimized in a relatively simpler information environment is likely to be costly. A failure to adapt conceptual taxonomies to new evidence can result in information overload, whereby distinctions between cases are missed or aggregated together due to superficial similarities (DeCanio and Watkins 1989). In particular, without proportional changes in the efficiency of organizing and composing human rights reports, an influx of new information will necessarily lead to underinterpretation of relevant evidence and relevant across-country distinctions. For example, if one is simply classifying images by color, greater detail and nuance, as compared with a constant hue across pixels, is likely to impair the coding process. For human rights monitoring agencies, any unintended mistakes are likely to lead to less accurate textual descriptions of states under evaluation and invite controversies. However, creating new categories from scratch is also costly, adding additional organization overhead to the information processing that would be necessary in any system.

One way to combat the risk of information overload is to efficiently encode this new information by extending an existing network of concepts with new nodes/taxons that describe the additional distinctions and details that are systematically available. These new nodes/taxons encode new distinctions within a class, but they retain similarities that conform to its parents or neighbors (DeCanio and Watkins 1989). For example, when pictures and video of paramilitary gangs intimidating political opposition become readily available, distinctions about whether those masked troops are using government-issued weapons and vehicles or not and their potential ties to the official security apparatus are newly visible. This information could be encoded into a new facet of physical integrity violations<sup>16</sup> focused on the role of the security apparatus in repression (Carey, Colaresi, and Mitchell 2015). Previously this level of detail would have been difficult to formulate across cases and thus would not be systematically included across cases.<sup>17</sup> Incomprehensible piles of evidence are avoided by having an appropriately complex and deep system of related concepts/piles.

In addition, creating new human rights concepts and standards is a complicated political process.<sup>18</sup> A group

<sup>12</sup> Amnesty International. “Twitter to the Rescue? How Social Media Is Transforming Human Rights Monitoring” February 20, 2013, available from <https://www.amnestyusa.org/twitter-to-the-rescue-how-social-media-is-transforming-human-rights-monitoring/>.

<sup>13</sup> See for example, “A New Fleet of Satellites is Detecting Human Rights Abuses From Space.” Australian Broadcasting Company September 4, 2018.

<sup>14</sup> See for example, “Use Your Phone to Stop Conflict and Abuse.” Young African Leaders Initiative: US State Department, available from <https://yali.state.gov/use-your-phone-to-promote-human-rights/>.

<sup>15</sup> See “Smart Phones New Tool to Capture Human Rights Violations.” UN Inter Press Service, June 23/2015. Further, applications such as Ushahidi also rely on smart devices to crowdsource the documenting of abuse and are used by monitoring teams in the field. Information on Ushahidi is available from <https://www.ushahidi.com>. For an example of Amnesty International using the Ushahidi platform, see <https://www.amnestyusa.org/dont-ignore-the-dire-human-rights-situation-in-sudan/>.

<sup>16</sup> Where physical integrity violations is assumed to be a preexisting category.

<sup>17</sup> This does not mean that the rights are new to everyone, only that there previously was not sufficient information to systematically compare countries on these aspects within an annual report.

<sup>18</sup> Because we believe changes in ICT are relevant to many human rights monitoring agencies, we are not making a decision between government agencies such as the State Department, intergovernmental organizations such as the UN, or nongovernmental organizations such as Amnesty International and Human Rights Watch. We explicitly compare and contrast several important and large monitoring organizations in the Online Appendix.

of scholars studying the emergence of new human rights norms/issues have focused on why some human rights issues are defined and adopted as new norms and others are not (Carpenter 2014; Keck and Sikkink 1998; Murdie 2014a, 2014b; Wilson, Davis, and Murdie 2016). First, the issues involving certain types of violations that identify and establish “clear causal chain assigning responsibility” are more likely to be recognized as new norms (Keck and Sikkink 1998, 27). Also, human rights issues that “resonate” within preexisting moral frames or advocacy networks tend to emerge as new international standards (Bob 2002; Florini 1996). Bob (2002) emphasizes that most HROs have paid attention to human rights issues that can be categorized under the International Covenant on Civil and Political Rights. New issues colliding with those already existing standards are less likely to be recognized and adopted as new norms. Third, the sociopolitical dynamics between human rights advocacy networks have a significant influence on issue adoption (Carpenter 2007, 2014). Carpenter (2007, 116) insists that defining and adopting new human rights norms is often the result of “bureaucratic and coalitional politics” within HRO networks. Thus, processing newer information about human rights is much more difficult and complicated than is often assumed.

Computational models of organizational structure and innovations consistently find that the optimal structure to adapt to new information will vary depending on costs of information processing (DeCanio and Watkins 1989). If information is simple and easy to understand without expertise, then flat structures without much hierarchical differentiation are ideal. On the other hand, if specific expertise is needed to make sense or judge the new evidence, as in the human rights context, then deeper hierarchies are especially useful.

### Observable Implications in the Texts

Our theory that increased ICT will lead to a deepening of human rights taxonomies guiding judgments in country reports in later as opposed to earlier years has important observable implications for the structure and content of the annual reports over time. Specifically, we describe three new testable predictions consistent with our proposed available-information density mechanism. We then turn to outlining the tools from machine learning and natural language processing that will allow us for the first time to extract the implicit taxonomies of human rights concepts from texts from multiple reporting agencies.

#### *Implication 1: An Increase in the Structural Depth of Taxonomies Judged across Countries*

We expect that the structure of human rights taxonomies will become increasingly hierarchical, with additional distinctions for newer concepts, as denser information is available to reveal these distinctions in evidence. Information effects and the concomitant higher-resolution signals will be most efficiently processed not by creating new orthogonal concepts out of

whole cloth but by leading to new distinctions between previously aggregated violations and protections (Barner and Baron 2016). Judgments differentiating killing as a part of the judicial system from extra-judicial killing, LGBT rights from religious protections, and freedom of information from press freedom might be consistently present only in later data, not in earlier texts where the technology for evidence collection was less available. Scholars note that more specific types of violations, such as sexual violence and the exploitation of child soldiers, were judged in recent reports but “overlooked” in older country texts (Clark and Sikkink 2013, 542; Keck and Sikkink 1998). Alternatively, it could be the case that the same rights are judged consistently across the years, without new concepts being compared across countries. Given the literature on information overload and efficient processing, this contradictory result would imply that more relevant information was not being processed into the text over time.

In addition, because the reports attempt to compare countries across time and space, there is likely to be some continuity to facilitate these interspatial and intertemporal comparisons.<sup>19</sup> Greater information on repression around the world, for example, might allow governments and HROs to observe both the age of those who are fighting as well as how and why they were forced to be a child soldier, where in the past only the occurrence of repression might have been observable, not the details.<sup>20</sup> A preexisting category of repression would have been carried over from the past, and then new conditions/distinctions added, creating a new deeper set of taxa.

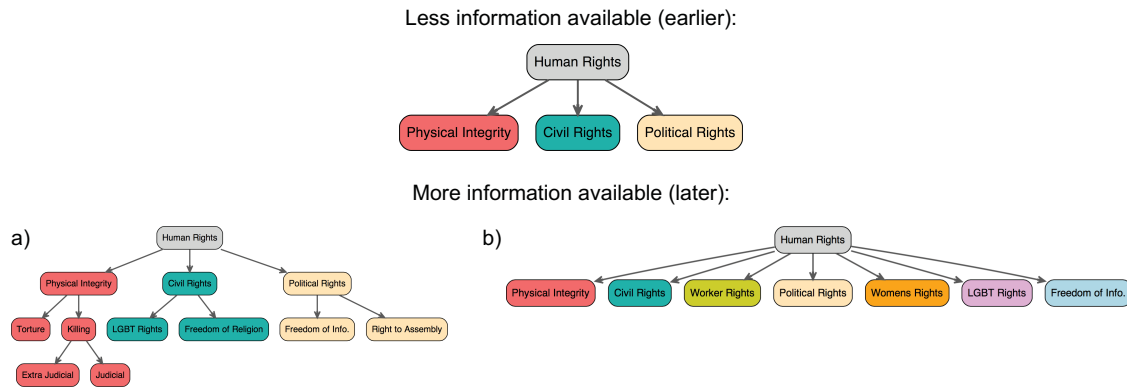
A competing possibility is that there is simply an increasing multitude of distinct, non-nested concepts in newer as opposed to older taxonomies that are being judged. If new types of violations are identified, they may have been added to the bundle of protected human rights at the top of the hierarchy. The reports would then be tracking more distinct aspects, which are only related to other sibling aspects by the fact that they are human rights. Instead of zooming in and increasing the clarity of existing pictures of human rights, new information can simply bring new concepts of rights.

The two different types of changes in taxonomies, from the sparse and flat (top) example taxonomy towards either the deeper multilevel hierarchy (left) or towards a wider, consistently flat grouping (right), are displayed in Figure 2.

These are ideal types, as it may be the case that both types of changes occur, new top-level concepts and new descendants of existing rights. In that mixed case, we are interested in the balance of new concepts and their position in the taxonomy. We will return to how we operationalize this balance below.

<sup>19</sup> The State Department itself conducts analyses to judge the consistency of sections over time. (see GAO-12-561R).

<sup>20</sup> For example, the US State Department Report, 2015, Burundi, Section 1.g.

**FIGURE 2. Two Types of Structural Changes to Human Rights Aspects Over Time**

*Note:* The top subplot is earlier, and it only has two levels of hierarchy. In subplot (a), which illustrates our prediction, the hierarchy has grown, unevenly, to three and four levels of specificity across different aspects of human rights. Subplot (b) illustrates another possibility, which would not be consistent with our predictions. Here, the hierarchy has grown horizontally but not vertically. New concepts are created from distinctions from the original overall concepts (root) instead of being more specific semantic concepts of preexisting non-root leaves. Thus, there are more nested and deeper concepts in (a) than in (b).

### *Implication 2: An Increase in the Systematic Attention to Deeper Rights*

Further, with greater access to additional dense information on finer-grained rights, the amount of attention systematically paid to those rights will grow. We do not believe that the effect of ICT changes will stop with the structure of rights. Denser depictions of rights and protections allow more paragraphs to be written about the evidence relevant to judgments on those rights.

### *Implication 3: An Increase in the Sharpness of Distinctions between Rights*

As our theory is about distinctions between rights, additional information collected by monitors from smart phones, the internet, social media, and satellites should more clearly delineate when an issue is being discussed. Just as an image at higher resolution reveals crisper distinctions between objects, so should increased information on events around the world clarify previously obscure situations and how they map to specific rights. For example, when it was only clear, given past limits to available information, that a civil war was ongoing, there may have been the possibility of sexual violence or the use of child soldiers. However, it would have been difficult to explicitly document and clearly discuss any of these specific potential violations. Alternatively, with additional information, the distinctions between situations when rights are being violated should be much sharper in discussions. Clear language and key phrases should signal when sexual violence has been detected in photographic stories and verified posts or pictures of child soldiers in the civil war shared. The inclusion of this evidence and judgment of these rights should allow for a sharper distinction than would be possible with less available information (in the past).

### **Detecting Explicit and Implicit Taxonomic Changes in Human Rights: Structure, Attention, and Sharpness**

Our research design is inspired by aspect-based sentiment analysis, which extracts signals and concepts from texts (Liu 2012, 2015). However, conventional applications of aspect-based sentiment analysis are limited to movie and product reviews, and they do not attempt, as we do, to learn the systematic evolution of taxonomies from text.<sup>21</sup> We are partially aided in our task by the fact that the corpora of the country reports that we begin with, the US State Department Reports, include explicit metadata on the rights that are being judged within sections, subsections, and further descendant classifications.<sup>22</sup> Define a given labeled taxonomy as  $G_t$  for year  $t$ . Thus, with a provided  $G_t$  we are able to visualize the structure of the expressed taxonomies presented from 1977 to 2016 for this set of reports. Having this structure and the connections and distinctions that define the taxonomy, we can compare the depth of branches down the taxonomy, measuring whether added distinctions are more specific and complex than those that came before and thus added further down the taxonomy, as expected by *implication 1*.

Yet, there are reasons to probe more deeply than just extracting these explicitly labeled taxonomic structures. Explicit labels might mask the actual distinctions used in the texts. For example, older texts might have

<sup>21</sup> In this work we focus on the taxonomy of aspects, as much of the previous literature has been focused on harsher judgments. We will return to this connection in future work as we highlight in the conclusion.

<sup>22</sup> We add corpora from other human rights monitoring organizations including Amnesty International and Human Rights Watch below. It is important to note that the State Department is unique out of these three organizations for the detail and consistency of their explicit taxonomy that we are able to leverage.



had implicit taxonomies that only become explicitly labeled in later reports. Thus, the judged human rights might have been consistently discussed in the text over time with only the labels becoming more specific recently. We solve this problem by using the later more detailed metadata as a target taxonomy and identifying high-performing classifiers that can detect when textual items are being judged in the text even if they are unlabeled. Specifically, we set  $G_{2015\_2016}$  as the training taxonomy. For 2015/2016, we have the most detailed explicit taxonomy in use in both years. Thus, each paragraph is labeled with a location  $y_d = \ell$  for  $\ell \in G_{2015\_2016}$  and  $d \in (1, \dots, D)$  in a  $D$ -length set of paragraphs within all country reports, and the input features (document-term matrix  $X$ ),  $X_{2015\_2016}$ , are generated from the text of these reports. We use the paragraph as the unit of analysis for annual reports. The topics of paragraphs are generally judgments on given rights. Sentences are not appropriate document lengths in these cases because many sentences do not contain judgments and many are ambiguous about the right being discussed. Likewise, using whole sections is not helpful because it is the implicit not explicit sections that we are attempting to infer.<sup>23</sup> In years before 2015, we have the text,  $X_t$ , where  $1976 < t < 2015$ , but not the labels from  $G_{2015\_2016}$ . Thus, we use supervised learning techniques to train a set of algorithms to learn the mapping from the documents in the 2015 and 2016 reports to the location in that explicit taxonomy. This mapping then supplies a prediction for the older documents across each potential label, allowing us to detect the implicit locations in the taxonomy that are being judged across all years.

We compute accuracy scores using cross-validation within our training window (The specifics are outlined further in the Online Appendix). One important step we take is to build an automated parser that detects the aspect and judgment phrases in text. Parsing Unstructured Language into Sentiment–Aspect Representations (PULSAR) is useful because it not only finds multiword expressions but also uses the syntactic structure of a sentence to identify likely judgments and the aspects that are being judged (Park, Colaresi, and Greene 2018). Without a tool that explicitly recognizes passive language and the importance of terms such as “reports,” the judgment and aspect would be omitted. In the reports for a particular country, statements such as “*There were no reports of politically motivated disappearances*” are highly informative and provide a clear positive judgment on an identified aspect. In this case, PULSAR recognizes that the phrase “*were no reports*” provides the judgment, and the aspect is within the expression “*politically motivated disappearances*.” Thus, PULSAR resolves the sentence to [“*politically motivated disappearances*,” “*NEG\_were*”].<sup>24</sup>

With a set of tools to identify which rights are discussed in each paragraph, we can measure our three important sets of patterns: structure, attention, and

sharpness. First, we can compare the implicit structure of taxonomies based on their number of branches and depth, as we can for the explicit taxonomies.<sup>25</sup>

Second, this research design allows us to summarize the amount of attention, in the form of the number of paragraphs per report, that we detect on each right. We can then look at not only when human rights appear in the taxonomy but also how much emphasis they receive over time. We measure the amount of attention to each potential rights concept as the expected number of paragraphs from the model predictions.

Third, we derive from information theory a direct measure of the sharpness of distinctions between predicted rights being discussed across all years in annual reports.<sup>26</sup> This statistic is the expectation over document set  $D$  of the relative entropy of our predictions for each document versus a uniform prior over labels. We refer to this statistic as the average sharpness, and it grows as our algorithms become more certain about classifying a piece of text (paragraph) in a specific location in the taxonomy. Our algorithms can only increase their certainty if there is information in the text that clearly identifies a distinction between one or some rights and others. When there are no relevant textual distinctions present, the algorithm is left with a flatter set of beliefs, the maximum-entropy distribution over beliefs being the limiting case, which leads to lower average sharpness. We track these values as available information density increases over time.<sup>27</sup>

In addition to detecting taxonomic changes in structure, attention, and sharpness, our research design provides additional tools to explore human rights reports across different dimensions. We can also now identify the judgment of rights that appear outside of their closest semantically labeled section. We will present this in an asymmetric confusion matrix, which will identify systematic patterns of when human rights are discussed under unexpected labels across years. This tool can guide future human and automated systems. Because our tool can be applied across reporting agencies, we also provide an example where we compare what rights are judged in a given report across monitoring agencies.

## US State Department Reports and the Taxonomy Metadata

To assess changes within human rights reports, we first leverage the unique document metadata within the text in the State Department Reports from 1977–2016 (See Online Appendix for details). The reports provide

<sup>23</sup> We return to this point further below.

<sup>24</sup> Examples of the input and output of PULSAR are provided in the Online Appendix.

<sup>25</sup> Note that this is distinct from the explicit taxonomy provided for a given year by the subheadings because there may be rights in the later 2015/2016 taxonomy that are not labeled explicitly in the past but are discussed. Our method can find those rights, where an explicit label, by definition, cannot.

<sup>26</sup> Again we use the most explicit specific taxonomy,  $G_{2015\_2016}$ , to define the set of possible rights.

<sup>27</sup> A description and proof that average sharpness is the expected KL divergence of the estimated probabilities from a uniform prior is in the Online Appendix.

coverage for nearly all countries in the world, recording instances of abuses of not only physical integrity rights but also economic, social, and political issues. In particular, the metadata provided in section, subsection, and related headings gives useful guidance on what is being discussed in specific sections of each state's report. Further, this explicit metadata forms a hierarchical taxonomy, with human rights at the root, sections as the first, most general (lowest resolution) branches, and subsequent descendant subsections as more detailed (at a higher resolution) concepts. The leaves of the taxonomy are the locations where specific paragraphs of the text appear. The State Department Reports have been instrumental to highly cited and influential measures of human rights including The Political Terror Scale and the Cingranelli and Richards (2010) Human Rights Data Project. However, because we believe information density will change not only State Department monitoring but also monitoring by other human rights-focused organizations, we extend our analysis to Amnesty International Annual Reports and Human Rights Watch press releases. While adding complexity, these extensions allow us to probe the generality of our argument across reporting organizations as well as across formats.

### Explicit Evidence of Changing Human Rights Taxonomies

Two snap shots from our interactive application, representing the number of terminal nodes/aspects and the associated hierarchies of the State Department reports are presented in Figure 3. These illustrate the rather dramatic changes in the taxonomies over time. In 1977, there were 11 total nodes/taxons, which had an average depth across sections of 2.7.

The explicit metadata is consistent with our argument on the incentives for taxonomic change in response to information effects. Both implication 1 and 2 largely hold, as there is a larger taxonomic structure in more recent years and that structure has grown in hierarchy and complexity. Recent country reports have zoomed in on more specific aspects of predefined concepts. In 1977, "Denial of Fair Public Trial" was its own leaf, the most specific aspect category defined in the hierarchy. By 2015, "Denial of Fair Public Trial" is a bundle of related leaves, with explicit judgments on aspects including "the impartial and fair public trial procedures" and "the detention of political prisoners or detainees." The increasing resolution of aspect categories within the hierarchy are suggestive of important changes in what is judged across the annual reports.

We summarize the changes in the depth of the leaves in the aspect-category hierarchies over each year in Figure 4. The y-axis plots the average depth (levels down from the root) across the top-level sections of the document in each year. To calculate the average depth of the explicit taxonomy for each year, we add the maximum depth reached by each of the seven sections and then divide by the total possible number of sections. The x-axis represents the number of final nodes/

concepts, the most specific concepts labeled in the texts, in a given year. This graph can tell us the relative push and pull for changes in the number and depth of new rights being explicitly judged over time. If the texts grew more specific while adding final nodes, then the path would proceed upwards and to the right over time. On the other hand, if final nodes increased, but average depth and precision remained constant, we would see a relatively flat horizontal line. Figure 4 illustrates the path of these explicit taxonomies across time. There is a clear movement upwards, slightly less than a 45 degree angle, as the later documents have both increasingly included new rights as distinctions within existing rights, adding complexity and depth to the previous taxonomy.

### ESTIMATING IMPLICIT HUMAN RIGHTS TAXONOMIES

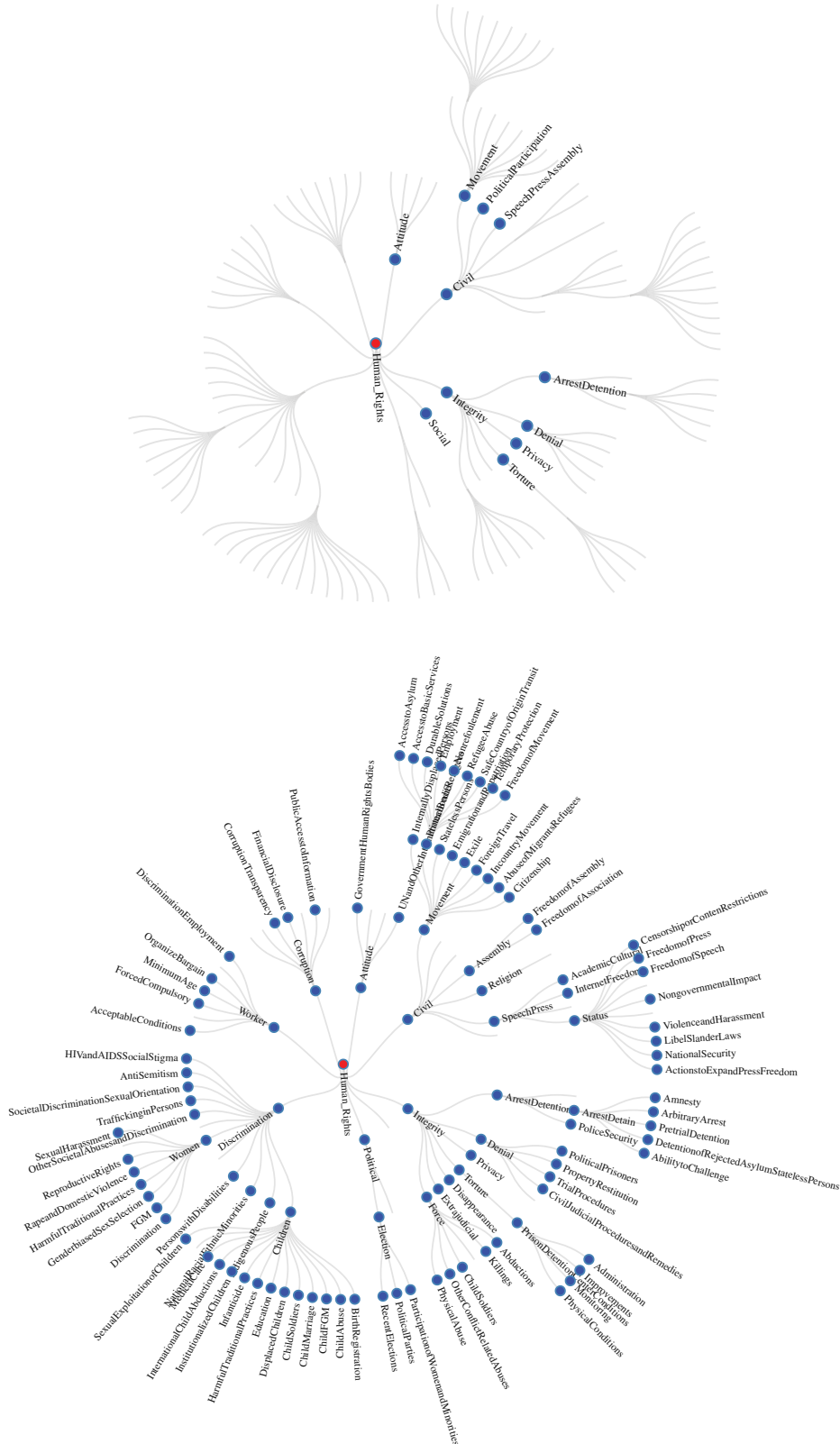
The previous section provides new evidence that is consistent with our first observable implication stated above. It appears that over time when more detailed information is available, the explicitly labeled taxonomy of human rights evaluated in the State Department Reports has grown more distinct concepts. Further, the structure of that taxonomy has evolved more complexity, with newer nodes appearing as children of existing nodes, thus introducing new deeper distinctions. Again, this is consistent with dense information that contains the more detailed information that is available over time.

Yet, as noted above, it is possible that the changes in the section labels are not reflective of the information in the actual texts underneath those labels. It might be the case that the explicit labels change over time, but the implicit lexical and semantic content of the text might be consistent. For example, the section on government corruption did not exist until 2005. This does not mean that there were no paragraphs on government/political corruption in any earlier reports. Instead, the problems of political corruption were often included in the "Arbitrary Arrest and Imprisonment" or "Denial of Fair Public Trial" in the "Respect for the Integrity of Persons" section. In later years, the "Government Corruption and Transparency" section was created, covering corruption committed by government officials in greater detail such as "Political Corruption," "Financial Disclosure," and "Public Access to Government Information." To further probe changes in the operable human rights taxonomies, we next need to analyze patterns in the natural language of the texts.

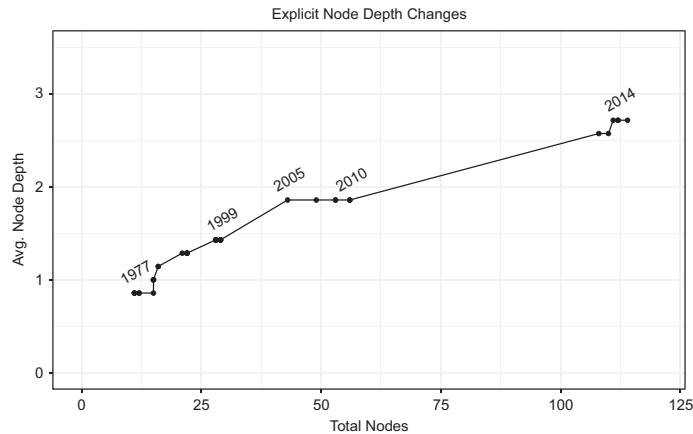
### Revealing the Evolution of the Implicit Human Rights Taxonomy

Our supervised learning approach (Kotsiantis 2007) to measuring the *implicit taxonomy* of rights is further detailed in the Online Appendix. With our computed model representation, we move on towards measuring and quantifying the information that we discover on the

FIGURE 3. US State Department Sections for 1977 and 2016



Note: Explicit US State Department sections for 1977 (top, lowest available information) and 2016 (bottom, highest available information). Each node (blue) represents a section that is explicitly covered by the report for a given year. The nodes are nested such that the Assembly node is a subsection of the main Civil Rights section.

**FIGURE 4. Total Number and Depth of Leaves in Each Aspect Hierarchy**

Note: A scatter plot of the total number of leaves in each annual aspect hierarchy (x-axis) and the average depth of leaves across the sections (first level below the root). The points are jittered slightly to avoid overplotting.

human rights being judged, first in the US State Department Reports going back from 2014 to 1977. We produce the three sets of measures denoted above. We begin by exploring the consistency of the implicit taxonomic structure through radial plots for specific years. These views allow us to present the structure of the implicit taxonomies of human rights over time, analogous to the explicit taxonomies we produced above. Using our model, we are now detecting evaluated rights, using the actual text and not just the explicit sections that were provided in a given year. Thus we can see what rights were *implicitly* judged in the text, even if labels were not provided in the explicit taxonomy.

### Structure

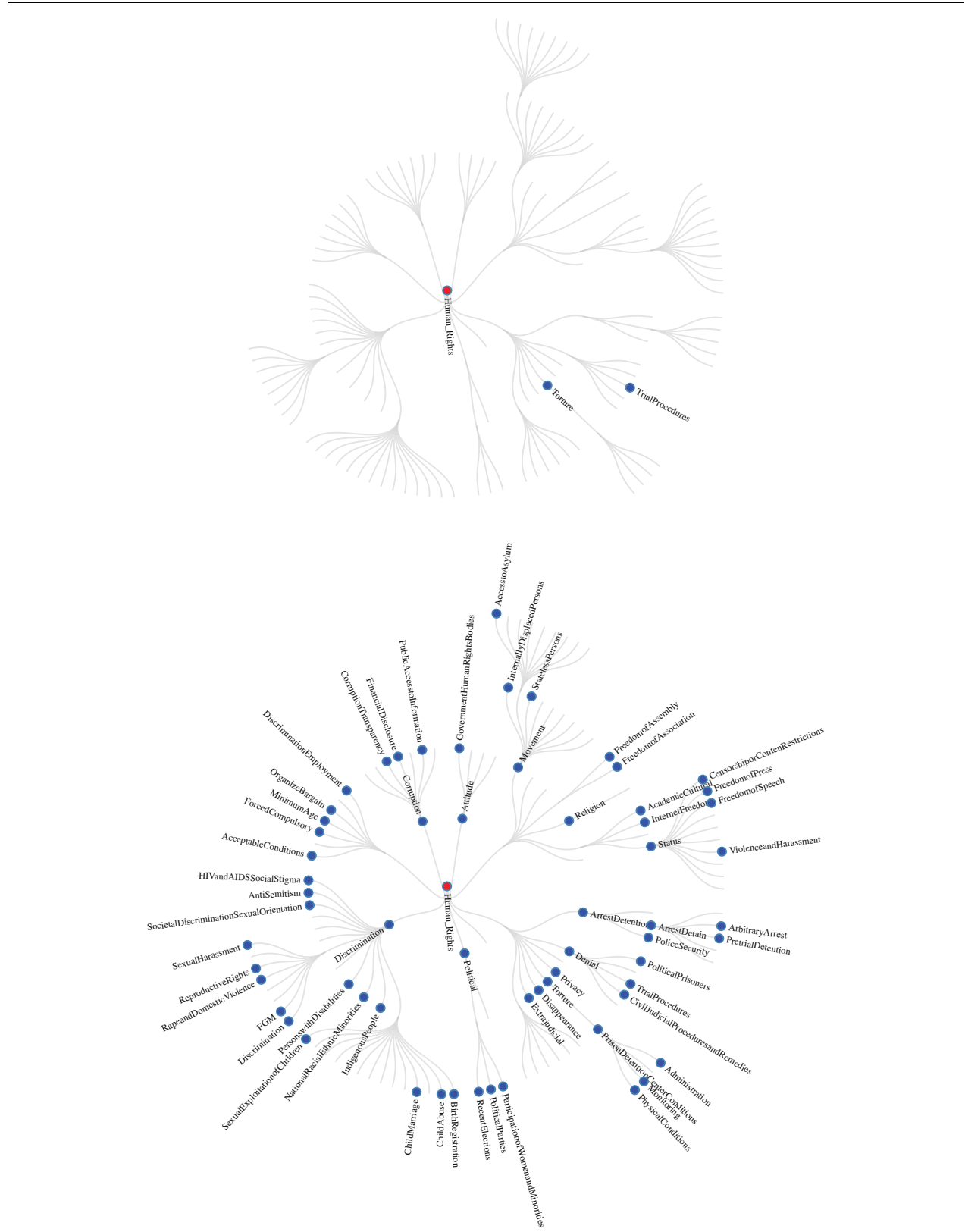
We use a rule that if there is one or more paragraphs per country report detected for a potential node in the set of reports at  $t$ , then that node exists in the structure of  $G_t$  for that year. This is a very low threshold, because it means that when five paragraphs are composed on a given rights concept in one country, the right would not have to be discussed in four other countries to meet our threshold. Further, paragraphs can be one sentence in length. These estimated implicit taxonomies for each year will allow us to calculate the number of nodes in each year and the average depth across sections, as we did in the explicit sections. Two snapshots from our interactive application representing the number of implicit terminal nodes/aspects and the associated hierarchies of the reports are shown in Figure 5. Similar to the explicit plots, these illustrate the changes in the taxonomies over time. In 1977, there were two sections that met our one-paragraph-per-country threshold, meaning that they could have been systematically judged across countries. This number increases dramatically to 62 sections in 2014.

We summarize the changes in the depth of the leaves in the implicit aspect-category hierarchies for each year

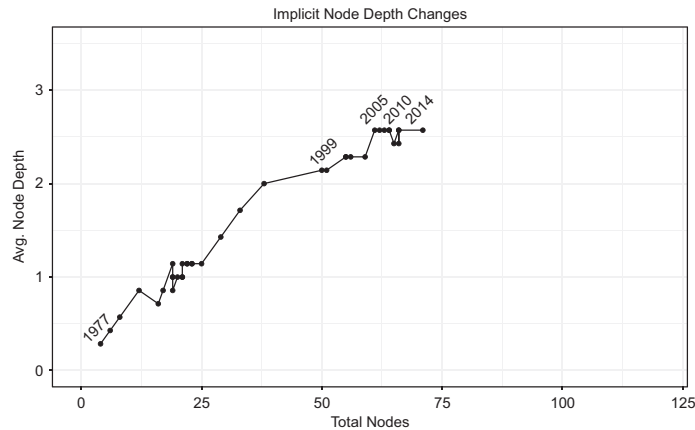
in Figure 6. The y-axis plots the average depth (levels down from the root) across the top-level sections of the document in each year. To calculate the average depth of the implicit taxonomy for each year, we first identify all of the nodes where the sum of the predicted probabilities for all texts in each year is larger than the total number of countries. Second, as in the case of the explicit average depth, we count the maximum number of all the nodes where they branch out to the next child nodes for each section and then average them by the total number of possible sections. The x-axis represents the number of final nodes/concepts, the most specific concepts labeled in the texts, in a given year. The plot illustrates the path of these implicit taxonomies across time. There is a clear movement upwards, as the later documents have both increasingly grown rights as distinctions within existing rights, adding complexity and depth to the previous taxonomy. In contrast to the explicit sections, the implicit contains fewer total nodes and in the middle years (1999–2010), but the implicit model shows greater average depth than the explicit model. This suggests that the reports actually contain information on violations and protection of a few specific human rights, even if they did not have an explicit section label in that year. We also affirm the inferences from the explicit taxonomy that judgments are being offered on more human rights and more hierarchically differentiated rights over time.

While modeling the exact mechanism by which information enters text is beyond the scope of our paper, we also explore the relationship between our measure of taxonomic depth in the US State Department Reports and available information density in comparison with two other potential conjectures. It could be the case that (a) there is a deterministic time trend that better captures changes in depth over time (compared with our measure of information) or (b) that bureaucratic changes in US Presidential administrations are driving taxonomic evolution (Cordell et al. 2020). We find that information availability provides a more accurate cross-validated prediction of future depth one and two years

**FIGURE 5. Implicit US State Department Sections for 1977 and 2014**



Note: Implicit US State Department sections for 1977 (top, least information) and 2014 (bottom, more information). Each node (blue) represents a section classified as being about a given section in the report for a given year.

**FIGURE 6. Total Number and Depth of Leaves in Each Aspect Hierarchy**

Note: A scatter plot of the total number of leaves in each annual aspect hierarchy (x-axis) and the average depth of leaves across the sections (first level below the root). The points are jittered to avoid overplotting.

into the future, compared with models that are fit with either features that represent a deterministic trend or shifts in administration. Interestingly, the best-fitting model includes both information and administrative indicators, suggesting that available information provides the opportunity to deepen coverage, but new bureaucratic players or a new agenda might be necessary to translate the information into the text systematically. We present these findings in the Online Appendix.

#### Attention

Next, we are interested in the evolution of the total amount of attention to all potential rights that we can detect across the structure of the reports.<sup>28</sup> Figure 7 presents the number of paragraphs per country report that were classified across each specific 2015/2016 aspect category over time. We present charts for the two sections with the largest coverage; the other sections can be found in the Online Appendix.<sup>29</sup> We also visually demarcate the rights that are systematically compared across countries that meet our threshold (one paragraph per country in a given year) with a dark outline. If a right's attention does not meet that threshold, it has a white border. These plots allow us to detect not only the increasing length in paragraphs of the reports over time but also the specific rights that receive increased attention in later (with higher availability of dense information) as opposed to earlier years (with lower availability of dense information).<sup>30</sup>

<sup>28</sup> As noted, we compare the expected number of paragraphs from our model on each concept across time.

<sup>29</sup> We also show that in later years more text is found in deeper subsections of the text.

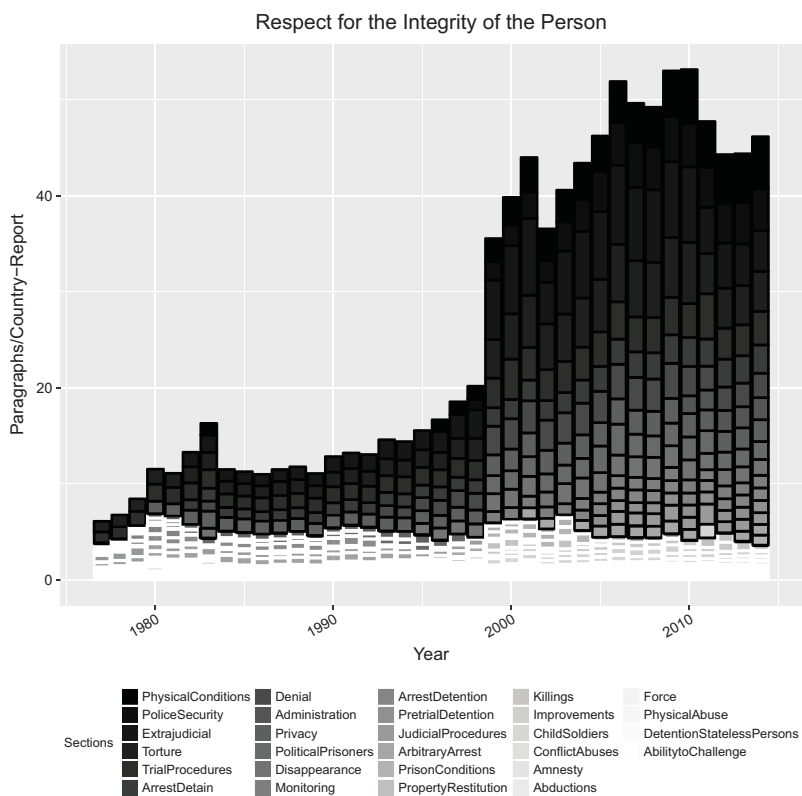
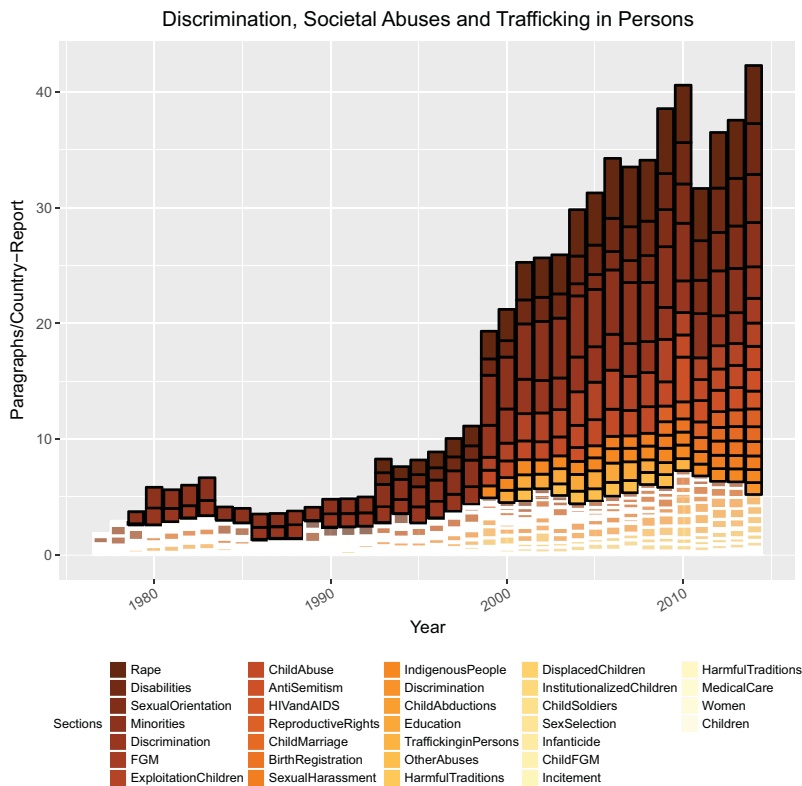
<sup>30</sup> Since our measure of available information density rises monotonically over time, the order of the bars would be identical if that indicator was binned and plotted on the x-axis.

There are substantial changes in several of the sections. The “*Physical Conditions*” aspect was the most discussed aspect of physical integrity rights in 2014, receiving roughly five paragraphs of coverage per country report, while in 1977 this right was hardly mentioned, receiving only 0.14 paragraphs per country report. The aspect “*Sexual Exploitation of Children*” in the Discrimination section showed a large increase in coverage from 2014, where it received roughly two paragraphs per country report, compared with 0.028 in 1977. In fact our model could barely detect any coverage before 1999.

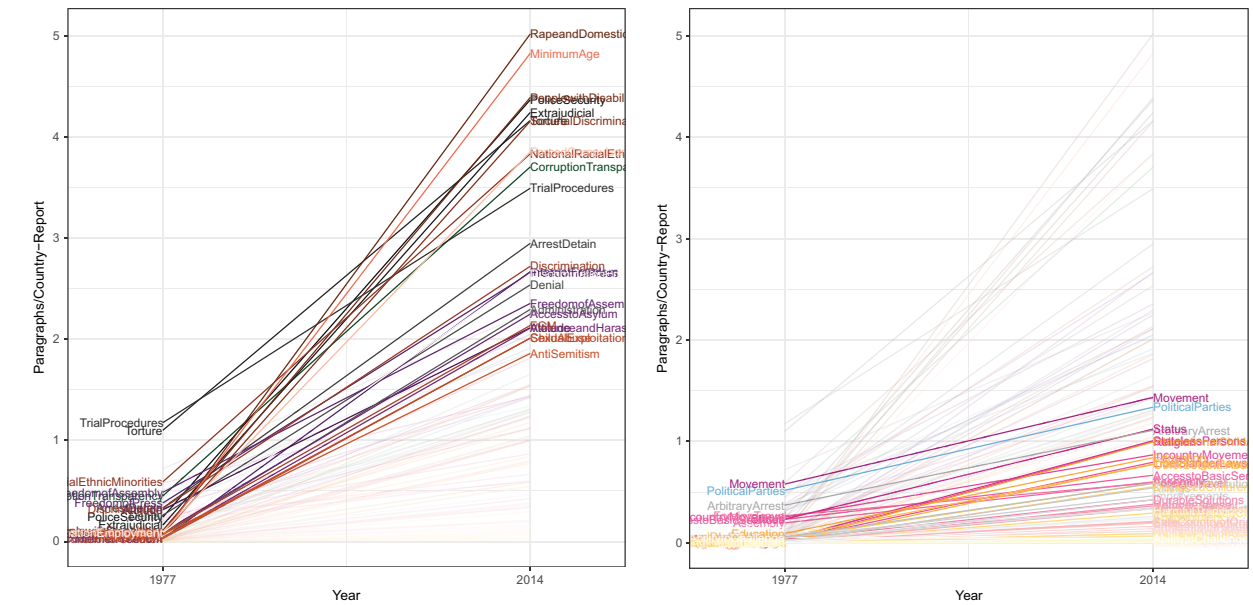
To clarify the changes over time, we provide a parallel coordinates plot in Figure 8 that presents the difference in the number of classified paragraphs per country report in 2014 and 1977 for the specific aspects that rose or declined by between one and four paragraphs per country report (left) and those that are more consistent (right).<sup>31</sup> The coverage of the “*Rape and Domestic Violence*” aspect sees one of the largest increases in coverage from older to more recent reports, as suggested by Clark and Sikkink (2013). In 2014, this right receives roughly five paragraphs per country report, while it receives almost no coverage in 1977. The coverage of the aspect “*Minimum Age*” of employment sees a similar increase in coverage. The plot on the right highlights aspects that have been consistently discussed across time. These include freedom of “*In-Country Movement*,” “*Political Parties and Political Participation*,” and “*Arbitrary Arrest or Detention*.”

<sup>31</sup> Three specific but notable outliers are omitted from this plot that were already discussed above to zoom into more nuanced shifts. “*Acceptable Condition at Work*” and “*Right to Association*,” both in the Workers’ Rights section, were found with much lower frequency in later as opposed to earlier reports. This coincides with the decline in the overall paragraphs on Worker’s-Rights aspects in later as opposed to earlier years. The aspect “*Institutionalized Children*” in the Discrimination and Societal Abuse section increased less than one in 1977 to almost 10 paragraphs in each country report in 2015.

**FIGURE 7. Number of Paragraphs on Human Rights in the 2015/2016 Implicit Taxonomy**



*Note:* Number of paragraphs on human rights with the 2015/2016 implicit taxonomy, classified per country report from a model that was trained on 112 leaf labels. The aspects are sorted from low to high based on their prevalence in 2015/2016. The outline/border of the bars is black where there is more than 1 expected paragraph per aspect per country report in that year and white otherwise.

**FIGURE 8. Change in the Paragraphs on High-resolution Aspect Categories**

Note: Change in the paragraphs on high-resolution aspect categories classified per country report from a model that was trained on the 2015/2016 leaf labels, comparing 1977 with 2014. Several examples of large changes are presented on the left, and stable aspect mentions are on the right. The colors are keyed to the sections.

### Sharpness

In our third conjecture, we suggested that increases in available information density will allow for sharper distinctions between concepts in the text. Our average sharpness<sup>32</sup> calculation carries information as defined by Shannon (1948), so it provides a nice connection to information effects and the denseness of distinctions that our model can detect in the texts.

Our evidence related to this question is presented in Figure 9. The maximum average sharpness that our predictions could theoretically hold is approximately 6.8 binary digits (bits).<sup>33</sup> The value in Figure 9 rises from 5.41 in 1977 to 6.43 in 2014. This one bit of sharpness information per paragraph has substantial meaning, as Shannon's (1948) information is additive. Thus, the same 1,861 paragraphs in the 1977 reports would carry 1,861 fewer bits, literally binary digits, than the same length of reports in 2014, due to the crisper language in the 2014 reports matching the detailed taxonomy. Moreover, since there are significantly more paragraphs in later reports, there are two sources of increasing information over time (a) longer reports that

(b) simultaneously carry more information per paragraph. Our analysis suggests that the previous research that counted words in reports over time possibly undercounted the increases in information because the paragraphs have grown more informative. Our approximate calculation is that the 2014 report contains over 193,000 more bits of information than the 1977 report does.<sup>34</sup> This suggests that our model is less able to make clear predictions in earlier years, relative to later reports, supporting our conjecture concerning more information being included on more rights in later than in earlier reports. In the Online Appendix, we also present the results from an additional metric, which we call best-case proportional reduction in error, as an accuracy measure.

### Evidence from Amnesty International and Human Rights Watch

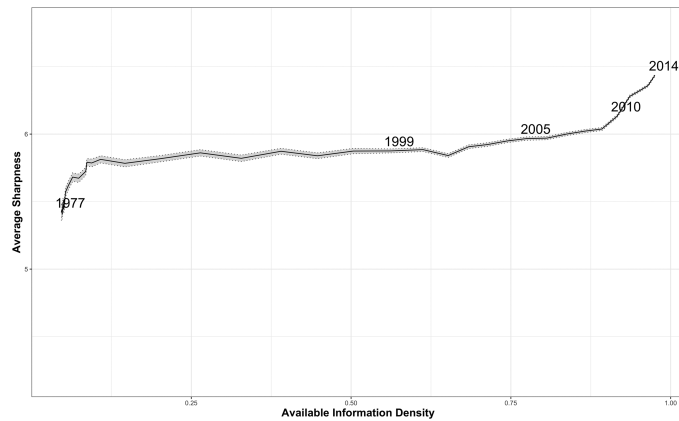
We do not believe that available-information density effects are only present in the US State Department Reports. As such, we generalize our research design and collect information from two other important corpora from monitoring agencies, Amnesty International Annual Reports (AI) (1977–2016) and press releases from Human Rights Watch (HRW) for the period 1997–2016. As discussed in the previous section, Amnesty International and Human Rights Watch are the most prominent international human rights nongovernmental organizations. However, AI and HRW do

<sup>32</sup> Average sharpness measures the additional expected bits of information in a paragraph in a given year over the information contained in the maximum-entropy message for that year. We use the log base 2 in the calculations.

<sup>33</sup> This is the log base 2 of the number of potential categories. Average sharpness would only reach this maximum if every paragraph received a prediction of 1 for some label and a 0 for all others. The practical minimum for our sharpness is given by the training-set proportions across the labels. We use this as the minimum for our plot.

<sup>34</sup> Specifically,  $(31,773 \text{ paragraphs} \times 6.4 \text{ bit/paragraph}) - (1,861 \text{ paragraphs} \times 5.4 \text{ bit/paragraph}) = 193,297.8 \text{ bits}$ .



**FIGURE 9. Average Sharpness of Our Predictions**

*Note:* The average sharpness of our predictions of the rights in every paragraph and available information density. Higher values on the y-axis reflect that our model was able to extract sharper distinctions between concepts, and lower values suggest that information on sharp distinctions per right across the taxonomy is missing. The maximum of the y-axis is set to the theoretical maximum average sharpness. The minimum is set to the average sharpness of a classifier that simply randomly assigns a label based on the relative frequency of the locations in the training set. The dotted lines represent plus or minus two standard errors from the calculated average sharpness.

not have a comprehensive and explicit taxonomy structure included in their text that contains the same level of detail as does that of the US State Department. Thus, we use the nested set of all possible nodes from the State Department in training and make implicit predictions across that taxonomy for AI and HRW.<sup>35</sup> Figure 10 shows how SD, AI, and HRW grew the number and depth of the rights, respectively, that they were judging as ICT evolved since 1977.<sup>36</sup> Figure 10 illustrates an increasing plurality of human rights over time for not only the State Department but also for AI and HRW as information communication technology has expanded. Interestingly, the slope is most dramatic for HRW. Results for attention and average sharpness for these agencies are provided in the Online Appendix.<sup>37</sup> The overall patterns are consistent with the State Department findings. As available information density has risen, it has changed the structure of the taxonomy of rights being judged as well as the amount of attention to specific rights and the sharpness of the distinctions between these concepts.

<sup>35</sup> Although AI started in 1961, the scale and availability are only consistent after 1977. Thus, we cover 1977–2016 for AI. Similarly, press releases were sparse for HRW before 1997. In addition, the unit of analysis for AI is paragraphs, but for HRW, unlike the other annual reports, press releases are very short and discuss a single topic. Thus the press releases are treated as documents themselves.

<sup>36</sup> For these calculations and the structural calculations in the appendix, we code rights as being systematically judged (and thus present in the implicit taxonomy for a given year) for AI if they are detected in 0.3 paragraphs per country report in that given year. If we used the same threshold as the State Department, almost no nodes would be present across time, as AI are sparser. Similarly, for HRW we use a threshold of five press releases in a given year to represent substantial coverage to a right.

<sup>37</sup> It is important to point out that our results for sharpness in particular are conservative, as noted in the Online Appendix.

## COMPARISONS ACROSS MONITORING AGENCIES AND IMPLICIT VERSUS EXPLICIT TAXONOMIES

Beyond demonstrating the ways in which changing information communication technologies have altered the taxonomies of implicit human rights being judged in texts over the last 40 years, our tools can help researchers probe the relative coverage of rights across monitoring agencies as well as compare explicit versus implicit coverage.

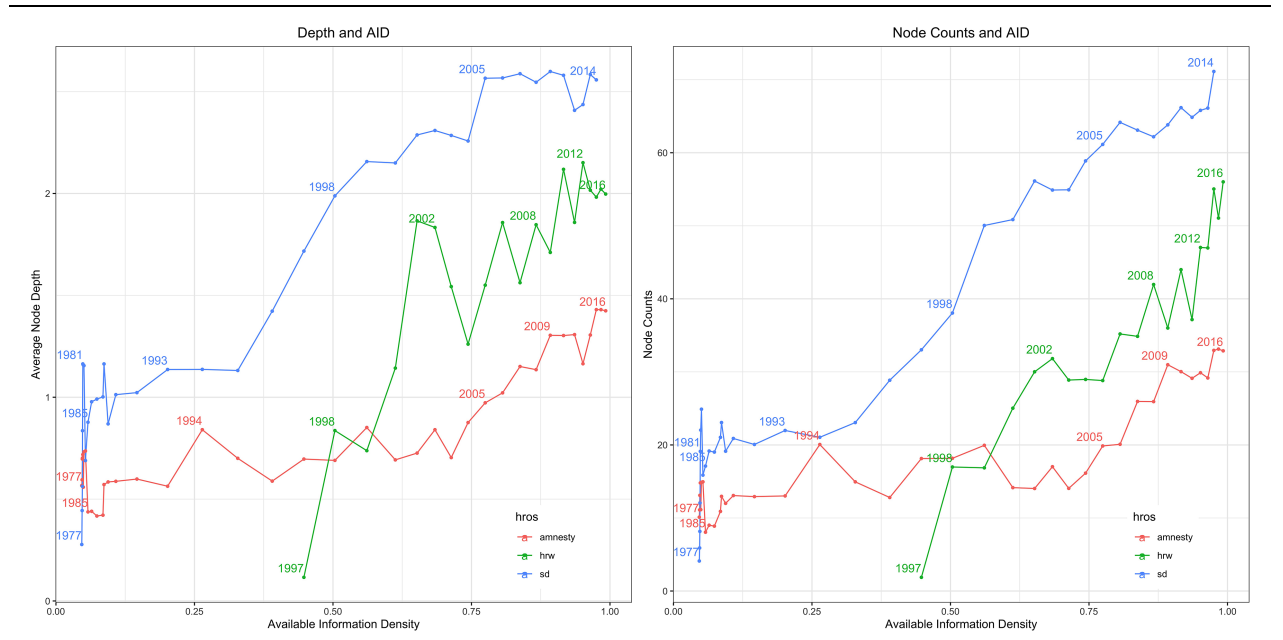
### An Example of Comparing Rights Coverage across Monitoring Agencies: Iran, 2014

According to the *Yearbook of International Organizations*, there are more than a thousand international human rights organizations. Existing studies on HROs find that because of greatly different institutional, country-level backgrounds, resources, and clout, these legacies lead to differences in investigating and documenting states' human rights practices and how they publish and release the collected information (Murdie 2014b; Stroup 2012; Stroup and Wong 2017).

Our approach allows researchers to systematically discover, for the first time, which aspects of human rights each HRO pays attention to in a given country report. As one example, Figure 11 shows differences in the proportion of attention to specific rights within the State Department (x-axis) versus Amnesty International (y-axis) texts for the Republic of Iran in 2014.<sup>38</sup> Amnesty pays a significant proportion of their

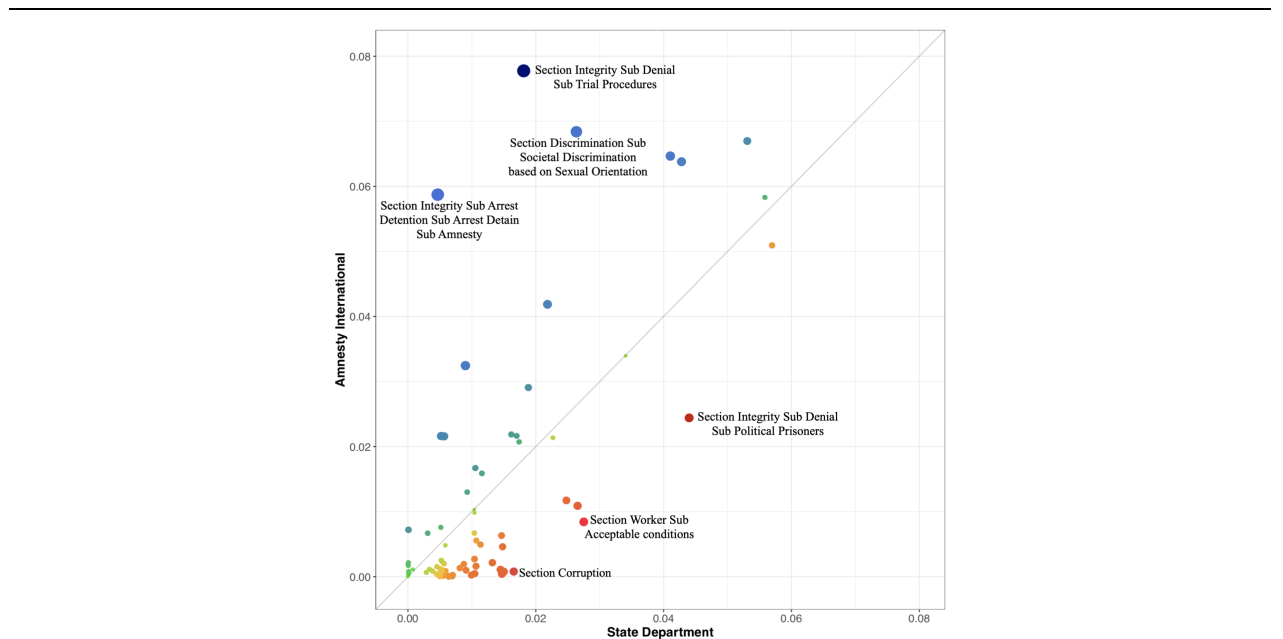
<sup>38</sup> We normalize by the number of paragraphs in each country report because, overall, the State Department Reports are much longer than are the Amnesty International reports, thus using absolute attention results in all of the rights receiving more attention in the State Department country report.

**FIGURE 10. Comparing the Taxonomic Structure of the State Department, Amnesty International, and Human Rights Watch Corpora**



Note: Left: the x-axis is available information density (AID) over time, and the y-axis is the implicit depth of human rights for State Department (sd), Amnesty International (amnesty), and Human Rights Watch (hrw). Right: the x-axis is the same as the left figure (AID), and the y-axis is the implicit number of nodes for these three monitoring agencies.

**FIGURE 11. Amnesty International and State Department: Iran**



Note: The x-axis and y-axis represent the expected proportion of paragraphs for all human rights in the 2015/2016 taxonomy for Iran in 2014 from the State Department and Amnesty International, respectively. If both sources pay the same proportion of attention to a given human rights aspect, then they would be on the diagonal line. The size and color are keyed to the signed difference in proportion, with Amnesty as blue and the State Department as red.

attention to the right to a fair trial and the administration and implementation of justice. In contrast, the State Department focused more of their attention on the issues of political prisoners or detainees as well as corruption. The tools we provide are available for researchers to explore other country reports and test new hypotheses about differences in attention over time, without manually coding all of the underlying documents.

### The Mismatch between Implicit and Explicit Taxonomies

A use case for our tools and data that goes beyond testing for information effects in text is to allow researchers to compare the explicit rights that are signaled in the section headings of country reports versus our estimated implicit rights aspects that are discussed in the text. We are uniquely able to identify a paragraph that is highly likely to semantically signal content on physical integrity rights but appears explicitly in the section on political rights. One can imagine this comparison as an asymmetric confusion matrix, where each observation is the text of a paragraph. In later years, the explicit taxonomy nearly matches each node in  $G_{2015, 2016}$ ; however, this does not mean that all of the text falls under the explicit section wherein it was labeled in the reports. We build a graphical tool that illuminates whether, for example, physical integrity rights judgments, often an important focus of human coding, can be found outside of that specific section. This has particularly important implications for the creation of human rights data because the codebooks for the Cingranelli and Richards (2010) Human Rights Data Project and the Political Terror Scale instruct coders to search only specific sections of the reports for codable information on given human rights violations, thus creating the potential to miss meaningful information. At the same time, as the reports become increasingly longer, having humans closely read the reports is not necessarily practical. Therefore, our modeling approach could be used to guide reading across sections so relevant information is not missed. This can be used to improve both the accuracy and efficiency of the human coding of human rights violations and protections.<sup>39</sup> More conceptually, our tools can illustrate semantic connections between rights across nodes of the taxonomy. Some implicit rights concepts might co-occur with other explicit section labels systematically. Looking only at explicit taxonomies would not be able to uncover these cross-taxon connections

Based on the model that was trained on the 2015–2016 data ( $G_{2015, 2016}$ ), we create an asymmetric confusion matrix that displays the explicit aspect labels on the  $x$ -axis and implicit aspect labels on the  $y$ -axis. The observations within the confusion matrix are the individual paragraphs from the text in that year. As with

<sup>39</sup> A similar approach, combining human and machine coding, is being used by the Sub-National Analysis of Repression Project (SNARP) team. See <http://snarpdata.org>.

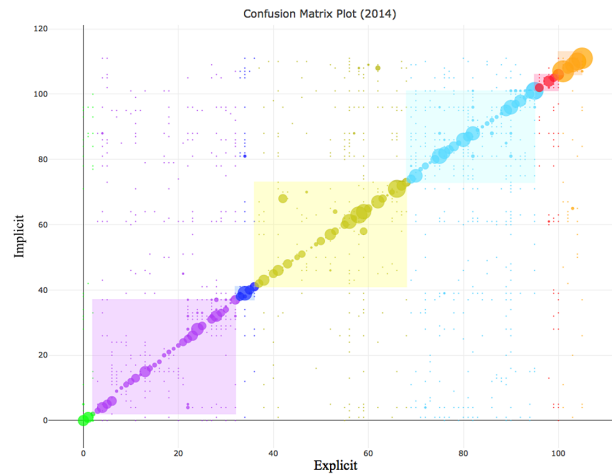
traditional confusion matrices, agreement between the implicit label (predicted label) and the explicit label (actual label) are found on the diagonal, while disagreements are found off the diagonal. An added benefit of our approach can be seen in the colored rectangles, each representing one of the seven main sections.<sup>40</sup> Looking for points off the diagonal (but within a colored area) can provide insights into semantic similarities between human rights aspects. In particular, this suggests that there are two distinct aspects of human rights within the same section that are conceptually similar.

Here, we present the asymmetric confusion matrix for the year 2014 ( $G_{2014}$ ). Figure 12 demonstrates how our approach can be used to explore where the implicit and explicit section labels for a given paragraph are congruent. In general, this information can be found by looking for paragraphs that fall on the diagonal. Within the application, a coder can simply mouse over a point to retrieve the explicit and implicit labels as well as the full text of the paragraph. Figure 13 presents a paragraph that is clearly about “*Libel and Slander Laws*,” and it categorized as such by both the implicit and explicit taxonomies.

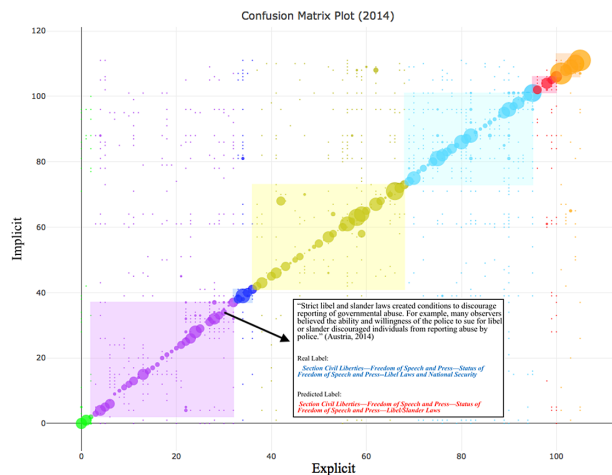
In Figure 14, we present the asymmetric confusion matrix for the year 1977 ( $G_{1977}$ ). The  $x$ -axis is much shorter than the  $y$ -axis is because there were only nine explicit section labels for the year 1977, while the  $y$ -axis displays our model’s classification of each paragraph based on the 112 implicit section labels in the 2015–2016 data ( $G_{2015, 2016}$ ). The plot displays an example of disagreement between the implicit and explicit labels. The explicit label for the text is “*Political Participation*” within the “*Civil Liberties*” section. However, rather than being about political participation broadly, this paragraph (text) is about the more specific political participation of women, which our model identifies. We are able to classify this text as being about “*Participation of Women and Minorities*,” despite the fact that this specific section did not exist in the explicit taxonomy until 2011.<sup>41</sup> This provides some evidence that information on more specific rights is contained in the reports even as far back as 1977. Rather than attempting to read the entirety of the report, using this approach, human coders can focus their attention to paragraphs that fall off the diagonal. By using machine learning to indicate possible missed information and then having this information verified by human coders, we may be able to build more efficient and accurate coding of human rights (Colaresi and Mahmood 2017).

<sup>40</sup> The green boxed area denotes “Governmental Attitude Regarding International and Nongovernmental Investigation of Alleged Violations of Human Rights”; the purple, “Respect for Civil Liberties”; the blue, “Corruption and Lack of Transparency in Government”; the yellow, “Discrimination, Societal Abuses, and Trafficking in Persons”; the light blue, “Respect for the Integrity of the Person”; the red, “Freedom to Participate in the Political Process”; and the orange, “Worker Rights”.

<sup>41</sup> Our model shows systematic implicit coverage of this right beginning in 1996.

**FIGURE 12. Comparison of Implicit and Explicit Locations of Text Within the State Department Reports**

Note: The x-axis and y-axis refer to the explicit aspect labels (actual labels) and the implicit aspect labels (predicted labels) respectively. The observations within the confusion matrix are the individual paragraphs from the text in that year. Agreement between the implicit label and the explicit label are found on the diagonal, while disagreements are found off the diagonal. The colored rectangles represent the seven sections in the State Department Reports.

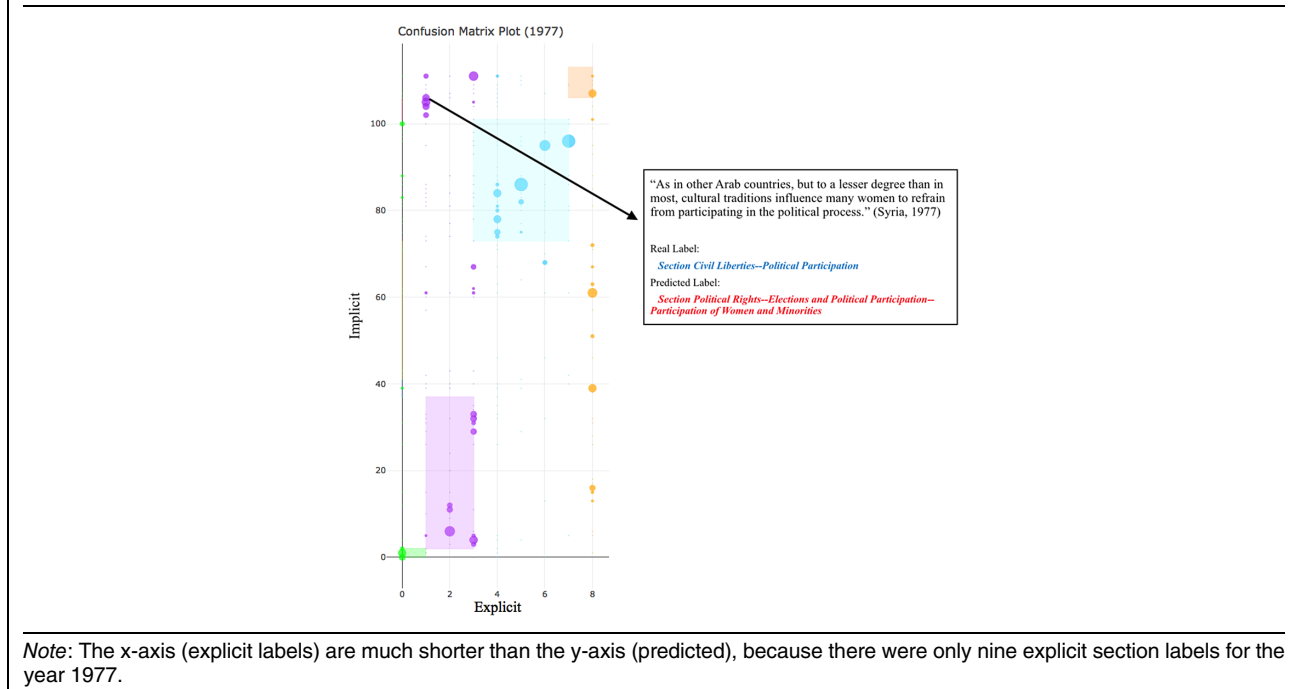
**FIGURE 13. Comparison of Implicit and Explicit Locations of Text Within the State Reports with an Example Showing Agreement Between the Implicit and Explicit Section Labels (2014 vs 2015/2016)**

## CONCLUSION: WHERE DO WE GO FROM HERE?

Our analysis is consistent with the explanation that increased access to denser information over time—spurred by developments in open-source satellite imagery, access to social media, and the internet as well as the spread of smart phones—has led to changes in the taxonomy of human rights being judged in later as opposed to earlier reports. The plausibility of information-driven taxonomic evolution—across its structure, attention to particular rights, and the

sharpness of the conceptual distinctions in the text—is crucial to moving the debate about the measurement of human rights forward (Cingranelli and Filippov 2018a; Fariss 2019b). If information effects have led to judgments on different and potentially increasingly specific aspects in human rights over time, then any human or machine coding would need to account for those changes in the rights being judged. Simultaneously, comparisons across time need to be identified so that empirical information in the text of reports, for example, can uniquely update inferences (Fariss 2019b). Carefully tracking taxonomic changes is likely to produce improvements in both the

**FIGURE 14. Comparison of Implicit and Explicit Locations of Text Within the State Reports with an Example Showing Disagreement Between the Implicit and Explicit Section Labels (1977 vs 2015/2016)**



validity and the reliability of downstream human rights scores and inferences.<sup>42</sup>

As Fariss (2019b) suggests, the validity concerns are clearest if we cast the measurement of human rights into an item-response theory framework (Clinton, Jackman, and Rivers 2004; Martin and Quinn 2002). Countries are analogous to respondents, each with an ability parameter or vector that would represent their latent protection of human rights. The human rights monitor sets “the test” for the respondent. To date, research has treated reports as though they were only asking a limited, consistent number of questions. For example, the Political Terror Scale (Gibney et al. 2015) focuses on *Physical Integrity Rights* and grades the monitors’ text of a country’s answer, similarly to grading an essay, on a five-point scale. Likewise, the Cingranelli and Richards (2010) human rights data grade human rights as though there were 15 questions, each being answered on a three-point scale.

The conventional approach to fixing human rights measurement over time assumes that the rights being judged in country reports across the years have remained static. This can be clearly seen in recent debates (Cingranelli and Filippov 2018a; Fariss 2019b), where the axis of contention revolves around the usefulness of time-varying difficulty parameters on a given right (or set of rights).<sup>43</sup> The use of human-coded scores and external event-based measures to estimate

evolving difficulty parameters ignores the possibility that the taxonomy of human rights that are monitored across time is also evolving. It may be not *only* the difficulty of a given right that is changing over time but also the differences in sets/taxonomies of rights that are judged in recent versus past reports.

In our reconceptualization of the data-generation process, each right that is being judged is an item or question, formulated by the monitor. The correctness of each country’s behavior related to a specific right defines whether there is a violation (wrong answer) or a protection (correct answer), perhaps with a signal of the intensity of the violation/protection, translated into language such as “widespread” or “systematic.” These written judgments, on specific items, represent the “grades” of responses in Item Response Theory (IRT) modeling. Our work highlights that judgments are not always offered on one general fixed right, like a right to physical integrity, but that there are different items noted across sections. Different subsections and paragraphs can speak to distinct rights within a larger taxonomy of human rights concepts.

Ignoring the taxonomy of rights in a given set of country reports is akin to assuming that all rights have the same difficulty.<sup>44</sup> However, if behavior is consistent across time but later taxonomies have more difficult items included within them, we would have observed worse judgments in later years, and not the consistent pattern in figure 1. Further, an analyst ignoring the differences across taxonomies, but only paying

<sup>42</sup> One important implication is that if rights that have been judged consistently in texts over time can be identified, then we can use this information to bridge inferences across years.

<sup>43</sup> This discussion leads to the related point about how to uniquely identify the set of parameters.

<sup>44</sup> Or, for a model of fixed items but varying difficulty parameters, where the difficulties for each right changes on average at the same rate each year.

attention to judgment terms, would invalidly score countries as worse when it was the items that were changing. The way to correct this validity issue is to use tools like ours to track the taxonomies of rights that are operable in human rights reports over time and identify consistently scored rights.<sup>45</sup>

Understanding the complexity and nested structure of human rights taxonomies over time also offers solutions to reliability concerns. Rights that share parents, for example “*Extrajudicial Killings*” and “*Torture*,” are likely to have similar judgments, all else being equal. Thus, taxonomic information supplies a hierarchical structure that can be used to adaptively shrink estimates back towards a common value (Gelman et al. 2013). Additionally, our open-source tools, when fed the same information, will output the same rights labels, which will not always be the case for human-coded rights taxonomies.<sup>46</sup>

Our research suggests that newly available dense information on human interactions raises not only privacy issues but also measurement concerns that are broadly applicable. The influence of social media, satellite imagery, internet adoption, and smart phones has changed the density of available information that is encoded in variety of texts, including human rights reports, but potentially also news reports, government cables, and other nongovernmental and intergovernmental organizations’ media. Thus, similar forms of information effects, with deepening attention to sharper taxonomies, are likely to be part of the data-generation process<sup>47</sup> across varied domains in the social sciences including event data, protests, international events, and the counting of battle-related deaths. However, we have demonstrated that changing information environments can leave important characteristic fingerprints within texts. These patterns can be detected and used to explore and correct for ICT changes over time.

## SUPPLEMENTARY MATERIALS

To view supplementary material for this article, please visit <http://dx.doi.org/10.1017/S0003055420000258>.

Replication materials can be found on Dataverse at: <https://doi.org/10.7910/DVN/CCGPUQ>.

## REFERENCES

- Bagozzi, Benjamin E., and Daniel Berliner. 2018. “The Politics of Scrutiny in Human Rights Monitoring: Evidence from Structural Topic Models of U.S. State Department Human Rights Reports.” *Political Science Research and Methods* 6 (4): 661–77.
- <sup>45</sup> These are often denoted as bridging observations. If we ignored the changing taxonomy of items, counting the uses of quantifiers and intensifiers like “many,” “frequent,” and “widespread” may be giving a false impression, as they are modifying different aspects in recent versus older reports.
- <sup>46</sup> This is akin to one of the benefits of automated event-data-coding systems, which are discussed in King and Lowe (2003)
- <sup>47</sup> See Weidmann (2015) for a careful analysis of the effect of information on measurement.
- Barner, David, and Andrew Scott Baron. 2016. *Core Knowledge and Conceptual Change*. Oxford: Oxford University Press.
- Baum, Matthew A., and Yuri M. Zhukov. 2015. “Filtering Revolution.” *Journal of Peace Research* 52 (3): 384–400.
- Bob, Clifford. 2002. “Globalization and the Social Construction of Human Rights Campaigns.” In *Globalization and Human Rights*, ed. Alison Brysk. University of California Press, Oakland, CA: 133–47.
- Carey, Sabine C., Michael P. Colaresi, and Neil J. Mitchell. 2015. “Governments, Informal Links to Militias, and Accountability.” *Journal of Conflict Resolution* 59 (5): 850–76.
- Carpenter, R. Charli. 2007. “Setting the Advocacy Agenda: Theorizing Issue Emergence and Nonemergence in Transnational Advocacy Networks.” *International Studies Quarterly* 51 (1): 99–120.
- Carpenter, Charli. 2014. *Lost Causes: Agenda Vetting in Global Issue Networks and the Shaping of Human Security*. Ithaca, NY: Cornell University Press.
- Cingranelli, David, and David L. Richards. 2010. “The Cingranelli and Richards (CIRI) Human Rights Data Project.” *Human Rights Quarterly* 32 (2): 401–24.
- Cingranelli, David, and Mikhail Filippov. 2018a. “Are Human Rights Practices Improving?” *American Political Science Review* 112 (4): 1083–9.
- Cingranelli, David, and Mikhail Filippov. 2018b. “Problems of Model Specification and Improper Data Extrapolation.” *British Journal of Political Science* 48 (1): 273–4.
- Clark, Ann Marie, and Kathryn Sikkink. 2013. “Information Effects and Human Rights Data: Is the Good News about Increased Human Rights Information Bad News for Human Rights Measures?” *Human Rights Quarterly* 35 (3): 539–68.
- Clinton, Joshua, Simon Jackman, and Douglas Rivers. 2004. “The Statistical Analysis of Roll Call Data.” *American Political Science Review* 98 (2): 355–70.
- Colaresi, Michael. 2014. *Democracy Declassified: The Secrecy Dilemma in World Politics*. New York: Oxford University Press.
- Colaresi, Michael, and Zuhair Mahmood. 2017. “Do the robot: Lessons from machine learning to improve conflict forecasting.” *Journal of Peace Research* 54 (2): 193–214.
- Cordell, Rebecca, K. Chad Clay, Christopher J. Fariss, Reed M. Wood, and Thorin M. Wright. 2020. “Changing Standards or Political Whim? Evaluating Changes in the Content of the US State Department Human Rights Reports.” *Journal of Human Rights* 19 (1): 3–18.
- DeCanio, Stephen J., and William E. Watkins. 1989. “Information Processing and Organizational Structure.” *Journal of Economic Behavior and Organization* 36 (2): 275–94.
- Fariss, Christopher J. 2014. “Respect for Human Rights Has Improved over Time: Modeling the Changing Standard of Accountability.” *American Political Science Review* 108 (2): 297–318.
- Fariss, Christopher. 2019a. “Supplementary Appendix: Yes, Human Rights Practices Are Improving over Time.” Available on <https://doi.org/10.1017/S000305541900025X>
- Fariss, Christopher. 2019b. “Yes, Human Rights Practices Are Improving over Time.” *American Political Science Review* 113 (3): 868–81.
- Finnemore, Martha, and Kathryn Sikkink. 1998. “International Norm Dynamics and Political Change.” *International Organization* 52 (4): 887–917.
- Florini, Ann. 1996. “The Evolution of International Norms.” *International Studies Quarterly* 40 (3): 363–89.
- Forsythe, David P. 2009. *Encyclopedia of Human Rights*. Vol. 1. Oxford: Oxford University Press.
- Forsythe, David P. 2017. *Human Rights in International Relations*. Cambridge, UK: Cambridge University Press.
- Gelman, Andrew, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. 2013. *Bayesian Data Analysis*. Boca Raton, FL: CRC press.
- Gibney, Mark, Linda Cornett, Reed Wood, Peter Haschke, and Daniel Arnon. 2015. “The Political Terror Scale 1976–2015.” West Lafayette, Indiana: Purdue University.
- Gillies, James M., and Robert Cailliau. 2000. *How the Web Was Born: The Story of the World Wide Web*. Oxford: Oxford University Press.
- Greene, Kevin, Baekkwon Park, and Michael Colaresi. 2018. “Machine Learning Human Rights and Wrongs: How the Successes

- and Failures of Supervised Learning Algorithms Can Inform the Debate About Information Effect.” *Political Analysis* 27: 223–30.
- Keck, Margaret E., and Kathryn Sikkink. 1998. “Transnational Advocacy Networks in the Movement Society.” In *The Social Movement Society: Contentious Politics for a New Century*, eds. David S. Meyer and Sidney Tarrow. Roman & Littlefield, 217–38.
- King, Gary, and Will Lowe. 2003. “An Automated Information Extraction Tool for International Conflict Data with Performance as Good as Human Coders: A Rare Events Evaluation Design.” *International Organization* 57 (3): 617–42.
- Kotsiantis, S. B. 2007. “Supervised Machine Learning: A Review of Classification Techniques.” *Informatica* 31 (1): 249–68.
- Landman, Todd. 2006. *Studying Human Rights*. Sussex, UK: Psychology Press.
- Levitsky, Steven, and Daniel Ziblatt. 2018. *How Democracies Die*. Crown.
- Liu, Bing. 2012. “Sentiment Analysis and Opinion Mining.” *Synthesis Lectures on Human Language Technologies* 5 (1): 1–167.
- Liu, Bing. 2015. *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge: Cambridge University Press.
- Martin, Andrew D., and Kevin M. Quinn. 2002. “Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the US Supreme Court, 1953–1999.” *Political Analysis* 10 (2): 134–53.
- Murdie, Amanda. 2014a. *Help or Harm: The Human Security Effects of International NGOs*. Redwood City, California: Stanford University Press.
- Murdie, Amanda. 2014b. “The Ties That Bind: A Network Analysis of Human Rights International Nongovernmental Organizations.” *British Journal of Political Science* 44 (1): 1–27.
- Park, Baekwan, Michael Colaresi, and Kevin Greene. 2018. “Beyond a Bag of Words: Using PULSAR to Extract Judgments on Specific Human Rights at Scale.” *Peace Economics, Peace Science and Public Policy* 24 (4).
- Poe, Steven C., Sabine C. Carey, and Tanya C. Vazquez. 2001. “How Are These Pictures Different? A Quantitative Comparison of The US State Department and Amnesty International Human Rights Reports, 1976–1995.” *Human Rights Quarterly* 23 (3): 650–77.
- Richards, David L. 2016. “The Myth of Information Effects in Human Rights Data: Response to Ann Marie Clark and Kathryn Sikkink.” *Human Rights Quarterly* 38 (2): 477–92.
- Shannon, C. E. 1948. “A Mathematical Theory of Communication.” *Bell System Technical Journal* 27: 379–423, 623–656, July and October.
- Sikkink, Kathryn. 2011. *The Justice Cascade: How Human Rights Prosecutions Are Changing World Politics*. New York: W. W. Norton.
- Stroup, Sarah S. 2012. *Borders among Activists: International NGOs in the United States, Britain, and France*. Ithaca, NY: Cornell University Press.
- Stroup, Sarah S., and Wendy H Wong. 2017. *The Authority Trap: Strategic Choices of International NGOs*. Ithaca, NY: Cornell University Press.
- Weidmann, Nils B. 2015. “On the Accuracy of Media-based Conflict Event Data.” *Journal of Conflict Resolution* 59 (6): 1129–49.
- Wilson, Maya, David R. Davis, and Amanda Murdie. 2016. “The View from the Bottom: Networks of Conflict Resolution Organizations and International Peace.” *Journal of Peace Research* 53 (3): 442–58.
- Wood, Reed M., and Mark Gibney. 2010. “The Political Terror Scale (PTS): A Re-Introduction and a Comparison to CIRI.” *Human Rights Quarterly* 32 (2): 367–400.